"Polar patch" proteases as glycopeptiligases†

Katie J. Doores and Benjamin G. Davis*

Received (in Cambridge, UK) 5th August 2004, Accepted 24th September 2004
First published as an Advance Article on the web 3rd December 2004
DOI: 10.1039/b412030b

The strategy of combined site directed mutagenesis and chemical modification with polar prosthetic groups was used to broaden substrate specificity of proteases resulting in the first successful formation of glycopeptides through the use of glucoamino acid acyl donors in yields of up to 90%.

Glycoproteins exist in different glycoforms each with different carbohydrate structures that can alter function.[1] Current methods for the isolation or creation of single glycoforms are difficult and often only small quantities of proteins can be obtained.[1] Therefore new methods for glycopeptide synthesis are desirable.

Enzymes are now widely accepted as useful catalysts in organic synthesis often with high stereo- and regiospecificity.[2] Their use in glycopeptide synthesis might therefore advantageously allow the coupling of carbohydrate and amino acid derivatives with reduced protecting group manipulation (Fig. 1).[1] Enzymatic conditions are also mild, so reactions can proceed without racemization of the amino acid α centre or β-elimination of the sugar. They are often carried out in water, a more desirable solvent for large peptide or oligosaccharide fragments, which have limited solubility in organic solvents. However, the use of enzymes may be limited by their typically stringent structural and stereospecificity: naturally occurring wild type (WT) enzymes cannot accept all the structures of synthetic chemical interest.

Enzymatic coupling of glycoamino acids with other amino acids has not yet been mastered. Wong *et al.*[3] have successfully coupled *N*- and *O*-linked glycoamino acid acyl donor fragments with nucleophilic amino acids and peptide acyl acceptors in DMF or aqueous environments using the serine protease subtilisin BPN′. However, this reaction was restricted to xylose containing glycoamino acids at the amide bond forming site (direct ligation). For more complex glycoamino acids, at least one amino acid spacer residue was required.[4] This potentially powerful strategy culminated in the enzyme catalysed ligation of glycopeptide fragments to create a variant of RNase bearing sialyl Lewis x.[5] The utility of peptide ligation has recently been emphasised further by its use in a combined native ligation–enzymatic ligation approach.[6]

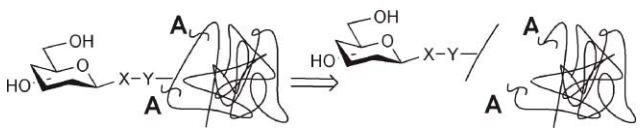


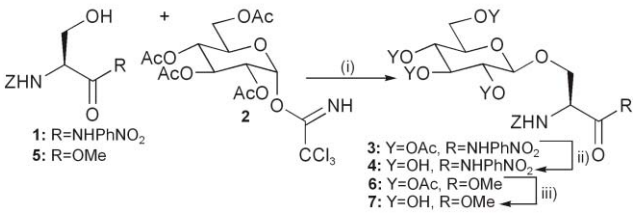
Fig. 1 Enzymatic ligation as one strategy for glycopeptide synthesis.[1]

† Electronic supplementary information (ESI) available: experimental details. See http://www.rsc.org/suppdata/cc/b4/b412030b/

*ben.davis@chem.ox.ac.uk

We have previously shown that a combination of site-directed mutagenesis and chemical modification of serine protease subtilisin *Bacillus lentus* (SBL) is a powerful technique for the efficient and rapid creation of new active site environments in enzymes with a broad substrate acceptance.[7] Introduction of a cysteine at position 166 at the base of the primary specificity determining S_1 pocket followed by addition of a "polar patch" (polar modification) was found to dramatically influence the specificity of SBL allowing D-amino acid ligation by chemically modified mutant enzymes (CMMs).

To probe the potential of these "polar patch" and other modified proteases in the direct ligation of glycoamino acids, we synthesised novel glycoamino acid substrate **4** containing a chromophoric moiety, which upon hydrolysis by the enzyme would be released and detected by UV spectroscopy, thus allowing the ready identification of this novel enzyme activity through screening. Glucose was chosen as an archetypal sugar. *N*-terminal Cbz (Z) protection was chosen on the basis of prior ligation reactions with simple amino acids[8] and this was introduced to serine using benzyl chloroformate.[9] The hydroxyl group was protected with TBDMS using TBDMSCl and imidazole in DMF. Reaction with *para*-nitroaniline (*p*NA) and phosphorous oxychloride in pyridine afforded the fully protected amino acid in a pleasing 80% yield.[10] The TBDMS group was removed using TBAF in THF[11] providing the *p*NA amino acid **1**. Glycosylation of the serine side chain is a known synthetic challenge.[12] Many conditions were explored for the glycosylation of **1**,[13] with most resulting in low yields and multiple by-products. The most efficient method used anomeric trichloroacetimidate **2** as the glycosyl donor[14] which was activated with catalytic TMSOTf in dry DCM resulting in the glucosyl amino acid **3** in 40% yield[15] which, to the best of our knowledge, is the first *p*NA glycoamino acid to be synthesised. Acetate protection of the glycosyl donor allowed absolute stereocontrol giving anomerically pure β-*O*-glycoside. The sugar–serine linkage is sensitive to β-elimination in basic environments[16] and this can cause difficulties in the removal of base-cleavable protecting



Scheme 1 (i) TMSOTf, DCM, (**3**: 40%, **6**: 46%), (ii) $NH_2NH_2{\cdot}H_2O$, MeOH, 79%, (iii) $NH_2NH_2{\cdot}H_2O$, MeOH, 33%.

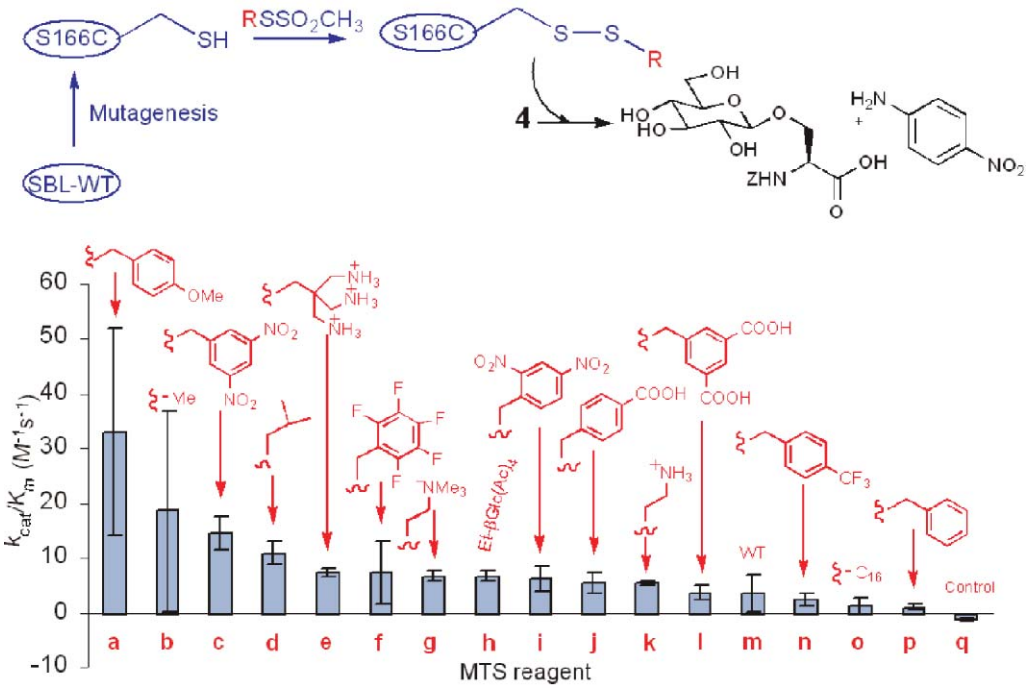


Fig. 2 Activities towards glycoamino acid **4** of a range of novel proteases created through the *in situ* chemical modification of S166C. Novel proteases and their creation are shown in blue, corresponding modifications of the enzymes to alter activity shown in red and the reaction catalyzed and used as a screen shown in black.

groups. However, after a survey of a range of methods, deprotection was pleasingly achieved using hydrazine monohydrate in 79% methanol resulting in deprotected *p*-nitroanilide glycoamino acid **4**.

Armed with this useful chromophoric probe for our desired enzyme activity we surveyed various protease catalysts. Potential structural variation through chemical modification in the active site of SBL is so vast that a combinatorial approach to enzyme screening and preparation was taken to identify the best CMMs for this transformation. The method developed by Plettner *et al.*[17] took advantage of the quantitative and rapid reaction of methanethiosulfonate (MTS) reagents with free protein cysteine residues (inset in Fig. 2).[18] Modifications of cysteine mutant SBL-S166C (Fig. 2) were carried out on a microscale in a 96-well plate and were monitored using Ellman's reagent.[19] Estimations for k_{cat}/K_M were then obtained directly from the rate of *p*NA released from **4** (ε_{414} = 2502 M^{-1} cm^{-1}) using the limiting case of the Michaelis–Menten equation at low substrate concentrations of 0.1 mM using these modified enzymes. The results are summarised in Fig. 2.

Promising catalysts, S166C-**a**, -**c**, -**e**, -**g**, were selected for more detailed evaluation on the basis of these results and previous significant substrate specificity broadening by polar aromatics[20] and charged polar patches.[7] Larger scale modification reactions in aqueous buffer were rapid and quantitative, as judged by titration of free thiols with Ellman's reagent.[19] Mass spectrometry was used to confirm CMM structure. The precise effects of the modifications to SBL were assessed by full Michaelis–Menten kinetics for the hydrolysis of **8** at pH 8.6 over a range of concentrations (0.02–8.0 mM) and the results are summarised in Table 1.

Table 1 Catalytic activity for the hydrolysis of glycoamino acid **4** by selected modified proteases

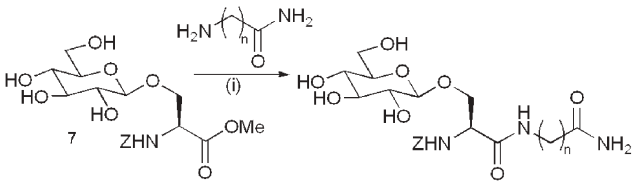
Enzyme	Est. k_{cat}/K_M/M^{-1} s^{-1}	K_M /M	k_{cat}/s^{-1}	k_{cat}/K_M/M^{-1} s^{-1}
WT	3.62	0.0023	0.007	3.04
S166C-**g**	6.79	0.0036	0.008	2.16
S166C-**e**	7.39	0.0020	0.009	4.47
S166C-**a**	14.6	0.0014	0.0005	0.39[a]
S166C-**c**	33.2	0.0054	0.008	1.51[a]

[a] Large difference from est. may be due to 'low substrate approxiamation' breakdown.

Having identified the modifications which gave the most desirable catalytic activity, peptide ligation reactions were attempted with the glycoamino acid methyl ester substrate **7** as a suitable acyl donor. The serine methyl ester was successfully prepared using thionyl chloride and methanol.[21] Glycosylation using the same conditions as for **3** yielded **6**, followed by hydrazine deprotection to afford **7** (Scheme 1).

With glycinamide as an acyl acceptor, the majority of the modifications showed an increase in efficiency compared to the wild type, with S166C-**e** yielding a pleasing 90% glycopeptide. Excitingly, using S166C-**e**, substrate specificity was also broadened to allow β-alaninamide and γ-butyricamide as acyl acceptors, albeit in lower yields than for glycinamide (Table 2). These are the first direct ligations of hexosyl amino acids.

Application to potentially therapeutic targets was also investigated. Compound **11** is an analogue of known active glycolipid galactosyl–ceramide mimics that are able to inhibit HIV uptake and infection of CD4-negative cells.[22] The desired glycolipid analogue **11** was formed under catalysis by S166C-**e** in 35% yield (Scheme 3).

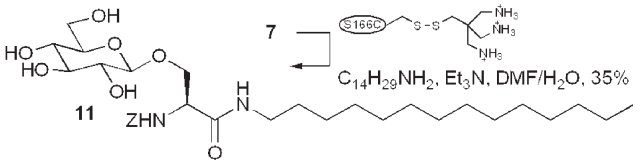


Scheme 2 (i) Enzyme (0.1 mol%), Et_3N, DMF–H_2O (1 : 1). Acyl donor : acceptor, 1 : 3.

Table 2 Results from enzyme catalysed ligation reactions according to Scheme 2 with amino acid acyl acceptors

Amino acid	*n*	Product	Yield (%)[a] (5 days) S166C			
			WT	-g	-e	-c
Gly–NH_2·HCl	1	Glc–Z–Ser–Gly–NH_2 **8**	60	73	90	20
β-Ala–NH_2·HCl	2	Glc–Z–Ser–β-Ala–NH_2 **9**	22	—[b]	53	—[b]
γ-Aba–NH_2·HCl	3	Glc–Z–Ser–γ-Aba–NH_2 **10**	15	—[b]	36	—[b]

[a] Yield based on compound **7**. [b] Ligation not attempted.



Scheme 3 Synthesis of glycolipid analogue *via* direct enzymatic ligation.

In summary, we have shown that combined site-directed mutagenesis and chemical modification can be used to broaden the substrate specificity of SBL and was utilised in the first direct ligations of hexosyl amino acids. A triamino substituent in the S_1 pocket leads to the greatest substrate broadening and was used in the synthesis of various glycopeptides and analogues of compounds showing anti-HIV activity.

We thank the Leverhulme Trust for funding.

Katie J. Doores and Benjamin G. Davis*
Department of Chemistry, University of Oxford, Chemistry Research Laboratory, Mansfield Road, Oxford, UK OX1 3TA.
E-mail: ben.davis@chem.ox.ac.uk; Fax: + 44 (0)1865 285002;
Tel: + 44 (0)1865 275652

Notes and references

1 B. G. Davis, *Chem. Rev.*, 2002, **102**, 579.
2 C.-H. Wong and G. M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, Pergamon, Oxford, 1994.
3 C.-H. Wong, M. Schuster, P. Wang and P. Sears, *J. Am. Chem. Soc.*, 1993, **115**, 5893.
4 K. Witte, O. Seitz and C.-H. Wong, *J. Am. Chem. Soc.*, 1998, **120**, 1979.
5 K. Witte, P. Sears, R. Martin and C.-H. Wong, *J. Am. Chem. Soc.*, 1997, **119**, 2114.
6 Z. Machova, R. Eggelkraut-Gottanka, N. Wehofsky, F. Bordusa and A. Beck-Sickinger, *Angew. Chem., Int. Ed.*, 2003, **42**, 4916.
7 K. Matsumato, B. G. Davis and J. Jones, *Chem. Eur. J.*, 2002, **8**, 4129.
8 K. Khumtaveeporn, G. DeSantis and J. B. Jones, *Tetrahedron: Asymmetry*, 1999, **10**, 2563.
9 A. Scheurer, P. Mosset, W. Bauer and R. W. Saalfrank, *Eur. J. Org. Chem.*, 2001, **16**, 3067.
10 D. Rijkers, H. Adams, H. C. Hemker and G. Tesser, *Tetrahedron*, 1995, **51**, 11235.
11 Y. Bilokin, A. Melman, V. Niddam, B. Benhamu and M. Bachi, *Tetrahedron*, 2000, **56**, 3425.
12 C. Taylor, *Tetrahedron*, 1998, **54**, 11317.
13 K. P. R. Kartha, L. Ballell, J. Bilke, M. McNeil and R. A. Field, *J. Chem. Soc., Perkin Trans. 1*, 2001, 770.
14 R. R. Schmidt, *Angew. Chem., Int. Ed. Engl.*, 1986, **25**, 212.
15 G. Grundler and R. R. Schmidt, *Liebigs Ann. Chem.*, 1984, 1826.
16 P. Sjolin and J. Kihlberg, *Tetrahedron Lett.*, 2000, **41**, 4435.
17 E. Plettner, K. Khumtaveeporn, X. Shang and J. Jones, *Bioorg. Med. Chem. Lett.*, 1998, **8**, 2291.
18 T. Bruice and G. Kenyon, *J. Protein. Chem.*, 1982, **1**, 47.
19 G. Ellman, K. Courtney, V. Andres and R. Featherstone, *Biochem. Pharmacol.*, 1961, **7**, 88.
20 K. Khumtaveeporn, A. Ullmann, K. Matsumoto, B. G. Davis and J. Jones, *Tetrahedron: Asymmetry*, 2001, **12**, 249.
21 D.-R. Hwang, P. Helquist and M. Shekhani, *J. Org. Chem.*, 1985, **50**, 1265.
22 B. Faroux-Corlay, L. Clary, C. Gadras, D. Hammache, J. Greiner, C. Santaella, A.-M. Aubertin, P. Vierling and J. Fantini, *Carbohydr. Res.*, 2000, **327**, 223.

