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

Kazutugu Matsumoto, a Benjamin G. Davis* 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 S1 pocket in the serine protease subtilisin Bacillus lenterus (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 stereospecificities 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 wild-type (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 auxiliaries3 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

Representative carbohydrates were attached to the interior of the primary S1,5 binding pocket, which modulates P1 specificity, of the serine protease subtilisin Bacillus lenterus (SBL, EC 3.4.21.14) by site-selective glycosylation2° 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 P1 specificity of S166C-S-a-d by examining the ligation of L-amino acids, (1)-(3) as acyl donors, with glycaminide (7) as an acyl acceptor (Scheme 2, Table 1).† using 1 mg ml⁻¹ of active enzyme 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

Table 1 Glyco-SBL catalyzed peptide coupling reactions

<table>
<thead>
<tr>
<th>Acyl donor</th>
<th>Acyl acceptor</th>
<th>Product</th>
<th>Yield (%) with</th>
<th>time/h</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z-L-PheOBn (1)</td>
<td>GlyNH 2 (7)</td>
<td>Z-L-PheGlyNH 2 (9)</td>
<td>SBL-WT</td>
<td>92</td>
</tr>
<tr>
<td>Z-L-AlaOBn (2)</td>
<td>7</td>
<td>Z-L-AlaGlyNH 2 (10)</td>
<td>S166C-a</td>
<td>93</td>
</tr>
<tr>
<td>Z-L-GluOMe (3)</td>
<td>7</td>
<td>Z-L-GluGlyNH 2 (11)</td>
<td>S166C-b</td>
<td>64</td>
</tr>
<tr>
<td>L-AlaNH 2 (8)</td>
<td>Z-L-Phe-L-AlaNH 2 (12)</td>
<td>24h</td>
<td>57</td>
<td>34</td>
</tr>
<tr>
<td>2</td>
<td>Z-L-Ala-L-AlaNH 2 (13)</td>
<td>24h</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Z-L-Glu-L-AlaNH 2 (14)</td>
<td>24h</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>Z-L-PheOBn (4)</td>
<td>GlyNH 2 (7)</td>
<td>Z-L-PheGlyNH 2 (15)</td>
<td>48h</td>
<td>0</td>
</tr>
<tr>
<td>Z-L-AlaOBn (5)</td>
<td>7</td>
<td>Z-L-AlaGlyNH 2 (16)</td>
<td>48h</td>
<td>80</td>
</tr>
<tr>
<td>Z-L-GluOBn (6)</td>
<td>7</td>
<td>Z-L-GluGlyNH 2 (17)</td>
<td>48h</td>
<td>63</td>
</tr>
</tbody>
</table>

† 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. ‡ Isolated yields. § Further 0.036 mol% added at 24 h.

Scheme 1

Scheme 2
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\textsubscript{166}C-S-glycosylation on P\textsubscript{t} stereospecificity using the \(\alpha\)-branched \(\alpha\)-alaninamide (8) as a challenging acyl acceptor probe. The restricted S\textsubscript{1}’ specificity of the unglycosylated SBL-WT enzyme is demonstrated by its ability to catalyze only the coupling of 8 with the preferred \(\alpha\)-phenylalaninate acyl donor 1. Remarkably, in spite of the small S\textsubscript{1}’ pocket\(^6\) all of the glyco-SBLs S166C-S-glycosylated were catalysts for the coupling of \(\alpha\)-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\textsubscript{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\textsubscript{t} towards \(\alpha\)-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.\(^7\)

Finally, we examined these powerful glyco-SBLs in the coupling of \(\alpha\)-amino acids (4)–(6) as acyl donors with acyl acceptor 1. The number of examples of enzyme catalyzed coupling of \(\alpha\)-amino acids at the C-terminus of peptides by using \(\alpha\)-amino acid acyl donor 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 coupling of \(\alpha\)-amino acids \(\alpha\)-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 of these three of \(\alpha\)-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 \(\alpha\)-dipeptide 17 (62–64%) were, surprisingly, superior to those found for coupling of \(\alpha\)-Glu. Indeed, S166C-S-e showed a 1:2:1 stereospecific preference for \(\alpha\)-glutamate 6 over \(\alpha\)-glutamate 3. Very high yields (up to 80%) of \(\alpha\)-Ala dipeptide (16) were also observed.

Such is the remarkable nature of these broadened P\textsubscript{t} stereospecificities that we speculated that these glyco-SBLs may bind \(\alpha\)-amino acids in a different mode from \(\alpha\)-amino acids. Molecular mechanics analyses of SBL-WT and S166C-S-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\textsubscript{1} pocket with the reacting acyl group covalently linked to O\textsubscript{4} of the side chain of Ser221. This binding mode is not available to \(\alpha\)-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 NH\textsubscript{2} group in the S\textsubscript{1} pocket. Thus, glyco-SBLs may operate either by binding mode (a) or (b). However, in these glyco-SBLs, neither \(\alpha\)-amino acids with the (a) mode nor \(\alpha\)-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 \(\alpha\)-stereospecificity and excellent efficiency (matched pair) into an enzyme that has broadened \(\alpha\)- and \(\alpha\)-sterosepecificity 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 \(k_{\text{cat}}\) and \(K_M\) using \(\alpha\)- and \(\alpha\)-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 \(\alpha\)-amino acids as acyl donors and an \(\alpha\)-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 \(\alpha\)-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 Buck and Kanji 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/b001021m/


8 For 3D-structure of SBL-WT see M. Knapp, J. Daubermann and R. R. Bött, RCSB-PDB entry 1jja.
