Site-Selective Glycosylation of Subtilisin *Bacillus lentus* Causes Dramatic Increases in Esterase Activity

Richard C. Lloyd,[a] Benjamin G. Davis[b,*] and J. Bryan Jones[a,†]

[a]*Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, ON, Canada M5S 3H6*
[b]*Department of Chemistry, University of Durham, South Road, Durham, DH1 3LE, UK*

Abstract—Using site directed mutagenesis combined with chemical modification, we have developed a general and versatile method for the glycosylation of proteins which is virtually unlimited in the scope of proteins and glycans that may be conjugated and in which the site of glycosylation and the nature of the introduced glycan can be carefully controlled. We have demonstrated the applicability of this method through the synthesis of a library of 48 glycosylated forms of the serine protease subtilisin *Bacillus lentus* (SBL) as single, pure species. As part of our ongoing program to tailor the activity of SBL for use in peptide synthesis, we have screened these enzymes for activity against the esterase substrate succinyl-Ala-Ala-Pro-Phe-S-benzyl. Gratifyingly, 22 enzymes displayed greater than wild type (WT) activity. Glycosylation at positions 62, in the S_2 pocket, resulted in five glycosylated forms of SBL that were 1.3- to 1.9-fold more active than WT. At position 217, in the S_1' pocket, all glycosylations increased k_{cat}/K_M up to a remarkable 8.4-fold greater than WT for the glucosylated enzyme L217C-S-β-Glc(Ac)$_3$. Furthermore, the ratio of amidase to esterase activity, $(k_{cat}/K_M)_{esterase}/(k_{cat}/K_M)_{amidase}$ (E/A), is increased relative to wild type for all 48 glycosylated forms of SBL. Again, the most dramatic changes are observed at positions 62 and 217 and L217C-S-β-Glc(Ac)$_3$ has an E/A that is 17.2-fold greater than WT. The tailored specificity and high activity of this glycoform can be rationalized by molecular modeling analysis, which suggests that the carbohydrate moiety occupies the S_1' leaving group pocket and enhances the rate of deacylation of the acyl-enzyme intermediate. These glycosylated enzymes are ideal candidates for use as catalysts in peptide synthesis as they have greatly increased $(k_{cat}/K_M)_{esterase}$ and severely reduced $(k_{cat}/K_M)_{amidase}$ and will favor the formation of the amide bond over hydrolysis. © 2000 Published by Elsevier Science Ltd. All rights reserved.

Introduction

We have recently reported a general method for the regio- and glycan-specific glycosylation of proteins.[1] This method provides a highly controlled and versatile route that is virtually unlimited in the scope of the sites and glycans that may be conjugated, and opens up hitherto inaccessible opportunities for the systematic determination of the properties of glycosylated proteins.[2] The broad scope of this method has been clearly demonstrated by the preparation of a library of neoglycosubtilisins from the serine protease subtilisin *Bacillus lentus* (SBL) using the carbohydrate methanethiosulfonate (MTS) reagents **1a–k** (Scheme 1). The glycosylated chemically modified mutant (glyco-CMM) proteins were extensively characterized, confirming that the glycan has

been introduced in a controlled fashion leading to single species of high purity. Four SBL sites at different locations and of different characteristics were chosen for glycosylation. S156 in the S_1 pocket[3] has a surface-exposed side chain, and mutation here allows the introduction of externally-disposed glycans, mirroring those found naturally in glycoproteins. N62 in the S_2 pocket, S166 in the S_1 pocket and L217 in the S_1' pocket all have internally oriented side chains and glycosylation at these sites has allowed the effects on catalysis of introducing sugars into the active site of an enzyme to be studied in a systematic manner for the first time.

Due to the heterogeneous nature of both natural and chemical protein glycosylation, a comprehensive study of the effects of site-selective glycosylation upon catalysis has been impossible until now. In the preceding paper,[2] we described the effects of site-selective glycosylation upon the amidase activity of SBL, and demonstrated how this method of glycosylation can be used for the determination of precise structure–activity relationships.

*Corresponding author. Tel.: +44-191-374-3136; fax: +44-191-384-4737; e-mail: ben.davis@durham.ac.uk
†Second corresponding author. Tel.: +1-416-978-3859; fax: +1-416-978-1553; e-mail: jbjones@alchemy.chem.utoronto.ca
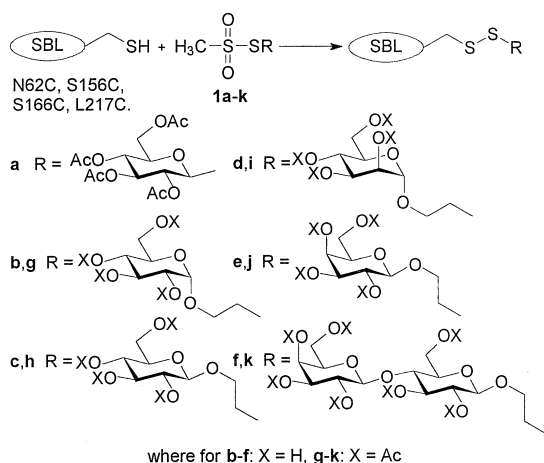
Scheme 1.

In ester hydrolysis, enzyme deacylation is generally rate-limiting,[4–6] leading to an accumulation of acyl-enzyme intermediate, which is required for attack by a nucleophile other than water in synthetic applications such as peptide ligation. However, under typical reaction conditions a high amidase activity will also lead to hydrolysis of both the acyl-enzyme intermediate and the product during coupling. As a current goal of our research program is to tailor SBL for use in peptide ligation reactions,[7] we desire enzymes with both increased esterase activities and lowered amidase activities. This paper describes the systematic tailoring of

SBL's esterase activity through site-selective glycosylation and discusses the esterase to amidase activity ratio (E/A) as a method for the evaluation of these catalysts for use in peptide synthesis.

Results and Discussion

Esterase activity screen: kinetic effects of glycosylation with unprotected carbohydrates

The glyco-CMMs shown in Table 1 were prepared by reacting reagents **1a–k** with the chosen cysteine mutants SBL-N62C, -S156C, -S166C and -L217C, purified and extensively characterized as described previously.[2] The kinetic parameters of esterase activity were determined at pH 8.6 by indirectly following the release of thiobenzyl alcohol from the substrate succinyl-Ala-Ala-Pro-Phe-S-benzyl (suc-AAPF-SBn) with Ellman's reagent.[8] To allow a rapid screen of esterase activity, a low substrate concentration ([S] ≪ K_M) was used that allowed k_{cat}/K_M to be determined directly from the initial rate of reaction. The results from the screen are shown in Table 1.

Gratifyingly, glycosylation at position 62, in the S_2 pocket, with deprotected sugar reagents **1b–f** increases k_{cat}/K_M, resulting in a series of five enzymes that have similar k_{cat}/K_Ms that are 1.3- to 1.9-fold greater than WT (Fig. 1(a)). The presence of an α-linkage is clearly deleterious to activity, as β-anomer N62C-S-Et-βGlc (**-c**) has a k_{cat}/K_M 1.2-fold greater than its α-anomer N62C-S-Et-α-Glc (**-b**) and 1.9-fold greater than WT.

**Table 1.** Esterase screen results for glycosylated CMMs[a,c]

R	N62C-R		L217C-R		S166C-R		S156C-R	
	k_{cat}/K_M (s^{-1} mM^{-1})	E/A[b]	k_{cat}/K_M (s^{-1} mM^{-1})	E/A[b]	k_{cat}/K_M (s^{-1} mM^{-1})	E/A[b]	k_{cat}/K_M (s^{-1} mM^{-1})	E/A[b]
H	4380	48	5540	109	350	4	—	—
S-β-Glc	—	—	6350	186	—	—	—	—
S-**b**	5327	37	7773	90	2688	37	1984	22
S-**c**	6625	62	9791	130	2732	40	2197	21
S-**d**	4627	36	8651	77	2924	45	2648	24
S-**e**	5729	48	11,923	163	3241	56	2505	28
S-**f**	5990	48	9080	137	2511	44	1692	24
S-β-Glc(Ac)$_2$	—	—	8876	86	—	—	—	–
S-β-Glc(Ac)$_3$	—	—	11,642	115	—	—	—	—
S-**a**	2502	19	—	—	2537	50	1356	17
S-Et-α-Glc(Ac)$_2$	—	—	7893	79	—	—	—	—
S-Et-α-Glc(Ac)$_3$	—	—	—	—	1829	45	—	—
S-**g**	6851	30	—	—	—	—	1887	34
S-Et-β-Glc(Ac)$_2$	1925	36	—	—	5058	46	—	—
S-Et-β-Gal(Ac)$_3$	2736	51	4935	156	6033	54	—	—
S-**h**	—	—	—	—	—	—	2321	26
S-Et-α-Man(Ac)$_3$	—	—	11,751	71	1537	37	—	—
S-**i**	5807	50	—	—	—	—	2629	31
S-Et-α-Man(Ac)$_3$	2169	50	3889	116	3157	38	—	—
S-**j**	—	—	—	—	—	—	2010	24
S-Et-Lac(Ac)$_5$	—	—	—	—	1543	38	—	—
S-Et-Lac(Ac)$_6$	—	—	5468	139	—	—	—	—
S-**k**	2135	29	—	—	—	—	1516	22

[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. [S] = 15 or 30 μM, [E] = 4.8×10^{-11} to 6.0×10^{-10} M.
[b]$E/A = (k_{cat}/K_M)_{esterase}/(k_{cat}/K_M)_{amidase}$.
[c]For SBL-WT: $k_{cat}/K_M = 3592.5$ mM^{-1} s^{-1}, $E/A = 17$.
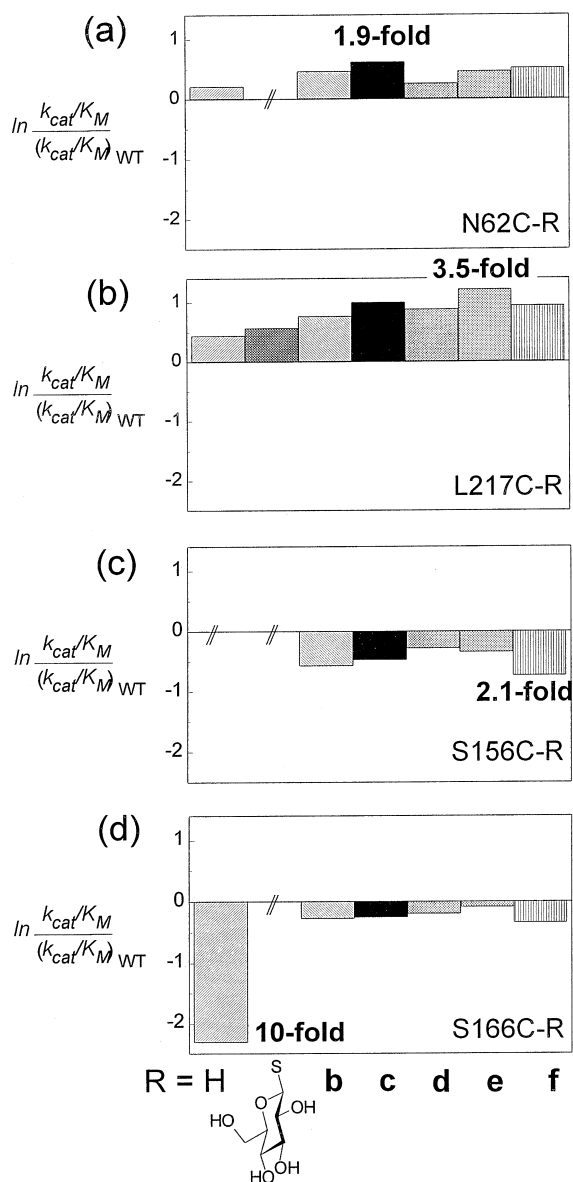
Figure 1. Esterase k_{cat}/K_Ms for deprotected glyco-CMMs relative to WT. (a) At position 62, in the S_1 pocket, glycosylation leads to a series of enzymes that have similar activities that are 1.3- to 1.9-fold greater than WT. (b) At position 217, in the S_1' pocket, glycosylation also increases k_{cat}/K_M, to a maximum 3.5-fold greater than WT for L217C-S-Et-β-Gal (-**e**). (c) At position 156, in the S_1 pocket, glycosylation leads to reduction in k_{cat}/K_M. (d) At position 166, in the S_1 pocket, the dramatic loss of activity upon mutation to cysteine (R = H) is restored by glycosylation. All five S166C deprotected glyco-CMMs have similar k_{cat}/K_Ms that are 1.1- to 1.4-fold lower than WT.

Furthermore, the α-linked N62C-S-Et-α-Man (-**d**) has the lowest k_{cat}/K_M in this group: 1.3-fold greater than WT.

As at position 62, the introduction of any of the deprotected sugar side chains **b**–**f** at position 217, in the S_1' pocket, increases k_{cat}/K_M (Fig. 1(b)). However, the effects of glycosylation at this site are far more dramatic as demonstrated by a k_{cat}/K_M for L217C-S-Et-β-Gal (-**e**) that is 3.4-fold greater than WT. By comparing the k_{cat}/K_Ms of L217C-S-β-Glc and L217C-S-Et-β-Glc (-**c**), it is

possible to gauge the effect on activity of introducing an ethyl tether at this position. This introduction increases k_{cat}/K_M, from 1.8-fold greater than WT for L217C-S-β-Glc to 2.7-fold greater than WT for L217C-S-Et-β-Glc (-**c**). As at position 62, β-linked glyco-CMMs (-**c**,-**e**,-**f**) have higher k_{cat}/K_Ms than those that are α-linked (-**a**, -**d**). For example, L217C-S-Et-β-Glc (-**c**) has a k_{cat}/K_M 1.3-fold greater than L217C-S-Et-α-Glc (-**b**).

Consistent with the surface exposed orientation of the S156 side chain, the effects of glycosylation at position 156 are slight. As a result, the S156C deprotected glyco-CMMs S156C-S-**b**–**f** have similar k_{cat}/K_Ms that are 1.3- to 2.1-fold lower than WT (Fig. 1(c)).

At position 166, in the S_1 pocket, mutation to cysteine results in an enzyme with a severely lowered k_{cat}/K_M that is 10-fold lower than WT. However, subsequent glycosylation with **1b**–**f** restores much of the catalytic activity (Fig. 1(d)), and the S166C deprotected glyco-CMMs S166C-S-**b**–**f** have similar k_{cat}/K_Ms, that vary from 1.1- to 1.4-fold lower than WT.

Kinetic effects of glycosylation with acetylated carbohydrates

At position 62, in the S_2 pocket, in contrast to the narrow trend observed for deprotected glycans, glycosylation with acetylated glycans creates N62C glyco-CMMs with a wide range of k_{cat}/K_Ms. In particular, the introduction of acetate groups increases the k_{cat}/K_Ms of α-linked glyco-CMMs relative to the corresponding deprotected glyco-CMMs (Fig. 2(a)). Thus, N62C-S-Et-α-Glc(Ac)$_4$ (-**b**) and N62C-S-Et-α-Man(Ac)$_4$ (-**d**) have k_{cat}/K_Ms 1.9- and 1.6-fold greater than WT, respectively. Acetylation is clearly deleterious for β-linked glyco-CMMs, as N62C-S-Et-β-Gal(Ac)$_3$, N62C-S-Et-β-Glc(Ac)$_2$ and N62C-S-Et-β-Glc(Ac)$_3$ all have k_{cat}/K_Ms lower than WT. Interestingly, increasing the number of acetates present in the glycan partially restores activity: N62C-S-Et-β-Glc(Ac)$_3$ has a k_{cat}/K_M only 1.3-fold lower than WT and 1.5-fold higher than N62C-S-Et-β-Glc(Ac)$_2$. In spite of the sterically bulky side chain in the lactosylated N62C CMMs N62C-S-Et-Lac (-**f**) and N62C-S-Et-Lac(Ac)$_7$ (-**k**), they have k_{cat}/K_Ms that are similiar to those of the CMMs derived from monosaccharides. This provides a clear example of the versatility of the glycosylation method illustrated in Scheme 1 and demonstrates that by using this method it is possible to introduce large glycan structures into the active site of SBL while maintaining catalytic competency.

Excitingly, modification with acetylated reagents **1a,g**–**k** at position 217, in the S_1' pocket, as with **1b**–**k**, leads to CMMs with greater than WT k_{cat}/K_Ms (Fig. 2(b)). For the untethered glyco-CMMs, increasing the number of acetates dramatically increases k_{cat}/K_M, from 1.8-fold greater than WT for L217C-S-β-Glc to 2.4-fold greater than WT for L217C-S-β-Glc(Ac)$_2$ and to 3.2-fold greater than WT for L217C-S-β-Glc(Ac)$_3$, and mirrors the trend seen in amidase kinetics.[2] For the ethyl linked L217C glyco-CMMs, the effect of acetylation is again dependent on anomeric stereochemistry, and similar to
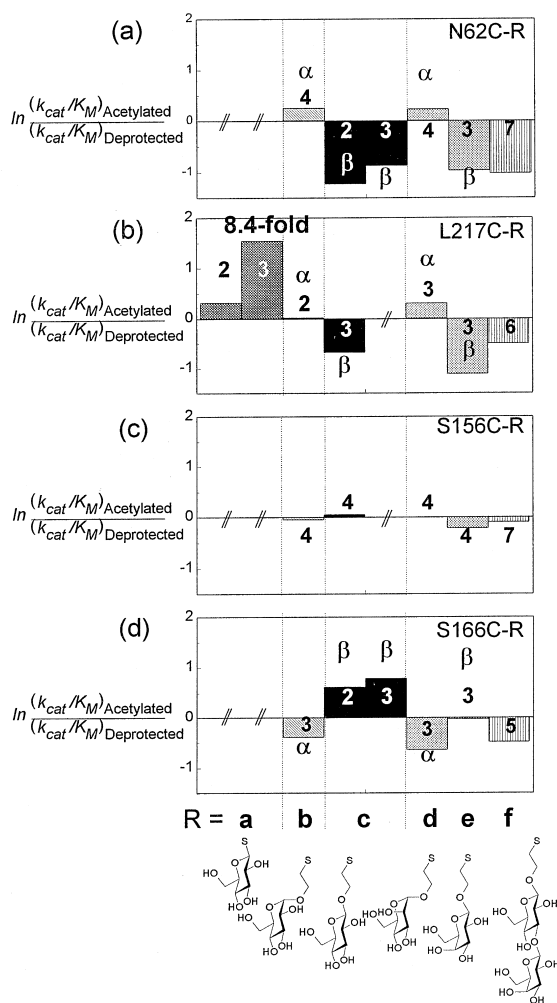
Figure 2. The effect of acetylation on esterase k_{cat}/K_Ms of glyco-CMMs. For each glycan the number of acetate groups present is indicated by a label on the corresponding bar. A break in the axis indicates that the value was not determined. (a), (b) At positions 62, in the S_2 pocket, and 217, in the S_1' pocket, the effect of acetylation is dependent on anomeric stereochemistry. At both sites, acetylation of α-linked sugars (**-b,-d**) leads to an increase in k_{cat}/K_M whereas k_{cat}/K_M is decreased for β-linked sugars (**-c,-e,-f**). k_{cat}/K_M also increases as the number of acetates increases. (c) Consistent with the surface-exposed nature of its side chain, acetylation at position 156, in the S_1 pocket, has very little effect on k_{cat}/K_M. (d) At position 166, in the S_1 pocket, the effects of acetylation are opposite to those at positions 62 and 217. Acetylation increases k_{cat}/K_M of β-linked glyco-CMMs (**-c**, **-e**), whilst causing a decrease for the α-linked glyco-CMMs (**-b**, **-d**).

that observed for N62C glyco-CMMs; it also mirrors the trend observed for the amidase kinetics of these L217C glyco-CMMs.[2] Thus, acetylation increases k_{cat}/K_Ms of α-linked CMMs but decreases k_{cat}/K_Ms of β-linked CMMs. This is most pronounced for L217C-S-Et-β-Gal(Ac)$_3$ which has a k_{cat}/K_M only 1.1-fold greater than WT and 3.1-fold lower than L217C-S-Et-β-Gal (**-e**). In contrast to the effect on deprotected L217C glyco-CMMs, the activity of acetylated L217C glyco-CMMs decreases upon the introduction of the ethyl linker. For example, L217C-S-β-Glc(Ac)$_3$ has a k_{cat}/K_M 3.2-fold greater than WT, as compared with the 1.4-fold greater than WT k_{cat}/K_M of L217C-S-Et-β-Glc(Ac)$_3$.

At position 156, in the S_1 pocket, the S156C acetylated glyco-CMMs display little difference in their kinetic constants from their unacetylated counterparts (Fig. 2(c)), an observation that is again consistent with the surface exposed nature of the position 156 side chain. Introducing the ethyl linker leads to a slight increase in k_{cat}/K_M from 2.6-fold lower than WT for S156C-S-β-Glc(Ac)$_4$ (**-a**) to 1.5-fold lower than WT for S156C-S-Et-β-Glc(Ac)$_4$ (**-h**).

In general, at position 166, in the S_1 pocket, the effect of acetylation on S166C ethyl linked glyco-CMMs is to reduce k_{cat}/K_Ms relative to their deprotected counterparts (Fig. 2(d)). The exceptions are the ethyl linked β-gluco-CMMs, S166C-S-Et-β-Glc(Ac)$_2$ and S166C-S-Et-β-Glc(Ac)$_3$, which display k_{cat}/K_Ms 1.4- and 1.7-fold greater than WT, respectively. These are the only two glyco-CMMs prepared at this site to show an enhancement in k_{cat}/K_M relative to WT, and these examples illustrate that the correct selection of sugar is crucial to the tailoring of enzyme activity. In contrast to the effects observed at positions 62 and 217, an α-linkage to the sugar moiety is deleterious to the activity of the acetylated CMMs and α-anomer S166C-S-Et-α-Glc(Ac)$_3$ has a k_{cat}/K_M 1.9-fold lower than WT, in direct contrast to β-anomer S166C-S-Et-β-Glc(Ac)$_3$. Introduction of the sterically bulky lactose moiety, in both acetylated and unacetylated forms, leads to CMMs S166C-S-Et-Lac (**-f**) and S166C-S-Et-Lac(Ac)$_5$ with low k_{cat}/K_Ms that are 1.4- and 2.3-fold lower than WT, respectively.

Full esterase kinetics

The three esterases with the highest k_{cat}/K_Ms, as determined by the above screen, L217C-S-β-Glc(Ac)$_3$, L217C-S-Et-α-Man(Ac)$_3$ and L217C-S-Et-β-Gal (**-e**), were selected and their individual k_{cat}s and K_Ms were determined by the initial rates method. The results of these full esterase kinetics determinations are shown in Table 2.

The results are in good agreement with those determined by the screen for L217C-S-Et-α-Man(Ac)$_3$ and L217C-S-Et-β-Gal (**-e**) and confirm the activity of these two enzymes to be 3.6- and 3.5-fold higher than WT, respectively. These increases in activity arise from both increased transition state stabilization, with k_{cat}s 2- and 2.3-fold greater than WT, respectively, and from enhanced substrate binding, with K_Ms 1.8- and 1.5-fold lower than WT, respectively.

Remarkable results were obtained for L217C-S-β-Glc(Ac)$_3$. Excitingly, this enzyme has a k_{cat} 2.3-fold greater than WT and a K_M 3.6-fold lower than WT, giving a k_{cat}/K_M 8.4-fold greater than WT and some 2.5-fold greater than the value estimated by the screen.[9] Gratifyingly, this is the largest enhancement of activity relative to WT we have achieved using the combined site-directed mutagenesis and chemical modification strategy.

Esterase activity versus amidase activity

The above results demonstrate that esterase k_{cat}/K_Ms can be controlled and enhanced by the introduction of

Table 2. Full esterase kinetic data for glycosylated CMMs[a]

Enzyme	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (mM^{-1} s^{-1})	×WT	E/A[b]
SBL-WT	1940.0±180	0.54±0.07	3592.5±572.7	1	17
L217C–S-β-Glc(Ac)$_3$	4427.5±100.9	0.15±0.01	29516±2079.0	8.4	293
L217C-S-Et-α–Man(Ac)$_3$	3827.0±59.5	0.30±0.01	12756.7±469.2	3.6	77
L217C-S-Et-β-Gal	4398.5±189.8	0.36±0.04	12218.1±1456.3	3.5	167

[a]Kinetic constants determined by method of initial rates in 0.1 M buffer, pH 8.6, 0.005% Tween 80, 1% DMSO with suc-AAPF-SBn as substrate. [S] = 31.25 μM to 2.0 mM, [E] = 9.6×10^{-11} to 1.1×10^{-10} M.
[b]$E/A = (k_{cat}/K_M)_{esterase}/(k_{cat}/K_M)_{amidase}$.

sugar moieties into the active site of SBL. In the preceding paper, we described how the amidase k_{cat}/K_M was also altered by these glycosylations.[2] The differing effects of glycosylation upon amidase and esterase k_{cat}/K_M can be compared in an informative manner using the $(k_{cat}/K_M)_{esterase}/(k_{cat}/K_M)_{amidase}$ ratio, E/A.

Gratifyingly, all N62C deprotected glyco-CMMs have E/As that are enhanced relative to WT. The increase in E/A is dependent on the presence of either an α- or a β-linkage with the β-linked CMMs N62C-S-Et-β-Glc (-**c**), -S-Et-β-Gal (-**e**) -S-Et-Lac (-**f**) having higher E/As than the α-linked CMMs N62C-S-Et-α-Glc (-**b**) and -S-Et-α-Man (-**d**) (Fig. 3(a)). These increased ratios are due to both increases in esterase k_{cat}/K_Ms and reductions in amidase k_{cat}/K_Ms.

As observed for the modifications made at position 62, glycosylation at position 217, in the S$_1'$ pocket, also leads to enzymes that have greatly increased E/A ratios relative to WT. Mutation to cysteine at position 217 increases E/A to 6.4-fold greater than WT. Modification with deprotected β-linked sugars -S-β-Glc,-S-Et-β-Glc (-**c**), -S-Et-β-Gal (-**e**), and -S-Et-Lac (-**f**) further increases E/A (Fig. 3(b)). In contrast, α-linked glyco-CMMs have lower E/A values than that of the mutant. These changes in E/As are a result of parallel changes in esterase k_{cat}/K_Ms that are further amplified by opposing changes in amidase k_{cat}/K_Ms.[2] Introduction of an ethyl linker reduces the E/A from 10.9-fold greater than WT for L217C-S-β-Glc to 7.6-fold greater than WT for L217C-S-Et-β-Glc (-**c**).

At position 156, in the S$_1$ pocket, E/A ratios for unacetylated glyco-CMMs are all similar to each other in the range 1.2- to 1.6-fold greater than WT (Fig. 3(c)).

The mutant S166C has an exceptionally low E/A that is 4.2-fold lower than WT (Fig. 3(d)) and this is largely a result of its very low esterase k_{cat}/K_M. Since glycosylation of position 166 restores esterase k_{cat}/K_Ms to levels approaching that of WT and because S166C glyco-CMMs are poor amidases relative to WT, the net result is a family of S166C glyco-CMMs with similar E/A ratios that are all enhanced relative to WT and significantly higher than the cysteine mutant (Fig. 3(d)).

Effects of glycosylation with acetylated carbohydrates on E/A

At position 62, in the S$_2$ pocket, with the exception of N62C-S-Et-β-Gal(Ac)$_3$ and N62C-S-Et-α-Man(Ac)$_4$ (-**i**),
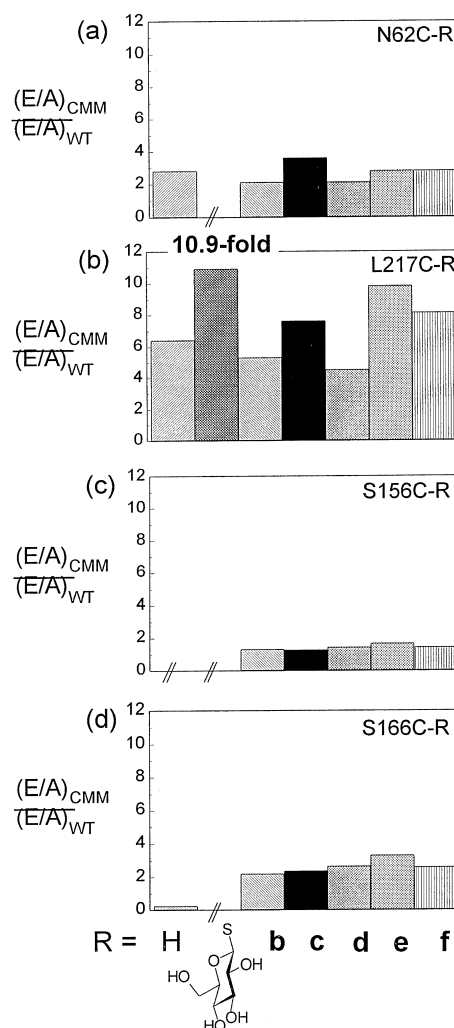
Figure 3. E/A of deprotected glyco-CMMs relative to WT. A break in the axis indicates that the value was not determined. Glycosylation with deprotected reagents **1b**–**f** increases the E/A ratio in all cases. The greatest effects are observed at positions 62, in the S$_2$ pocket (a), and 217, in the S$_1'$ pocket (b), where the largest increases in E/A result in values up to 10.9-fold greater than WT. At both sites E/A is higher for the β-linked glyco-CMMs than the α-linked ones. At positions 156, (c) and 166 (d), in the S$_1$ pocket, there is little variation in E/A.

acetylation of deprotected N62C glyco-CMMs leads to a reduction in E/A. However, like their deprotected counterparts, the acetylated N62C glyco-CMMs all have larger E/As than WT (Fig. 4(a)). Increasing the level of acetylation increases E/A from 2.1-fold greater than WT for N62C-S-Et-β-Glc(Ac)$_2$ to 3.0-fold greater
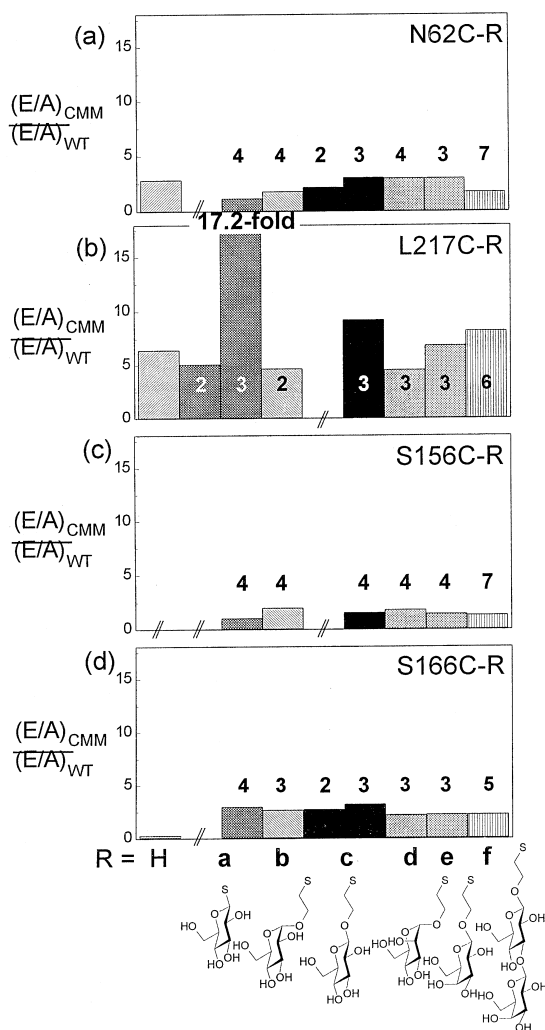
than WT for N62C-S-Et-β-Glc(Ac)$_3$. Interestingly, in spite of the general increases in E/A observed for the ethyl linked glyco-CMMs, the untethered CMM N62C-S-β-Glc(Ac)$_4$ (**-a**) has an E/A very similar to WT.

At position 217, in the S$_1'$ pocket, L217C acetylated glyco-CMMs all have enhanced E/As relative to WT (Fig. 4(b)). As for their deprotected counterparts, β-linked acetylated glyco-CMMs have higher E/As than that of the L217C mutant, whereas the E/As of the α-linked acetylated glyco-CMMs are lower than that of the mutant. In fact, glycosylation at this site produces enzymes with by far the greatest enhancement of this ratio and the E/A for L217C-S-β-Glc(Ac)$_3$ is a dramatic 17.2-fold greater than WT. In contrast to the L217C deprotected glyco-CMMs, introducing an ethyl linker lowers E/A from 17.2-fold greater than WT for L217C-S-β-Glc(Ac)$_3$ to 9.2-fold greater than WT for L217C-S-Et-β-Glc(Ac)$_3$. For the untethered L217C acetylated glyco-CMMs, increasing the number of acetates also

increases E/A, from 5.0-fold greater than WT for L217C-S-Et-β-Glc(Ac)$_2$ to 17.2-fold greater than WT for L217C-S-Et-β-Glc(Ac)$_3$.

At position 156, in the S$_1$ pocket, glycosylation with acetylated glycans yields glyco-CMMs with similar E/As to those formed by glycosylation with unprotected glycans (Fig. 4(c)). For example, S156C-S-Et-β-Glc(Ac)$_4$ (**-h**) has an E/A that is 1.5-fold greater than WT, S156C-S-Et-α-Glc(Ac)$_4$ (**-g**) has an E/A 2.0-fold greater than WT and S156C-S-Et-α-Man(Ac)$_4$ (**-i**) has an E/A only 1.3-fold greater than that of its unacetylated counterpart, S156C-S-Et-α-Man (**-d**).

At position 166, in the S$_1$ pocket, acetylation causes an increase in E/A for the glucosylated CMMs, irrespective of the anomeric stereochemistry (Fig. 4(d)). Acetylation of all other sugar moieties leads to a reduction in E/A. For these glucosylated enzymes, increasing the number of acetates increases E/A from 2.7-fold greater than WT for S166C-S-Et-β-Glc(Ac)$_2$ to 3.2-fold greater than WT for S166C-S-Et-β-Glc(Ac)$_3$.

Molecular modeling

The molecular basis for the vastly improved esterase activities caused by glycosylation was analyzed by molecular modeling of the peptidyl product inhibitor AAPF bound to the glyco-CMM L217C-S-β-Glc(Ac)$_3$.[10]

As shown in Figures 5 and 6, molecular modeling analysis of L217C-S-β-Glc(Ac)$_3$ reveals that the observed k_{cat}/K_M changes correlate with the occupation of the S$_1'$ pocket of SBL by the glucosylated-S-β-Glc(Ac)$_3$ side chain. In the minimized structure (Fig. 5), the position of the glucose moiety is fixed by a network of hydrogen-bonding interactions (shown as white dotted lines)
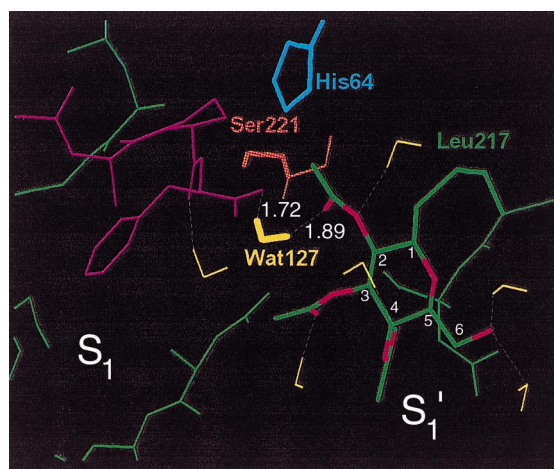
Figure 4. E/A of acetylated glyco-CMMs relative to WT. For each glycan the number of acetate groups present is indicated by a label on the corresponding bar. A break in the axis indicates that the value was not determined. Most notably (b) glycosylation at position 217 creates an enzyme L217C-S-β-Glc(Ac)$_3$ with an E/A that is 17.2-fold greater than WT.

Figure 5. Modeling the high esterase activity of L217C-S-β-Glc(Ac)$_3$. A minimized structure of the active site of glyco-CMM L217C-S-β-Glc(Ac)$_3$ showing the catalytic residues Ser221 (orange) His64 (blue). The carbon atoms of the triacetylated D-glucosyl moiety, which is bound to Leu217 via a disulfide bond, are numbered. The phenyl ring of product AAPF (pink) occupies the S$_1$ binding site and forms a crucial hydrogen bond (1.72 Å) to water molecule Wat127 (yellow). This water molecule is further stabilized by a second hydrogen bond (1.89 Å) to the carbonyl O of the C-2 acetate group of glucose.
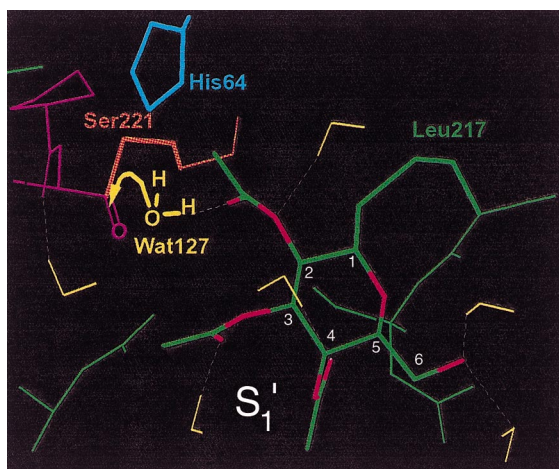
Figure 6. Proposed acyl-enzyme intermediate of L217C-β-Glc(Ac)₃. The carboxy terminus of AAPF forms a bond to the O_γ atom of Ser221. Water molecule Wat127, acting as the crucial deacylating nucleophilic water molecule, is stabilized through a key hydrogen bond to the carbonyl group of the C-2 acetate group.

between the oxygen atoms on the C-3, 4 and 6 substituents of glucose and water molecules in the surrounding external solvent. This extensive solvation directs the C-2 substituent of glucose internally towards the catalytic triad. In this orientation, the carbonyl oxygen atom of the C-2 acetate group acts as a hydrogen bond (1.89 Å) acceptor and stabilizes water molecule 127 (Wat127) in close proximity to the carboxy terminus of AAPF, used here as a substrate analogue for modeling, and is shown in Figure 5 hydrogen-bonding to the O atom of the carboxylic acid.

Excitingly, these results suggest that firstly the low amidase activity of L217C-S-β-Glc(Ac)₃, which is 2-fold lower than WT, is a result of the S_1' pocket being occupied by the glucose moiety at position 217. This prevents efficient binding of the *p*NA leaving group and therefore decreases the rate of acyl-enzyme intermediate formation, which is the rate-determining step for amidase activity. Secondly, after the *p*NA has been displaced to form the covalent acyl-Ser221 intermediate the glucose moiety stabilizes a crucial, nucleophilic water molecule (Wat 127) in close proximity to the carbonyl carbon atom, through a hydrogen-bond to the oxygen of the C-2 acetate of glucose, as illustrated in Figure 6. This facilitates hydrolysis of the acyl-enzyme intermediate and therefore increases the rate of deacylation, which is the rate-limiting step for esterase activity.[4–6]

Conclusions

Glycosylation of SBL at key positions within the active site can dramatically enhance its esterase activity. The library of glycosylated CMMs that we have synthesized using the combined site directed mutagenesis and chemical modification strategy contains 22 enzymes with greater than WT activity. Glycosylation at positions 62, in the S_2 pocket, and 217, in the S_1' pocket, gave the greatest increases in k_{cat}/K_M. The most active glyco-CMM L217C-S-β-Glc(Ac)₃ has a k_{cat}/K_M that is 8.4-fold

greater that WT and is the most active esterase we have synthesized using this approach. When surface exposed position 156 was glycosylated, there was little alteration in activity, and this demonstrates that the introduction of sugars at such sites has little effect on the catalytic activity of SBL. This discovery may be of use in the synthesis of specific glyco-CMMs with externally-oriented carbohydrate recognition determinants for targeting purposes.

In addition to the tailoring of esterase k_{cat}/K_M values, glycosylation also led to enormous improvements in specificity for ester versus amide hydrolysis, as determined by measurement of the ratio $(k_{cat}/K_M)_{esterase}/(k_{cat}/K_M)_{amidase}$, E/A. Excitingly, this ratio has been increased to 17.2-fold greater than WT for L217C-S-β-Glc(Ac)₃. Such enzymes are very attractive candidates for use in peptide synthesis, where a high esterase to amidase ratio is desirable. Furthermore, the CMMs described in this paper have an even greater potential for this purpose, as the increases in E/As have been achieved by an absolute increase in the catalytic efficiency of these enzymes towards ester substrates, in addition to a reduction in the amidase activity. Experiments are currently underway to confirm the preparative utility of these enzymes.

Experimental

Mutants and chemically modified mutants of subtilisin *B. lentus* were prepared and purified as previously described.[2]

Esterase screen

Specificity constants determined using the low substrate approximation were measured indirectly using Ellman's reagent[8] ($\epsilon_{412} = 13,600 \, M^{-1} \, cm^{-1}$) 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.

Full esterase kinetics measurements

Michaelis–Menten constants were measured at 25 °C by curve fitting (Grafit® 3.03) of the initial rate data determined at eight concentrations (31.25 μM–2.0 mM) of the succinyl-AAPF-SBn substrate, followed indirectly using Ellman's reagent 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.

Molecular modeling

The X-ray structure of subtilisin *B. lentus* with the peptide inhibitor AAPF bound (Brookhaven database entry 1JEA) was used as the starting point for calculations on wild type and CMMs. The enzyme setup was performed with Insight II, version 2.3.0.[11] To create initial coordinates for the minimization, hydrogens were added at the pH 8.6 used for kinetic measurements. This protonated all Lys and Arg residues and the N-terminus and deprotonated all Glu and Asp residues and the

C-terminal carboxyl group. The protonated form of His 64 was used in all calculations. The model system was solvated with a 5 Å layer of water molecules. The total number of water molecules in the system was 1143. The overall charge of the enzyme–inhibitor complex resulting from this setup was +4 for the WT enzyme. Energy simulations were performed with the DISCOVER program, Version 2.9.5 on a Silicon Graphics Indigo computer, using the consistent valence force field (CVFF) function. A non-bonded cutoff distance of 18 Å with a switching distance of 2 Å was employed. The non-bonded pair list was updated every 20 cycles and a dielectric constant of 1 was used in all calculations. The WT enzyme was minimized in stages, with initially only the water molecules being allowed to move, followed by water molecules and the amino acid side chains, and then finally the entire enzyme. The mutated and chemically modified enzymes were generated by modifying the relevant amino acid using the Builder module of Insight. These structures were then minimized in a similar manner. Initially the side chain of the mutated residue and the water molecules were minimized. Then all side chains and the water molecules were minimized while the backbones of the residues were constrained, then all of the atoms were minimized. The AAPF inhibitor was free to move throughout all stages of the minimization. Each stage of energy minimization was conducted by means of the method of steepest descents without Morse or cross terms until the derivative of energy with respect to structural perturbation was less than 5.0 kcal/Å; then the method of conjugate gradients, without Morse or cross terms until the derivative of energy with respect to structural perturbation was less than 1.0 kcal/Å; and finally the method of conjugate gradients *with* Morse and cross terms until the final derivative of energy with respect to structural perturbation was less than 0.1 kcal/Å.

References and Notes

1. Davis, B. G.; Lloyd, R. C.; Jones, J. B. *J. Org. Chem.* **1998**, *63*, 9614.
2. Davis, B. G.; Lloyd, R. C.; Jones, J. B. *Bioorg. Med. Chem.* **2000**, *8*, 1527.
3. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157 (nomenclature).
4. Zerner, B.; Bond, R. P. M.; Bender, M. L. *J. Am. Chem. Soc.* **1964**, *86*, 3674.
5. Whitaker, J. R.; Bender, M. L. *J. Am. Chem. Soc.* **1965**, *87*, 27728.
6. Berezin, I. V..; Kazanskaya, N. F.; Klysov, A. A. *FEBS Lett.* **1971**, *15*, 121.
7. Plettner, E.; DeSantis, G.; Stabile, M. R.; Jones, J. B. *J. Am. Chem. Soc.* **1999**, *121*, 4977.
8. Ellman, G. L.; Courtney, K. D.; Andres, V.; Featherstone, R. M. *Biochem. Pharmacol.* **1961**, *7*, 88.
9. The difference in parameters obtained from the screen and the full kinetic analysis exposes a limitation of the low substrate screen. For the low substrate approximation to be accurate, the substrate concentration must be small compared to K_M. The K_M of L217C-SβGlc(Ac)$_3$ (0.15 mM) is evidently so small that the approximation does not hold in this case.
10. While the substrate employed for kinetic analysis is succinylated, we did not include this moiety in molecular modeling since its orientation was not reported in the X-ray structure of SBL, suggesting high mobility in the crystal.
11. Biosym Technologies, Inc., San Diego, CA, USA.