CHEMCOMM Communication
www.rsc.org/chemcomm

Glycosylation of the primary binding pocket of a subtilisin protease causes a remarkable broadening in stereospecificity in peptide synthesis

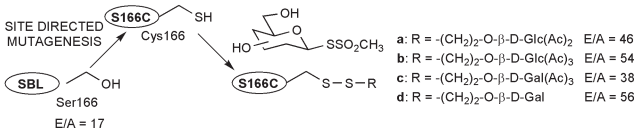
Kazutsugu Matsumoto,[a] Benjamin G. Davis*[b] and J. Bryan Jones*[c]

[a] *Department of Applied Chemistry and Biotechnology, Faculty of Engineering, Fukui University, Bunkyo 3-9-1, Fukui 910-8507, Japan*
[b] *Department of Chemistry, University of Durham, South Road, Durham, UK DH1 3LE. E-mail: Ben.Davis@durham.ac.uk*
[c] *Department of Chemistry, University of Toronto, 80 St. George St, Toronto, Ontario, Canada, M5S 3H6*

Received (in Liverpool, UK) 8th December 2000, Accepted 8th March 2001
First published as an Advance Article on the web 1st May 2001

Site-selective glycosylation at position 166 at the base of the primary specificity S_1 pocket in the serine protease subtilisin *Bacillus lentus* (SBL) created glycoproteins that are capable of catalyzing the coupling reactions of not only L- amino acid esters but also D-amino acid esters to give the corresponding dipeptides in good yields as a result of greatly broadened substrate specificities that can be rationalized by the interaction of the glycans acting as chiral auxiliaries in stereochemically mismatched pairs.

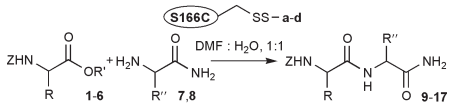
Enzymatic peptide coupling requires minimal protection of the substrate, proceeds under mild condition and without racemization.[1] In spite of these advantages, two major problems have limited the use of serine proteases in peptide synthesis. One is their efficient proteolytic (amidase) activity which causes hydrolysis of the coupled peptide product, and the other is their stringent structural and stereo-specificity which typically confines their use in synthesis to a limited range of L-amino acid substrates. Controlled site-selective glycosylation can be used to create glycosylated proteases with greatly improved esterase activities and enhanced esterase–amidase activity ratios (E/As) that are up to 8.4- and 17.2-fold enhanced over their unglycosylated wildtype (WT) counterparts, respectively.[2] By virtue of their higher esterase and lower amidase activities, these glycosylated enzymes 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. Moreover, these catalysts can be constructed on scales of hundreds of milligrams, thereby allowing their use in multigram scale syntheses. We reasoned that such glycosylated enzymes would not only be efficient catalysts for peptide ligation but that the presence of an internally bound carbohydrate might influence the stereospecificity of such ligations. Carbohydrates are effective chiral auxiliaries[3] and we wished to probe the effect of introducing the homochiral groups **a**–**d** (Scheme 1) into the primary binding region of an enzyme as a tactic for broadening substrate stereospecificity, perhaps by increasing the potential for stereochemically mismatched pairs.[4]



Scheme 1

Representative carbohydrates were attached to the interior of the primary S_1[5] binding pocket, which modulates P_1 specificity, of the serine protease subtilisin *Bacillus lentus* (SBL, EC 3.4.21.14) by site-selective glycosylation[2,6] at position 166. Gratifyingly, this also resulted in enhanced E/A values relative to unglycosylated SBL-WT (Scheme 1).[2]

Firstly, we probed the structural breadth of the P_1 specificity of S166C-S-**a**–**d** by examining the ligation of L-amino acids, (**1**)–(**3**) as acyl donors, with glycinamide (**7**) as an acyl acceptor (Scheme 2, Table 1)† using 1 mg ml^{-1} of active enzyme[7] in simple 1 : 1 water–DMF solutions. In accord with our goals of not reducing the existing substrate breadth of SBL, these results indicated that the introduction of carbohydrate groups **a**–**d** did not affect the essential, inherent ability of SBL to accept L-amino acids as acyl donors in peptide coupling reactions. Good



Scheme 2

Table 1 Glyco-SBL catalyzed peptide coupling reactions[a]

				Yield[b] (%) with				
Acyl donor	Acyl acceptor	Product	time/h	SBL-WT	S166C-**a**	S166C-**b**	S166C-**c**	S166C-**d**
Z-L-PheOBn (**1**)	$GlyNH_2$ (**7**)	*Z*-L-PheGlyNH_2 (**9**)	1	92	95	93	91	95
Z-L-AlaOBn (**2**)	**7**	*Z*-L-AlaGlyNH_2 (**10**)	5	91	85	77	92	83
Z-L-GluOMe(**3**)	**7**	*Z*-L-GluGlyNH_2 (**11**)	5	62	58	65	54	67
1	L-AlaNH_2 (**8**)	*Z*-L-Phe-L-AlaNH_2 (**12**)	24[c]	57	28	34	31	32
2	**8**	*Z*-L-Ala-L-AlaNH_2 (**13**)	24[c]	0	15	16	22	11
3	**8**	*Z*-L-Glu-L-AlaNH_2 (**14**)	24[c]	0	48	50	51	55
Z-D-PheOBn (**4**)	**7**	*Z*-D-PheGlyNH_2 (**15**)	48[d]	0	6	8	7	8
Z-D-AlaOBn (**5**)	**7**	*Z*-D-AlaGlyNH_2 (**16**)	48[d]	0	80	77	72	70
Z-D-GluOBn (**6**)	**7**	*Z*-D-GluGlyNH_2 (**17**)	48[d]	0	63	62	64	64

[a] DMF–water, 1 : 1, 0.1 M donor, 0.3 M acceptor, 0.036 mol% enzyme. Under the same conditions, spontaneous reaction did not occur. [b] Isolated yields. [c] 0.2 M acceptor. [d] Further 0.036 mol% added at 24 h.

DOI: 10.1039/b010021h

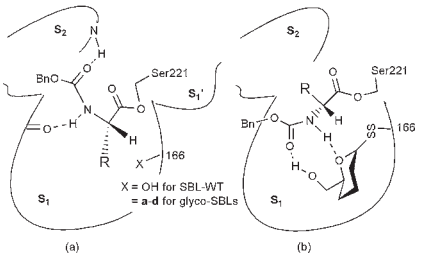


Fig. 1 Summarized representation of molecular mechanics analysis which showed different binding modes of (a) L-amino acids in SBL-WT or glyco-SBLs and (b) D-amino acids only in glyco-SBLs

to excellent yields of **9**, (91–95%), **10**, (77–92%), and fair yields of **11**, (54–67%) compared well, or were superior to, those for SBL-WT.

Our next goal was to investigate any secondary effects of S_1 pocket glycosylation upon P_1' specificity using the α-branched L-alaninamide (**8**) as a challenging acyl acceptor probe. The restricted S_1' specificity of the unglycosylated SBL-WT enzyme is demonstrated by its ability to catalyze only the coupling of **8** with the preferred L-phenylalaninate acyl donor **1**. Remarkably, in spite of the small S_1' pocket[8] all of the glyco-SBLs S166C-S-**a**–**d** were catalysts for the coupling of L-amino acids **1**–**3** with **8**. In the synthesis of **13**, low absolute yields were observed yet these represent a significant breakthrough in comparison with the absence of coupling activity prior to glycosylation. Excitingly, the yields of the coupling of **3** with **8** were similar to those obtained using unhindered **7** (48–55%). These results represent a dramatic improvement of the S_1' specificity of SBL. Interestingly, chemically-modified mutant enzymes bearing non-polar modifications at position 166 do not catalyze this ligation.[9] Possibly by exploiting polar interactions, we have obtained our goal of broadening substrate specificity, here of P_1' towards L-Ala to allow the synthesis of **12**–**14**, without diminishing the natural breadth of SBL. Indeed, previous molecular mechanics analysis of similarly internally glycosylated enzymes have indicated that such internally-bound carbohydrates may act as hydrogen-bonding motifs that give rise to enhanced kinetic parameters.[2]

Finally, we examined these powerful glyco-SBLs in the coupling of D-amino acids (**4**)–(**6**) as acyl donors with acyl acceptor **1**. The number of examples of enzyme catalyzed coupling of D-amino acids at the *C*-terminus of peptides by using D-amino acid acyl donors is vanishingly small and even then proceed with typically low efficiencies.[10] For example, to the best of our knowledge, yields above 10% for the incorporation of D-Glu have never been achieved.[9] Consistent with this, SBL-WT did not accept D-amino acids as acyl donors and starting materials **4**–**6** were recovered. In dramatic contrast, *all* of the glyco-SBLs S166C-S-**a**–**d** were able to catalyze the coupling of all three of these D-amino acid esters with acyl acceptor **1**. The reactions of **4** in all cases were slow and gave **15** in low yields, the best being 8% using S166C-S-**b**,**d**, and starting material **4** was recovered after 48 h in all cases. Peptide couplings of **5**, **6** with **7** proceeded more rapidly, and the yields were dramatically improved. In fact, the good yields of the D-dipeptide **17** (62–64%) were, surprisingly, superior to those found for coupling of L-Glu. Indeed, S166C-S-**c** showed a 1.2 : 1 stereochemical preference for D-glutamate **6** over L-glutamate **3**. Very high yields (up to 80%) of D-Ala dipeptide (**16**) were also observed.

Such is the remarkable nature of these broadened P_1 stereospecificities that we speculated that these glyco-SBLs may bind D-amino acids in a different mode from L-amino acids. Molecular mechanics analyses of SBL-WT and S166C-**a**–**d** with substrates **1**–**6** resulted in the models summarized in Fig. 1. In SBL-WT the normal binding mode for the acyl-enzyme intermediate, which is shown in Fig. 1(a), operates. The amino acid side chain (R) binds in the S_1 pocket with the reacting acyl group covalently linked to O_γ of the side chain of Ser221. This binding mode is not available to D-amino acids. In contrast, in the glyco-SBLs, Fig. 1(b), the presence of a homochiral auxiliary at position 166 with correct stereochemical disposition of its C-6 hydroxymethyl group and ring oxygen creates a hydrogen bond acceptor and donor motif which allows binding of the NHZ group in the S_1 pocket. Thus, glyco-SBLs may operate either by binding mode (a) *or* (b). However, in these glyco-SBLs, neither L-amino acids with the (a) mode nor D-amino acids with the (b) mode represent optimally matched diastereomeric pairs. This is consistent with their broadened specificity. By introducing a carbohydrate moiety at position 166 we have therefore tailored an enzyme with previously exclusive L-stereospecificity and excellent efficiency (matched pair) into an enzyme that has broadened L- and D-stereospecificity and reduced, but still good, efficiency due to mismatched pairing. This modified enzyme clearly has greater synthetic utility. This molecular mechanics analysis therefore raises the interesting possibility that the stereochemistry of the sidechain at 166 may influence stereospecificity through a double diastereoselective process. Further studies, including the synthesis of diastereomeric protein glycosylating reagents, to explore whether this is correct or whether location of just polar achiral groups at this position will have the same effect are now in progress. In addition, we are probing the precise effects of these modifications upon kinetic parameters k_{cat} and K_M using L- and D-amino acid substrates as both acceptors and donors in order to correlate the observed yield enhancements with true kinetic specificity.

In conclusion, glyco-SBLs S166C-S-**a**–**d** 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 ability of SBL to handle L-amino acids. It therefore represents a true broadening of synthetic utility and demonstrates that this site-selective glycosylation technology is a powerful tool for enhancing the application of enzymes in organic synthesis.

We acknowledge NSERC (Canada) and Genencor International, Inc 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 EPSRC (UK) for access to the Mass Spectrometry Service at Swansea and the Chemical Database Service at Daresbury.

References and notes

† Electronic supplementary information (ESI) available: detailed experimental procedures of peptide couplings and selected chemical data of the products. See http://www.rsc.org/suppdata/cc/b0/b010021m/

1 C.-H. Wong and G. M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Pergamon Press, Oxford, 1994, pp 41–130; K. Faber, *Biotransformations in Organic Chemistry*, 2nd ed. Springer-Verlag, Heidelberg, 1995, pp. 298–305 and references cited therein.
2 R. C. Lloyd, B. G. Davis and J. B. Jones, *Bioorg. Med. Chem.*, 2000, **8**, 1537.
3 H. Kunz and K. Ruck, *Angew. Chem., Int. Ed. Engl.*, 1993, **32**, 336.
4 S. Masamune, W. Choy, J. S. Peterson and L. R. Sita, *Angew. Chem., Int. Ed. Engl.*, 1985, **24**, 1.
5 Nomenclature according to I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun.*, 1967, **27**, 152.
6 B. G. Davis, R. C. Lloyd and J. B. Jones, *J. Org. Chem.*, 1998, **63**, 9614; B. G. Davis and J. B. Jones, *Synlett*, 1999, 1495; B. G. Davis, M. A. T. Maughan, M. P. Green, A. Ullman and J. B. Jones, *Tetrahedron: Asymmetry*, 2000, **11**, 245; B. G. Davis, R. C. Lloyd and J. B. Jones, *Bioorg. Med. Chem.*, 2000, **8**, 1527; B. G. Davis, *Chem. Commun.*, 2001, 351.
7 Determined by titration with PMSF: C. Y. Hsia, G. Granshaw, C. Paech and C. J. Murray, *Anal. Biochem.*, 1996, **242**, 221.
8 For 3D-structure of SBL-WT see M. Knapp, J. Daubermann and R. R. Bott, RCSB-PDB entry *1jea*.
9 K. Khumtaveeporn, G. DeSantis and J. B. Jones, *Tetrahedron: Asymmetry*, 1999, **10**, 2563.
10 Y. Asano and T. L. Lübbehüsen, *J. Biosci. Bioeng.*, 2000, **89**, 295.