



**HAL**  
open science

## Functional analysis of a group A streptococcal glycoside hydrolase Spy1600 from family 84 reveals it is a beta-N-acetylglucosaminidase and not a hyaluronidase

William L Sheldon, Matthew S Macauley, Edward J Taylor, Charlotte E Robinson, Simon J Charnock, Gideon J Davies, David J Vocadlo, Gary W Black

### ► To cite this version:

William L Sheldon, Matthew S Macauley, Edward J Taylor, Charlotte E Robinson, Simon J Charnock, et al.. Functional analysis of a group A streptococcal glycoside hydrolase Spy1600 from family 84 reveals it is a beta-N-acetylglucosaminidase and not a hyaluronidase. *Biochemical Journal*, 2006, 399 (2), pp.241-247. 10.1042/BJ20060307 . hal-00478542

**HAL Id: hal-00478542**

**<https://hal.science/hal-00478542>**

Submitted on 30 Apr 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Full Title:** Functional analysis of a group A streptococcal glycoside hydrolase from family 84 reveals it is a  $\beta$ -*N*-acetylglucosaminidase and not a hyaluronidase.

**Authors:** William L. Sheldon\*<sup>†</sup>, Matthew S. Macauley<sup>‡</sup>, Edward J. Taylor<sup>§</sup>, Charlotte E. Robinson\*, Simon J. Charnock\*<sup>1</sup>, Gideon J. Davies<sup>§</sup>, David J. Vocadlo<sup>‡</sup>, Gary W. Black\*<sup>2</sup>

**Addresses:** \*Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, U.K., <sup>†</sup>School of Health, Natural and Social Sciences, University of Sunderland, Sunderland SR1 3SD, U.K., <sup>‡</sup>Department of Chemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada, and <sup>§</sup>York Structural Biology Laboratory, Department of Chemistry, University of York, York YO31 5YW, U.K.

**Short (page heading) title:** Family 84 glycosidases are GlcNAcases and lack hyaluronidase activity

**Key words:** GlcNAcase; mammalian glycoproteins; NAG-thiazoline; *O*-GlcNAc; *O*-GlcNAcase; substrate-assisted catalysis

**Abbreviations used:** GlcNAc, 2-acetamido-2-deoxy- $\beta$ -D-glucopyranose; GlcNAcase,  $\beta$ -*N*-acetylglucosaminidase; MU, 4-methylumbelliferyl; MU-GlcNAc, 4-methylumbelliferyl 2-acetamido-2-deoxy- $\beta$ -D-glucopyranoside; NAG-thiazoline, 1,2-dideoxy-2'-methyl- $\alpha$ -D-glucopyranoso-[2,1-d]- $\Delta$ 2'-thiazoline; *O*-GlcNAc, *O*-glycoprotein 2-acetamido-2-deoxy- $\beta$ -D-glucopyranoside; *O*-GlcNAcase, *O*-glycoprotein 2-acetamido-2-deoxy- $\beta$ -D-glucopyranosidase

<sup>1</sup>Present address: Megazyme International Ireland Limited, Bray, County Wicklow, Republic of Ireland.

<sup>2</sup>To whom correspondence should be addressed: Biomolecular and Biomedical Research Centre, School of Applied Sciences, Northumbria University, Newcastle

upon Tyne NE1 8ST, U.K.; Tel 01912273550; Fax 01912273519; Email  
gary.black@northumbria.ac.uk

## Synopsis

Group A streptococcus (*Streptococcus pyogenes*) is the causative agent of severe invasive infections such as necrotizing fasciitis (the so-called “flesh eating disease”) and toxic shock syndrome. Spy1600, a glycoside hydrolase from family 84 of the large superfamily of glycoside hydrolases, has been proposed to be a virulence factor. Here we show that Spy1600 has no activity toward galactosaminides or hyaluronan, but does remove  $\beta$ -*O*-linked *N*-acetylglucosamine from mammalian glycoproteins; an observation consistent with the inclusion of eukaryotic *O*-GlcNAcases within Glycoside Hydrolase family 84. Proton NMR studies, structure-reactivity studies for a series of fluorinated analogues, and analysis of NAG-thiazoline as a competitive inhibitor reveals that Spy1600 uses a double-displacement mechanism involving substrate-assisted catalysis. Family 84 glycoside hydrolases are therefore comprised of both prokaryotic and eukaryotic  $\beta$ -*N*-acetylglucosaminidases using a conserved catalytic mechanism involving substrate-assisted catalysis. Since these enzymes do not have detectable hyaluronidase activity, many family 84 glycoside hydrolases are most likely incorrectly annotated as hyaluronidases.

## Introduction

Family 84 glycoside hydrolases [1] are a group of enzymes found in higher eukaryotes and pathogenic or symbiotic bacteria. The first family 84 glycoside hydrolase cloned was identified as a tumour associated antigen termed MGEA5 [2]. This enzyme had most likely been earlier isolated as hexosaminidase C [3] and shown to have exo- $\beta$ -*N*-acetylglucosaminidase activity. Further studies [4] using protein purified from rat brain defined this enzyme as *O*-GlcNAcase, a key regulatory enzyme responsible for mediating the removal of *O*-linked *N*-acetylglucosamine (*O*-GlcNAc) from serine and threonine residues of post-translationally modified proteins [5]. The rat brain *O*-GlcNAcase was later cloned and found to be identical to MGEA5 in sequence and it has been proposed that this enzyme has both *O*-GlcNAcase and hyaluronidase activity [6]. Preliminary studies of the family 84 NagH enzyme from *Clostridium perfringens* have also suggested that this enzyme is a hyaluronidase [7]. Based on these studies, several members of this family of enzymes have been annotated as hyaluronidases. Human family 84 *O*-GlcNAcase has been shown to use a substrate-assisted catalytic mechanism to effect removal of *O*-GlcNAc residues from proteins [8], a mechanism in which the 2-acetamido group of the substrate acts as a nucleophile to form a bicyclic oxazoline intermediate (Figure 1). This is in contrast to some other hexosaminidases, such as the bacterial  $\beta$ -*N*-acetylglucosaminidases from glycoside hydrolase family 3 that, instead, use an enzymatic nucleophile [9]. Recently, X-ray structural studies of two bacterial family 84 enzymes have given strong support to the substrate-assisted catalytic mechanism for family 84 enzymes [10, 11] yet despite genome annotations, no unambiguous studies have addressed whether family 84 glycosidases do genuinely process hyaluronan [6].

As part of an ongoing study into virulence factors from the human pathogen group A streptococcus, we are interested in Spy1600, a family 84 glycoside hydrolase from this organism. Group A streptococcus causes severe invasive infections such as necrotizing fasciitis and toxic shock syndrome [12], and it has recently been estimated that these organisms cause 0.5 million deaths each year [13]. The streptococcal family 84 enzyme has been shown to be upregulated during phagocytosis and has,

accordingly, been implicated as a virulence factor [14]. To better understand this putative virulence factor as well as to facilitate the design of potential inhibitors of this enzyme that could eventually prove useful in a clinical setting, we have undertaken a series of studies of this enzyme. Here we describe the cloning, recombinant overexpression and purification, studies on the stereochemical outcome of the enzyme catalyzed reaction, substrate structure-function studies, and inhibition studies. Collectively, these studies show that Spy1600, in common with eukaryotic *O*-GlcNAcases, uses substrate-assisted catalysis. This enzyme is also capable of cleaving *O*-GlcNAc from post-translationally modified proteins, opening the possibility that it may encounter this modification during the lifecycle of the bacterium. Furthermore, despite annotation as a putative hyaluronidase, we show that Spy1600 displays no detectable hyaluronidase activity, suggesting that Glycoside Hydrolase Family 84 is comprised of  $\beta$ -*N*-acetylglucosaminidases and not, as implied by genomic annotations, hyaluronidases.

## Experimental

### Materials

All materials were purchased from Sigma-Aldrich (Poole, Dorset, U.K.), unless stated otherwise. All methods followed the manufacturer's instructions, where given, or were performed as described by Sambrook et al. [15]

### Methods

#### **Amplification and cloning of ORF encoding Spy1600, a family 84 glycoside hydrolase from *Streptococcus pyogenes* M1 GAS SF370**

The amino acid sequence of the family 84 glycoside hydrolase from *Streptococcus pyogenes* M1 GAS SF370 (Spy1600) was checked for the presence of a signal peptide, using the SignalP 3.0 server [16], and common protein domains, using the Pfam database [17]. The coding sequence, *spy1600*, was amplified from genomic DNA from *S. pyogenes* SF370 (ATCC 700294, [18]) using the following primers: 5'-CAT ATG ACC ATT TAT CAC GCC TTA AAG -3' and 5'-GGA TCC GAT TCG TAT CAG AAG GCC AGA AC -3'. The amplified product was cloned into pCR-Blunt (Invitrogen), then subcloned into pET-28a (Novagen) on an *Nde* I – *Bam* HI fragment, and the resulting plasmid designated pSpy1600. The conditions used for amplification were as specified for use with Platinum *Pfx* DNA polymerase (Invitrogen).

#### **Production and purification of Spy1600**

*N*-terminally hexahistidine-tagged Spy1600 was produced via induction of *Escherichia coli* BL21 (DE3; Novagen) cultures carrying pSpy1600 following a protocol described in Brown *et al.* for expression of a different protein [19], except

cultures were grown at 20 °C post-induction. The recombinant protein product was purified via affinity chromatography using nickel-charged Sepharose™ chelating fast flow resin (GE Healthcare) as described in the manufacturer's instructions. The fractions containing pure protein were concentrated, exchanged into 18.2 MΩ cm<sup>-1</sup> H<sub>2</sub>O using a 30 kDa cut-off concentrator unit (Viva Science), and the protein concentration determined according to the method of Bradford, using BSA as a standard [20]. The purity of Spy1600 was judged by SDS-PAGE [21] and static nanospray ion-trap mass spectrometry (ThermoFinnigan LCQ Advantage) of a tryptic digest was used to confirmed the identity of the protein.

### Spy1600 assays

Assays using 4-methylumbelliferyl (MU) glycosides, i.e. 2-(MU) α-D-N-acetylneuraminic acid; MU-α-D-galactoside; MU-α-D-mannopyranoside; MU-β-D-galactoside; MU-β-D-glucuronide; MU-β-D-mannopyranoside; MU-N-acetyl-β-D-glucosaminide (MU-2-acetamido-2-deoxy-β-D-glucopyranoside or MU-GlcNAc); MU-N-acetyl-β-D-galactosaminide; and the three N-fluoroacetyl derivatives of MU-GlcNAc (Figure 2A), MU-2-deoxy-2-fluoroacetamido-β-D-glucopyranoside (MU-GlcNAcF<sub>1</sub>), MU-2-deoxy-2-difluoroacetamido-β-D-glucopyranoside (MU-GlcNAcF<sub>2</sub>) and MU-2-deoxy-2-trifluoroacetamido-β-D-glucopyranoside (MU-GlcNAcF<sub>3</sub>), were performed in triplicate and liberated 4-methylumbelliferone was measured, in real-time, using a Bio-Tek FL600 Fluorescence Microplate Reader. Fluorinated substrates were prepared as previously described [8]. Excitation and emission wavelengths of 360 and 460 nm, respectively were used. Standard assay were performed at 37 °C in a total volume of 100 μl of 20 mM Hepes buffer, pH 7.6, containing 1.0 mg ml<sup>-1</sup> BSA and 2.5 x 10<sup>-4</sup> mg Spy1600. The range of substrate concentrations used for determining *k*<sub>cat</sub> and *K*<sub>M</sub> values was 0.5 – 7.0 mM for MU-GlcNAc, 0.7 – 3.7 mM for MU-GlcNAcF<sub>1</sub>, 0.7 – 3.7 mM for MU-GlcNAcF<sub>2</sub>, and 0.3 – 1.3 mM for MU-GlcNAcF<sub>3</sub>. For the latter three substrates, 5.0 x 10<sup>-4</sup> mg, 1.0 x 10<sup>-3</sup> mg and 1.0 x 10<sup>-1</sup> mg of enzyme was added per 100 μl reaction, respectively. Goodness-of-fit statistical analyses of linear trendlines of the resulting Lineweaver-Burke plots yielded *R*<sup>2</sup> values of ≥ 0.9744. For inhibition assays the range of concentrations of MU-GlcNAc and



NAG-thiazoline (Figure 2B) used were 0.1, 0.3, 0.5, 0.7, 1.0, 2.0 mM, and 0.41, 1.23, 3.70, 11.00, 33.00  $\mu$ M, respectively. Goodness-of-fit statistical analysis of linear trendlines of the resulting Lineweaver-Burke plots produce  $R^2$  values of  $\geq 0.9744$ . The buffer used for determination of the pH-activity optimum was 20 mM Hepes (pH 6.8, 7.0, 7.2, 7.4, 7.5, 7.6, 7.8, 8.0, 8.2). The effect of divalent ions on enzyme activity was determined by adding  $\text{CoCl}_2$ ,  $\text{MgCl}_2$ ,  $\text{MnCl}_2$ ,  $\text{BaCl}_2$ ,  $\text{NiCl}_2$  and  $\text{CaCl}_2$  separately to reactions to yield a final metal ion concentration of 1 mM. Assays for hyaluronidase activity were performed according to the method of Miller [22]. Reactions were incubated for up to 16 h and contained 2 mg  $\text{ml}^{-1}$  Na hyaluronate (Fisher Scientific, UK), 1 mg  $\text{ml}^{-1}$  BSA, 20 mM Hepes at a range of pHs (6.5 to 8.5) and Spy1600 at a range of concentrations ( $7.5 \times 10^{-6}$  to  $5.0 \times 10^{-1}$  mg  $\text{ml}^{-1}$ ). Reactions containing  $5.0 \times 10^{-1}$  mg  $\text{ml}^{-1}$  of Spy1600 were also analysed using high-performance anion-exchange chromatography as described by Price *et al.* [23]. An identical concentration of bovine testes hyaluronidase (Sigma, UK) was used according to the manufacturer's instructions as a positive control.

### **Proton NMR studies of hydrolysis of MU-GlcNAc**

$^1\text{H}$  NMR spectra were collected at ambient probe temperature with a JEOL Eclipse spectrometer operating at 270 MHz. All reactions were performed in  $\text{D}_2\text{O}$  using 5 mm NMR tubes containing 1 ml total volume. Both Spy1600 and MU-GlcNAc were lyophilised and resuspended in  $\text{D}_2\text{O}$  prior to monitoring the reactions. MU-GlcNAc ( $3.3 \text{ mg ml}^{-1}$ ) was hydrolysed using  $2.5 \text{ mg ml}^{-1}$  of Spy1600.

### **Western blot analyses of human proteins treated with Spy1600**

COS-7 cells were cultured to 75 % confluence and cell extract obtained as previously described [8]. The cell extract was incubated at  $37^\circ\text{C}$  for 36 h in the presence of 1 mM phenylmethylsulphonylfluoride, 1 mM  $\beta$ -mercaptoethanol, 10 mM imidazole (to reduce nonspecific binding to Ni-Agarose) and  $2 \text{ mg ml}^{-1}$  of the purified Spy1600 enzyme, where appropriate. Spy1600 was removed prior to Western blotting by two consecutive affinity precipitations using Ni-agarose beads (Qiagen). The samples

were electrophoresed through a 10 % (w/v) polyacrylamide gel, blotted onto nitrocellulose, and probed as previously described [8].

## Results and Discussion

### **Spy1600 is produced in large quantities in *E. coli* and shares a high degree of sequence similarity with the *N*-terminal domain of the human *O*-GlcNAcase**

Spy1600 is predicted to be a non-secretory protein by the SignalP server [16] and contains a single PfamA domain, annotated as Hyaluronidase\_2 [17]. We recombinantly expressed full length ORF, *spy1600*, in *E. coli* and its product, Spy1600, was obtained in large quantities ( $\sim 200 \text{ mg L}^{-1}$ ) as a *N*-terminally hexahistidine-tagged protein. The enzyme was readily purified to homogeneity as judged by SDS-PAGE (Supplementary Figure at <http://www.BiochemJ.org/...>). Spy1600 is significantly smaller than the human family 84 *O*-GlcNAcase, however, they are very similar over a  $\sim 300$  residue stretch that coincides with the location of the Hyaluronidase\_2 PfamA domain, and this level of identity includes two aspartates that have been shown to be the catalytic residues in human *O*-GlcNAcase [24] as well as within the structurally characterised bacterial family 84 enzymes [10, 11] (Figure 3). This observation suggests that the Hyaluronidase\_2 PfamA domain, which these enzymes share in common, likely performs related functions in nature as was first proposed by Hanover [25] and more recently demonstrated by kinetic studies [24, 26, 27].

### **Spy1600 shows no hyaluronidase activity and is only active against $\beta$ -*N*-acetylglucosaminides**

Spy1600 and many other family 84 glycoside hydrolases are annotated in GenBank, GenPept and SWISPROT as putative hyaluronidases. Indeed, eukaryotic *O*-GlcNAcases have been proposed to have hyaluronidase activity on the basis of zymogram-like assays [2]. We, however, were unable to detect any hyaluronidase activity when Spy1600 was incubated with hyaluronan using a reducing sugar assay [22] or using high-performance anion-exchange chromatography [23] to look for expected oligosaccharide breakdown products (Figure 4). Given the detection limits

of these assays, the upper limit of hyaluronidase activity that could be present without being observed here is  $\sim 1.0 \times 10^{-6} \mu\text{mol min}^{-1} \text{mg}^{-1}$ . We also find that Spy1600 shows activity only against  $\beta$ -*N*-acetylglucosaminides, with no detectable activity against  $\beta$ -*N*-acetylgalactosaminides. This difference in epimeric specificity contrasts with the family 20  $\beta$ -hexosaminidases that have nearly equivalent activity towards these substrates [28] and is consistent with other family 84 glycoside hydrolases [4]. We also assayed Spy1600 with several other  $\beta$ - and  $\alpha$ -glycosides of methylumbelliferone (see list in Experimental section) and found that none of these are substrates. Owing to the marked sequence similarity between Spy1600 and the *N*-terminal domain of human family 84 *O*-GlcNAcase we expected that Spy1600 might have the ability to remove *O*-GlcNAc from eukaryotic glycoproteins. Using a COS-7 cell extract and incubating Spy1600 with these glycoproteins we were able to show, using an anti-*O*-GlcNAc antibody in a Western blot analysis, that this enzyme effectively removed *O*-GlcNAc from eukaryotic *O*-GlcNAc modified glycoproteins (Figure 5). This *O*-GlcNAcase activity may be of importance since Spy1600 may, at some point in the life cycle of *S. pyogenes*, encounter *O*-GlcNAc and process this modification either for purposes of foraging or for interfering with the host cellular machinery.

### **Biochemical and biophysical parameters of Spy1600 are similar to those of other GlcNAcases**

The pH-activity optimum for Spy1600 towards MU-GlcNAc is 7.4. When subjected to a 20 min exposure to various temperatures, Spy1600 showed no significant reduction in activity after pre-incubations at 37.6, 46.9, 49.1, 51.4 and 53.1 °C, but was only 33 % active after a pre-incubation at 57 °C, and was completely inactive after a pre-incubation at 68 °C. The percent inhibition of enzyme activity from inclusion of 1 mM divalent metal ions in the assays was as follows: 28 % ( $\text{Ba}^{2+}$ ), 3 % ( $\text{Ca}^{2+}$ ), 28 % ( $\text{Co}^{2+}$ ), 3 % ( $\text{Mg}^{2+}$ ), 20 % ( $\text{Mn}^{2+}$ ) and 46 % ( $\text{Ni}^{2+}$ ). The kinetic parameters,  $k_{\text{cat}}$  and  $K_{\text{M}}$ , for Spy1600 against MU-GlcNAc were  $32 \pm 3 \text{ s}^{-1}$  and  $3.6 \pm 0.3 \text{ mM}$ , respectively, with a  $k_{\text{cat}}/K_{\text{M}}$  value of  $9.0 \text{ s}^{-1} \text{ mM}^{-1}$ . The  $k_{\text{cat}}$  and  $K_{\text{M}}$  values for

Spy1600 against MU-GlcNAc are in the same order of magnitude as those for several other GlcNAcases [8, 9, 29].

### **Spy1600 uses substrate-assisted catalysis**

#### *Stereochemical outcome*

To ascertain the stereochemical outcome of the Spy1600 catalyzed hydrolysis of  $\beta$ -D-glucosaminides the progress of the early stages of the enzyme catalyzed reaction was followed using  $^1\text{H}$  NMR. Because *N*-acetyl-D-glucosamine undergoes mutarotation only slowly in water [30] it is possible to monitor the stereochemistry of the anomeric centre of the first formed product of the reaction. The equilibrium state in water at 20 °C is a 2:1 ratio of  $\alpha$ - to  $\beta$ -anomers for *N*-acetyl-D-glucosamine [30].

The anomeric region of the  $^1\text{H}$  NMR spectrum of a reaction containing MU-GlcNAc at several time points after adding Spy1600 is shown in Figure 6. The data shows that the first product formed is, unambiguously, the  $\beta$ -anomer and only over time does this species slowly mutarotate to form the  $\alpha$ -anomer. The assignment of the resonance at higher chemical shift as the  $\alpha$ -anomer and the lower chemical shift resonance as the  $\beta$ -anomer is consistent with literature values [31]. The  $\alpha$ - and  $\beta$ -anomers are also identifiable from the magnitudes of the vicinal coupling constant, which is larger for the *trans*-diaxial arrangement of the H-1 and H-2 protons of the  $\beta$ -anomer (11 Hz) than that observed for the  $\alpha$ -anomer (4.5 Hz). Further, comparison to a standard of *N*-acetyl-D-glucosamine in solution in which the known equilibrium position is obtained is consistent with these assignments. Spy1600 therefore catalyzes the reaction using a retaining catalytic mechanism, consistent with the mechanism proposed for the human family 84 enzyme [8, 24] as well as with  $\beta$ -hexosaminidases from family 20 of glycoside hydrolases [32].

#### *Substrate-assisted catalysis from the 2-acetamido group*

To provide direct support for a role of the 2-acetamido group in catalysis a series of fluorinated substrate analogues were studied. Four substrates with different levels of fluorine on the *N*-acetyl group were used (Figure 2A). Since increasing the number of fluorine withdrawing groups decreases the nucleophilicity of the carbonyl oxygen it is expected that a commensurate decrease in the rate constant governing the enzyme catalyzed reaction should result only if substrate-assisted catalysis is operative. Assays using these compounds as substrates for Spy1600 revealed Michaelian saturation kinetics for MU-GlcNAc, MU-GlcNAcF<sub>1</sub>, and MU-GlcNAcF<sub>2</sub>, however, the enzymatic activity towards the trifluorinated derivative was too low to be accurately measured using the assay conditions described here (Table 1; Figure 7A). A plot of the log  $V_{\max}[E]_0/K_M$  against the Taft electronic parameter ( $\sigma^*$ ) of the *N*-acetyl substituent shows a steep negative linear correlation on increasing fluorine substitution (Figure 7B). Such negative correlations have been previously observed for glycoside hydrolases from families 20 and 84 that are known to use substrate-assisted catalysis [8, 9, 33], yet remarkably absent for family 3 glycoside hydrolases that use an enzymic nucleophile [8]. Accordingly, these data strongly implicate the 2-acetamido group of the substrate as an intramolecular nucleophile.

#### *Inhibition of Spy1600 using NAG-thiazoline*

NAG-thiazoline (Figure 2B) is a potent inhibitor of human *O*-GlcNAcase [8] as well as  $\beta$ -hexosaminidases from family 20 of glycoside hydrolases [8, 34]. The potency of this inhibitor for these enzymes is attributed to the similarity of this molecule to the intermediate or a structurally related transition state. We expected that if, as we propose, Spy1600 uses a mechanism involving substrate-assisted catalysis NAG-thiazoline should act as a good competitive inhibitor. Indeed, when the compound was assayed as an inhibitor of Spy1600, we observed a clear pattern of competitive inhibition (Figure 8). Non-linear regression analyses reveal a  $K_I$  value for NAG-thiazoline of  $5.82 \pm 0.01 \mu\text{M}$ . Although the inhibitor binds approximately 80-fold less tightly to the streptococcal enzyme than it does to human *O*-GlcNAcase ( $K_I = 0.070 \mu\text{M}$ ) it remains an effective inhibitor of Spy1600. Nevertheless, the observation that

NAG-thiazoline is a good inhibitor of Spy1600 is consistent with the substrate-structure studies described above as well as with the stereochemical outcome experiment and strongly suggests that all members of Glycoside Hydrolase family 84, both eukaryotic and prokaryotic, use a catalytic mechanism involving anchimeric assistance.

## Summary

Spy1600 has been shown to be a family 84 glycoside hydrolase with activity towards *N*-acetylglucosaminides and no activity towards hyaluronan. Catalysis occurs via a substrate-assisted catalytic mechanism in which the *N*-acetamido group of the substrate acts as the nucleophile in a double-displacement reaction. By virtue of the fact that Spy1600 is upregulated during phagocytosis, it has been implicated as a virulence factor [14]. Since this enzyme is unlikely to be secreted, as indicated by the SignalP server [16], we propose that it is involved in the removal of GlcNAc from a variety of glycoconjugates that are imported into the bacterial cell during pathogenesis. Whilst it is formally possible that the enzyme is involved in processing *O*-GlcNAc from endogenous group A streptococcal proteins within the bacterium, this appears unlikely given that no *O*-GlcNAc transferase-like sequence is present in the genome sequence of this strain of group A streptococcus. It is thus far more likely that Spy1600 is involved in simple carbohydrate metabolism; a proposal which is supported by the observation that ORFs adjacent to *spy1600*, and therefore likely to be co-expressed with *spy1600*, encode (a) a putative transcription regulator with similarity to the transcriptional repressor of the arabinose operon, ORF *spy1602*; (b) other glycosidases including a putative  $\beta$ -glucosidase, ORF *spy1599*, and a putative  $\alpha$ -mannosidase, ORF *spy1604*; and (c) putative sugar-binding transport proteins, ORFs *spy1593* and *spy1595*. This genomic organisation strongly supports a role for Spy1600 in glycan foraging, although given its up-regulation during phagocytosis we cannot rule out a direct role in the de-glycosylation of human *O*-GlcNAc proteins for the purpose of compromising the host cell machinery.

### **Acknowledgments**

We thank Justin Perry for help with the analysis of the NMR spectra; Iain Sutcliffe for discussions about the possible role of Spy1600; and Dave Wealleans and Anna Lindsay for technical help with NMR and hyaluronan assays, respectively. We also thank Danielle Chin and Garrett Whitworth for reagents. This research was supported by a National Sciences and Engineering Research Council of Canada (NSERC) Discovery grant to DJV. MSM thanks NSERC and the Michael Smith Foundation for Health Research for fellowships. DJV is supported as a Canada Research Chair in Chemical Glycobiology and as a scholar of the Michael Smith Foundation for Health Research.



## References

- [1] Coutinho, P.M. and Henrissat, B. (1999) Carbohydrate-Active Enzymes server at URL: <http://afmb.cnrs-mrs.fr/CAZY/>.
- [2] Heckel, D., Comtesse, N., Brass, N., Blin, N., Zang, K.D. and Meese, E. (1998) Novel immunogenic antigen homologous to hyaluronidase in meningioma. *Hum. Mol. Genet.* **7**, 1859-1872.
- [3] Braidman, I., Carroll, M., Dance, N., Robinson, D., Poenaru, L., Weber, A., Dreyfus, J.C., Overdijk, B. and Hooghwin, G.J. (1974) Characterization of Human *N*-Acetyl- $\beta$ -Hexosaminidase-C. *FEBS Lett.* **41**, 181-184.
- [4] Dong, D.L. and Hart, G.W. (1994) Purification and characterization of an *O*-GlcNAc selective *N*-acetyl- $\beta$ -D-glucosaminidase from rat spleen cytosol. *J. Biol. Chem.* **269**, 19321-19330.
- [5] Torres, C.R. and Hart, G.W. (1984) Topography and polypeptide distribution of terminal *N*-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for *O*-linked GlcNAc. *J. Biol. Chem.* **259**, 3308-3317.
- [6] Comtesse, N., Maldener, E. and Meese, E. (2001) Identification of a nuclear variant of MGEA5, a cytoplasmic hyaluronidase and a  $\beta$ -*N*-acetylglucosaminidase. *Biochem. Biophys. Res. Commun.* **283**, 634-640.
- [7] Canard, B., Garnier, T., Saint-Joanis, B. and Cole, S.T. (1994) Molecular genetic analysis of the *nagH* gene encoding a hyaluronidase of *Clostridium perfringens*. *Mol. Gen. Genet.* **243**, 215-224.
- [8] Macauley, M.S., Whitworth, G.E., Debowski, A.W., Chin, D. and Vocadlo, D.J. (2005) *O*-GlcNAcase uses substrate-assisted catalysis: kinetic analysis and development of highly selective mechanism-inspired inhibitors. *J. Biol. Chem.* **280**, 25313-25322.
- [9] Vocadlo, D.J. and Withers, S.G. (2005) Detailed comparative analysis of the catalytic mechanisms of  $\beta$ -*N*-acetylglucosaminidases from families 3 and 20 of glycoside hydrolases. *Biochemistry* **44**, 12809-12818.
- [10] Rao, F.V., Dorfmüller, H.C., Villa, F., Allwood, M., Eggleston, I.M. and van Aalten, D.M. (2006) Structural insights into the mechanism and inhibition of eukaryotic *O*-GlcNAc hydrolysis. *EMBO J.* **25**, 1569-1578.
- [11] Dennis, R.J., Taylor, E.J., Macauley, M.S., Stubbs, K.A., Turkenburg, J.P., Hart, S.J., Black, G.W., Vocadlo, D.J. and Davies, G.J. (2006) Structure and mechanism of a bacterial  $\beta$ -glucosaminidase having *O*-GlcNAcase activity. *Nat. Struct. Mol. Biol.* **13**, 365-371.
- [12] Cunningham, M.W. (2000) Pathogenesis of group A streptococcal infections. *Clin. Microbiol. Rev.* **13**, 470-511.
- [13] Carapetis, J.R., Steer, A.C., Mulholland, E.K. and Weber, M. (2005) The global burden of group A streptococcal diseases. *Lancet Infect. Dis.* **5**, 685-694.
- [14] Voyich, J.M., Sturdevant, D.E., Braughton, K.R., Kobayashi, S.D., Lei, B., Virtaneva, K., Dorward, D.W., Musser, J.M. and DeLeo, F.R. (2003) Genome-wide protective response used by group A Streptococcus to evade destruction by human polymorphonuclear leukocytes. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 1996-2001.
- [15] Sambrook, J. and Russell, D.W. (2001) *Molecular Cloning: A Laboratory Manual* (3rd Ed.). Cold Spring Harbor Laboratory Press, New York.
- [16] Bendtsen, J.D., Nielsen, H., von Heijne, G. and Brunak, S. (2004) Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783-795.
- [17] Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L., Studholme, D.J., Yeats, C.

- and Eddy, S.R. (2004) The Pfam protein families database. *Nucleic Acids Res.* **32**, D138-141.
- [18] Ferretti, J.J., McShan, W.M., Ajdic, D., Savic, D.J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A.N., Kenton, S., Lai, H.S., Lin, S.P., Qian, Y., Jia, H.G., Najar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A. and McLaughlin, R. (2001) Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 4658-4663.
- [19] Brown, I.E., Mallen, M.H., Charnock, S.J., Davies, G.J. and Black, G.W. (2001) Pectate lyase 10A from *Pseudomonas cellulosa* is a modular enzyme containing a family 2a carbohydrate-binding module. *Biochem. J.* **355**, 155-165.
- [20] Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248-254.
- [21] Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- [22] Miller, G.L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31**, 426-428.
- [23] Price, K.N., Tuinman, A., Baker, D.C., Chisena, C. and Cysyk, R.L. (1997) Isolation and characterization by electrospray-ionization mass spectrometry and high-performance anion-exchange chromatography of oligosaccharides derived from hyaluronic acid by hyaluronate lyase digestion: Observation of some heretofore unobserved oligosaccharides that contain an odd number of units. *Carbohydr. Res.* **303**, 303-311.
- [24] Cetinbas, N., Macauley, M.S., Stubbs, K.A., Drapala, R. and Vocadlo, D.J. (2006) Identification of Asp174 and Asp175 as the key catalytic residues of human *O*-GlcNAcase by functional analysis of site-directed mutants. *Biochemistry* **45**, 3835-3844.
- [25] Hanover, J.A. (2001) Glycan-dependent signaling: *O*-linked *N*-acetylglucosamine. *FASAB J.* **15**, 1865-1876.
- [26] Ficko-Blean, E. and Boraston, A.B. (2005) Cloning, recombinant production, crystallization and preliminary X-ray diffraction studies of a family 84 glycoside hydrolase from *Clostridium perfringens*. *Acta Cryst.* **F61**, 834-836.
- [27] Kim, E.J., Kang, D.O., Love, D.C. and Hanover, J.A. (2006) Enzymatic characterization of *O*-GlcNAcase isoforms using a fluorogenic GlcNAc substrate. *Carbohydr. Res.* **341**, 971-982.
- [28] Wollen, J.W., Heyworth, R. and Walker, P.G. (1961) Studies on glucosaminidase. 3. Testicular *N*-acetyl- $\beta$ -glucosaminidase and *N*-acetyl- $\beta$ -galactosaminidase. *Biochem. J.* **78**, 111-116.
- [29] Li, H., Morimoto, K., Katagiri, N., Kimura, T., Sakka, K., Lun, S. and Ohmiya, K. (2002) A novel  $\beta$ -*N*-acetylglucosaminidase of *Clostridium paraputrificum* M-21 with high activity on chitobiose. *Appl. Microbiol. Biotechnol.* **60**, 420-427.
- [30] Kuhn, R. and Haber, F. (1953) About the  $\beta$ -form of *N*-acetylglucosamine. *Chem. Ber.* **86**, 722.
- [31] Finar, I.L. (1973) *Organic Chemistry. Fundamental Principles* (Vol. 1, 6th Ed.) Longman, London
- [32] Drouillard, S., Armand, S., Davies, G.J., Vorgias, C.E. and Henrissat, B. (1997) *Serratia marcescens* chitobiase is a retaining glycosidase utilizing substrate acetamido group participation. *Biochem. J.* **328**, 945-949.

- [33] Jones, C.S. and Kosman, D.J. (1980) Purification, properties, kinetics, and mechanism of  $\beta$ -*N*-acetylglucosamidase from *Aspergillus niger*. *J. Biol. Chem.* **255**, 11861-11869.
- [34] Knapp, S., Vocadlo, D., Gao, Z., Kirk, B., Lou, J. and Withers, S. (1996) NAG-thiazoline, An *N*-Acetyl- $\beta$ -hexosaminidase Inhibitor That Implicates Acetamido Participation. *J. Am. Chem. Soc.* **118**, 6804-6805.
- [35] Hansch, C. and Leo, A. (1979) Substituent constants for correlation analysis in chemistry and biology. Wiley, New York.

## Figure legends

**Figure 1. Schematic representation of the retentive double displacement catalytic mechanism used by family 20 and family 84 GlcNAcases.**

**Figure 2. The fluorinated substrates and inhibitor used in this study.** A, structure of the *N*-fluoroacetyl derivatives of MU-GlcNAc ( $n = 3 - 0$ ,  $m = 3 - 0$ ); B, structure of the NAG-thiazoline inhibitor.

**Figure 3. Sequence alignments of characterized family 84 enzymes.** Boxshaded ClustalW 1.82 alignment of the Hyaluronidase\_2 PfamA domain of the structurally characterised family 84 bacterial enzymes, BT4395 from *Bacteroides thetaiotaomicron* VPI-5482 and CPE1234 from *Clostridium perfringens* 13, and the human family 84 enzyme with the *Streptococcus pyogenes* strain SF370 family 84 enzyme, Spy1600. Black background, identical residues; grey background, similar residues; white background, different residues; asterixes, catalytic residues of BT4395, CPE1234 and the human *O*-GlcNAcase. Identities: *B. thetaiotaomicron* v *C. perfringens*, 45 %; *B. thetaiotaomicron* v *S. pyogenes*, 32 %; *B. thetaiotaomicron* v *H. sapiens*, 33 %; *C. perfringens* v *S. pyogenes*, 30 %; *C. perfringens* v *H. sapiens*, 31 %; *S. pyogenes* v *H. sapiens*, 30 %.

**Figure 4. High-performance anion-exchange chromatograms of Na hyaluronate digestions with Spy1600 and bovine testes hyaluronidase.** A, 0 min incubation with bovine testes hyaluronidase; B, 16 h incubation with bovine testes hyaluronidase; C, 0 min incubation with Spy1600; D, 16 h incubation with Spy1600. The hyaluronan tetrasaccharide and hexasaccharide peaks can be seen on the chromatogram of the 16 h Na hyaluronate incubation with bovine testes hyaluronidase. However no such peaks can be seen on the chromatogram of the 16 h Na hyaluronate incubation with Spy1600.

**Figure 5. Western blot of proteins from COS-7 cells incubated in presence and absence of Spy1600.** Top panel, probed with anti-*O*-GlcNAc monoclonal IgM antibody; bottom panel, probed with anti- $\beta$ -actin IgG; lane 1, COS-7 cell extract +

Spy1600; lane 2, COS-7 cell extract only; lane 3, Spy1600 only; lane 4, COS-7 cell extract + boiled Spy1600. The sizes of the marker proteins are indicated (in kDa).

**Figure 6.  $^1\text{H}$  NMR spectra of the hydrolysis of MU-GlcNAc by Spy1600.** The anomeric resonances of MU-GlcNAc,  $\alpha$ -*N*-acetylglucosamine and  $\beta$ -*N*-acetylglucosamine are centered (relative to  $\text{H}_2\text{O}$  at  $\delta$  4.78 ppm) on  $\delta$  5.21 ppm ( $J$  10 Hz),  $\delta$  5.14 ppm ( $J$  4.5 Hz) and  $\delta$  4.65 ppm ( $J$  11 Hz) respectively. A, MU-GlcNAc only; B, MU-GlcNAc + Spy1600 at  $t = 10$  min; C, MU-GlcNAc + Spy1600 at  $t = 1$  h; D, MU-GlcNAc + Spy1600 at  $t = 2$  h; E, MU-GlcNAc + Spy1600 at  $t = 16$  h.

**Figure 7. Activity of Spy1600 with *N*-fluoroacetyl derivatives of MU-GlcNAc.** A, Initial velocity of Spy1600 catalysed hydrolysis of *N*-fluoroacetyl derivatives of MU-GlcNAc: MU-GlcNAc ( $\bullet$ ); MU-GlcNAcF<sub>1</sub> ( $\blacksquare$ ), MU-GlcNAcF<sub>2</sub> ( $\blacktriangle$ ). B, Linear free energy analysis plotting the Taft parameter ( $\sigma^*$ ) of the *N*-fluoroacetyl substituent of MU-GlcNAc substrate analogues against the  $\log V_{\max}[\text{E}]_0 / K_M [\log (\mu\text{mol min}^{-1} \text{mg}^{-1} K_M^{-1})]$  values measured from the substrates as shown in panel A with Spy1600.

**Figure 8. Lineweaver-Burke plot of Spy1600 hydrolysis of MU-GlcNAc in the presence of NAG-thiazoline.** Concentration of NAG-thiazoline: 0.033 ( $\blacklozenge$ ), 0.011 ( $\blacktriangledown$ ), 0.0037 ( $\blacktriangle$ ), 0.00123 ( $\blacksquare$ ), and 0 ( $\bullet$ ) mM.

**Table 1. Michaelis-Menten parameters for Spy1600 a series of 4-methylumbelliferyl 2-*N*-acyl-2-deoxy- $\beta$ -D-glucopyranosides**

Substrate	$\sigma^{*a}$	$K_M$ (mM)	$V_{\max}[E]_0$ ( $\mu\text{mol min}^{-1} \text{mg}^{-1}$ )	$V_{\max}[E]_0/K_M$ ( $\mu\text{mol mM}^{-1} \text{min}^{-1} \text{mg}^{-1}$ )
MU- <i>N</i> -acetyl- $\beta$ -D-glucosaminide	0.0	$3.58 \pm 0.34$	$28.78 \pm 2.70$	8.04
MU-2-deoxy-2-fluoroacetamido- $\beta$ -D-glucopyranoside	0.8	$5.21 \pm 0.17$	$4.95 \pm 0.16$	0.95
MU-2-deoxy-2-difluoroacetamido- $\beta$ -D-glucopyranoside	2.0	$14.07 \pm 0.70$	$0.60 \pm 0.03$	0.04
MU-2-deoxy-2-trifluoroacetamido- $\beta$ -D-glucopyranoside	2.8	ND	ND	ND

<sup>a</sup>The Taft electronic parameters ( $\sigma^*$ ) used for each *N*-acyl substituent were obtained from [35]. ND, no activity detected

Figure 1

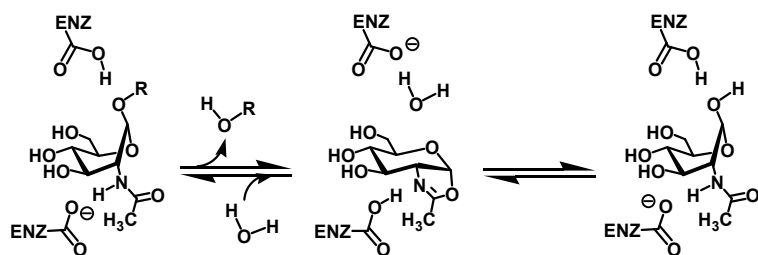


Figure 2

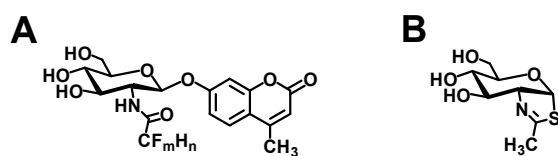


Figure 3

```

B. thetaiotaomicron 151 RGVVEGFYGTPTWVSHQARLSQLRFYGNKKNNTYTYGPKDDPYHSAPNWRLLPYPDKAAQLQELMVAVANENE
C. perfringens      182 RGVVEGFYGTPTWTHKDRLDQIRFYGENKKNNTYTYAPKDDPYHREK-WREPYPENEMQRMQELLDASAEENK
H. sapiens        62  CGVVEGFYGRPWVMEQRKELFRRLQKWEINTYLYAPKDD-YKHRMFWREMYSVEEAEQLMTLLISAAREYE
S. pyogenes      132 RGLIEGFYGTPTWTREEERLDCLRFIGNKRNTYMYAPKDDYQRKL-WRDLYPEDWVTYFKELLAVAKKEEG

B. thetaiotaomicron 221 VDFVWATHHPGQDIKWN---KEERDLRLAKFEKMYQLCVRSFAVFFDDISGEGTNPQKQ----AELLNY
C. perfringens      251 VDFVFGTSPGIDIRFDGEAGEEDFNHLLAKAESLYDMCVRSFAIYWDDIQDK--SAKH----AQVLNR
H. sapiens         131 IEEIYATSPGLDITFSN---PKFVSTLKRKLDQVSQFQCRSFALLFDDIDHNMCAADKEVSSFAHAQVS
S. pyogenes       201 LDFWYMTSPGLDFDYTK---EALYQLLYQKLQQLLALGVCHFGLLLDDIDYQIVDAVERRFKKTAYAQAH
                                     **

B. thetaiotaomicron 282 IDEKFAQVKPD---INQLVMCPTEYNKSWSNPNGN---YLTLGDKLNPSIQIMWTGDRVLS-DITRDGI
C. perfringens      314 FNEEFVKAKGD---VKPLITVPTFYDTGAMVSNQPRTYTRIFAETVDPSLEVMWTGPGVVTNEIPLSDA
H. sapiens         198 ITNEIYQYLGE---PETFLFCPTFYCGTFCYPNVSQSPYLRVTGKLLPGLEVLWTGPKVVSKEIPVESI
S. pyogenes       268 LATEVHHFLNQQAAPELVICPTEYDNHDS-----IYLQELSERLPKEVAFFWTGPTLASQISQADI

B. thetaiotaomicron 345 SWINERIKRPAYIWNVFPVSDYVRD--RLLLGFEVYGNDDTLAK---EMSGFVTNPMHAESSKIAIYSVA
C. perfringens      381 QLISGIYRNMAVWVWNYPVTDYFKG--KLALGEMHGLDKGLNQ---YVDFFTVNPMEHAELSKISHTAA
H. sapiens         265 EEWSKIIRAPVITWDNIHANDYDQK--RLFLGFEYKGRSTELIP---RIKGVLTNPNCFEANYVAIHTLA
S. pyogenes       332 ETMAAVYQRPIITWDNIPVNDYQKDPERLFLTEFANRSPFLCQPDYQVKGIVSNPMSWELSKLITDMS

B. thetaiotaomicron 410 SWAWN-----PAKYDTWQTKDAIRTIILPS--AAEELECFAMHNS---
C. perfringens      446 DWISWN-----MDNYDYDKAWNRAIDMLYGD--LAEDMKVFANHST---
H. sapiens         330 TWYKSNMNGVRKDVVMTDSEDSTVSIQIKLENEG--SDEDETDVLVSPQMA
S. pyogenes       402 HMLWD-----ANRYQPSHSWLETITDYTEDELALALQAFWHNG---
```



Figure 4

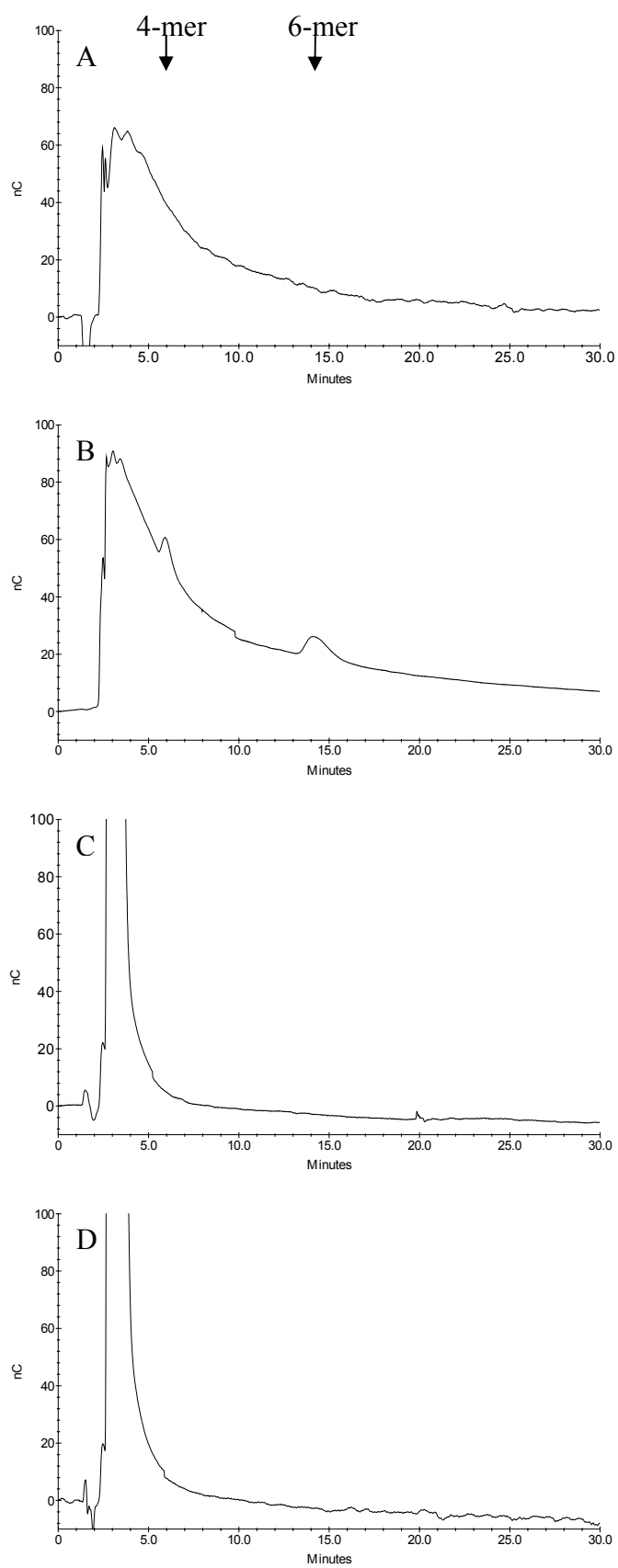


Figure 5

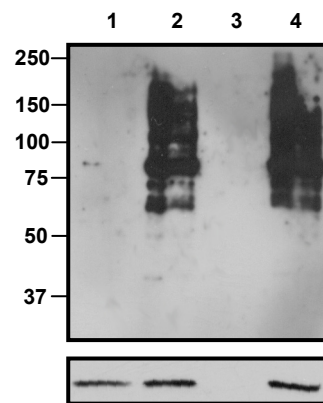


Figure 6

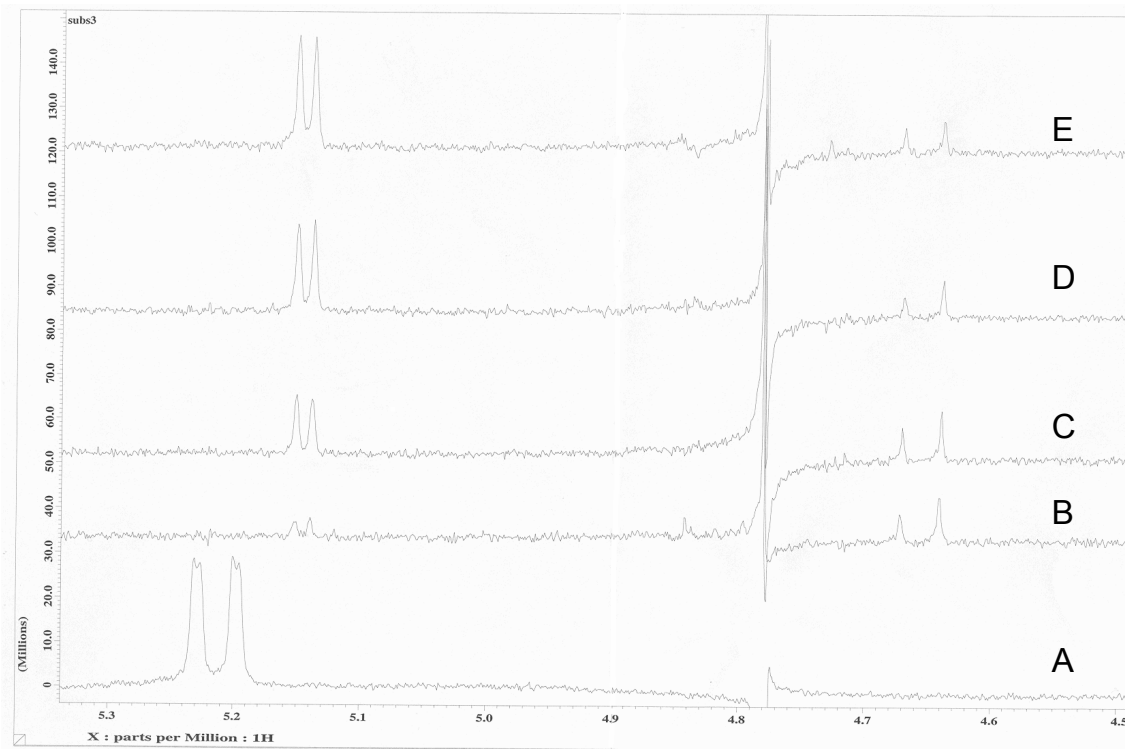


Figure 7B

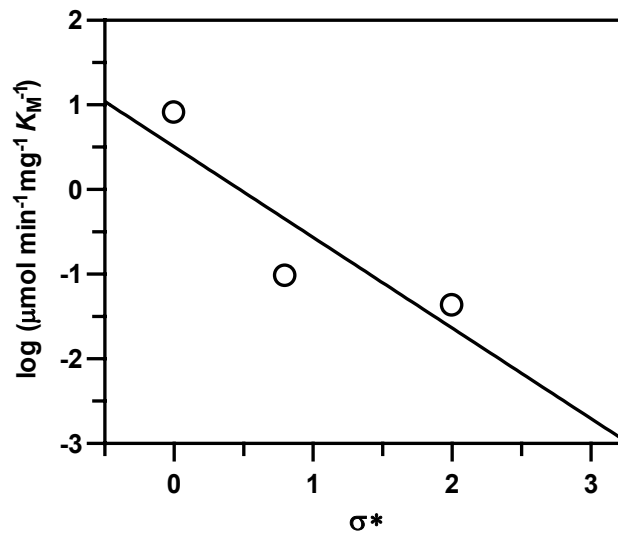


Figure 7A

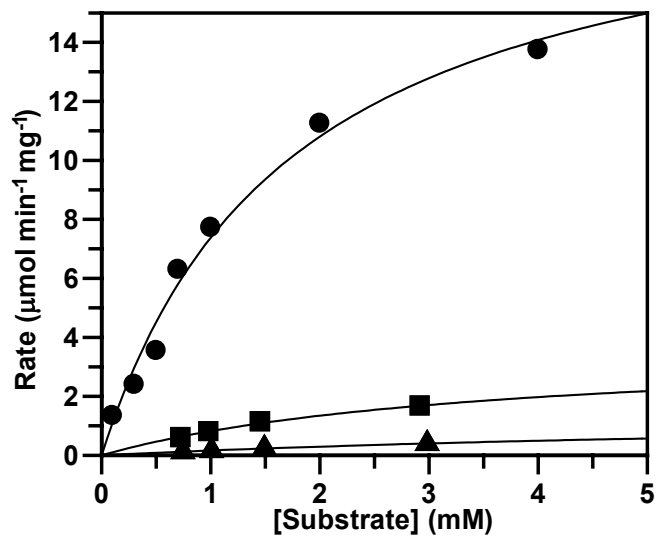


Figure 8

