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

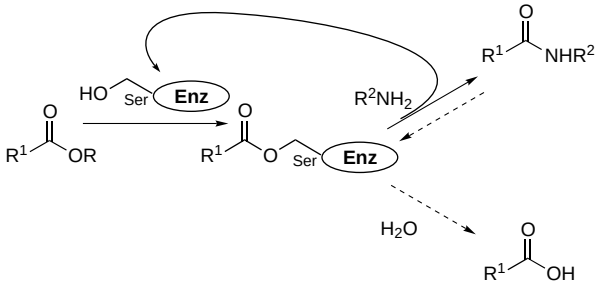
Kazutsugu Matsumoto,[b] Benjamin G. Davis,*[a] and J. Bryan Jones[c]

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 = oxazolidinones, alkylammonium groups, and carbohydrates at position 166 at the base of the primary specificity S_1 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 glycinamide 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-GlyNH_2 using SBL-WT → 74% using S166C-S-$(CH_2)_2$ NMe_3^+) 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-GlyNH_2 and 86% yield of Z-D-Ala-GlyNH_2 using S166C-S-(3*R*,4*S*)-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).



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-

[a] Dr. B. G. Davis
Dyson Perrins Laboratory, Department of Chemistry
University of Oxford
South Parks Road, Oxford, OX1 3QY (UK)
Fax: (+44) 1865-275674
E-mail: ben.davis@chem.ox.ac.uk

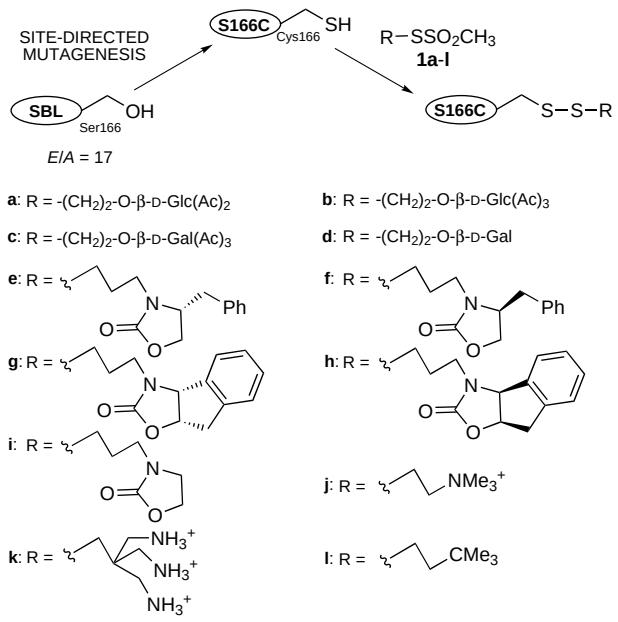
[b] Dr. K. Matsumoto
Department of Chemistry
College of Science & Technology
Meisei University
Hodokubo 2-1-1, Hino, Tokyo 191-8506 (Japan)

[c] Prof. J. B. Jones
Department of Chemistry
University of Toronto
80 St. George St., Toronto, ON, M5S 3H6 (Canada)

 © 2002 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim 0947-6539/02/0818-4129 $ 20.00+.50/0

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.[4] We have previously exploited this method to improve enzyme activity and alter substrate specificity of the subtilisin from *Bacillus lentus* (SBL, EC 3.4.21.62 formerly EC 3.4.21.14).[5, 6] 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.



Scheme 2.

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.[7] Alternatively, elegant substrate engineering or mimcry[8–12] 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[13, 14] or forced evolution[15–17] 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 lentus* (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[18] 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.[19]

Using this technique polar groups intended to influence the specificity of SBL were selected and introduced to position 166[20] at the base of the primary specificity determining S_1[21] pocket (Figure 1) to create S166C-CMMs.[22] These were systematically surveyed for their synthetic utility in representative and challenging peptide ligation reactions. Key amongst these were reactions that SBL-WT cannot perform, for example, ligations of D-amino acids as acyl donors and ligations of α-branched amino acid L-alanine.

Primarily, groups were chosen that might reasonably be expected to influence the stereochemical P_1 specificity of such ligations. Carbohydrates[23] and oxazolidinones[24] are used widely as effective chiral auxiliaries and we wished to probe the effect of introducing the homochiral groups **a**–**h** 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[25] 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 S_1 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 S_1 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 S_1 pocket which modulates P_1 substrate specificity (Figure 1) and is therefore the pocket that determines acyl donor selectivity in amino acid ester peptide

© 2002 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
0947-6539/02/0818-4130 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 18

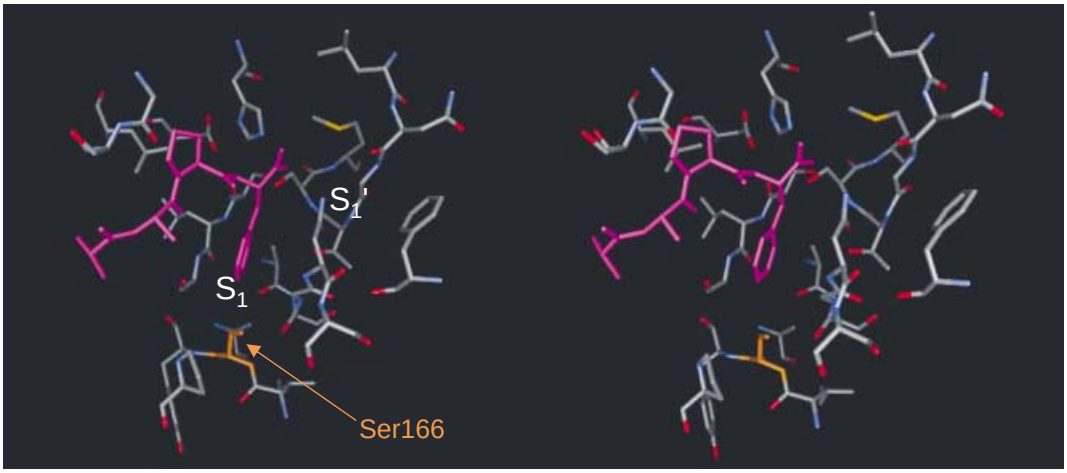


Figure 1. Stereoview of the active site region of SBL-WT [RCSB *1jea*] indicating the location of the S166 residue (orange) at the base of the S_1 primary specificity determining pocket. The binding mode of substrates is illustrated by the substrate mimic AAPF (magenta).

ligations. Initial aspects of this work have been described in a previous communication.[26]

In short, carbohydrate methanethiosulfonates were synthesized[27] from the parent carbohydrates D-glucose (Glc) and D-galactose (Gal) and used to chemoselectively modify[28] the thiol side chain in the position 166 Ser → Cys mutant SBL-S166C (Scheme 2). In all cases these modification reactions in aqueous buffer were rapid and quantitative, as judged by monitoring of changes in specific activity and by titration of free thiols with Ellman's reagent.[29] The glyco-CMMs were purified by size-exclusion chromatography and dialysis, and their structures were confirmed by rigorous TOF-ES-MS analyses (±4 Da). These CMMs each appeared as a single band on non-denaturing gradient PAGE, thereby establishing their high purities. The active enzyme concentration of the resulting CMM solutions (typically 5 mg mL^{-1}) was determined by active site titration with α-toluenesulfonyl fluoride (PMSF) using a fluoride ion-sensitive electrode.[30] These revealed typical protein yields of >80%. Modification with the fully deprotected galactose reagent **1d** led to site-specific glycosylation at position 166 and the formation of a single glycoform: S166C-S-**d**. Furthermore, modification with acetyl protected glyco-MTS reagents gave products with controllable levels of acetylation. Through careful adjustment of pH and appropriate selection of the glycosylation site, we were able to prepare the partially acetylated glycoforms of SBL: S166C-S-**a-c**. The effects of these modifications upon SBL were assessed by the determination of k_{cat}/K_M for the hydrolysis at pH 8.6 of the amide substrate succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-*p*NA), as a measure of amidase activity[28] $A = (k_{cat}/K_M)_{\text{amidase}}$ and the ester substrate succinyl-Ala-Ala-Pro-Phe-S-benzyl (Suc-AAPF-SBn), as a measure of esterase activity[31] $E = (k_{cat}/K_M)_{\text{esterase}}$. These kinetic parameters are compared with those of SBL-WT in Table 1. Gratifyingly, in all cases glycosylation resulted in enhanced E/A values relative to unglycosylated SBL-WT (Table 1). By virtue of their higher esterase and lower amidase activities ($E/A = 38-75$) as compared to the values for SBL-WT (E/A = 17), such glyco-CMMs S166C-S-**a-d** are excellent candidates for efficient amide bond formation as they possess enhanced acylating properties and yet substantially reduced hydrolytic activity towards the peptide products of coupling. It should also be noted that such was the convenience of this method that these CMMs were prepared on up to a gram scale.

Use of carbohydrate-modified CMMs in peptide ligations: Firstly, we probed the structural breadth of the P_1 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 glycinamide (**8**) as an acyl acceptor (Table 2). These reactions were carried out by using 0.5 mg[32] 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-

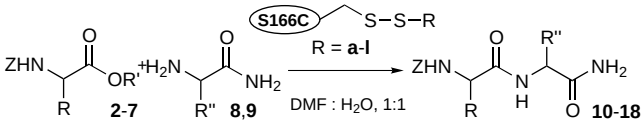
Table 1. Kinetic parameters of WT and CMMs of SBL.

	Enzyme												
Kinetic parameter	SBL-WT	S166C-S-**a**	**-b**	**-c**	**-d**	**-e**	**-f**	**-g**	**-h**	**-i**	**-j**	**-k**	**-l**
$(k_{cat}/K_M)_{\text{esterase}} = E$[a] [s^{-1}mM^{-1}]	3593	5058	4264	6033	3241	1495	3277	1488	4556	4898	2400	1800	4900
$(k_{cat}/K_M)_{\text{amidase}} = A$[b] [s^{-1}mM^{-1}]	209	109	112	81	58	22	52	20	47	75	23	8	14
E/A	17	46	38	75	56	68	63	74	97	65	104	225	350

[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 suc-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, suc-AAPF-*p*NA as the substrate.

 © 2002 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim 0947-6539/02/0818-4131 $ 20.00+.50/0

Table 2. Peptide coupling catalyzed by WT and CMMs of SBL.[a]



Acyl donor	Acyl acceptor	Product	Time [h]	SBL-WT	Isolated yield [%] with S166C-S-**a**	**-b**	**-c**	**-d**	**-e**	**-f**	**-g**	**-h**	**-i**	**-j**	**-k**	**-l**
Z-L-PheOBn (**2**)	$GlyNH_2 \cdot HCl$ (**8**)	Z-L-PheGlyNH_2 (**10**)	1	92	95	93	91	95	86	88	82	100	92	74	75	92
Z-L-AlaOBn (**3**)	**8**	Z-L-AlaGlyNH_2 (**11**)	5	91	85	77	92	83	87	88	91	95	82	99	91	88
Z-L-GluOMe (**4**)	**8**	Z-L-GluGlyNH_2 (**12**)	5	62	58	65	54	67	60	54	68	56	67	63	71	64
2	L-Ala$NH_2 \cdot HCl$ (**9**)	Z-L-Phe-L-AlaNH_2 (**13**)	24[b]	57	28	34	31	32	31	30	33	37	50	44	36	42
3	**9**	Z-L-Ala-L-AlaNH_2 (**14**)	24[b]	0	15	16	22	11	12	19	21	20	10	14	11	0
4	**9**	Z-L-Glu-L-AlaNH_2 (**15**)	24[b]	0	48	50	51	55	60	59	61	59	64	58	60	0
Z-D-PheOBn (**5**)	**8**	Z-D-Phe-GlyNH_2 (**16**)	48[c]	0	6	8	7	8	8	12	7	14	9	4	4	5
Z-D-AlaOBn (**6**)	**8**	Z-D-Ala-GlyNH_2 (**17**)	48[c]	0	80	77	72	70	88	80	88	80	61	79	73	38
Z-D-GluOBn (**7**)	**8**	Z-D-Glu-GlyNH_2 (**18**)	48[c]	0	63	62	64	64	62	60	62	52	64	74	64	8

[a] Reactions were performed in DMF/water (1/1, v/v) at a concentration of 0.1 M acyl donor, 0.3 M acyl acceptor, and 0.3 M Et_3N 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, 0.2 M of **9** and 0.2 M of Et_3N were used. [c] After 24 h, 1 mg mL^{-1} of active enzyme was added and the mixture was stirred for another 24 h.

Gly-NH_2 (**10**, 91–95%), Z-L-Ala-Gly-NH_2 (**11**, 77–92%), Z-L-Glu-Gly-NH_2 (**12**, 54–67%) compared well, and in some cases were superior to, those for SBL-WT (92, 91, 62% for **10**, **11**, **12**, respectively).

Our next goal was to investigate any secondary effects of these modifications in the S_1 pocket upon the P_1' specificity that is controlled by the neighbouring S_1' acyl acceptor pocket. For this purpose, we used the α-branched L-alaninamide (**9**) as a challenging acyl acceptor probe. Remarkably, in spite of the usually small S_1' pocket (Figure 1) all of the CMMs S166C-S-**a**–**d** were catalysts for the coupling of L-amino acids **1**–**3** with α-branched acyl acceptor, L-alaninamide (**9**). The strict specificity of the SBL-WT enzyme is demonstrated by its ability to catalyze the coupling of **9** as an acyl acceptor with only preferred acyl donor **2**. However, glyco-CMMs S166C-S-**a**–**d** not only catalyzed the coupling of **9** with **2** to afford slightly reduced yields of Z-L-Phe-L-AlaNH_2 (**13**), but also with **3** and **4** to afford the corresponding dipeptides Z-L-Ala-L-AlaNH_2 (**14**), and Z-L-Glu-L-AlaNH_2 (**15**), respectively, in greatly enhanced yields. In the synthesis of **14**, lower absolute yields than for **13** and **15** were observed (the highest yield of 22% was obtained by using S166C-S-**c**). However, even this, in comparison with the absence of coupling observed for SBL-WT, represents an immeasurable relative enhancement of S_1' specificity. Excitingly, the yields of the coupling of **4** with **9** to give **15** were as good as those obtained using unhindered **8** as an acyl acceptor (48–55%), and here the use of S166C-S-**d** gave the best result (55%).

Finally, we focused on the extension of the use of these powerful glyco-CMMs S166C-S-**a**–**d** to the coupling of D-amino acids as acyl donors by using Z-D-PheOBn (**5**), Z-D-AlaOBn(**6**), and Z-D-GluOBn (**7**) with acyl acceptor **8**. The enzyme-catalyzed coupling of D-amino acids at the N-terminus of oligopeptide products by using D-amino acid acyl donors has been achieved only rarely and typically with low efficiencies.[33, 34] For example, to the best of knowledge, yields above 10% for the coupling of D-Glu have never been successfully achieved.[35] Typical difficulties encountered with the use of unmodified L-amino acid specific amidases with standard substrates are highlighted by examples where even with a 40-fold excess of acyl acceptor only 31% yields of D-Ala coupling were observed.[36] One notable exception has been the impressive use by Bordusa and co-workers of substrate mimicry.[37] Even specifically isolated D-amino acid specific amidases do not allow the ready use of D-amino acid donors to form dipeptides in synthesis.[38–41] Consistent with these difficulties, SBL-WT did not accept D-amino acids as acyl donors and starting materials **5**–**7** were recovered.

Excitingly, all of S166C-S-**a**–**d** were catalysts for the synthesis of **16**–**18**. Although, the reactions of **5** in all cases were slow to give Z-D-PheGlyNH_2 (**15**) in low yield (the best was 8% by using S166C-S-**b**,**d**). Starting material **5** was recovered in all cases indicating that this is due to a low rate of conversion and not competitive hydrolysis. Peptide couplings of **6** and **7** with **8** proceeded more rapidly and no starting materials remained. Clearly, the yields of products are dramatically improved compared with SBL-WT. In particular, good yields of **18** (62–64%) using S166C-S-**a**–**d** for coupling of D-Glu are in some cases superior to those found for coupling of L-Glu, with the same catalysts. High yields were also observed for the synthesis of D-Ala containing substrates; for example, S166C-S-**a** gave Z-D-AlaGlyNH_2 (**17**) in 80% yield.

The construction of CMMs bearing oxazolidinones as chiral auxiliaries in their binding pockets, S166C-S-e–h: In the light of these early exciting indications obtained through the introduction of homochiral substituents to the S_1 binding pocket of SBL, we wished to probe the nature of this useful broadening in stereoselectivity. Interestingly, both the introduction of group **b** and its C-4 epimer substituent **c** had similarly allowed broadening but in some cases to quite different extents e.g. 92% and 77% yield of **11** from S166C-S-**c** and -**b**, respectively. However, to establish the significance of

© 2002 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
0947-6539/02/0818-4132 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 18

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. L-galactose and L-glucose, which would be required for the synthesis of the enantiomers of **1a–d**, are not readily available chirons and we chose instead to create two enantiomeric pairs of substituents based on the well-established oxazolidinone auxiliaries.[24, 42] S166C was modified with **1e–h**[42] 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-**e** cf. -**f** and S166C-S-**g** cf. -**h**) in the hydrolysis of chiral substrates suc-AAPF-*p*NA and suc-AAPF-SBn and this indicated the possibility of pair mismatching with L-Phe-containing substrates.[42] For example, $(k_{cat}/K_M)_{esterase}$ 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 P_1 specificity of S166C-S-**e–h** by examining the ligation of L-amino acids, Z-L-PheOBn (**2**), Z-L-AlaOBn (**3**), and Z-L-GluOMe (**4**) as acyl acceptors, with glycinamide (**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 $(k_{cat}/K_M)_{esterase}$ values of S166C-S-**g** and -**h** with substrates containing L-Phe as the P_1 substituent (*see above*, Table 1), were partially reflected in the yields of Z-L-PheGly-NH_2 (**10**) obtained from their use as ligation catalysts: 82% using S166C-S-**g** yet 100% from S166C-S-**h** under identical conditions. Moreover, converse yield differences were observed with these same two catalysts in the synthesis of Z-L-GluGly-NH_2 (**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 S_1 pocket upon P_1' specificity in ligations of L-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 L-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 D-amino acid-ligating 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-D-AlaGly-NH_2 (**17**) in very good yields of 88%, and are very close to the best reported yields for any enzyme catalyzed C-terminal D-amino acid ligation.[37] 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 L-amino acids with **8**, these ligations of D-amino acids with **8** showed some variations according to the configuration of the oxazolidinonyl substituent. For example, S166C-S-**h** bearing the *S*,*R*-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 noticeable 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-D-AlaGlyNH_2 (**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 S166-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 -$SCH_2CH_2NMe_3^+$ and triply charged -$SCH_2C(CH_2NH_3^+)_3$, respectively were constructed[5h] to examine the effect of simply introducing different levels of charge into the S_1 pocket. Finally, non-polar neohexyl side chain -$SCH_2CH_2CMe_3$ (-**l**), which is near isosteric to side chain -**j**, was introduced to create S166C-S-**l** 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).

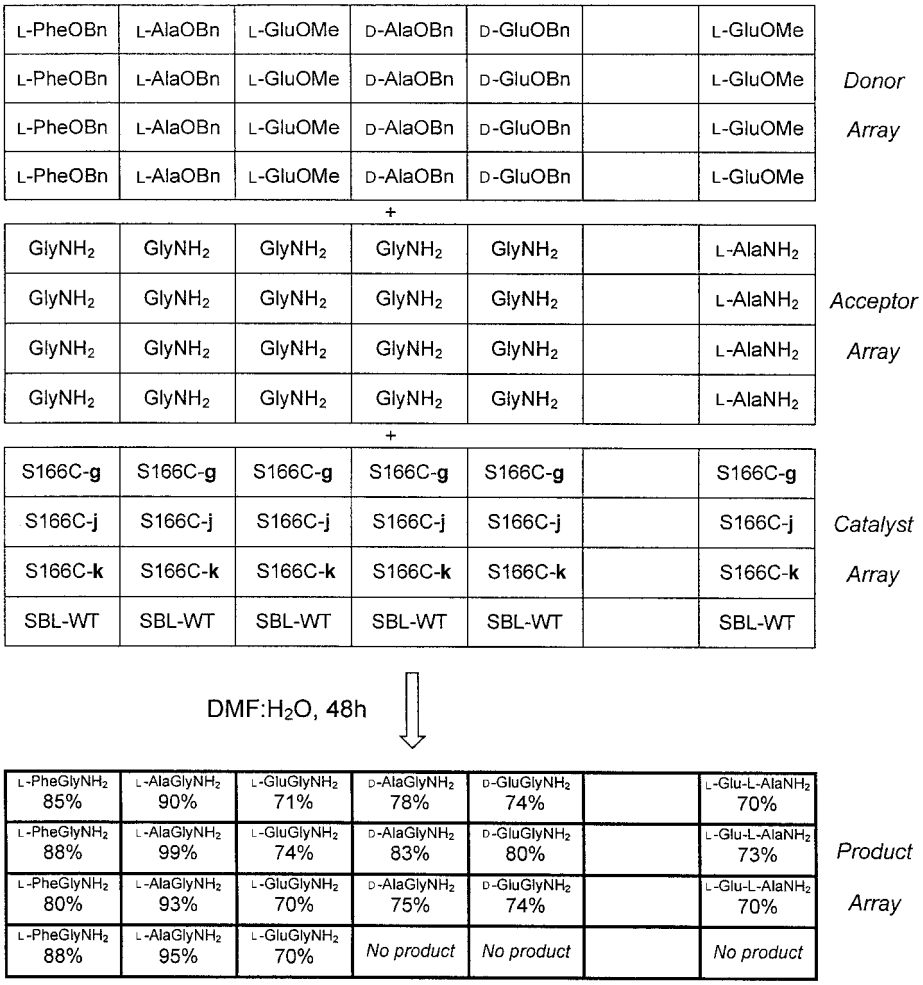
Use of CMMs modified with polar achiral groups in peptide ligations: Although, the use of CMM S166C-S-**i**, bearing an achiral oxazolidinonyl side chain at position 166, as a catalyst resulted in slight reductions in yield as compared with those with a chiral oxazolidinonyl side chain CMMs S166C-S-**e–h** 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$(CH_2)_2NMe_3^+$ (-**j**) and S166C-$SCH_2C(CH_2NH_3^+)_3$ (-**k**). For example, with S166C-S-**j**,**k** as catalysts, hydrophobic P_1 substrate **2** gave distinctly lower yields (by 17–18%) of **10**

 © 2002 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim 0947-6539/02/0818-4133 $ 20.00+.50/0

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 S_1 pocket, gave the highest yield in the coupling of **4**, which bears a negatively charged P_1 glutamate side chain, to give **12** in 71% yield. These opposing effects suggest the creation of a more highly charged S_1 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 +1 to +3 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-**l** with its less polar S_1 pocket behaves more like SBL-WT than broad-specificity polar CMMs S166C-S-**a**–**k**.



Scheme 3.

The use of CMMs in trial library syntheses: Valuably, the greatly broadened specificity of these CMMs and therefore their potential as combinatorial biocatalysts[2, 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 multiwell 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 S_1 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- over D-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-

© 2002 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
0947-6539/02/0818-4134 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 18

peptide formation"[40] for L-amino acid specific peptidases—these results show that this is no longer true. The yields obtained here are vastly superior to any other D-amino acid ligations from our laboratories and to the best of our knowledge the yields of **17** and **18** described here are the best yet obtained in enzyme-catalyzed peptide ligations of D-Glu acyl donors and very close to the best[37] for D-Ala acyl donors, *including those catalyzed by D-amino acid specific enzymes.*[38–40]

Remote effects of these position 166 S_1 modifications upon the neighbouring S_1' 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**–**l** 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 S_1' pocket. From inspection of Figure 1 it can be seen that residues 154–156 in the wall of the S_1 pocket encroach on the wall of the S_1' pocket. Any bulging or distortion in this wall might therefore restrict the S_1' pocket. This is consistent with the observed drop in yield of **13** (Table 2) as the size of the position 166 side chain increases: $-CH_2OH$ (57%) $\rightarrow -CH_2S$-**i**,**j**,**l** (42–50%) $\rightarrow -CH_2S$-**a**–**h**,**k** (28–37%).

We have also established that a CMM bearing a non-polar modification, that is, S166C-S-**l**, 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 S_1 pocket or even disrupt its role by exposing it to solvent: perhaps effectively turning the 152-167 loop of SBL that forms the base of the S_1 pocket inside-out or at least imparting additional flexibility.[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 P_1 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-AAPE*p*NA and a CMM S166C-S-$(CH_2)_2NH_3^+$ resulted in an 19-fold increase in k_{cat}/K_M over SBL-WT[5f] and hydrogen bonding interactions between substrate and the D-glucosyl residue of a CMM L217C-S-$Glc(Ac)_3$ resulted in an 8.4-fold increase in esterase activity.[31] 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 D-amino acids in a different mode from L-amino acids, that is, the α-carbobenzoxyamine group (NHZ) of may bind in the S_1 pocket in place of the amino acid side chain. If this is the case, repulsion between the benzyl group of **5**, which has the largest substituent among the three D-amino acid substrates **5**–**7**, and other groups in the active 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 requre. 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 D-amino acids as acyl donors and an α-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 represents 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 ^{13}C NMR spectra were measured on a Varian Unity 400 (400 MHz for ^1H and 100 MHz for ^{13}C) spectrometer in [D_6]DMSO with resdual 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 (Merck). 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 lentus* (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-**e-l**: 140 mM CHES, 2 mM $CaCl_2$, pH 9.5; for S166C-**a**,**d** 140 mM MES, 2 mM $CaCl_2$ pH 6.5; for S166C-**b**,**c** 140 mM MES, 2 mM $CaCl_2$). To this solution was added 100 µL of a 0.2 M MTS reagent solution for every 2.5 mL of the resulting enzyme solution (**1a**–**c**,**e**–**h**,**l** in MeCN, **1d**,**i**,**j** 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-AlaAlaProPhe-*p*-nitroanilide ($\varepsilon_{410} = 8800$ M^{-1} cm^{-1})[13] as substrate in 0.1 M Tris-HCl buffer containing 0.005% Tween 80, 1% DMSO, pH 8.6 showing constant activity and by titration with Ellman's reagent[29] ($\varepsilon_{412} = 13600$ M^{-1} cm^{-1}) showing no free thiol present in solution. A further solution (50 µL for every 2.5 mL 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 Quench Buffer (5 mM MES 1 mM $CaCl_2$, pH 5.5). The eluant was dialysed at 4 °C against 10 mM MES, 1 mM $CaCl_2$ pH 5.8 (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[29] in phosphate buffer 0.25 M, pH 8.0. In all cases no free thiol was detected. Modified enzymes were analyzed by nondenaturing gradient (8–25%) gels at pH 4.2, run towards the cathode, on the Pharmacia Phast-system and appeared as a single band. 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, 0.01% TFA as the running buffer and eluted with 80% acetonitrile, 0.01% TFA in a one step gradient. MS (ES-MS) *m/z*: S166C-S-**a**: calcd 27036, found 27040; S166C-S-**b**: calcd 27078, found 27078; S166C-S-**c**: calcd 27078, found 27079; S166C-S-**d**: calcd 26952, found 26950; S166C-S-**e**: calcd 26963, found 26967; S166C-S-**f**: calcd 26963, found 26964; S166C-S-**g** calcd 26961, found 26958; S166C-S-**h** calcd 26961, found 26962; S166C-S-**i** calcd 26846, found 26850; S166C-S-**j** calcd 26832, found 26835; S166C-S-**k** calcd 26861, found 26861; S166C-S-**l** calcd 26859, found 26855. The active enzyme concentration was determined as previously described[30] by monitoring fluoride release upon enzyme reaction with α-toluenesulfonyl fluoride (PMSF) as measured by a fluoride ion sensitive electrode (Orion Research 96–09).

Kinetic measurements: For amidase activity Michaelis-Menten constants were measured at 25(±0.2) °C by curve fitting (GraFit 3.03) of the initial rate data determined at nine concentrations (0.125 mM–3.0 mM) of succinyl-AAPF-*p*NA substrate in 0.1 M Tris–HCl buffer containing 0.005% Tween 80, 1% DMSO, pH 8.6 ($\varepsilon_{410} = 8800\ M^{-1}\,cm^{-1}$).[13] For esterase activity specificity constants determined using the low substrate approximation were measured indirectly using Ellman's reagent[29] ($\varepsilon_{412} = 13600\ M^{-1}cm^{-1}$) using 15 or 30 μM succinyl-AAPF-SBn as substrate in 0.1 M Tris.HCl, containing 0.005 vol% Tween 80, 1 vol% 37.5 mM Ellman's reagent in DMSO, pH 8.6. Selected kinetic data was also taken from references 5 h,28,31,42.

Synthesis of 3,3-dimethylbutyl methanethiosulfonate (11): A solution of 3,3-dimethylbutylbromide (1 g, 6.1 mmol) and $NaSSO_2CH_3$ (0.9 g, 6.7 mmol) in DMF (10 mL) was heated at 50 °C under N_2. After 50 h the solvent was removed and the resulting residue purified by flash chromatography (1:4, EtOAc/hexane) to give **11** as a colourless oil: IR (film): $\tilde{\nu}$ = 1280, 1115 cm^{-1} (S-SO_2); ^{1}H NMR ($CDCl_3$, 400 MHz): $\delta = 0.88$ (s, 9H; $C(CH_3)_3$), 1.46 (m, 2H; H-2), 3.48 (s, 3H; CH_3SO_2); ^{13}C NMR ($CDCl_3$, 100 MHz): $\delta = 27.0$, 28.4, 33.4, 46.1, 51.1 (CH_3SO_2); (ES-MS) *m/z* (%): 219 $[M+Na]^+$, (100), 197 $[M+H^+]$ (70); HRMS *m/z* (ES): Found 196.0588; $C_7H_{16}O_2S_2$ requires 196.0592.

Parallel array synthesis: Acyl donor (0.05 mmol), acyl acceptor (0.15 mmol) and enzyme (0.5 mg) were arrayed in DMF/water (1:1, 200 μL) as shown in Scheme 3 in 24 wells of a 96-well quartz plate and gently agitated for 48 h at room temperature. At this time the contents were analysed by HPLC-MS using a C-18 column (5:95–80:20 CH_3CN:H_2O gradient) on a Micromass ES-ZMD system employing a Waters 600 HPLC system with Waters 2700 autosampler.

General procedure for peptide coupling: To a solution of Z-L-PheOBn (**2**, 19.2 mg, 0.05 mmol) in DMF (0.25 mL) and water (0.144 mL), glycinamide hydrochloride (**8**, 17 mg, 0.15 mmol) and Et_3N (0.15 mmol, 0.0625 mL) were added, followed by addition of S166C-S-**a** (0.106 mL), 0.5 mg of active enzyme in MES buffer (10 mM MES, 1 mM $CaCl_2$, pH 5.8). The reaction mixture was stirred for 1 h at room temperature. The mixture was diluted with EtOAc and washed with 1 M $KHSO_4$ (1 mL) and brine (1 mL), and the organic layer was dried over $MgSO_4$. After evaporation, the residue was purified by preparative TLC (CH_2Cl_2/MeOH = 90:10) to afford Z-L-PheGlyNH_2 (**10**, 16.9 mg, 95%). Peptide ligations of other substrates using other enzymes were carried out following the same procedure for the reaction times indicated in Table 2. In the case of D-amino acids as acyl donors, 0.5 mg more active enzyme was added to the reaction vessel after 24 h, and then the mixture was stirred for another 24 h.

Z-L-PheGlyNH_2 (10): $[\alpha]_D^{21} = -3.9$ ($c = 1.04$ in MeOH) (literature value $[\alpha]_D^{20} = -3.2$ ($c = 1$ in MeOH)[46]); ^{1}H NMR ([D_6]DMSO): $\delta = 2.74$ (dd, J = 11.0, 14.0 Hz, 1H; C*H*H′Ph), 3.04 (dd, $J = 4.0$, 14.0 Hz, 1H; C H*H*′Ph), 3.59–3.72 (m, 2H; NHCH_2CO), 4.21–4.35 (m, 1H; NHC*H*CO), 4.93 (d, $J = 12.5$ Hz, 1H; OC*H*H′Ph), 4.94 (d, $J = 12.5$ Hz, 1H; OCH*H*′Ph), 7.12 (brs, 2H; NH), 7.16–7.38 (m, 10H; Ph), 7.60 (d, $J = 8.5$ Hz, 1H; NH), 8.27 ppm (t, $J = 5.5$ Hz, 1H; NH); ^{13}C NMR ([D_6]DMSO): $\delta = 37.3$, 42.0, 56.3, 65.3, 126.3, 127.5, 127.8, 128.1, 128.4, 129.3, 137.0, 138.2, 156.0, 170.8, 171.8 ppm; HRMS (FAB+): calcd for $C_{19}H_{22}N_3O_4$ $[M + H]^+$ 356.1610, found 356.1639.

Z-L-AlaGlyNH_2 (11): $[\alpha]_D^{25} = -8.4$ ($c = 0.64$ in MeOH) (literature values $[\alpha]_D^{23} = -7.5$ ($c = 2$ in MeOH); [47] $[\alpha]_D^{25} = -6.3$ ($c = 0.80$ in MeOH)[48]); ^{1}H NMR ([D_6]DMSO): $\delta = 1.20$ (d, $J = 7.0$ Hz, 3H; CH_3), 3.60 (dd, $J = 5.5$, 16.0 Hz, 1H; C*H*H′NH), 3.62 (dd, $J = 5.5$, 16.0 Hz, 1H; CH*H*′NH), 4.03 (dq, $J = 7.0$, 7.0 Hz, 1H; CH_3C*H*NH), 5.00 (d, $J = 12.5$ Hz, 1H; OC*H*H′Ph), 5.03 (d, $J = 12.5$ Hz, 1H; OCH*H*′Ph), 7.11 (brs, 1H; NH_2), 7.18 (brs, 1H; NH_2), 7.27–7.42 (m, 5H; Ph), 7.57 (d, $J = 7.0$ Hz, 1H; NH), 8.11 ppm (t, J = 5.5 Hz, 1H; NH); ^{13}C NMR ([D_6]DMSO): $\delta = 17.9$, 42.0, 50.3, 65.5, 127.85, 127.89, 128.4, 136.9, 155.9, 170.9, 172.7 ppm; HRMS (FAB$^+$) calcd for $C_{13}H_{18}N_3O_4$ $[M + H]^+$ 280.1297, found 280.1307.

Z-L-GluGlyNH_2 (12): $[\alpha]_D^{25} = -9.3$ ($c = 0.69$ in MeOH) (literature value $[\alpha]_D^{25} = -10.2$ ($c = 1.0$ in MeOH)[6b]); ^{1}H NMR ([D_6]DMSO): $\delta = 1.66–1.79$ (m, 1H; C*H*H′CH_2COOH), 1.83–1.95 (m, 1H; CH*H*′CH_2COOH), 2.26 (t, $J = 7.5$ Hz, 2H; CH_2COOH), 3.62 (d, $J = 5.5$ Hz, 2H; NHCH_2CO), 3.95–4.05 (m, 1H; NHC*H*CO), 5.01 (d, $J = 12.5$ Hz, 1H; OC*H*H′Ph), 5.03 (d, J = 12.5 Hz, 1H; OCH*H*′Ph), 7.07 (brs, 1H; NH_2), 7.20 (brs, 1H; NH_2), 7.25–7.40 (m, 5H; Ph), 7.55 (d, $J = 7.5$ Hz, 1H; NH), 8.11 (t, $J = 5.5$ Hz, 1H; NH), 12.20 ppm (brs, 1H; COOH); ^{13}C NMR ([D_6]DMSO): $\delta = 27.0$, 30.2, 41.9, 54.1, 65.6, 127.8, 127.9, 128.4, 136.9, 156.2, 170.8, 171.7, 174.0 ppm; HRMS (FAB$^+$) calcd for $C_{15}H_{20}N_3O_6$ $[M + H]^+$ 338.1352, found 338.1364.

Z-L-Phe-L-AlaNH_2 (13): $[\alpha]_D^{24} = -9.2$ ($c = 0.56$ in MeOH) (literature value $[\alpha]_D^{29} = -8.86$ ($c = 0.57$ in MeOH)[6b]); ^{1}H NMR ([D_6]DMSO): $\delta = 1.22$ (d, $J = 7.0$ Hz, 3H; CH_3), 2.71 (dd, $J = 13.5$, 13.5 Hz, 1H; C*H*H′Ph), 3.03 (dd, $J = 3.5$, 13.5 Hz, 1H; CH*H*′Ph), 4.18–4.31 (m, 2H; NHC*H*CO × 2), 4.93 (s, 2H; OCH_2Ph), 7.04 (brs, 1H; NH), 7.14–7.22 (m, 1H; NH), 7.24–7.42 (m, 10H; Ph), 7.55 (d, $J = 8.5$ Hz, 1H; NH), 8.08 ppm (d, $J = 7.5$ Hz, 1H; NH); ^{13}C NMR ([D_6]DMSO): $\delta = 18.5$, 37.4, 48.1, 56.2, 65.2, 126.3, 127.4, 127.7, 128.1, 128.4, 129.3, 137.1, 138.2, 155.9, 171.1, 174.1 ppm; HRMS (FAB$^+$) calcd for $C_{20}H_{24}N_3O_4$ $[M + H]^+$ 370.1767, found 370.1769.

Z-L-Ala-L-AlaNH_2 (14): $[\alpha]_D^{21} = -21.8$ ($c = 0.89$ in MeOH) (literature values $[\alpha]_D^{23} = +8.30$ ($c = 5$ in DMF);[49] $[\alpha]_D^{25} = -20.4$ ($c = 0.77$ in MeOH); [6b] $[\alpha]_D^{25} = -29.1$ ($c = 0.4$ in MeOH)[50]); ^{1}H NMR ([D_6]DMSO): $\delta = 1.19$ (d, $J = 7.0$ Hz, 3H; CH_3), 1.24 (d, $J = 7.5$ Hz, 3H; CH_3), 3.90–4.26 (m, 2H; NHC*H*CO × 2), 5.01 (s, 2H; CH_2OPh), 7.02 (brs, 1H; NH_2), 7.13 (brs, 1H; NH_2), 7.25–7.45 (m, 5H; Ph), 7.51 (d, $J = 6.5$ Hz, 1H; NH), 7.88 ppm (d, J = 7.5 Hz, 1H; NH); ^{13}C NMR ([D_6]DMSO): $\delta = 18.1$, 18.5, 47.9, 50.2, 65.4, 127.78, 127.84, 128.4, 137.1, 155.8, 172.0, 174.2 ppm; HRMS (FAB$^+$) calcd for $C_{14}H_{20}N_3O_4$ $[M + H]^+$ 294.1454, found 294.1457.

Z-L-Glu-L-AlaNH_2 (15): $[\alpha]_D^{25} = -15.9$ ($c = 0.68$ in MeOH) (literature value $[\alpha]_D^{25} = -16.7$ ($c = 0.76$ in MeOH)[6b]); ^{1}H NMR ([D_6]DMSO): δ = 1.20 (d, $J = 8.0$ Hz, 3H; CH_3), 1.68–1.82 (m, 1H; C*H*H′CH_2COOH), 1.82–2.03 (m, 1H; CH*H*′CH_2COOH), 2.21–2.40 (m, 2H; CH_2CH_2COOH), 3.93–4.25 (m, 2H; NHC*H*CO × 2), 5.02 (s, 2H; OCH_2Ph), 7.02 (brs, 1H; N*H*H′), 7.18–7.46 (m, 5H; Ph), 7.51 (brs, 1H; NH*H*′), 7.54 (d, $J = 7.5$ Hz, 1H; NH), 7.92 (d, $J = 7.5$ Hz, 1H; NH), 12.40 ppm (brs, 1H; COOH); ^{13}C NMR ([D_6]DMSO): $\delta = 18.4$, 26.2, 30.3, 48.0, 53.1, 65.4, 127.0, 127.7, 129.3, 137.0, 156.2, 173.7, 173.8, 174.1 ppm; HRMS (FAB$^+$) calcd for $C_{16}H_{22}N_3O_6$ $[M + H]^+$ 352.1509, found 352.1502.

Z-D-PheGlyNH_2 (16): $[\alpha]_D^{21} = +3.4$ ($c = 1.17$ in MeOH); ^{1}H and ^{13}C NMR spectral data identical to **10**. HRMS calcd for $C_{19}H_{22}N_3O_4$ $[M + H]^+$ 356.1610, found 356.1592.

Z-D-AlaGlyNH_2 (17): $[\alpha]_D^{24} = +8.5$ ($c = 0.86$ in MeOH) (literature value $[\alpha]_D^{25} = +9.3$ ($c = 1.1$ in MeOH)[50]); ^{1}H and ^{13}C NMR spectral data identical to **11**. HRMS calcd for $C_{13}H_{18}N_3O_4$ $[M + H]^+$ 280.1297, found 280.1303.

Z-D-GluGlyNH_2 (18): $[\alpha]_D^{24} = +9.1$ ($c = 1.08$ in MeOH); ^{1}H and ^{13}C NMR spectral data identical to **12**. HRMS calcd for $C_{15}H_{20}N_3O_6$ $[M + H]^+$ 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 Rick Bott and Kanjai Khumtaveeporn for helpful discussions. We also thank the

 © 2002 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim 0947-6539/02/0818-4136 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 18

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.

[1] a) C.-H. Wong, G. M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Pergamon Press, Oxford, **1994**, pp 41—130; b) K. Faber, *Biotransformations in Organic Chemistry*, 3rd ed., Springer, Heidelberg, **1999**, and references therein.

[2] P. C. Michels, Y. L. Khmelnitsky, J. S. Dordick, D. S. Clark, *Trends Biotechnol.* **1998**, *16*, 210.

[3] J. L. Krstenansky, Y. Khmelnitsky, *Bioorg. Med. Chem.* **1999**, *7*, 2157.

[4] a) G. DeSantis, J. B. Jones, *Curr. Opin. Biotechnol.* **1999**, *10*, 324; b) D. Qi, C.-M. Tann, D. Haring, M. D. Distefano, *Chem. Rev.* **2001**, *101*, 3081; c) C.-M. Tann, D. Qi, M. D. Distefano, *Curr. Opin. Chem. Biol.* **2001**, *5*, 696.

[5] a) M. R.Stabile, G. W. Lai, G. DeSantis, M. Gold, J. B. Jones, C. Mitchinson, R. R. Bott, T. P. Graycar, C.-C. Liu, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2501; b) P. Berglund, M. R. Stabile, M. Gold, J. B. Jones, C. Mitchinson, R. R. Bott, T. P. Graycar, *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2507; c) P. Berglund, G. DeSantis, M. R. Stabile, X. Shang, M. Gold, R. R. Bott, T. P. Graycar, T. H. Lau, C. Mitchinson, J. B. Jones, *J. Am. Chem. Soc.* **1997**, *119*, 5265; d) G. DeSantis, P. Berglund, M. R. Stabile, M. Gold, J. B. Jones, *Biochemistry* **1998**, *37*, 5968; e) G. DeSantis, J. B. Jones, *Bioorg. Med. Chem.* **1999**, *7*, 1381; f) G. DeSantis, X. Shang, J. B. Jones, *Biochemistry* **1999**, *38*, 13391; g) B. G. Davis, X. Shang, G. DeSantis, R. R. Bott, J. B. Jones, *Bioorg. Med. Chem.* **1999**, *7*, 2293; h) B. G. Davis, K. Khumtaveeporn, R. R. Bott, J. B. Jones, *Bioorg. Med. Chem.* **1999**, *7*, 2303; i) E. Plettner, G. DeSantis, M. R. Stabile, J. B. Jones, *J. Am. Chem. Soc.* **1999**, *121*, 4977; j) G. DeSantis, C. Paech, J. B. Jones, *Bioorg. Med. Chem.* **2000**, *8*, 563.

[6] For application of CMMs in synthesis, see: a) M. Dickman, R. C. Lloyd, J. B. Jones, *Tetrahedron: Asymmetry* **1998**, *9*, 4099; b) K. Khumtaveeporn, G. DeSantis, J. B. Jones, *Tetrahedron: Asymmetry* **1999**, *10*, 2563; c) A. Ullman, K. Matsumoto, B. G. Davis, J. B. Jones, *Tetrahedron: Asymmetry* **2001**, *12* , 249.

[7] V. Gotor, *Bioorg. Med. Chem.* **1999**, *7*, 2189.

[8] V. Schellenberger, U. Schellenberger, H. D. Jakubke, N. P. Zapevalova, Y. V. Mitin, *Biotechnol. Bioeng.* **1991**, *38*, 319.

[9] F. Bordusa, D. Ullmann, C. Elsner, H. D. Jakubke, *Angew. Chem.* **1997**, *109*, 2583; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2473.

[10] R. Gunther, F. Bordusa, *Chem. Eur. J.* **2000**, *6*, 463.

[11] M. Thormann, S. Thust, H-J. Hofmann, F. Bordusa, *Biochemistry* **1999**, *38*, 6056.

[12] R. Günther, S. Thust, J-J. Hofmann, F. Bordusa, *Eur. J. Biochem.* **2000**, *267*, 3496.

[13] P. R. Bonneau, T.P Graycar, D. A. Estell, J. B. Jones, *J. Am. Chem. Soc.* **1991**, *113*, 1026.

[14] D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom, J. A. Wells, *Science* **1994**, *266*, 243.

[15] K. Q. Chen, F. H. Arnold, *Proc. Natl Acad. Sci. USA* **1993**, *90*, 5618.

[16] K. P. Hopfner, E. Kopetzki, G. B. Kresse, W. Bode, R. Huber, R. A. Engh, *Proc. Natl Acad. Sci. USA* **1998**, *95*, 9813.

[17] J. E. Ness, M. Welch, L. Giver, M. Bueno, J. R. Cherry, T. V. Borchert, W. P. C. Stemmer, J. Minshull, *Nat. Biotechnol.* **1999**, *17*, 893.

[18] a) G. L. Kenyon, T. W. Bruce, *Methods Enzymol.* **1977**, *47*, 407; b) R. Wynn, F. M. Richards, *Methods Enzymol.* **1995**, *251*, 351.

[19] E. Plettner, K. Khumtaveeporn, X. Shang, J. B. Jones, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2291.

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

[21] Nomenclature according to: I. Schechter, A. Berger, *Biochem. Biophys. Res. Commun.* **1967**, *27*, 152.

[22] For 3D structure of SBL-WT see: M. Knapp, J. Daubermann, R. R. Bott, RCSB Entry *1jea*.

[23] H. Kunz, K. Ruck, *Angew. Chem.* **1993**, *105*, 355; *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 336.

[24] a) J. R. Gage, D. A. Evans, *Org. Synth.* **1990**, *68*, 77; b) D. J. Ager, I. Prakaash, D. R. Schaad, *Aldrichemica Acta* **1997**, *30*, 3.

[25] S. Masamune, W. Choy, J. S. Peterson, L. R. Sita, *Angew. Chem.* **1985**, *97*, 1; *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 1.

[26] K. Matsumoto, B. G. Davis, J. B. Jones, *Chem. Commun.* **2001**, 903.

[27] B. G. Davis, M. A. T. Maughan, M. P. Green, A. Ullman, J. B. Jones, *Tetrahedron: Asymmetry* **2000**, *11*, 245.

[28] B. G. Davis, R. C. Lloyd, J. B. Jones, *Bioorg. Med. Chem.* **2000**, *8*, 1527.

[29] G. L. Ellman, K. D. Courtney, V. Andres, R. M. Featherstone, *Biochem. Pharmacol.* **1961**, *7*, 88.

[30] C. Y. Hsia, G. Ganshaw, C. Paech, C. J. Murray, *Anal. Biochem.* **1996**, *242*, 221.

[31] R. C. Lloyd, B. G. Davis, J. B. Jones, *Bioorg. Med. Chem.* **2000**, *8,* 1537.

[32] This corresponds to about 0.036 mol% (3.6×10^{-5} M) of catalyst.

[33] a) A. L. Margolin, D.-F. Tai, A. M. Kilbanov, *J. Am. Chem. Soc.* **1987**, *109*, 7885; b) J. B. West, J. Scholten, N. J. Stolowich, J. L. Hogg, A. I. Scott, C.-H. Wong, *J. Am. Chem. Soc.* **1988**, *110*, 3709; c) C. F. Barbas, J. R. Matos, J. B. West, C.-H. Wong, *J. Am. Chem. Soc.* **1988**, *110*, 5162; d) N. Chinsky, A. L. Margolin, A. M. Klibanov, *J. Am. Chem. Soc.* **1989**, *111*, 386; e) X-Z. Zhang, X. Wang, S. Chen, X. Fu, X. Wu, C. Li, *Enzyme Microb. Technol.* **1996**, *19*, 538.

[34] For a review see: Y. Asano, T. L. Lübbehüsen, *J. Biosci. Bioeng.* **2000**, *89*, 295.

[35] These were in fact achieved using less efficient CMMs. See ref. [6b].

[36] K. Itoh, H. Sekizaki, E. Toyota, K. Tanizawa, *Bioorg. Chem.* **1997**, *25*, 307.

[37] Excellent yields of 50–95% (as determined by HPLC ± 5%) for the ligation of the acyl-4-guanidinophenyl esters of D-Phe, D-Leu and D-Ala using α-chymotrypsin have been obtained by using a 10-fold excess of acyl acceptor, see ref. [12].

[38] a) Y. Asano, A. Nakazawa, Y. Kato, K. Kondo, *Angew. Chem.* **1989**, *101*, 511; *Angew. Chem. Int. Ed. Engl.* **1989**, *28*, 450; b) Y. Kato, Y. Asano, A. Nakazawa, K. Kondo, *Tetrahedron* **1989**, *45*, 5743: while an amide was formed using a simple amine acyl acceptor, in these examples, the use of an amino acid acceptor was unsuccessful.

[39] A dipeptide donor Ac-L-Lys(Ac)-D-Ala-D-Lac-OH has been used in transpepdtidation to form tripeptides with the D-specific muramoyl-pentapeptide carboxypeptidase from *B. subtilis* in approximate yields of up to 40%: a) M. Kempe, B. Ekberg, K. Mosbach, *Int. J. Peptide Protein Res.* **1989**, *34*, 147; b) B. Ekberg, C. Lindbladh, M. Kempe, K. Mosbach, *Tetrahedron Lett.* **1989**, *30*, 583.

[40] The synthesis of D-Phe oligomers catalyzed by a D-amino acid transferase using D-Phe-OMe as an acyl donor has been detected by HPLC but no yields reported: Y. Watanabe, T. Muro, A. Sugihara, Y. Shimada, T. Nagao, S. Takenishi, Y. Tominaga, *Biochim. Biophys. Acta* **1997**, *1337*, 40. In a more recent study using a similar alkaline D-peptidase to catalyze the same reaction yields of up to 25% of D-Phe-D-Phe-OMe dimer were reported: H. Komeda, Y. Asano, *J. Mol. Catal. B: Enzym.* **1999**, *6*, 379.

[41] An amino acid esterase has very recently been used to used to catalyse the synthesis of Ac-D-Phe-D-Phe-NH_2, Ac-D-Phe-L-Phe-NH_2, Ac-D-Phe-D-Leu-NH_2 and Ac-D-Phe-L-Leu-NH_2 from Ac-D-Phe-OMe in yields of 5–21%: A. Sugihara, Y. Shimada, S. Sugihara, T. Nagao, Y. Watanabe, Y. Tominaga, *J. Biochem.* **2001**, *130*, 119.

[42] M. Dickman, J. B. Jones, *Bioorg. Med. Chem.* **2000**, *8*, 1957.

[43] For modelling and discussion of the role of active site side chain flexibility and its dependency on polarity see: D. S. Hartsough, K. M. Merz, *J. Am. Chem. Soc.* **1992**, *114*, 10113.

[44] P. P. Wangikar, T. P. Graycar, D. A. Estell, D. S. Clark, J. S. Dordick, *J. Am. Chem. Soc.* **1993**, *115*, 12231.

[45] K. Corbett, A. P. Fordham-Skelton, J. A. Gatehouse, B. G. Davis, *FEBS Lett.* **2001**, *509*, 355.

[46] J. Gante, U. Kalthof, F.-G. Klaerner, T. Weber, *Angew. Chem.* **1990**, *102*, 1081; *Angew. Chem. Int. Ed. Engl.* **1990**, *29*, 1025.

[47] P. A. Jaquenod, *Helv. Chim. Acta* **1965**, *48*, 1899.

[48] A. E. Sutton, J. Clardy, *Org. Lett.* **2000**, *2*, 319.

[49] R Katakai, M. Oya, F. Toda, K. Uno, Y. Iwakura, *Macromolecules* **1973**, *6*, 827.

[50] A. E. Sutton, J. Clardy, *J. Am. Chem. Soc.* **2001**, *123*, 9935.

Received: February 22, 2002
Revised: May 10, 2002 [F3896]