

Endohexosaminidase-catalyzed synthesis of glycopeptides and proteins*

Antony J. Fairbanks

Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Abstract: The synthetic application of endohexosaminidase enzymes (e.g., Endo A, Endo M, Endo D) promises to allow ready access to a wide variety of defined homogeneous glycoproteins and glycopeptides. The use of *N*-glycan oligosaccharides that are activated at the reducing terminus as oxazolines allows their high-yielding attachment to almost any amino acid, peptide, or protein that contains a GlcNAc residue as an acceptor. A wide variety of oxazoline donors are readily available, either by total synthesis or by isolation of the corresponding oligosaccharide from natural sources and then conversion to the oxazoline in water. The synthetic potential of the enzymes is particularly augmented by the production of mutant glycosynthases, the use of which allows the synthesis of a wide variety of glycopeptides and glycoproteins bearing defined homogeneous *N*-glycan structures.

Keywords: carbohydrates; endohexosaminidases; enzyme catalysis; *N*-glycans; glycopeptides; glycoproteins; peptides; proteins.

INTRODUCTION

Glycosylation of proteins is the most diverse form of post-translational modification. It is now well established that glycosylation can play a key role in protein folding and can crucially affect important protein properties such as conformation and stability, susceptibility to proteases, and circulatory lifetime. Moreover, protein glycosylation is also fundamental to other key biological processes, and is widely implicated in many aspects of cell–cell signalling, development, and immune response [1]. Protein glycosylation occurs most commonly at the amide nitrogen of particular asparagine (Asn) residues, so-called *N*-linked glycosylation. In the alternative *O*-linked glycosylation, the attachment of glycans occurs through the side chain of serine or threonine residues. Amongst the many important mammalian proteins that are *N*-glycosylated, antibodies are of particular therapeutic and commercial importance. Indeed, monoclonal antibodies (mAbs) are certainly the most important class of therapeutic protein. Human IgG antibodies contain *N*-linked glycans at the two Asn-297 residues of the dimeric Fc (constant fragment) region, and extensive work has confirmed that the precise identity of the glycans at these sites is key for modulating antibody-mediated responses; thus, correct glycosylation of mAbs is important for therapeutic efficacy. *N*-Linked glycosylation can also be of particular significance for mammalian enzymes. In the context of therapeutic agents, recombinant glycoproteins used for enzyme replacement therapies are of particular significance [2], and again glycosylation can play an important role in the efficacy of treatment.

*Pure Appl. Chem. **85**, 1759–1900 (2013). A collection of invited papers based on presentations at the 26th International Carbohydrate Symposium (ICS 2012), Madrid, Spain, 22–27 July 2012.

Since the biosynthesis of the glycan portions of these glycoconjugates is not under direct genetic control, then mAbs and therapeutic glycoproteins are produced as heterogeneous and frequently inseparable mixtures of glycoforms—materials in which different oligosaccharide structures are linked to the same peptide chain. Access to pure single glycoforms of glycoproteins has therefore now become a major scientific objective, not only to facilitate more precise biological investigations into the effects different glycans have on peptide and protein properties, but also for the production and administration of optimized mAbs and therapeutic glycoproteins. Additionally, glycosylation may beneficially affect peptide properties, such as increasing stability to protease cleavage and reducing aggregation. Therefore, the production and biological investigation of specifically glycosylated peptides as improved versions of current peptide therapeutics has also recently become an area of scientific endeavor [3].

There are numerous approaches that can be taken to attempt to access glycoproteins and glycopeptides in homogeneous form, ranging from bioengineering of protein expression systems to total synthesis, and many of the developments in these areas have been covered in recent reviews [4]. The focus of our work in the area has been the use of endohexosaminidases, or more specifically endo- β -*N*-acetylglucosaminidases (ENGases), to attach glycan structures to peptides and proteins that bear a GlcNAc residue at *N*-linked glycosylation sites (Fig. 1).

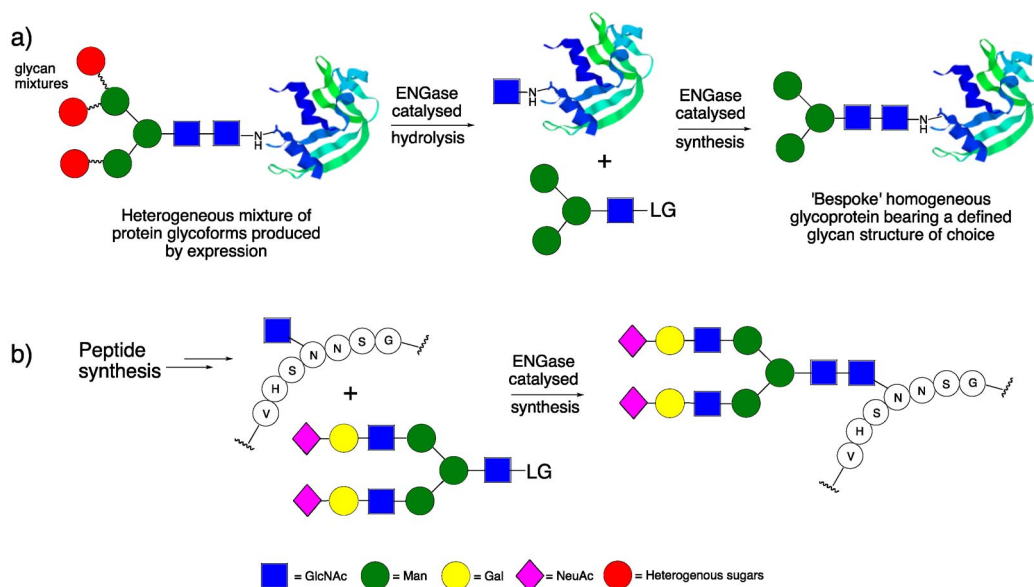


Fig. 1 The use of endohexosaminidase (ENGase) enzymes to catalyze: (a) glycoprotein remodeling; (b) convergent glycopeptide synthesis.

Endohexosaminidases are a class of enzyme first identified in the early 1970s, the members of which specifically cleave the chitobiose core [GlcNAc β (1-4)GlcNAc] of *N*-linked glycans between the two *N*-acetyl glucosamine residues. It was established in the 1990s that these enzymes could in fact be used to synthesize *N*-glycan structures that they normally hydrolyze by transglycosylation reactions using full-length *N*-glycans as glycosyl donors. However, the efficiency of these processes was limited, and typically required a very large excess of acceptor. Taking this idea as a starting point one could imagine that a substantial advantage could be gained by the use of a truncated donor activated at the reducing terminus with a suitable anomeric leaving group (LG) in order to increase the efficiency of the kinetically controlled reaction. One could therefore imagine a two-step procedure for in vitro glycoprotein remodeling (Fig. 1a) in which firstly the oligosaccharide chains of a heterogeneous mixture of

protein glycoforms produced by expression are “trimmed” back to single GlcNAc residues at *N*-linked glycosylation sites using an endohexosaminidase-catalyzed hydrolysis. Then in a second step a defined glycan is attached to this GlcNAc “handle”, again in a process catalyzed by an endohexosaminidase, but using an activated *N*-glycan as the glycosyl donor. In a similar fashion, the same procedure could be applied to synthetic peptides that bear GlcNAc residues (Fig. 1b). Herein, endohexosaminidase-catalyzed addition of glycan structures should lead to a highly efficient and convergent method to access defined glycopeptides. In order to develop this idea further, access is required to both endohexosaminidases in order to catalyze the process, and a variety of activated glycosyl donors as substrates to undergo glycosylation.

ENDOHEXOSAMINIDASES

Wild-type (WT) enzymes

The endohexosaminidases fall into one of two of the CAZy families [5] of the glycosyl hydrolases, namely, family GH18 or family GH85. Some of the enzymes that have so far been studied for synthetic potential are listed in Table 1.

Table 1 Endohexosaminidases examined for synthetic utility.

Enzyme	Source	GH family	Glycan structure hydrolyzed	Core fucose?	Catalytic residues	Oxazolines processed?	Core fucose tolerated for oxazoline transfer?
Endo H	<i>Streptomyces plicatus</i>	18	High mannose/hybrid	Yes	E132 D130	No	No
Endo F ₁	<i>Flavobacterium meningosepticum</i>	18	High mannose/hybrid	No	E132 D130	Yes	Marginally
Endo F ₂	<i>Flavobacterium meningosepticum</i>	18	Biantennary complex	Yes	–	Yes	Required
Endo F ₃	<i>Flavobacterium meningosepticum</i>	18	Bi-/triantennary complex	Yes	E128 D126	Yes	Required
Endo S	<i>Streptococcus pyogenes</i>	18	Complex (IgG Fc specific)	Yes	E235 D233	Yes	Yes
Endo A	<i>Arthrobacter protophormiae</i>	85	High mannose/hybrid	No	E173 N171	Yes	No
Endo M	<i>Mucor hiemalis</i>	85	High mannose/hybrid and biantennary complex	No	E177 N175	Yes	No
Endo D	<i>Streptococcus (Diplococcus) pneumoniae</i>	85	Core structures only	Yes	E324 N322	Yes	No

Although each enzyme basically carries out the same hydrolytic reaction, they have different tolerances as to the precise structure of the *N*-glycan that can be hydrolyzed. Several of these enzymes, such as Endo H [6], found widespread utility as biochemical tools to carry out this “normal” hydrolytic process [7]. However, studies in the 1990s revealed that in particular two of the family GH85 endo-

hexosaminidases, Endo M from *Mucor Hiemalis* [8] and Endo A from *Arthrobacter protophormiae* [9], showed useful *synthetic* transglycosylation activity [10,11].

All of the enzymes are retaining glycosidases that hydrolyze substrates via a two-step mechanism involving general acid/base catalysis. The family GH18 enzymes all contain two carboxylic acid residues in the active site, a glutamic acid and an aspartic acid. In contrast, the family GH85 enzymes contain only one carboxylic acid residue, a glutamic acid; the other key catalytic residue is an Asn. Regardless of whether the active site contains one or two carboxylic acids, the hydrolytic mechanism is now generally considered to proceed via neighboring group participation (NGP) of the 2-acetamide residue, as shown in Fig. 2, which depicts a hydrolytic mechanism for the family GH85 enzymes. Either an oxazoline or an oxazolinium ion is implicated as a high-energy intermediate in this process.

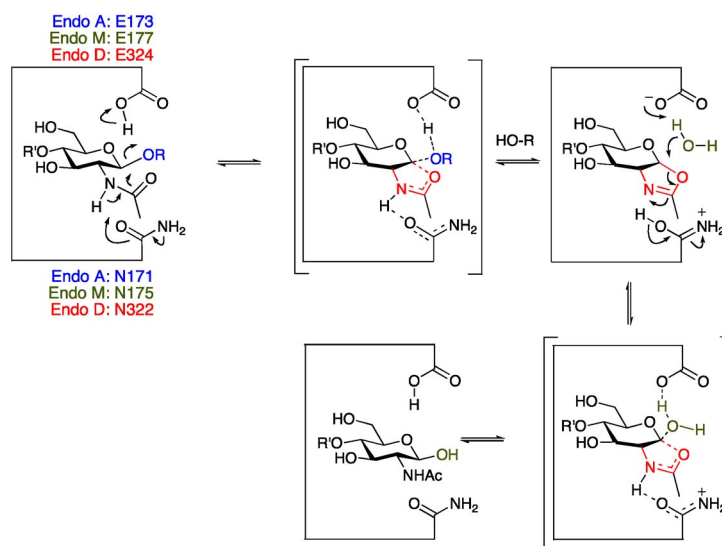
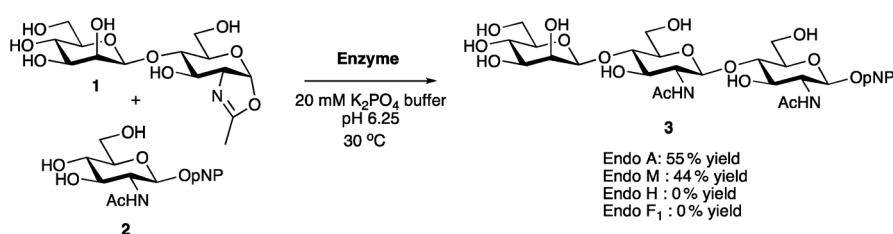


Fig. 2 Catalytic mechanism for family GH85 enzymes involving NGP.

As far back as 1998 we had proposed that *N*-glycan glycosyl oxazolines may serve as substrates for these enzymes. However, a dearth of research funding meant that this hypothesis could not be tested. It was only in 2001 that Shoda and co-workers published their first seminal report in the field [12] and demonstrated that a disaccharide oxazoline **1** could be used as an activated glycosyl donor substrate for some of these enzymes (Scheme 1). In particular, Endo A and Endo M catalyzed the synthetic glycosylation of both GlcNAcOpNP **2** and a dansyl-Asn(GlcNAc)-OH glycosyl amino acid (not shown) by the disaccharide oxazoline **1**, to produce the corresponding trisaccharide, in the former case, **3**. However, neither Endo H nor Endo F₁, both family GH18 enzymes, catalyzed this synthetic reaction.

This initial report immediately implied to us [13] that other oxazolines should also be substrates for these enzymes. It also stimulated the interest of others active in the field, most notably Lai-Xi Wang, who simultaneously began a similar research program, and who has subsequently been the most prolific proponent of the approach [14]. These initial results also seemed to perhaps imply that only family GH85 enzymes would be useful synthetically. Indeed, later studies performed by us on Endo H, and also on a variety of Endo H mutants, indicated that this enzyme did not catalyze any synthetic reactions using oxazolines as donors. It was therefore the family GH85 enzymes that attracted the vast majority of attention from the synthetic community for a considerable period of time. However, very recently it



Scheme 1 Shoda's demonstration of the use of oxazolines as glycosyl donors for endohexosaminidase-catalyzed glycosylation.

has become clear that family GH18 enzymes will actually also catalyze synthetic reactions using oxazolines as donors, and consequently attention is now being diverted back towards them [15].

Mutant enzymes

Whenever glycosidases are used for synthetic purposes, either in “true” transglycosylation reactions, or when using activated glycosyl donors, such as oxazolines, then product hydrolysis can represent a significant limitation. For all of these kinetically controlled processes the key to synthetic success is ensuring that the rate of synthesis is much higher than that of product hydrolysis. However, it generally holds true for the endohexosaminidases that as the size of the oligosaccharide is increased then the rate of product hydrolysis increases, whilst there is correspondingly little increase in the rate of the synthetic reaction. A point is therefore reached at which no effective synthetic reaction can be achieved using WT enzymes, even using activated donors such as oxazolines.

A potential solution to this problem is to access mutant enzymes by site-directed mutagenesis, in which the hydrolytic activity of the enzymes has been curtailed, but which actually still display synthetic ability using oxazolines as activated glycosyl donors. Work in this respect builds upon the use of “glycosynthases”, as originally developed by Withers [16] and Planas [17]. Two approaches have been developed with respect to the family GH85 endohexosaminidases—by the groups of Wang/Yamamoto, and also by ourselves, and each is shown in Fig. 3.

The term “glycosynthase” was first applied by Withers to retaining glycosidases, in which the catalytic nucleophile had been replaced with a nonparticipating residue, for example, alanine. The Wang/Yamamoto approach to producing mutant glycosynthases of the endohexosaminidases is in essence the same. However, a difference in this case is that family GH85 enzymes do not contain a nucleophilic residue as the mechanism operates by NGP. Wang and Yamamoto therefore chose to mutate out the Asn residue that stabilizes/deprotonates the oxazolinium intermediate, reasoning that removal of this residue would effectively achieve the same end; without this Asn, NGP could not operate, and so the enzyme could not effect product hydrolysis. The first mutant they reported was a mutant of Endo M in which residue Asn175 was replaced by alanine (Fig. 3a), and this was a reasonably effective synthase enzyme [14f]. Later studies in which all other possible mutations were made at this site revealed that the Asn-to-glutamine mutation was the most effective in terms of synthetic efficiency (Fig. 3b), and this E175Q mutant is now available commercially. It is clear that a simple one-carbon extension effectively curtails hydrolytic activity whilst maintaining the ability of these enzymes to accept oxazolines as donors. Wang has subsequently performed the analogous mutations on a variety of other family GH85 enzymes including Endo A [14i] and Endo D [18], to produce a variety of synthetically useful synthase mutants. Very recently, Wang has extended these studies to family GH18 enzymes, and has shown that replacement of the aspartic acid active site residue in Endo S by either alanine or glutamine, also produces useful glycosynthase mutants [15c].

Our own work, originally performed simultaneously as that of Wang and Yamamoto, focused on the production of mutants of Endo A, since the Endo A plasmid was available. Rather than try to cur-

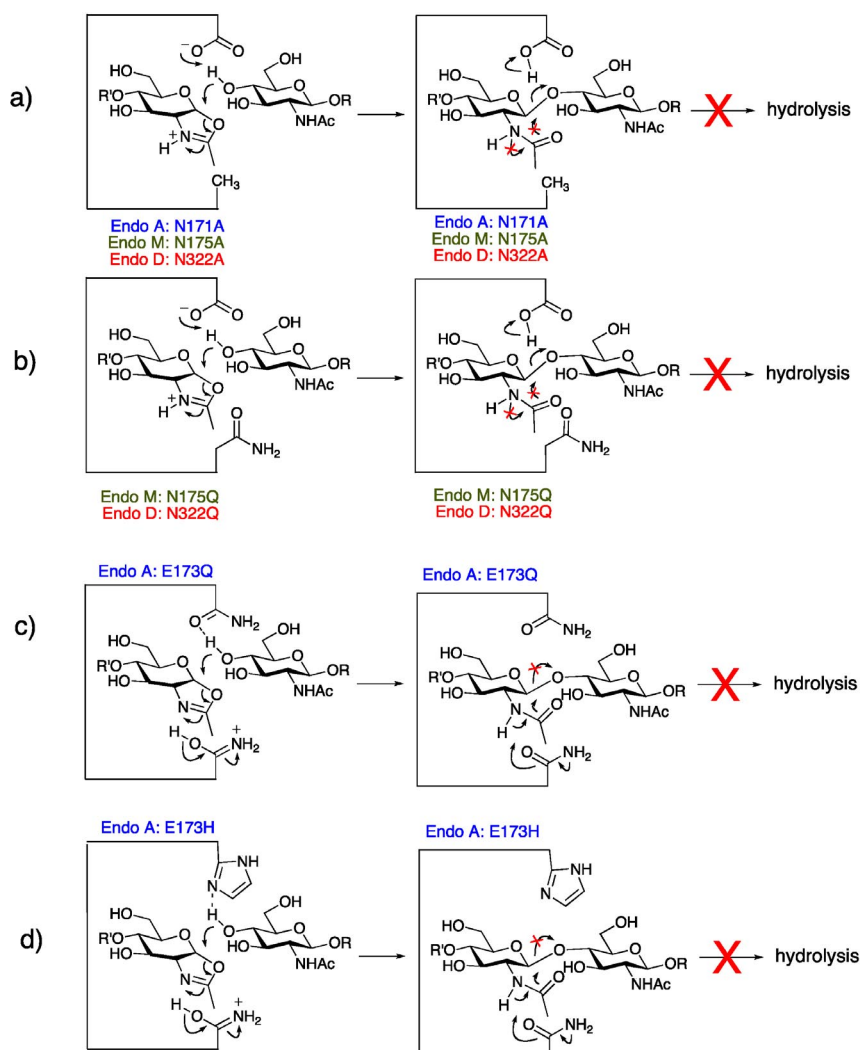


Fig. 3 Design of mutant glycosynthases of family GH85 endohexosaminidases.

tail hydrolytic activity by interfering with the NGP we chose to produce mutant enzymes in which the key catalytic acid residue of Endo A, glutamic acid 173, was replaced (Figs. 3c and 3d). Our reasoning was that such mutants may still be able to process oxazolines as activated glycosyl donors. However, without a proton donor residue at position 173 all hydrolytic activity should be curtailed since glycosidic bonds cannot be broken without prior protonation of the anomeric oxygen. Two mutants of Endo A were made in which E173 was replaced by glutamine and histidine; potential hydrogen bond acceptors were chosen as the replacements for glutamic acid as this residue normally acts as a general base to direct the incoming nucleophile during the second step of the “normal” hydrolytic reaction. The E173Q mutant of Endo A was successfully crystallized and its structure solved [19]; this was the first structure to be solved [20] for a family GH85 enzyme and deposited at the PDB. This structure, along with another structure solved for WT Endo A [20a], provided confirmation of the identity and spatial arrangement of essential active site amino acids, such as residues E173Q, N171, and the proximity of hydrophobic residues Y205, W216, and W244 that had previously been identified as being important

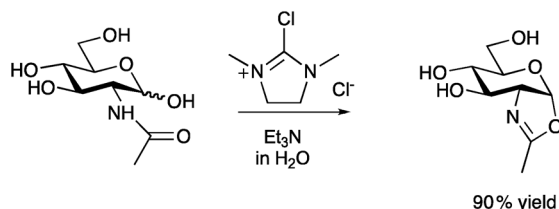
for transglycosylation activity. Moreover, comparison of the mutant and WT crystal structures revealed that the E-to-Q mutation had in fact done little to alter the active site geometry. Another GH85 structure, that for Endo D, was solved very shortly afterwards [20b], and although no crystal structure of Endo M is yet available, these structures provide other key insights that may be used to optimize the synthetic activity of the GH85 enzymes by future protein engineering. Further site-directed mutagenesis studies have also revealed that mutations to other active residues can increase the synthetic efficiency of these enzymes, primarily by decreasing the rate of direct competitive oxazoline hydrolysis. For example, the Y217F mutant of Endo M and the Y205F mutant of Endo A both show increased rates of synthetic glycosylation [14f]; however, such mutants are actually still capable of hydrolyzing the product.

ACCESS TO *N*-GLYCAN OXAZOLINES

In order to extend and develop the synthetic utility of the endohexosaminidase enzymes beyond the original Shoda report [12], access is required to a wide variety of *N*-glycan oxazolines. These can be produced either by semisynthesis using *N*-glycan oligosaccharides derived from natural sources, or by total chemical synthesis.

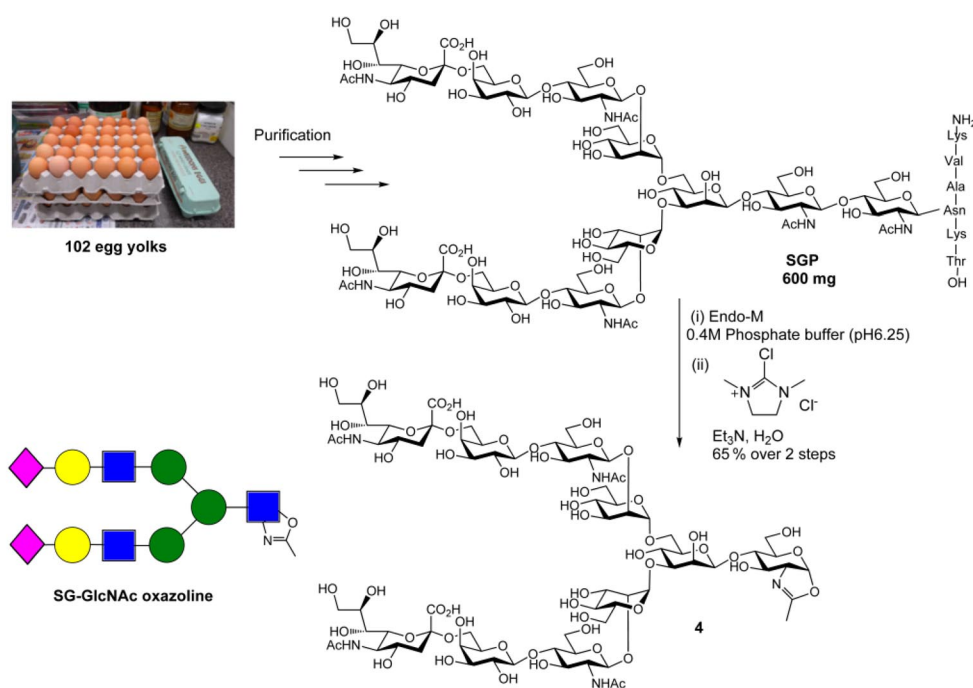
Isolation from natural sources and semisyntheses

Access to *N*-glycan oxazolines by semisynthesis has the clear advantages that significant amounts of material can be made available, provided that an abundant source of the oligosaccharide can be found, and that an efficient purification protocol can be established. Furthermore, a recent key methodological development, again reported by Shoda and co-workers [21], has demonstrated that any oligosaccharide terminating in a GlcNAc residue may be directly converted into the corresponding oxazoline in water in very good yield. As shown in Scheme 2, treatment of the free reducing sugar with 2-chloro-1,3-dimethylimidazolinium chloride (DMC) in the presence of triethylamine in water directly produces the corresponding oxazoline. The efficiency of this transformation means that this has now become the method of choice for oxazoline formation, regardless of whether the oligosaccharide itself has been produced by isolation or by total synthesis.



Scheme 2 Shoda's methodology for directly accessing glycosyl oxazolines from reducing sugars in water [21].

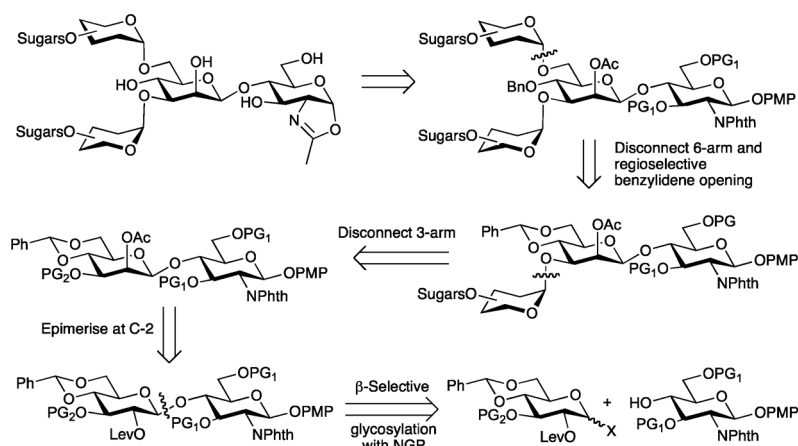
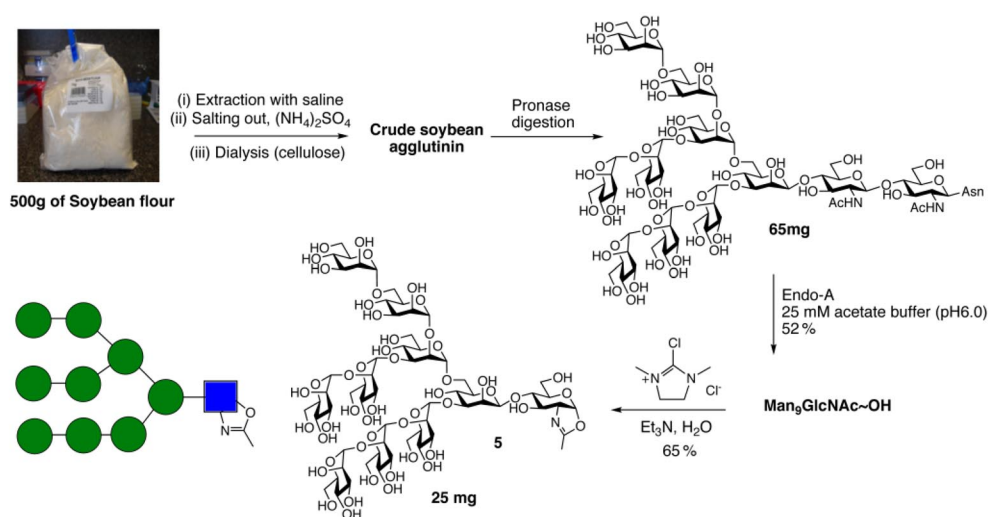
At this point in time, two *N*-glycan oxazolines are readily available by semisynthesis. Firstly, a complex biantennary *N*-glycan can be isolated from egg yolks, as shown in Scheme 3. The purification processes are extensive [22] and involve deproteinization (90 % phenol, Et₂O wash) followed by two rounds of size exclusion chromatography (Sephadex G-50 and then Sephadex G-25) collecting sialic acid positive fractions. This is then followed firstly by anion-exchange chromatography (Sephadex DEAE eluting with NaCl), which removes non-sialylated glycans, and then by cation-exchange chromatography (Sephadex C-25). Finally, a further round of size exclusion chromatography is performed (Sephadex G-25) to remove salt, to yield the pure sialylglycopeptide (SGP). Approximately 600 mg of



Scheme 3 Isolation and production of a sialic acid terminated complex biantennary oxazoline **4** from egg yolks.

material (commercial value ~US\$60 K) is available from ~100 eggs, making this route very practicable to produce gram quantities of oligosaccharide if required. The free oligosaccharide can then be released from SGP by enzymatic hydrolysis using WT Endo M; the product is purified by size exclusion chromatography (Sephadex G-25). Finally, the free oligosaccharide can be converted into the corresponding oxazoline **4** using the Shoda methodology by reaction with DMC in water.

A high-mannose glycan containing nine mannose units can be readily isolated from soybean flour [23], and converted into the corresponding oxazoline. As shown in Scheme 4, this process produces smaller amounts of material than the egg isolation procedure. Crude soybean flour is treated with a mixture of acetic acid and saline (pH 4.6), centrifuged, and the insoluble residue is removed. Saturated ammonium sulfate is added to reach a total of 40 % v/v, causing precipitation. The supernatant is taken, and further saturated ammonium sulfate is added to reach a total of 55 % v/v, again producing an insoluble residue. The supernatant is removed and discarded, and water is added to the residue. Dialysis (cellulose) is then used to remove remaining salts to produce crude soybean agglutinin (~9 g from 500 g of soybean flour). Digestion by pronase [24], and purifying the product by size exclusion chromatography (Sephadex G-50) and then high-performance liquid chromatography (HPLC) gives the pure high-mannose-type Asn-linked oligosaccharide (~65 mg from 8.75 g of crude agglutinin). Hydrolysis by treatment with WT Endo A produces the Man₉GlcNAc oligosaccharide, which can then be converted into the corresponding oxazoline **5** with DMC.



removal of protecting groups. Secondly, separate disconnections of the 3- and 6-branches of the central mannose unit allows the synthesis of nonsymmetrical glycans, and can be readily accomplished by the use of 4,6-benzylidene protection. Finally, formation of the β -manno linkage can be readily achieved by the C-2 epimerization approach, and in our hands the most efficient and generally applicable procedure for this has been that reported by Schmidt [27] and Boons [28], in which Lev protection is used for the 2-position of a *gluco* donor. The great utility of the 2-O-Lev protection is that Lev can be efficiently removed in a highly chemoselective fashion, something which did not prove possible for alternative ester protection of the 2-position due to competitive partial cleavage of the phthalimido protection of the glucosamine unit. Other important factors are the use of triflate as the LG, and sonication to promote the reaction as reported by Fürstner [29]. When all of these things are taken together the procedure is high yielding and very reliable.

Using these strategies, the *N*-glycans shown in Fig. 5 have been accessed by our group. Furthermore, “non-natural” epimeric versions of these glycans, in which the C-2 epimerization process was not performed, gave access to non-natural *N*-glycan structures containing a central glucose residue. It latterly transpired that these non-natural structures were accepted as substrates by a variety of the endohexosaminidases, and it was generally found that the products of these reactions were not hydrolyzed by the enzymes [13b]. The utility of synthetic approach for the total synthesis of a truncated complex biantennary oxazoline is shown in Scheme 5.

The key *gluco* 4,6-benzylidene protected building block **6**, in which the 2-OH is Lev protected and the 3-OH is allyl protected, was glycosylated with a GlcNAc acceptor **7**. Lev removal and epimerization gave the desired β -manno disaccharide **8**. Removal of the allyl group allowed glycosylation at OH-3 with the disaccharide thioglycoside donor **9**. Regioselective reductive ring opening of the 4,6-benzylidene with triethylsilane and dichlorophenylborane revealed the 6-hydroxyl group, which was then followed by a second glycosylation with disaccharide donor **9** to give the truncated complex *N*-glycan hexasaccharide **10**. Obviously, the use of a different glycosyl donor in this second glycosylation step allows the synthesis of *N*-glycans bearing different structures on the 3- and 6-branches of the central mannose. Protecting group manipulations, including replacement of phthalamide by acetamide, hydrogenation and acetylation, removal of the *para*-methoxyphenyl anomeric protection, oxazoline formation, and finally base-catalyzed deacetylation, yielded *N*-glycan oxazoline **11**.

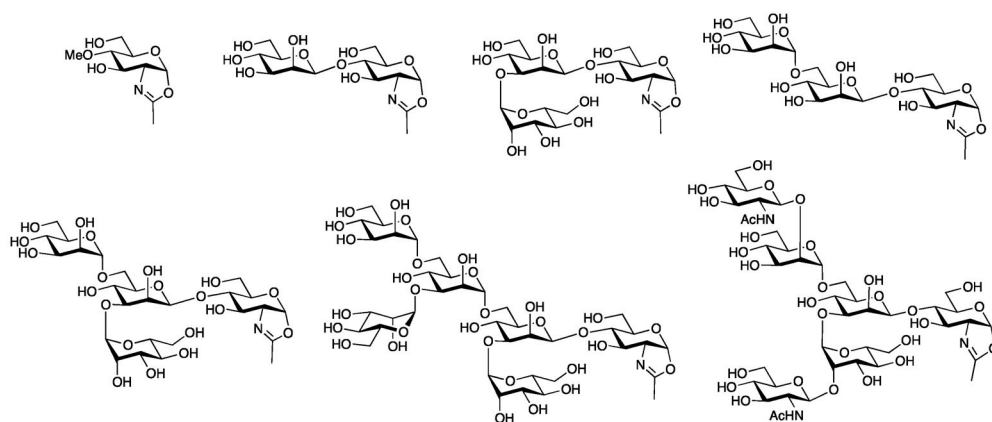
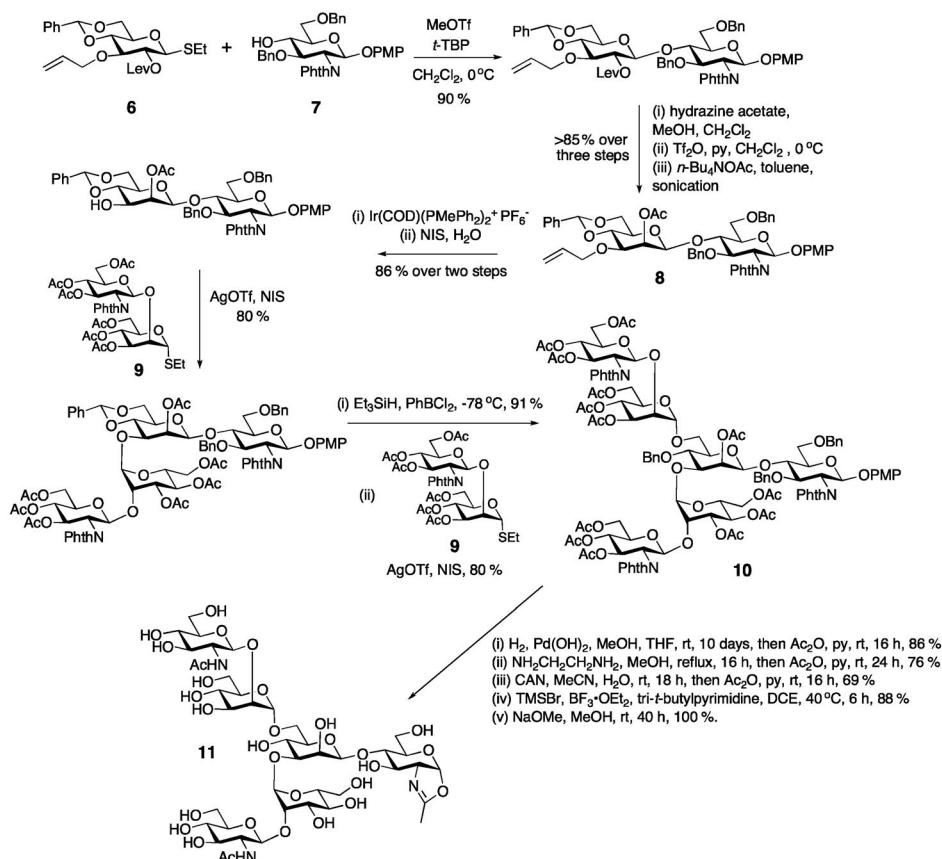


Fig. 5 *N*-Glycans accessed by total synthesis.



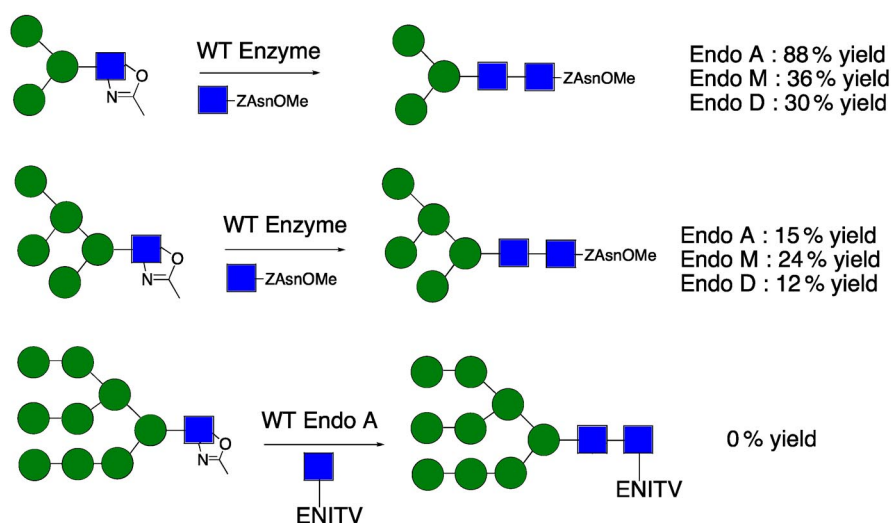
Scheme 5 Total synthesis of a truncated complex biantennary *N*-glycan oxazoline.

GLYCOSYLATION REACTIONS WITH OXAZOLINES AS DONORS

Wild-type enzymes

The family GH85 WT enzymes display a useful range of synthetic abilities with oxazolines as donors. Acceptor substrates that can be used may range enormously in size from monosaccharides, to glycosyl amino acids, to glycopeptides, even to intact glycoproteins bearing a GlcNAc residue. A survey of the numerous publications in the field reveals that the following general conclusions can be reached. With small oligosaccharide structures, the rate of product hydrolysis may be so slow as to be negligible, and so excellent synthetic efficiency can be achieved in many cases simply by running the reaction for an appropriate length of time (Scheme 6). This generally holds true up to and even slightly above the tetrasaccharide level; for example, WT Endo A was used to remodel the Fc region of human IgG1 with tetrasaccharides [14i]. However, as the oligosaccharide structure is increased in size, then product hydrolysis becomes more problematic. With high-mannose hexasaccharides as donors, only modest yields of product can be obtained using either Endo A or Endo M. Furthermore, using WT Endo A and the Man₉ decasaccharide donor then essentially no product can be isolated.

Another competing reaction that decreases synthetic yield is that of direct oxazoline hydrolysis. The rate of this reaction is highly enzyme-dependent, and there is little one can do to curtail its effect other than to use an excess of donor. Thus, with Endo D reaction yields are generally poorer than those using Endo A or Endo M when a limited quantity of donor is used; a large excess of donor needs to be added to obtain acceptable yields. Thankfully, this excess of donor is not wasted as the hydrolysis prod-



Scheme 6 Examples of glycosylation reactions with WT endohexosaminidases and high-mannose-type oxazolines.

uct can be reconverted into the oxazoline using the Shoda methodology [21]. Broadly speaking, the enzymes will only perform synthetic reactions on donors and acceptors that correspond to hydrolytic products that are produced by that enzyme. Thus, Endo D will only transfer oxazolines with an exposed 2-hydroxyl group in the α -mannose residue which is linked by an $\alpha(1-3)$ linkage to the inner mannose of the core structure. This effectively limits the utility of Endo D to the synthesis of glycoconjugates bearing the core *N*-glycan pentasaccharide. Similarly, Endo A will not process oxazolines derived from full-length complex *N*-glycans. Likewise, neither Endo A nor Endo M will glycosylate acceptor substrates in which a “core” fucose is present on the 6-OH of the GlcNAc acceptor. However, it should be noted that there are some discrepancies. For example, WT Endo D will not glycosylate fucosylated acceptors using oxazolines as donors, even though the WT enzyme does hydrolyze the products of these reactions [15a]. This perplexing result is actually the reverse of the “normal exception”, i.e., the finding that oxazolines can be processed to give products which that particular enzyme cannot hydrolyze. This discrepancy is something that can be used advantageously from a synthetic perspective. For example, as mentioned previously, a number of non-natural glycans containing a central glucose unit are processed by these enzymes to give synthetic products that are not hydrolyzed [13b]. A caveat to this apparently highly useful substrate controlled “irreversibility” is that as glycans are made larger these *gluco*-containing materials do become hydrolytic substrates. Likewise, WT Endo A will process non-natural *N*-glycan oxazolines that mimic truncated complex biantennary structures to synthesize products that the enzyme cannot hydrolyze [14e]; the natural hydrolytic activity of the enzyme is only towards high-mannose and hybrid *N*-glycans. WT Endo A will also process oxazolines containing a bisecting GlcNAc [14o]. Thus, by using an oxazoline as a highly activated glycosyl donor a significant synthetic advantage can be gained in terms of achieving an “irreversible” reaction. Finally, a comment is worth making about the GH18 enzymes. Although Shoda et al. had reported [12] that Endo H was not capable of effecting glycosylation using the disaccharide donor **1**, we investigated whether larger oxazolines could act as donor substrates. Indeed, the transglycosylation activity of Endo H had been previously demonstrated with full-length *N*-glycans [30]. As mentioned previously, we found Endo H to be inactive with all of the oxazolines that we studied. However, recent publications by the groups of Wang [15a,c] and Davis [15b] indicate that several of the other GH18 enzymes will indeed use oxazolines as substrates for glycosylation (Table 1). It may therefore eventually transpire that Endo H is a somewhat exceptional member of family GH18.

Mutant enzymes

The real synthetic utility of the endohexosaminidases becomes obvious when the mutant enzymes are used. For example, Fig. 6 shows a comparison of the synthetic glycosylation of dRNase B, a commercially available bovine RNase B that had been enzymatically trimmed back by treatment with Endo H to a homogeneous protein glycoform bearing a GlcNAc residue at the sole *N*-linked glycosylation site, with the Man₃GlcNAc tetrasaccharide oxazoline donor using either WT Endo A, or the E173H mutant. In both cases, the enzyme was able to effect production of a single glycoform, Man₃GlcNAc₂ RNase B. However, the time course study indicates that not only did the WT Endo A catalyze hydrolysis of the product glycoprotein, but also that the maximum obtainable yield was better when the E173H mutant was used.

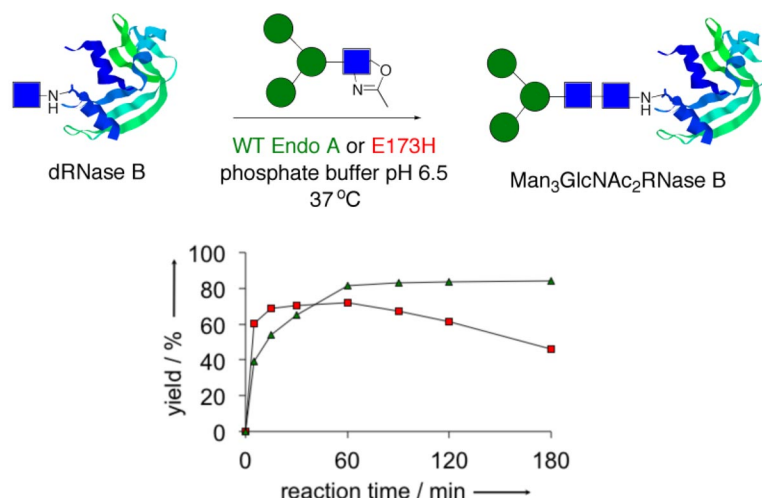
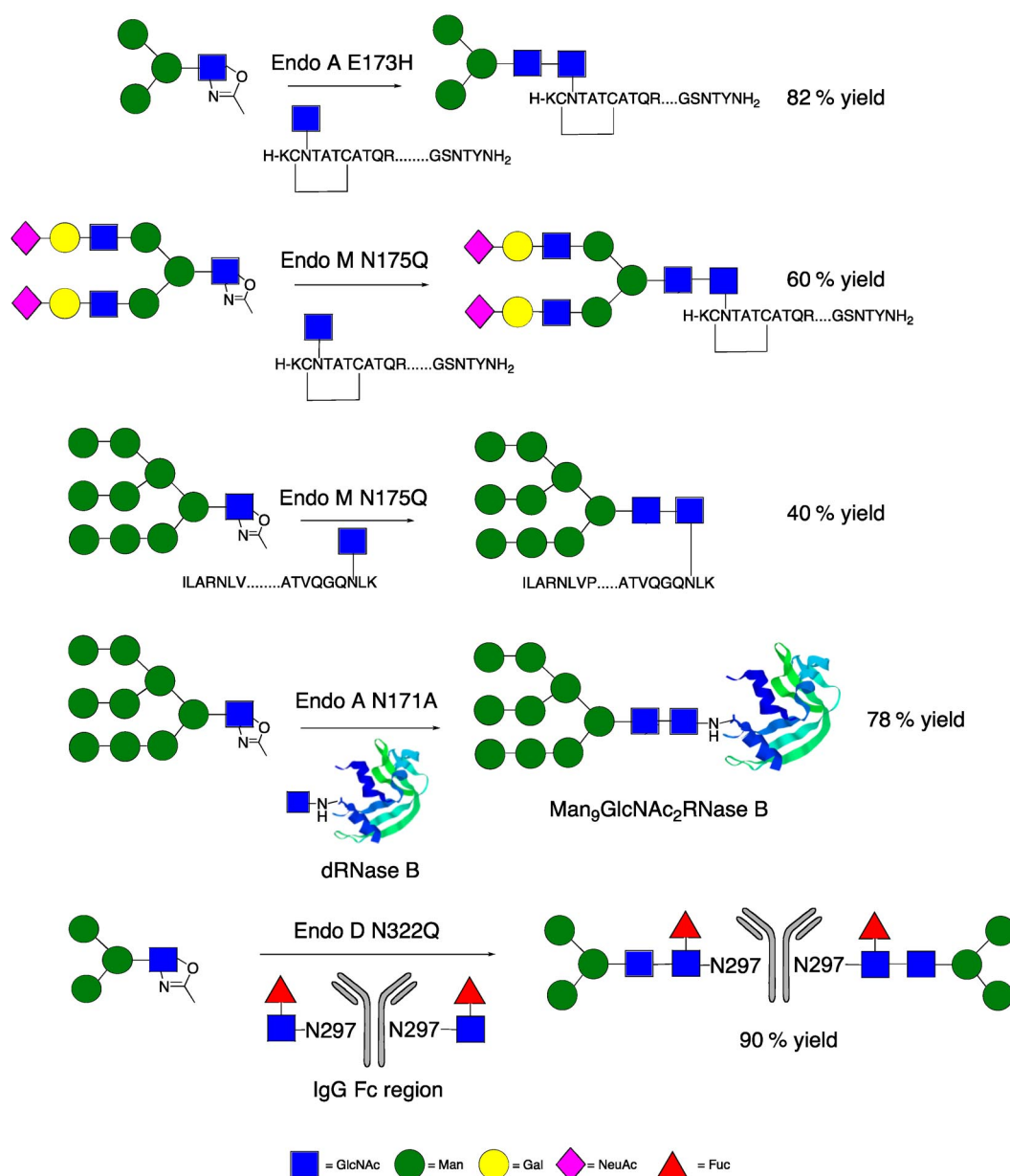


Fig. 6 Production of a Man₅GlcNAc₂ glycoform of RNase B by WT Endo A (■) and the E173H mutant (▲).

We have subsequently used the E173H mutant of Endo A as a catalyst for the attachment of the Man₃GlcNAc tetrasaccharide oxazoline to a variety of synthetic glycopeptides containing up to 37 amino acids, in yields that range from 70 to 95 % (Scheme 7). However, interestingly the Man₉GlcNAc oxazoline was not found to be a substrate for the E173H mutant of Endo A, thus limiting application of this enzyme to truncated high-mannose glycans. In contrast, the N171A mutant of Endo A shows real synthetic utility for the synthesis of glycopeptides and proteins bearing full-length high-mannose glycans [14l,p] (Scheme 7). The corresponding N175A mutant of Endo M will catalyze the transfer of biantennary complex glycans to a range of acceptor substrates [14l]. However, it is the N175Q mutant of Endo M that appears to currently be the most versatile enzyme; this enzyme possesses enhanced synthetic activity as compared to the N175A mutant, and is able to transfer both full-length high-mannose and full-length complex biantennary glycans onto a range of acceptors [14n]. Indeed, we have used this now commercially available N175Q Endo M mutant to produce a range of glycopeptides bearing either complex biantennary or Man₉ high-mannose glycans (Scheme 7).

Although Endo A, Endo M, and Endo D are all family GH85 enzymes, there are subtle differences between the activities of the WT and mutant enzymes of each. For example, a report by Yamamoto and Wang indicated that the E177H and E177Q mutants of Endo M, which are analogous to the mutants of E173H and E173Q mutants of Endo A made by us, were not synthetically active [14n]. Additionally, Wang reported attempted application of the Endo M N175A and N175Q mutants to the Fc region of IgG antibodies. However, in this case it was found that neither of the Endo M mutants was



Scheme 7 Some examples of glycopeptide and glycoprotein synthesis using mutant endohexosaminidases.

capable of transferring even truncated oxazolines to the Fc domain; in contrast, it was WT Endo A that was actually found to be the enzyme of choice to achieve such remodeling, but only with truncated structures [14o]. In subsequent studies, Wang has gone on to demonstrate that the N322Q mutant of Endo D was capable of remodeling an IgG-Fc domain with the Man₃GlcNAc tetrasaccharide (Scheme 7) [18]. However, the narrow substrate specificity of Endo D precludes its application to larger oligosaccharide structures. Therefore, the most recent developments in the field of mAb remodeling have focused on WT, D233A, and D233Q mutants of Endo S, a member of the family GH18 hydrolases, which appears to be highly specific for the Fc region of IgGs. In these most recent studies, Wang

has now demonstrated remodeling of rituximab, a therapeutic mAb, with full-length complex biantennary structures [15c].

CONCLUSIONS AND FUTURE PERSPECTIVES

The upshot of these developments is that it is now possible to routinely add full-length high-mannose or complex biantennary glycans to almost any peptide in which a GlcNAc residue has been installed to act as an acceptor. This methodology can be extended to fully folded proteins, and very recent studies have even successfully remodeled an intact therapeutic mAb with full-length glycans in this manner. However, future challenges remain. For example, none of the enzymes so far isolated are able to hydrolyze tetraantennary complex glycans, and so one can imagine they similarly are not able to catalyze their attachment to peptides or proteins; yet these glycans are optimal for the *in vivo* function of some glycoprotein therapeutics such as EPO. Other challenges remain around access to *N*-glycan oxazolines—total synthesis is a logistically challenging endeavor, and so the development of a semi-synthetic route to tetraantennary complex glycans from natural sources would also be highly beneficial.

ACKNOWLEDGMENTS

I sincerely thank all of the members of my group, especially Thomas Rising, Thomas Parsons, and Yusuke Tomabechei, together with all of my collaborators, especially James Moir, Ed Taylor, Al Boraston, David Vocadlo, and Margaret Brimble, for their industry and invaluable assistance. I also gratefully acknowledge financial support from the Royal Society of New Zealand (Marsden Fund), the BBSRC (UK), and the EPSRC (UK).

REFERENCES

1. (a) D. F. Wyss, J. S. Choi, J. Li, M. H. Knoppers, K. J. Willis. *Science* **269**, 1273 (1995); (b) J. D. Aplin, J. C. Wriston. *CRC Crit. Rev. Biochem.* **10**, 259 (1981); (c) T. Misaizu, S. Matsuki, T. W. Strickland, M. Takeuchi, A. Kobata, S. Takasaki. *Blood* **86**, 4097 (1995); (d) A. Varki. *Glycobiology* **3**, 97 (1993); (e) H.-J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, D. Solís. *Trends Biochem. Sci.* **36**, 298 (2011).
2. M. Rohrbach, J. T. R. Clarke. *Drugs* **67**, 2697 (2007).
3. (a) T. Ueda, T. Ito, K. Tomita, H. Togame, M. Fumoto, K. Asakura, T. Oshima, S.-I. Nishimura, K. Hanasaki. *Bioorg. Med. Chem. Lett.* **20**, 4631 (2010); (b) T. Ueda, K. Tomita, Y. Notsu, T. Ito, M. Fumoto, T. Takakura, H. Nagatome, A. Takimoto, S.-I. Mihara, H. Togame, K. Kawamoto, T. Iwasaki, K. Asakura, T. Oshima, K. Hanasaki, S.-I. Nishimura, H. Kondo. *J. Am. Chem. Soc.* **131**, 6237 (2009).
4. (a) J. R. Rich, S. G. Withers. *Nat. Chem. Biol.* **5**, 206 (2009); (b) L.-X. Wang, J. V. Lomino. *Chem. Biol.* **7**, 110 (2012).
5. <http://www.cazy.org/fam/GH85.html>
6. (a) A. L. Tarentino, F. Maley. *J. Biol. Chem.* **249**, 811 (1974); (b) A. L. Tarentino, T. H. Plummer, F. Maley. *J. Biol. Chem.* **249**, 818 (1974); (c) R. J. Trumbly, P. W. Robbins, M. Belfort, F. D. Ziegler, F. Maley, R. B. Trimble. *J. Biol. Chem.* **260**, 5683 (1985).
7. F. Maley, R. B. Trimble, A. L. Tarentino, T. H. Plummer. *Anal. Biochem.* **180**, 195 (1989).
8. (a) S. Kadowaki, K. Yamamoto, M. Fujisaki, H. Kumagai, T. Tochikura. *Agric. Biol. Chem.* **52**, 2387 (1988); (b) S. Kadowaki, K. Yamamoto, M. Fujisaki, K. Izumi, T. Tochikura, T. Yokoyama. *Agric. Biol. Chem.* **54**, 97 (1990); (c) K. Yamamoto, S. Kadowaki, J. Watanabe, H. Kumagai. *Biochem. Biophys. Res. Commun.* **203**, 244 (1994); (d) K. Haneda, T. Inazu, K. Yamamoto, H. Kumagai, Y. Nakahara, A. Kobata. *Carbohydr. Res.* **292**, 61 (1996); (e) K. Yamamoto, K. Fujimori, K. Haneda, M. Mizuno, T. Inazu, H. Kumagai. *Carbohydr. Res.* **305**, 415 (1998); (f)

- M. Mizuno, K. Haneda, R. Iguchi, I. Muramoto, T. Kawakami, S. Aimoto, K. Yamamoto, T. Inazu. *J. Am. Chem. Soc.* **121**, 284 (1999); (g) K. Fujita, K. Kobayashi, A. Iwamatsu, M. Takeuchi, H. Kumagai, K. Yamamoto. *Arch. Biochem. Biophys.* **432**, 41 (2004).
9. (a) K. Takegawa, M. Nakoshi, S. Iwahara, K. Yamamoto, T. Tochikura. *Appl. Environ. Microbiol.* **55**, 3107 (1989); (b) K. Takegawa, S. Yamaguchi, A. Kondo, I. Kato, S. Iwahara. *Biochem. Int.* **25**, 829 (1991); (c) K. Takegawa, M. Tabuchi, S. Yamaguchi, A. Kondo, I. Kato, S. Iwahara. *J. Biol. Chem.* **270**, 3094 (1995); (d) J.-Q. Fan, K. Takegawa, S. Iwahara, A. Kondo, I. Kato, C. Abeygunawardana, Y. C. Lee. *J. Biol. Chem.* **270**, 17723 (1995); (e) K. Takegawa, K. Yamabe, K. Fujita, M. Tabuchi, M. Mita, H. Izu, A. Watanabe, Y. Asada, M. Sano, A. Kondo, I. Kato, S. Iwahara. *Arch. Biochem. Biophys.* **338**, 22 (1997); (f) K. Fujita, K. Takegawa. *Biochem. Biophys. Res. Commun.* **283**, 680 (2001).
 10. For some other leading references on synthetic applications of Endo A see: (a) J.-Q. Fan, M. S. Quesenberry, K. Takegawa, S. Iwahara, A. Kondo, I. Kato, Y. C. Lee. *J. Biol. Chem.* **270** 17730 (1995); (b) K. Takegawa, K. Fujita, J.-Q. Fan, M. Tabuchi, N. Tanaka, A. Kondo, H. Iwamoto, I. Kato, Y. C. Lee, S. Iwahara. *Anal. Biochem.* **257**, 218 (1998); (c) K. Fujita, N. Tanaka, M. Sano, I. Kato, Y. Asada, K. Takegawa. *Biochem. Biophys. Res. Commun.* **267**, 134 (2000).
 11. For some other leading references on synthetic applications of Endo M, see: (a) K. Yamamoto, K. Fujimori, K. Haneda, M. Mizuno, T. Inazu, H. Kumagai. *Carbohydr. Res.* **305** 415 (1997); (b) K. Haneda, T. Inazu, M. Mizuno, R. Iguchi, K. Yamamoto, H. Kumagai, S. Aimoto, H. Suzuki, T. Noda. *Bioorg. Med. Chem. Lett.* **8**, 1303 (1998); (c) T. Yamanoi, M. Tsutsumida, Y. Oda, E. Akaike, K. Osumi, K. Yamamoto, K. Fujita. *Carbohydr. Res.* **339**, 1403 (2004); (d) K. Osumi, Y. Makino, E. Akaike, T. Yamanoi, M. Mizuno, M. Noguchi, T. Inazu, K. Yamamoto, K. Fujita. *Carbohydr. Res.* **339**, 2633 (2004); (e) T. Yamanoi, N. Yoshida, Y. Oda, E. Akaike, M. Tsutsumida, N. Kobayashi, K. Osumi, K. Yamamoto, K. Fujita, K. Takahashi, K. Hattori. *Bioorg. Med. Chem. Lett.* **15**, 1009 (2005); (f) K. Haneda, M. Takeuchi, M. Tagashira, T. Inazu, K. Toma, Y. Isogai, M. Hori, K. Kobayashi, M. Takeuchi, K. Takegawa, K. Yamamoto. *Carbohydr. Res.* **341**, 181 (2006); (g) Y. Makimura, S. Watanabe, T. Suzuki, Y. Suzuki, H. Ishida, M. Kiso, T. Katayama, H. Kumagai, K. Yamamoto. *Carbohydr. Res.* **341**, 1803 (2006); (h) D. Murakami, Y. Shimada, S. Kamiya, K. Yamazaki, Y. Makimura, K. Ito, N. Minamiura, K. Yamamoto. *Arch. Biochem. Biophys.* **477**, 299 (2008).
 12. M. Fujita, S.-i. Shoda, K. Haneda, T. Inazu, K. Takegawa, K. Yamamoto. *Biochim. Biophys. Acta* **1528**, 9 (2001).
 13. (a) T. W. D. F. Rising, T. D. W. Claridge, N. Davies, D. P. Gamblin, J. W. B. Moir, A. J. Fairbanks. *Carbohydr. Res.* **341**, 1574 (2006); (b) T. W. D. F. Rising, T. D. W. Claridge, J. W. B. Moir, A. J. Fairbanks. *ChemBioChem* **7**, 1177 (2006); (c) T. W. D. F. Rising, C. D. Heidecke, J. W. B. Moir, Z. Ling, A. J. Fairbanks. *Chem.—Eur. J.* **14**, 6444 (2008); (d) C. D. Heidecke, Z. Ling, N. C. Bruce, J. W. B. Moir, T. B. Parsons, A. J. Fairbanks. *ChemBioChem* **9**, 2045 (2008); (e) T. B. Parsons, J. W. B. Moir, A. J. Fairbanks. *Org. Biomol. Chem.* **7**, 3128 (2009); (f) C. D. Heidecke, T. B. Parsons, A. J. Fairbanks. *Carbohydr. Res.* **344**, 2433 (2009); (g) T. B. Parsons, D. J. Voadlo, A. B. Boraston, A. J. Fairbanks. *Org. Biomol. Chem.* **8**, 1861 (2010); (h) A. J. Fairbanks. *Comptes Rendus Chimie* **14**, 44 (2011).
 14. (a) B. Li, Y. Zeng, S. Hauser, H. J. Song, L.-X. Wang. *J. Am. Chem. Soc.* **127**, 9692 (2005); (b) H. Li, B. Li, H. Song, L. Breydo, I. V. Baskakov, L.-X. Wang. *J. Org. Chem.* **70**, 9990 (2005); (c) L.-X. Wang, H. J. Song, S. W. Liu, H. Lu, S. B. Jiang, J. H. Ni, H. G. Li. *ChemBioChem* **6**, 1068 (2005); (d) Y. Zeng, J. S. Wang, B. Li, S. Hauser, H. G. Li, L.-X. Wang. *Chem.—Eur. J.* **12**, 3355 (2006); (e) B. Li, H. Song, S. Hauser, L.-X. Wang. *Org. Lett.* **8**, 3081 (2006); (f) M. Umekawa, W. Huang, B. Li, K. Fujita, H. Ashida, L.-X. Wang, K. Yamamoto. *J. Biol. Chem.* **283**, 4469 (2008); (g) L.-X. Wang. *Carbohydr. Res.* **343**, 1509 (2008); (h) W. Huang, H. Ochiai, X. Zhang, L.-X. Wang. *Carbohydr. Res.* **343**, 2903 (2008); (i) Y. Wei, C. Li, W. Huang, B. Li, S. Strome, L.-X. Wang. *Biochemistry* **47**, 10294 (2008); (j) H. Ochiai, W. Huang, L.-X. Wang. *J. Am. Chem.*

- Soc.* **130**, 13790 (2008); (k) H. Ochiai, W. Huang, L.-X. Wang. *Carbohydr. Res.* **344**, 592 (2009); (l) W. Huang, C. Li, B. Li, M. Umekawa, K. Yamamoto, X. Zhang, L.-X. Wang. *J. Am. Chem. Soc.* **131**, 2214 (2009); (m) W. Huang, X. Zhang, T. Ju, R. D. Cummings, L.-X. Wang. *Org. Biomol. Chem.* **8**, 5224 (2010); (n) M. Umekawa, C. Li, T. Higashiyama, W. Huang, H. Ashida, K. Yamamoto, L.-X. Wang. *J. Biol. Chem.* **285**, 511 (2010); (o) G. Zou, H. Ochiai, W. Huang, Q. Yang, C. Li, L.-X. Wang. *J. Am. Chem. Soc.* **133**, 18975 (2011); (p) M. N. Amin, W. Huang, R. M. Mizanur, L.-X. Wang. *J. Am. Chem. Soc.* **133**, 14404 (2011).
15. (a) W. Huang, J. Li, L.-X. Wang. *ChemBiochem* **12**, 932 (2011); (b) J. J. Goodfellow, K. Baruah, K. Yamamoto, C. Bonomelli, B. Krishna, D. J. Harvey, M. Crispin, C. N. Scanlan, B. G. Davis. *J. Am. Chem. Soc.* **134**, 8030 (2012); (c) W. Huang, J. Giddens, S.-Q. Fan, C. Toonstra, L.-X. Wang. *J. Am. Chem. Soc.* **134**, 12308 (2012).
16. L. F. Mackenzie, Q. R. Wang, R. A. J. Warren, S. G. Withers. *J. Am. Chem. Soc.* **120**, 5583 (1998).
17. C. Malet, A. Planas. *FEBS Lett.* **440**, 208 (1998).
18. S.-Q. Fan, W. Huang, L.-X. Wang. *J. Biol. Chem.* **287**, 11272 (2012).
19. M. D. L. Suits, Z. Ling, R. J. Bingham, N. C. Bruce, G. J. Davies, A. J. Fairbanks, J. W. B. Moir, E. J. Taylor. *J. Mol. Biol.* **389**, 1 (2009); PDB accession number 2vtf, Deposition Date 14 May 2008.
20. Two other family GH85 structures were reported almost simultaneously: (a) WT Endo A: J. Yin, L. Li, N. Shaw, Y. Li, J. K. Song, W. Zhang, C. Xia, R. Zhang, A. Joachimiak, H.-C. Zhang, L.-X. Wang, Z.-J. Liu, P. Wang. *PLoS ONE* **4**, e4658 (2009); (b) Endo D: D. W. Abbott, M. S. Macauley, D. J. Vocadlo, A. B. Boraston. *J. Biol. Chem.* **284**, 11676 (2009).
21. M. Noguchi, T. Tanaka, H. Gyakushi, A. Kobayashi, S.-i. Shoda. *J. Org. Chem.* **74**, 2210 (2009).
22. A. Seko, M. Koketsu, M. Nishizono, Y. Enoki, H. R. Ibrahim, L. R. Juneja, M. Kim, T. Yamamoto. *Biochim. Biophys. Acta* **1335**, 23 (1997).
23. (a) D. L. Evers, R. L. Hung, V. H. Thomas, K. G. Rice. *Anal. Biochem.* **265**, 313 (1998); (b) I. E. Liener. *Arch. Biochem. Biophys.* **54**, 223 (1955); (c) L.-X. Wang, J. Ni, S. Singh, H. Li. *Chem. Biol.* **11**, 127 (2004).
24. H. Lis, N. Sharon. *J. Biol. Chem.* **253**, 3468 (1978).
25. (a) C. Unverzagt. *Angew. Chem., Int. Ed. Engl.* **36**, 1989 (1997); (b) C. Unverzagt, J. Seifert. *Tetrahedron Lett.* **41**, 4549 (2000); (c) C. Unverzagt. *Chem.—Eur. J.* **9**, 1369 (2003); (d) X. Schratt, C. Unverzagt. *Tetrahedron Lett.* **46**, 691 (2005); (e) R. Schuberth, C. Unverzagt. *Tetrahedron Lett.* **46**, 4201 (2005).
26. D. J. Cox, T. B. Parsons, A. J. Fairbanks. *Synlett* 1315 (2010).
27. M. V. Chiesa, R. R. Schmidt. *Eur. J. Org. Chem.* **65**, 3541 (2000).
28. G. M. Watt, G.-J. Boons. *Carbohydr. Res.* **339**, 181 (2004).
29. A. Fürstner, I. Konetzki. *Tetrahedron Lett.* **39**, 5721 (1998).
30. R. B. Trimble, P. H. Atkinson, A. L. Tarentino, T. H. Plummer, F. Maley, K. B. Tomer. *J. Biol. Chem.* **261**, 12000 (1986).