

# Glycoprotein Synthesis: From Glycobiological Tools to Tailor-made Catalysts

Benjamin G. Davis<sup>a\*</sup> and J. Bryan Jones<sup>b\*</sup>

<sup>a</sup>Department of Chemistry, University of Durham, South Road, Durham DH1 3LE, UK

Fax: +44 191 384 4737; E-mail: Ben.Davis@durham.ac.uk

<sup>b</sup>Department of Chemistry, University of Toronto, 80 St. George Street, Toronto, Ontario, M5S 3H6, Canada

Fax: +1 416 978 1553; E-mail: jbjones@alchemy.chem.utoronto.ca

Received 1 July 1999

new tools

**Abstract:** The synthesis of pure, well-defined glycoproteins is essential to the study of the host of biological processes that they mediate. The use of a combined site-directed mutagenesis and chemical modification strategy allows the site-selective synthesis of homogenous glycoproteins, the determination of precise structure-activity relationships and the design of new synthetic catalysts. This general and powerful method is reviewed in the context of existing preparative techniques and the wide-ranging applications of glycoproteins are highlighted

**Key words:** glycoproteins, carbohydrates, glycosylations, chemoselectivity, enzymes, biocatalysis

## 1. Introduction

It is becoming ever clearer that the presence of carbohydrate units in naturally occurring structures and their mimetics has a dramatic effect on their physical, chemical and biological properties. The ubiquity of glycoproteins in Nature reflects their broad functions as markers in cell-cell communication events that determine microbial virulence,<sup>1</sup> inflammation<sup>2,3</sup> and host immune responses.<sup>4,5</sup> In addition, the correct glycosylation of proteins is critical to their expression and folding,<sup>6</sup> and increases their thermal and proteolytic stability.<sup>7</sup>

Access to well-defined scaffolds to probe the nature of these processes is essential. Their manipulation is a dominant primary goal in glycoscience and has driven and continues to drive the synthesis of glycoconjugates and in particular glycoproteins. They are the tools of the glycobiology trade.

Carbohydrate structures are unrivalled in the density of information that they can convey. Precise differences in the nature of the linkages between two residues e.g., 1-2, 1-3, 1-4, 1-6 for two pyranoses, contrast with the linear nature of proteins and nucleic acids. A comparison of the permutations of hexamer formation illustrates this point well. Whereas, DNA (with a basis set of 4) and amino acids (with a basis set of 20) may construct a biological language for information transfer of 4096 and  $6.4 \times 10^7$  'words', respectively, carbohydrates have access to greater than  $1.05 \times 10^{12}$  variations.<sup>8</sup> Add to this the additional variety afforded by anomeric stereochemistry, ring size and sub-unit modification (e.g., sulfation, phosphorylation or acylation) and it can be quickly seen that this greater variety of possible combinations gives the language of carbo-

hydrates exquisite eloquence. This language has been christened *glycocode* - a term that well represents the potential level of complex information that carbohydrate structures are able to convey. It should also be noted that this vast number of potential permutations represents a technological barrier and means that no longer can the oligosaccharidic portions of glycoproteins be made on an iterative basis since there are far too many possible synthetic targets. It is therefore crucial that the design of new glycoproteins is guided by the identification of the associated functions and activities of existing structures.

The decipherers of glycocode are typically sugar-binding proteins called lectins which, despite their very shallow binding sites, show a remarkable specificity in their binding of multivalent complex carbohydrate structures.<sup>1</sup> The carbohydrate-lectin interaction stands out as an unusually weak and relatively indiscriminating one ( $K_d$  in the order of mM for monosaccharides)<sup>9</sup> when compared to others in Nature. This is largely due to the shallow, solvent-exposed nature of the lectin binding sites, which make few direct ligand contacts. The large difference in affinity shown by these shallow sites as compared with deep is amply illustrated by the influenza haemagglutinin lectin, which binds sialic acids with an approximately 1000-fold lower affinity than is shown by a neuraminidase found in the same virus.<sup>10</sup> However, when more than one saccharide of the right type and in the right orientation are clustered together there is a rapid increase in both affinity and specificity by the corresponding lectin.<sup>11</sup> This increase is more than would be expected due to the increase in local concentration (statistical effect) alone and has been termed the "cluster" or "multivalent effect".<sup>12</sup>

The reasons for the cluster effect are yet to be rigorously determined but their implications are profound. Firstly, the steady biological reservoirs of soluble monosaccharides are negligible inhibitors of any process that lectins mediate. Secondly, the specificity of this type of binding is exquisitely fine-tuned. It relies not only on the complementarity of the individual binding sites with a particular sugar ligand but also on the relative arrangement of the binding sites to each other in space and therefore by necessity the corresponding display of each sugar ligand relative to the next. Thirdly, the kinetics of such binding are different to those of monovalent binding and may afford faster "on rates". Fourthly, multipoint attachment is more resistant to shear stresses.

Nature ably exploits the tertiary structure of proteins as a scaffold for multivalent display. Yet more complex third order patterns are then in turn the product of the arrangement of these glycoproteins on cell surfaces.

There are numerous examples of the important role that the protein that displays a glycan has in determining activity. As early as 1929 it was appreciated that immunological activity toward sugars may be greatly enhanced through conjugation to proteins (for further details see Section 4).<sup>13</sup> Furthermore, the specificity of such immunological responses to sugars varies greatly with the nature of the protein to which they are conjugated. For example, synthetically prepared glycoproteins bearing the oligosaccharide blood group determinant Lewis-y ( $\text{Le}^y$ ) do not generate antibodies that can react with naturally occurring  $\text{Le}^y$  bearing structures. This could be due to a number of factors such as different densities of carbohydrate on the conjugate, as well as the influence of the protein or the linker upon conformation or accessibility. However, as yet, the reasons for such striking differences are undefined - an effect is observed and it is one that requires the preparation of glycoproteins for further investigation.<sup>14</sup>

Conversely there are also many examples of proteins and peptides whose biological activities are enhanced by conjugation to carbohydrates. For example, the activity of the anti-diuretic nonapeptide arginine-Vasopressin is almost doubled through galactosylation.<sup>15</sup> Also, different ribonuclease B (RNase-B) glycoforms that were carefully separated using capillary electrophoresis show 4-fold different hydrolysis activities. They also show decreased flexibility and greater protease resistance, possibly through the ac-

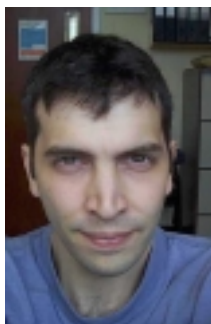
tion of the glycan as a "steric shield" for protease cleavage sites.<sup>16</sup>

The link between the glycan and the protein to which it is to be conjugated can be of virtually infinite variety and therefore a detailed discussion of different spacer arms is outside the scope of this review. However certain important structural features should be borne in mind. Several studies have investigated the effect of spacer arm length upon the affinity of some proteins (e.g., selectins<sup>17</sup> and the asialoglycoprotein receptor<sup>18</sup>) for glycoconjugates. The clear consensus result is that an optimal length is required that is long enough to allow accessibility but short enough that the loss of entropy upon binding is not a prohibitive cost in the binding equilibrium.

### 1.1 The need for pure, well-defined glycoproteins

Unlike the biosynthesis of proteins and nucleic acids there appears to be no associated mechanism for proofreading and correcting differently glycosylated biomolecules - the result is mixtures. Therefore, glycoproteins occur naturally in a number of forms (glycoforms)<sup>19</sup> that possess the same peptide backbone, but differ in both the nature and site of glycosylation. The different properties exhibited<sup>16,20</sup> by each component within these microheterogeneous mixtures present regulatory difficulties<sup>21</sup> and problems in determining exact function through structure-activity relationships. It has even been suggested that these naturally-occurring mixtures of glycoforms provide a spectrum of activities that can be biased in one direction or another as a means of fine-tuning.<sup>16</sup> Consequently, the few studies that have compared single glycoforms successfully<sup>16</sup> have required abundant sources and exten-

## Biographical Sketches



**Ben Davis** got his B.A. (1993) and D.Phil. (1996) from the University of Oxford. During this time he learnt the beauty of carbohydrate chemistry under the supervision of Professor George Fleet. He then spent two years as a post doctoral research assistant in the lab-

oratory of Professor Bryan Jones at the University of Toronto. In 1998 he returned to the UK to take up a lectureship at the University of Durham. His current research interests include protein engineering, biocatalysis, and all aspects of glycoscience.



**J. Bryan Jones** got his Ph.D. from the University of Wales and a D.Phil. from Oxford and then did post-doctoral work at MIT, and at Caltech - where he first began work on enzymes. He has been at Toronto since 1963 and is currently Uni-

versity Professor of Chemistry.

sive chromatographic separation. There is therefore an urgent need for alternative sources of homogenous glycoproteins. The development of highly successful small-molecule carbohydrate containing ligands has often involved careful structure-activity relationship (SAR) refinements.<sup>22</sup> If we are to achieve the same successes with glycoproteins, then homogeneity must be one of the first priorities. This goal is one that the field of glycoprotein synthesis is now beginning to address.

Typically syntheses of glycoproteins adopt one of two strategies. The first is the formation of the putative glycan-protein link early to form glycopeptide building blocks that may then be assembled. The second is the formation of the link late on in the synthesis once the protein scaffold for its presentation is in place. Given the instability that may be associated with the link<sup>23</sup> and the requirements for protection that need to be considered in the use of glycosylated building blocks, it is clear why the latter has often seemed the most attractive option. Whilst the construction of the protein-carbohydrate link is the focus of this review, the importance of a well-defined homogeneous source of glycan should not be understated. The indispensable syntheses of oligosaccharides<sup>24,25</sup> must continue hand-in-hand with methods for their conjugation to proteins and peptides.

## 2. Glycoprotein synthesis

### 2.1 Glycopeptide assembly

The need for homogenous samples (single glycoforms) that was outlined above has resulted in great effort in the field of *de novo* synthesis of glycoproteins. The linear assembly of glycosylated amino acids, has from the very first examples, such as the use of *N*-acetylglucosaminyl asparagine in the synthesis of a partial sequence of fibroblast interferon,<sup>26</sup> provided well-defined products. Thus, the required carbohydrate structure is attached to an amino acid residue (typically serine and threonine for *O*-linked glycopeptides and asparagine for *N*-linked glycopeptides). An excellent review of methods for the formation of the glycosidic link between peptides and glycans has very recently been published.<sup>27</sup> Suitably protected, the glycopeptide is then used as a building block in strategies that often rely heavily on standard peptide synthesis techniques. That these elegant approaches are still some way off synthesizing peptides of lengths approaching those of proteins is a testament to the inherent difficulties of this approach. Two factors limit the work; firstly the need not only for extensive carbohydrate protection but also amino acid protection regimes; and secondly the acid and base lability of glycosylated amino acid residues.<sup>23</sup> Indeed, this feature of glycoproteins has long been exploited to strip glycans from protein surfaces. The necessary protection and deprotection regimes, the use of solid phase techniques, including the introduction of specific linkers, and coupling methods have all been tailored to be compatible

with the presence of carbohydrates. Several excellent reviews<sup>23,28-32</sup> cover these aspects in detail.

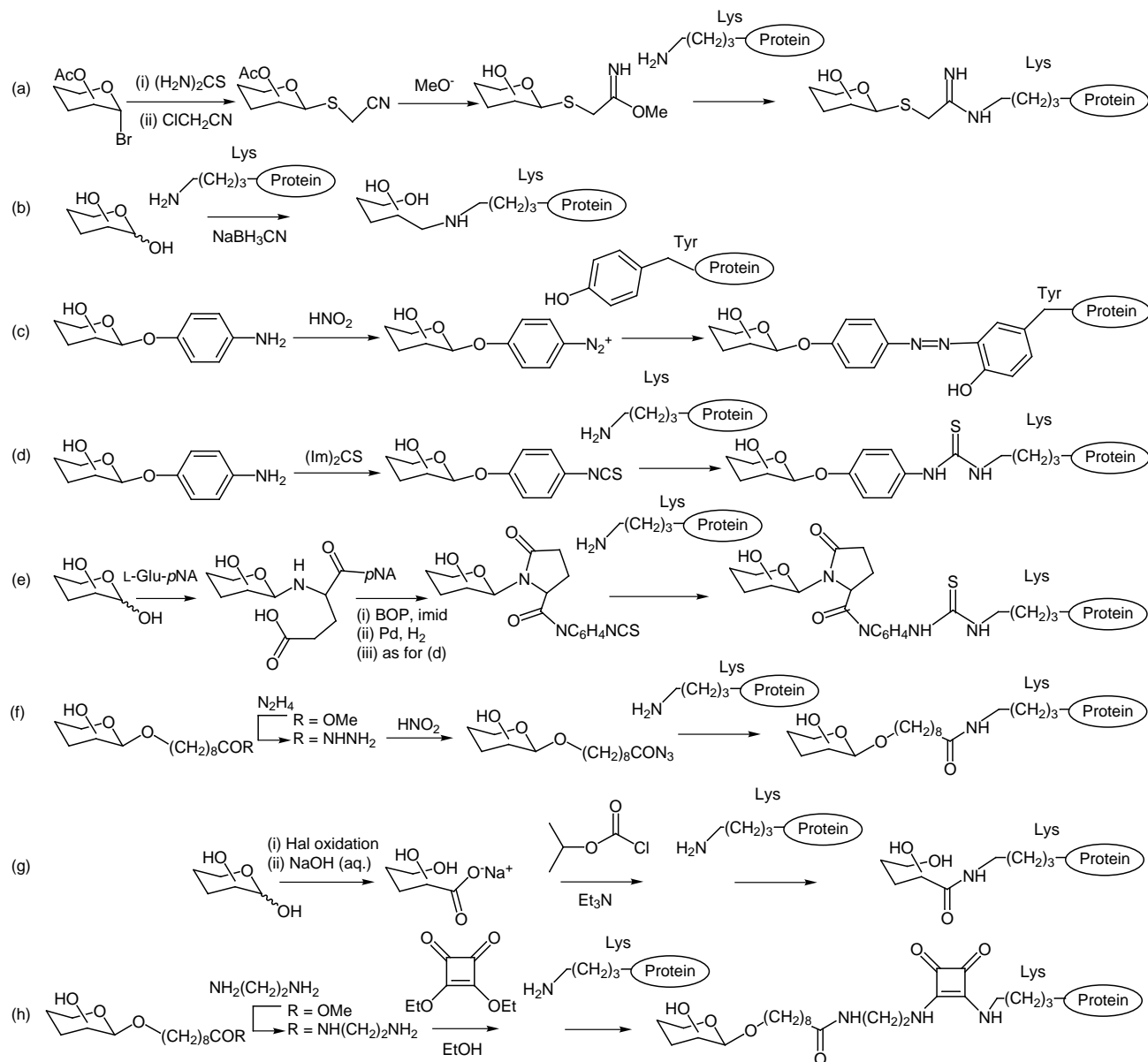
A linear strategy in glycopeptide synthesis is more usual since direct peptide glycosylation is often unsuccessful, given the variety of functional groups that would be required to protect a given oligopeptide of any significant length. However, Lansbury and co-workers have pioneered the use of glycosylamines in a convergent approach to glycopeptide synthesis. For example, HBTU-mediated coupling of GlcNAc glucosamine with the side chain aspartate carboxylate in a pentapeptide allowed the formation of an Asn linked *N*-acetylglucosaminyl containing glycopeptide (Scheme 1).<sup>33</sup> This method was successfully extended to peptides containing more complex glycans such as the high-mannose core of *N*-linked glycoproteins, Man<sub>5</sub>(GlcNAc)<sub>2</sub>.<sup>34</sup> Furthermore, this method has been expanded to encompass solid-phase bound glycosylamines, which are then coupled to side chain carboxylates in pentapeptides before the peptide chain is further extended.<sup>35</sup>

### 2.2 Chemical glycoprotein synthesis

Whilst the convergent glycosylation of oligopeptides is successful, it is limited by a lack of suitable functional groups when applied to proteins. For this reason alternative glycoprotein synthesis techniques have been more widely applied.<sup>21,36-38</sup>

The use of 2-iminomethoxymethyl thioglycosides<sup>39</sup> and reductive amination methods<sup>40</sup> are still after 20 years the most frequently used strategies for glycoprotein preparation. The former may be readily prepared by the action of methoxide on cyanomethyl thioglycosides that may in turn be derived from 1-thioaldoses (Scheme 1a). For the latter, Gray originally modified albumin with lactose through NaBH<sub>3</sub>CN-mediated reduction<sup>40</sup> (Scheme 1b) although borane may also be used. This method is amenable to other sources of aldehyde functionality such as those generated by ozonolysis of unsaturated spacer arms,<sup>41</sup> through periodate cleavage of diols, or the hydrolysis of acetal-containing spacer arms.<sup>42</sup> Conjugations through reductive amination are often accompanied by low protein loading levels, that in some cases are due to steric hindrance caused by short spacer arms. In a pragmatic approach to circumventing this problem, a second hydrazide spacer arm can be used to extend an existing aldehyde terminated spacer.<sup>43</sup> Reaction of the maleimido terminus of the resulting longer spacer arm with thiols introduced to the surface of the protein keyhole limpet haemocyanin (KLH) allowed 5-fold greater loading of the sialyl-GalNAc disaccharide, sTn.

The use of glycosidic aromatic diazonium salts, derived from the corresponding *p*-aminoaryl glycosides, as electrophiles to functionalize proteins was first demonstrated as early as 1929.<sup>13</sup> They modify a wide range of electron rich side chains within protein structures, such as those of aromatic tyrosinyl and tryptophanyl or nucleophilic lysinyl and histidinyl residues (Scheme 1c).<sup>44</sup> *p*-Aminoaryl



**Scheme 1** Non-selective Chemical Methods for Protein Glycosylation

glycosides may also be elaborated to phenylisothiocyanates, which react more selectively with amino groups alone (Scheme 1d).<sup>45</sup> A one-pot, two step preparation of anomeric *para*-nitroanilide (*pNA*) pyroglutamates from unprotected carbohydrates also provides an alternative route to aromatic isothiocyanates.<sup>46</sup> Following glycosylamine formation with the  $\alpha$ -amino group of *pNA* glutamic acid, the side-chain  $\gamma$ -carboxylic acid readily reacts with the resulting secondary amine to give a pyroglutamate which may then be elaborated (Scheme 1e).

In 1975, Lemieux and co-workers described the use of highly activated acyl azides for the formation of amides from proteinaceous amines and carboxylate ester-termi-

nus spacer arm carbohydrates.<sup>47</sup> The esters were converted to acyl hydrazides before oxidation with nitrous acid to give the corresponding acyl azides (Scheme 1f). Mixed anhydride methods are well established for the activation of carboxylic acids to form carboxyl derivatives and in this way aldonic acids may similarly be coupled to proteinaceous amines (Scheme 1g).<sup>48,49</sup> Similarly, carbodiimide chemistry<sup>50</sup> and the use of *N*-carboxyanhydrides<sup>51</sup> has also allowed the coupling of aldonates and glycosylated amino acids as sources of glycans bearing carboxylic acids. Hindsgaul and co-workers have described the use of diethyl squarate for the coupling of carbohydrates bearing amino-terminus spacer arms to amines in BSA (Scheme 1h).<sup>52</sup> 2-chloroethyl-1-thioglycosides have also been used

to indiscriminately alkylate protein amino and hydroxyl groups.<sup>53</sup>

A novel high temperature 'baking' method has been described for the modification of lyophilized proteins with reducing oligosaccharides.<sup>54</sup> Remarkably, despite being heated with the carbohydrate in air at 95–120 °C for up to 40 minutes both trypsin and an IgG antibody survived with little loss of biological activity. The results of tryptic digests and conjugate hydrolyses suggest that the mechanism of conjugation involves an Amadori rearrangement with protein lysines which destroys the integrity of the reducing end residue.

Although high levels of functionalization are thus easily accessible using the above methods, a lack of residue selectivity is often a drawback in the synthesis of well-defined conjugates. In addition, these techniques may alter the overall charge of the protein or destroy the cyclic nature of glycans introduced.

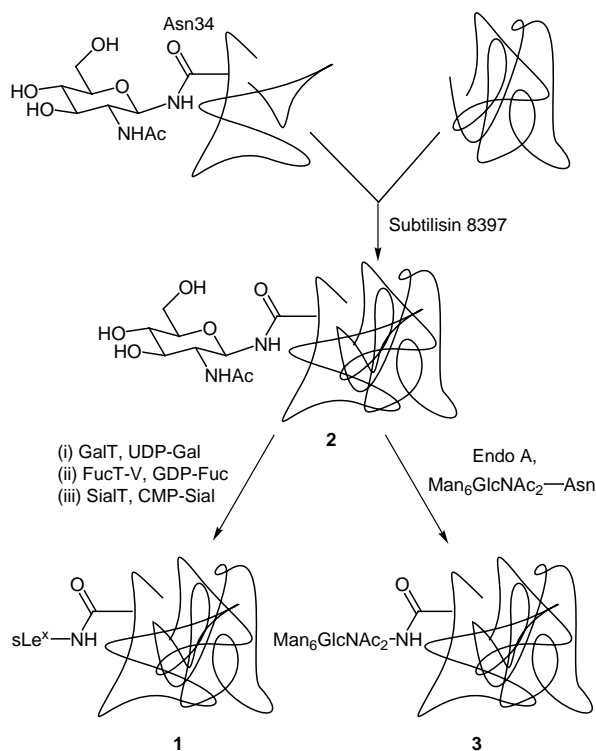
### 2.3 Enzymatic glycoprotein synthesis

The two distinct methods outlined in the two previous sections are equally available to enzyme-catalyzed techniques. For example, subtilisin peptidases have been elegantly used to catalyze the synthesis of glycopeptides,<sup>55,56</sup> in spite of the fact that the natural specificity of these enzymes has limited these peptide ligations to those in which the glycosylated residues are typically at least one residue distant ( $P_2, P_3, \dots$  or  $P_2', P_3', \dots$ ) from the amide bond formed. Thus, while ligation of Z-Gly-OBz with H-Gly-Ser(Ac<sub>3</sub>GlcNAc $\beta$ )-NH<sub>2</sub> was successful, no yield of product was obtained with H-Ser(Ac<sub>3</sub>GlcNAc $\beta$ )-NH<sub>2</sub>.

In a rare use of enzymatic glycosylation on the solid phase, the sLe<sup>x</sup>-Asn-Phe dipeptide was synthesized using aminopropyl silica as a support and subsequently cleaved from a glycine linker by the peptidase chymotrypsin.<sup>57</sup> In this context, it is interesting to note that the ability of glycosyltransferases to modify glycans on glycopeptides is greatly influenced by the peptide backbone and the sites that are glycosylated within it.<sup>58</sup> For example, prolines, negatively charged residues and the presence of a disaccharide all inhibited the glycosylation of an adjacent GlcNAc-Thr residue.

Convergent enzyme-catalyzed techniques have proved more adaptable than linear. One of the very first examples of the use of glycosyltransferases in glycoprotein synthesis was demonstrated by Paulson and co-workers who used a sialyltransferase and CMP-*N*-Ac-neuraminic acid to restore 95% of the sialic acids to a fully desialylated protein.<sup>59</sup> The use of enzymes in glycoproteins synthesis has recently culminated in the elegant synthesis of a single unnatural glycoform of ribonuclease B (RNase-B) **1** using a series of protease and glycosyltransferase catalyzed reactions (Scheme 2).<sup>60</sup> Takegawa and co-workers have applied endoglycosidase mediated transglycosylation to the same partially deglycosylated RNase-B **2** in the synthesis of the Man<sub>6</sub>GlcNAc<sub>2</sub> glycoprotein **3**.<sup>61</sup> Interestingly, Na-

ture employs a similar mechanism in the case of *Trypanosoma cruzi*, the protozoon that causes Chagas' disease. This parasite does not synthesize sialic acid, but instead expresses a transsialidase that catalyzes the transfer of sialic acid from glycoconjugates found in the host to its own surface proteins. The resulting sialylated glycoproteins are then bound by host sialic acid binding receptors thereby allowing cellular invasion.<sup>62</sup> Enzymatic approaches have even been used to directly modify cell surface proteins.<sup>63</sup> The use of a milk fucosyltransferase that displays a broad substrate specificity allowed modification of glycans with fucosyl residues bearing a range of substituents at C-6.



**Scheme 2** Enzymatic Synthesis of Glycoforms of Rnase-B

An attractive approach to enzymatic glycoprotein synthesis is to exploit the biosynthetic mechanism for the formation of the sugar-protein link in *N*-linked glycoproteins. The enzyme responsible, oligosaccharyltransferase, transfers a high mannose core oligosaccharide from a fatty acid pyrophosphate carrier to the side chain amide of an asparagine (Asn) residue in the consensus sequence Asn-X-Thr/Ser of the nascent glycoprotein. The use of this enzyme in isolated form in glycoprotein synthesis has however met with only modest success. Whilst transfer of carbohydrates to a 17-residue peptide containing an unusual Asn-Asn-Thr-Ser sequence was possible, direct glycan transfer to RNase-A failed.<sup>64</sup> Thus, such site-specific glycosylation of proteins still remains an elusive goal and as a result there is still no general enzymatic method for the synthesis of homogeneous glycoproteins. However, through the use of enzymes the unification of both the

strategies outlined chemically in Sections 2.1 and 2.2 can be imagined: enzyme catalyzed peptide ligation to construct a (glyco)protein scaffold, perhaps bearing single glycan tags, before convergent glycosylation by oligosaccharyltransferase or glycosyltransferases.

## 2.4 Molecular and cell biological techniques

*In vivo* methods, that alter the natural machinery of glycosylation, offer promising opportunities<sup>65</sup> but, as yet, still lead to heterogeneous products.<sup>66</sup> The task is made difficult by the daunting array of biosynthetic glycosylation products and thus the corresponding array of pathways that need to be controlled or adapted. By expressing a particular glycosyltransferase or glycosidase in one organism different glycosylation patterns may arise to those found by expression of the same enzyme in another. In this way glycosylation patterns may be guided in a particular direction. For example, mutant cells lines in which extra glycosyltransferase expressing genes have been introduced may be used to enhance the presence of particular sugars in glycan structures. The addition of a sialyltransferase to a Chinese hamster ovary (CHO) cell line resulted in the increased "misglycosylation" of *N*-linked glycoproteins to give glycans bearing  $\alpha(2-6)\text{Gal-}$  instead of  $\alpha(2-3)\text{Gal-}$  linked sialic acid terminated residues.<sup>67</sup>

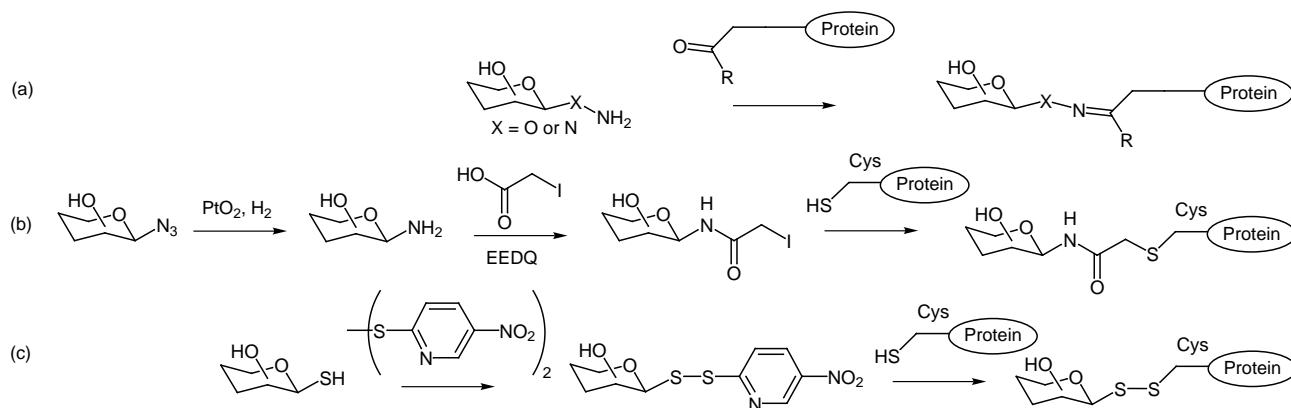
The inhibition of enzymes involved in the biosynthesis of glycoproteins offers an alternative way of controlling their structure. For example, the glycosyltransferase inhibitor tunicamycin inhibits the synthesis of the lipid-linked pyrophosphate oligosaccharide precursor that is used as a glycosyl donor in the formation of *N*-linked glycoproteins.<sup>68</sup> The resulting lack of donor prevents formation of Asn-linked glycans and results in only *O*-glycosylated proteins. Less drastic inhibition of later trimming steps that are mediated by glycosidases can be used to create smaller than natural ranges of Asn-linked glycoforms rather than none at all. For example, the use of the glucosidase inhibitor *N*-butyl deoxynojirimycin resulted in a reduction in the number of glycoforms of the HIV surface

protein gp120 that were produced from more than 100 to 3.<sup>5</sup>

The prospects are also good for the glycosylation of larger biomolecular complexes by taking advantage of the often-relaxed specificities of biosynthetic pathways. Indeed, the use of an unnatural *N*-levulinoylmannosamine as a precursor simply by feeding it to cells has allowed the introduction of a unique ketone tag into sialic acid residues found at cell surfaces.<sup>69</sup> This allowed the selective introduction of further glycans through reaction with aminoxy and hydrazide-functionalized carbohydrates to form imines and hydrazones according to the manner described in Section 2.5 (Scheme 3a).<sup>70</sup> This strategy is similar to one previously applied to aldehydes introduced chemically to cell surfaces.<sup>71</sup> In a similar manner, neural cell surfaces have been also engineered by introducing an unnatural *N*-propanoylneuraminic acid precursor.<sup>72</sup>

## 2.5 Selective Protein Glycosylation

In an attempt to increase the selectivity and predictability of protein glycosylation, various novel approaches have been described all of which exploit the chemoselectivities of different enzymatic and traditional methods. As Scheme 3a illustrates, Bertozzi and co-workers have employed the selectivity of galactose oxidase to introduce an aldehyde tag to the C-6 of a GalNAc residue in the antimicrobial 19-residue peptide drosocin.<sup>73</sup> (This type of strategy has been christened chemoselective ligation and its application to a wide range of bioconjugates has recently been reviewed.<sup>74</sup>) This tag was then selectively reacted with aminoxy glycosides to introduce further saccharides via the formation of an imine, in a manner previously demonstrated for the conjugation of spacer-arm hydrazides with cell surface aldehydes.<sup>71</sup> That this non-native glycopeptide shows comparable biological activity to the native form illustrates that certain unnatural linkages can in certain circumstances be tolerated. Unfortunately, this approach still requires the linear construction of an initial glycopeptide and so suffers from the same disadvantages of protection and lability outlined in section 2.1.



**Scheme 3** Selective and Site-Specific Protein Glycosylation Methods

However, it does hold the advantage that its application may be based on other methods for the introduction of a ketone tag (see Section 2.4). A similar chemoselective ligation approach has been applied to glycopeptide synthesis. Both  $\alpha$ - and lysine  $\epsilon$ -amino groups were derivatized with aminooxyacetyl groups before reaction of the free amine introduced with reducing sugars. The use of orthogonal *N*-protection also allowed regioselective glycosylation.<sup>75</sup>

## 2.6 Site-Specific Protein Glycosylation

In a sense, these selective approaches have avoided the crux of glycoprotein synthesis - the formation of the carbohydrate-protein link - by relying on the presence of an existing glycan in the protein structure as a tag for reaction. Similarly, the elegant enzymatic methods of Wong and co-workers and of Takegawa and co-workers described in Section 2.3 still require that a protein-*N*-glycan link be present from the start as a point of recognition for the enzymes concerned. As these methods alter one glycan structure for another they are therefore better described as glycoprotein remodelling (GPR) and they afford the glycoscientist no choice over the site of glycosylation.

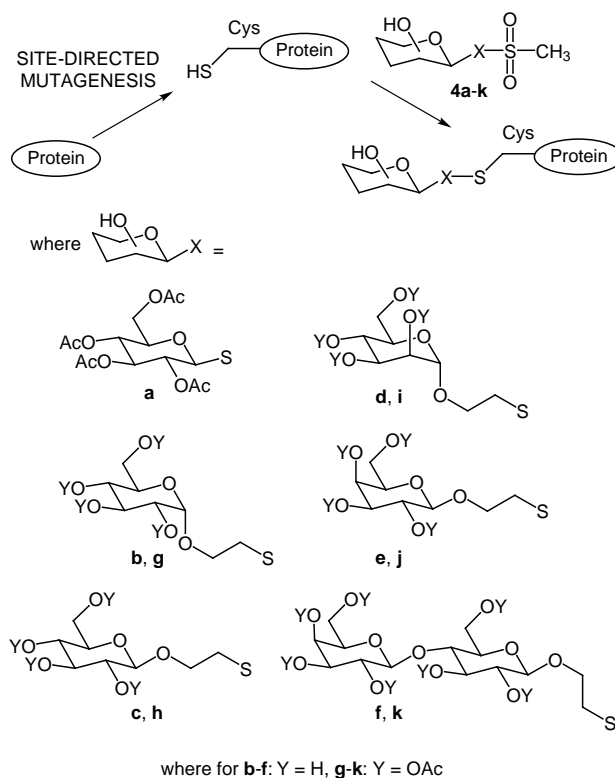
Several methods have been proposed that tackle this central issue. Among the first approaches was that of Flitsch and co-workers, who reacted the  $\alpha$ -iodoacetamide of *N*-acetyl-D-glucosamine with oxidized bovine serum albumin (BSA) to modify the single free cysteine present (Scheme 3b).<sup>76</sup> Later this method was applied by Dwek to introduce chitotriose and a heptasaccharide stripped from the surface of horseradish peroxidase to BSA.<sup>77</sup> Boons has used dithiopyridyl methodology to make disulfide linked BSA-*N*-acetyl-D-glucosamine constructs (Scheme 3c).<sup>78</sup>

For full control of glycosylation both choice of site (site-selectivity) and glycan are needed. Recently, using a combined site-directed mutagenesis and chemical modification approach we have solved this problem for the first time.<sup>79, 80</sup>

### 3. Site-Selective Glycosylation Using a Combined Site-Directed Mutagenesis and Chemical Modification Approach

#### 3.1 Site-Selective Glycosylation

This approach provides a general method that allows both regio- and glycan-specific glycosylation of proteins. This method is rapid, utilizes reagents that may be prepared in a facile manner and is unlimited in the scope of sites and glycans that may be conjugated. The strategy involves the introduction of cysteine as a chemoselective tag at preselected positions within a given protein and then reaction of its thiol residue with glycomethanethiosulfonate reagents (Scheme 4). Methanethiosulfonate (MTS) reagents react specifically and quantitatively with thiols<sup>81</sup> and allow the controlled formation of neutral disulfide linkages.

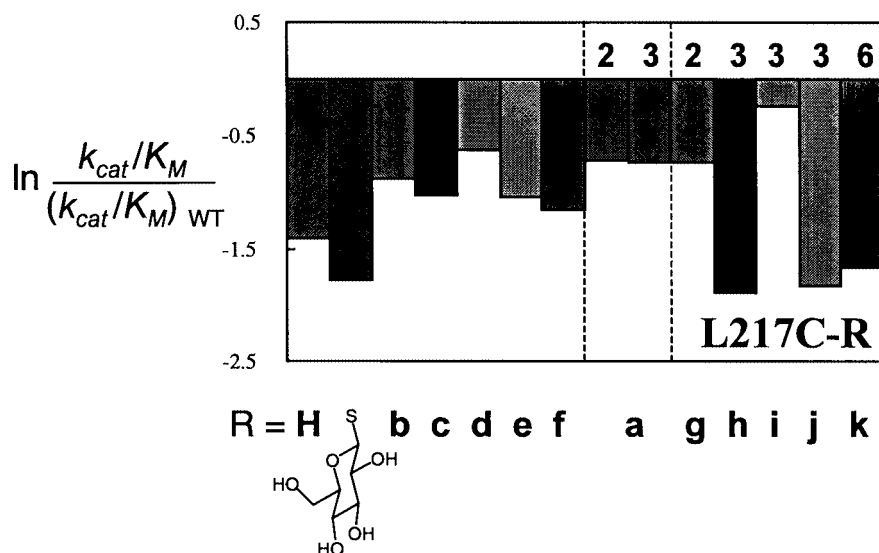


**Scheme 4** The Combined Site-Directed Mutagenesis and Chemical Modification Approach to Site-Selective Glycosylation

As our vehicle we have identified the representative serine protease subtilisin *Bacillus lentus* (SBL). SBL is an ideal model protein for evaluating the validity of this strategy as it does not contain a natural cysteine and is not naturally glycosylated. Four SBL sites at different locations and of different characteristics were selected for mutation to cysteine in order to provide a broad test of the glycosylation methodology. S156 of the  $S_1$ <sup>82</sup>-pocket is a surface-exposed residue that permits the introduction of externally-disposed glycans mirroring those found naturally in glycoproteins.<sup>83</sup> In contrast, N62 in the  $S_2$  pocket, S166 in the  $S_1$  pocket, and L217 in the  $S_1'$  pocket have side chains which are internally oriented and tested the applicability of the method for introducing sugars at hindered locations.

Broad applicability with respect to the sugar moiety was evaluated by using the representative library of protected and deprotected, mono- and disaccharide methanethiosulfonates **4a-k**. Two types of glycosylating reagents, the untethered methanethiosulfonate **1a** and the ethyl-tethered methanethiosulfonates **1b-k**, were prepared from their parent carbohydrates in good to excellent yields. The preparation of these reagents in fully protected **1a,g-k** and deprotected **1b-f** forms allowed the effects of increased steric bulk and hydrophobicity to be assessed.

The glyco-MTS reagents **1a-k** were reacted with SBL-N62C, -S156C, -S166C and -L217C in aqueous buffer.<sup>84</sup> These reactions were rapid and quantitative, as judged by



**Figure 1** Precise Glycan Structure-Hydrolytic SARs for Glycosylated forms of SBL Relative to Unglycosylated Wild-Type (WT) Using suc-AAPF-*p*NA at pH 8.6. For each glycan the number of acetate groups present is indicated by a label above the corresponding bar. A value of zero indicates that the value was not determined.

monitoring changes in specific activity and by titration of residual free thiols with Ellman's reagent.<sup>85</sup> The glycosylated proteins were purified by size-exclusion chromatography and dialysis, and their structures were confirmed by rigorous ES-MS analysis. Each of the glycosylated enzymes appeared as a single band on non-denaturing gradient PAGE, thereby establishing their high purities. Moreover, through adjustment of pH and appropriate selection of the glycosylation site, we were able to prepare differently acetylated glycoforms of SBL. We attribute this valuable site-dependent deacetylation to a novel intramolecular SBL-catalyzed process. In all cases, modification with the fully deprotected reagents **1b–f** led to site-specific glycosylations and the formation of single glycoforms. These therefore represented the first examples of homogenous glycoproteins in which both the site of glycosylation and the structure of the glycan introduced had been predetermined.

### 3.2 Precise SARs through Site-Selective Glycosylation

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. This potential was clearly demonstrated by the determination of detailed glycan structure-hydrolytic activity relationships for a library of 48 glycosylated forms of SBL.<sup>86</sup>

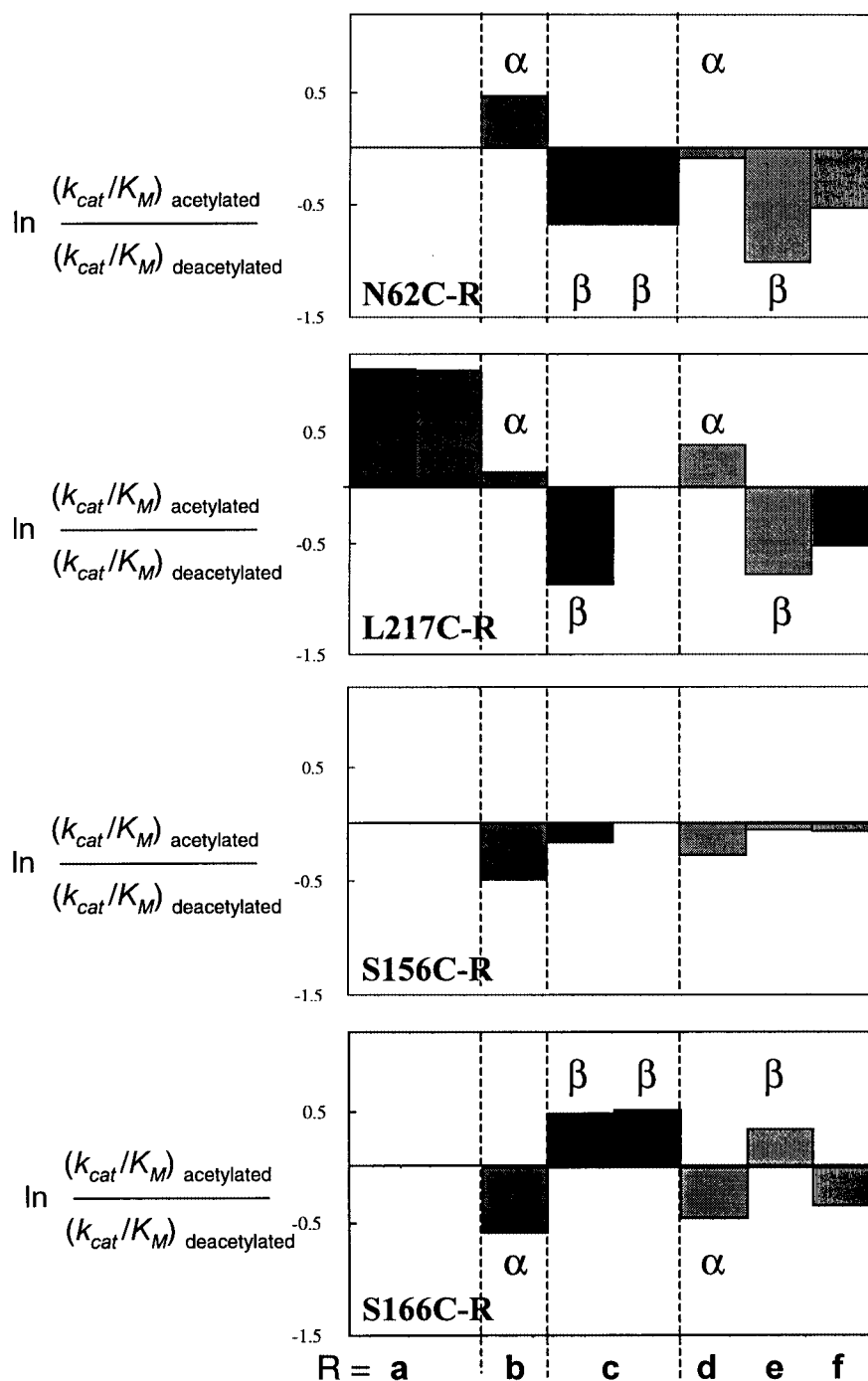
The effects of glycosylation upon SBL were assessed by the determination of  $k_{cat}$  and  $K_M$  for the hydrolysis an amide bond in the standard substrate succinyl-Ala-Ala-Pro-Phe-*p*-nitroanilide (Suc-AAPF-*p*NA) at pH 8.6. The

kinetic parameters for the amidase activity of some of the 48 enzymes generated are compared with those of WT and unmodified mutants in Figure 1. The site-selective glycosylation method shown in Scheme 4 allowed the introduction of structurally related monosaccharides, D-glucose, D-galactose and D-mannose, in addition to the disaccharide lactose.

Some clear features emerge from the precise set of structure-activity relationships that were generated. Subtle configurational changes in the carbohydrate moiety had a profound effect upon the catalytic activity of SBL. For example, the D-mannosylated enzyme L217C-S-Et- $\alpha$ -Man(Ac)<sub>3</sub> displayed a 5-fold higher amide bond cleaving catalytic activity than the D-glucosylated enzyme L217C-S-Et- $\beta$ -Glc(Ac)<sub>3</sub> and illustrates clearly the remarkable effect of the glycan introduced.

To examine the cause of such interesting variations, the  $k_{cat}/K_M$ s of acetylated glycosylated CMMs were also compared with those for deprotected glycosylated CMMs with the same glycan structure and stereochemistry (Figure 2). This separated the effects of acetylation from the effects of glycosylation and allowed the underlying effects of modification to be dissected. It is clear from Figure 2 that the anomeric stereochemistry of the acetylated glycans modulates  $k_{cat}/K_M$ . Thus, at positions 62 and 217 acetylation enhanced the activity of  $\alpha$ -glycosylated SBLs but decreased that of  $\beta$ -glycosylated. This trend is reversed at position 166 where, in contrast, acetylation enhanced the  $k_{cat}/K_M$ s of  $\beta$ -glycosylated SBLs but decreased those of  $\alpha$ -glycosylated. Consistent with its surface exposed nature changes at position 156 are more modest, but still allowed control of activity, particularly through glycosylation with disaccharide lactose.





**Figure 2** Variation in Hydrolytic Amidase Activity of Glycosylated CMMs of SBL upon Acetylation of Glycans. Comparison of the activity of acetylated with fully deprotected glycoproteins shows that anomeric stereochemistry of the acetylated glycans modulates  $k_{cat}/K_M$ .

### 3.3 Using Glycosylation to Tailor the Synthetic Utility of Enzymes

As part of our ongoing program to tailor the activity of SBL for use in synthesis,<sup>87-89</sup> we screened these enzymes for activity against the esterase substrate succinyl-Ala-Ala-Pro-Phe-S-benzyl.<sup>90</sup> Gratifyingly, 22 enzymes dis-

played greater than wild type (WT) activity. In particular, at position 217, in the  $S_1'$  pocket, all glycosylations increased  $k_{cat}/K_M$  up to 8.4-fold greater than WT for L217C-S $\beta$ Glc(Ac)<sub>3</sub> - the most active enzyme we have synthesized using this approach.

Furthermore, the ratio of amidase to esterase activity,  $(k_{cat}/K_M)_{\text{esterase}} / (k_{cat}/K_M)_{\text{amidase}}$  (E/A), was increased relative to wild type for all 48 glycosylated forms of SBL. The most dramatic changes were observed at positions 62 and 217 and L217C-S $\beta$ Glc(Ac)<sub>3</sub> displayed an E/A that was 17.2-fold greater than WT. The high activity of this glycoform was rationalized by molecular modelling analysis, which suggested that the carbohydrate moiety in this enzyme occupies the S<sub>1</sub>' 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  $(k_{cat}/K_M)_{\text{esterase}}$  and severely reduced  $(k_{cat}/K_M)_{\text{amidase}}$  and therefore favour the formation of the amide bond over hydrolysis. In preliminary investigations, these powerful catalysts have allowed syntheses of dipeptides that were not possible using the unglycosylated catalyst.<sup>91</sup>

The extension of this method to the glycosylation of other synthetically-useful enzymes is now in progress.<sup>92</sup>

## 4 Applications of Glycoprotein

The future of carbohydrate science will be honed by the application of its products; the applications of glycoproteins are therefore an important backdrop to this review and are the context in which their synthesis should rightly be judged.<sup>93,94</sup> The functions and versatility of glycosylated biomolecules are, as a consequence of their ability to transmit sophisticated information, incredibly broad and the potential applications of glycoproteins are therefore correspondingly wide. For example, glycoproteins have been implicated in physiological processes ranging from receptor mediated endocytosis, protein quality-control, the interaction and subsequent invasion of pathogens, and the triggering of effects that lead to the release of biomodulators. A thorough understanding of these processes is essential to their successful exploitation in pharmaceutical therapies that either block lectin-mediated process through inhibition or that exploit this binding to target designed glycoconjugates to lectin-expressing cells. Indeed, the recent failure of a number of carbohydrate-based drugs may be attributed to a poor understanding of their supposed mechanism of action rather than due to any inherent flaws associated with carbohydrate therapeutics.<sup>95</sup>

### 4.1 Probing protein activity, function and mechanism

In addition to their critical role in communication events, glycosylated proteins have long been known to have greater resistance to thermolysis and proteolysis.<sup>37</sup> A much less widely explored aspect, which we have already illustrated (Section 3.2, 3.3) is the use of glycosylated enzymes as tailor-made catalysts. Other examples include carbohydrate-protease conjugates, which show greater stabilities at high temperatures<sup>96</sup> and in organic solvents,<sup>97,98</sup> and also catalyze high yielding peptide syntheses.<sup>99</sup>

RNase-A which has been mono- and diglycosylated by EDC-mediated coupling of D-glucosamine to carboxylates at Asp 53 and Glu 49 displays lower catalytic activity but greatly enhanced thermal stability.<sup>100</sup>

An interesting aspect of *N*-linked protein glycosylation is its role in the "quality control" of proteins that are formed during translation. Without correct glycosylation many proteins fail properly to fold. This suggests a novel role for added glycans as indicators of correct protein structure. In *N*-linked glycoprotein biosynthesis a 14-residue oligosaccharide core is added as a first step and then trimmed down to size. It has been suggested that, if nascent protein fails to fold properly, these glycans are incorrectly displayed and cannot be processed in these trimming steps, leading to rejection and degradation. Therefore, these apparently superfluous trimming steps may not simply be a means to glycan structure but steps along a "quality controlled" protein production line.<sup>6</sup> It has also been suggested that glycans aid the folding and transport of proteins by protecting them from proteolysis.<sup>7</sup>

In aqueous solution a turn is induced in a SYSPTSPSYS segment of the C-terminal domain of RNA polymerase II when the threonine side chain is *N*-acetyl-D-glucosaminylated, whereas the corresponding non-glycosylated peptide adopts a randomly-coiled structure.<sup>101</sup> This striking difference in structure in aqueous solution is a clear indication of the importance of glycosylation, even by a single saccharide, upon local peptidic structure and opens the door to a host of such crucial structural subtleties in glycosylated proteins. Moreover, this result tallies well with recent suggestions that reversible glycosylation of this site might act as a regulatory mechanism for the control of transcription akin to phosphorylation.<sup>102</sup>

Glycoproteins which act as antifreeze in the serum of deep-sea fish allow them to survive at temperatures as low as -2 °C. Their ability to lower freezing point is not proportional to concentration and is not accompanied by altered melting points. This non-colligative effect is thought to arise from a mechanism that inhibits ice nucleation and crystal growth. Oligomeric glycopeptide analogues of such proteins have been prepared and show significant antifreeze properties.<sup>103</sup>

### 4.2 Drug delivery

The specificity of the hepatic asialoglycoprotein Gal/Gal-Nac-specific receptor<sup>104</sup> has been widely exploited for liver targetting and in many cases has been the model for targetted drug delivery<sup>105</sup> and gene delivery<sup>106,107</sup> therapies. For example, covalent attachment of drugs through various degradable linkers to lactosaminylated and galactosylated human serum albumin has allowed the targetted delivery of anti-inflammatory agents, such as Naproxen,<sup>108,109</sup> and anti-virals, such as arabanoside-AMP,<sup>110</sup> which are absorbed into hepatic cells through receptor-mediated endocytosis.<sup>111</sup> Furthermore, the immunogenic-

ities of such glycoproteins are low if prepared with a high degree of homogeneity.<sup>108</sup>

### 4.3 Anti-infectives

*Helicobacter pylori* are the bacteria that cause gastric ulcers. They attach themselves to gut cells by binding to extracellular sialylated glycoproteins. This adhesion has been effectively inhibited, as part of an anti-ulceric strategy, by albumin glycosylated with 3'-sialyllactosyl residues.<sup>112</sup> It is interesting to think that the array of sialylated conjugates in human milk might serve the same anti-adhesive function as these synthetic glycoconjugates and once again highlights the importance of studying Nature's tactics as models for our own.

### 4.4 Vaccines

The use of complex oligosaccharides as haptens for the induction of antibodies has a rich history dating back to the discovery in 1929 that oligosaccharides may be rendered immunogenic through their attachment to proteins.<sup>13,113</sup> As early as 1936 their use as a strategy to combat pneumococci was described.<sup>114</sup> However, until the 1970's this work was limited by the often-minute amounts of oligosaccharides available from natural sources. In a seminal series of four papers in 1975,<sup>47</sup> Lemieux and co-workers completed the total synthesis of the Lewis-a (Le<sup>a</sup>) trisaccharide using newly developed glycosidation techniques and conjugated it to BSA using the acyl azide method (see section 4.2) which they developed for the purpose.

As an approach to a potential anticancer vaccine, Danishefsky and co-workers have used reductive amination to construct a hexasaccharide-keyhole limpet haemocyanin (KLH) protein conjugate, at an approximate ratio of sugar : protein of 150:1. The hexasaccharide moiety, termed globo H, which was constructed using glycal methodology, was originally isolated from a glycosphingolipid associated with breast cancer. The synthetic globo H-KLH conjugate was successfully used to induce high anti-globo H antibody titers and induced cell lysis, in the presence of human complement, at levels approaching those of monoclonal antibodies raised against cancerous cells.<sup>115</sup> Other potential anticancer vaccines have been reported based on other tumour-associated epitopes,<sup>116</sup> such as the sialyl-Tn motif; which as a KLH-conjugate has shown promisingly higher survival rates in clinical trials.<sup>117</sup>

T-cells mediate intercellular immune responses and are particularly useful as they allow even infants under the age of two to be effectively vaccinated. Whilst oligosaccharides alone do not typically elicit T-cell responses, the discovery<sup>118</sup> that glycopeptides do, has allowed immune responses to sugars to be probed systematically. The recent developments in the synthesis of more complex glycans bound to peptides and proteins in a site-specific manner described in this review should allow these aspects to be probed in ever-greater detail in the future.

### 4.5 Other therapeutic strategies and uses

The approaches used to target drugs described in section 4.2 may also be used to target physiologically beneficial enzymes. For example, the mannosylation of enzymes has allowed the targeting of enzymes to particular diseased cells. Replacement  $\beta$ -glucocerebrosidase, an enzyme which is lacking in Gaucher disease,<sup>119</sup> and the beneficial anti-oxidant effects of superoxide dismutase (SOD)<sup>120</sup> have both been directed to macrophages. SOD has also been conjugated with sodium hyaluronate, a polymer of the dimeric motif GlcNAc-glucuronic acid. This combined the ability of SOD to catalyze superoxide anion decomposition with the hydroxyl radical scavenging capacity of hyaluronate in a potentially dual-action anti-inflammatory.<sup>121</sup>

Novel analytical techniques have also utilized glycoproteins. For example, layer-by-layer deposition of a mannose-specific lectin and mannosylated enzymes (glucose oxidase and lactate oxidase) on a platinum surface allowed the preparation of a sensitive active-enzyme electrode.<sup>122</sup>

## 5. Future directions

The work described in this review shows that the glycosylation of enzymes allows the development of novel synthetic catalysts. The ability of such site-selective glycosylation methods to glycosylate the binding pockets of synthetically useful enzymes also creates further opportunities to broaden substrate specificity. For instance, such an array of hydrogen bonding hydroxyl groups may enhance the specificity of peptidases towards hydrogen bonding substrates such as glycosylated amino acids and overcome some of the shortcomings of existing enzyme-catalyzed glycopeptide formation.<sup>55,56</sup>

Finally, in order that the application of glycoconjugates in therapeutic strategies becomes more widespread certain features must be addressed. Any of the glycoprotein structures described above represent potential sources of immunogenicity and hence much attention has been paid to the development of biocompatible polymers as alternatives. By building on the advances of glycoprotein applications, certain large (> 100kDa) biocompatible polymers, may offer all the advantages of glycoproteins but with reduced immunogenicity.<sup>123</sup> Unfortunately many glycoproteins are also poorly characterized and so future efforts must also focus on homogeneity as a goal. In this context, the precision and rigour of synthetic chemistry has a crucial role to play. Collaborations between those making glycoproteins and those applying them will no longer be viewed as multi- or inter-disciplinary - they are part of a new unified discipline intent on making exciting, well-defined molecules with enormous potential for the treatment of disease, the elucidation of biochemical and physiological mechanisms, and as powerful finely-tuned catalysts.

## 6. Acknowledgements

The work from our laboratories described in this review would not have been possible without the dedication and enthusiasm of Martin Green, Michael Maughan, Drs Richard Lloyd, Kazu Matsumoto, Kanjai Khumtaveeporn, Astrid Ullman, Tony Fordham-Skelton, John Gatheouse and Rick Bott. We also acknowledge the generous financial support of the Natural Sciences and Engineering Research Council of Canada, the University of Durham, the Nuffield Foundation and Genencor International, Inc.

## 7. References and notes

- (1) Sharon, N.; Lis, H. *Essays Biochem.* **1995**, *30*, 59.
- (2) Lasky, L. A. *Annu. Rev. Biochem.* **1995**, *64*, 113.
- (3) Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem.* **1996**, *65*, 441.
- (4) Varki, A. *Glycobiology* **1993**, *3*, 97.
- (5) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.
- (6) Helenius, A. *Molec. Biol. Cell* **1994**, *5*, 253.
- (7) Opdenakker, G.; Rudd, P. M.; Ponting, C. P.; Dwek, R. A. *FASEB J.* **1993**, *7*, 1330.
- (8) Laine, R. A. *Glycobiology* **1994**, *4*, 759.
- (9) Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, *26*, 323.
- (10) Janakiraman, M. N.; White, C. L.; Laver, W. G.; Air, G. M. A.; Luo, M. *Biochemistry* **1994**, *33*, 8172.
- (11) Kiessling, L. L.; Pohl, N. L. *Chem. Biol.* **1996**, *3*, 71.
- (12) Lee, Y. C.; Townsend, R. R.; Hardy, M. R.; Lönngren, J.; Arnarp, J.; Haraldsson, M.; Lönn, H. *J. Biol. Chem.* **1983**, *258*, 199.
- (13) Goebel, W. F.; Avery, O. T. *J. Exp. Med.* **1929**, *50*, 521.
- (14) Kitamura, K.; Stockert, E.; Garinchesa, P.; Welt, S.; Lloyd, K. O.; Armour, K. L.; Wallace, T. P.; Harris, W. J.; Carr, F. J.; Old, L. J. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12957.
- (15) Susaki, H.; Suzuki, K.; Ikeda, M.; Yamada, H.; Watanabe, H. *Chem. Pharm. Bull.* **1998**, *46*, 1530.
- (16) Rudd, P. M.; Dwek, R. A. *Biochemistry* **1994**, *33*, 17.
- (17) Kretzschmar, G.; Sprengard, U.; Kunz, H.; Bartnik, E.; Schmidt, W. *Tetrahedron* **1995**, *51*, 13015.
- (18) Kichler, A.; Schuber, F. *Glycoconj. J.* **1995**, *12*, 275.
- (19) Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Annu. Rev. Biochem.* **1988**, *57*, 785.
- (20) Parekh, R. B.; Dwek, R. A.; Rudd, P. M.; Thomas, J. R.; Rademacher, T. W.; Warren, T.; Wun, T. C.; Hebert, B.; Reitz, B.; Palmier, M.; Ramabhadran, T.; Tiemeier, D. C. *Biochemistry* **1989**, *28*, 7670.
- (21) Bill, R. M.; Flitsch, S. L. *Chem. Biol.* **1996**, *3*, 145.
- (22) Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C. H. *Chem. Rev.* **1998**, *98*, 833.
- (23) Kunz, H. *Angew. Chem., Int. Ed. Engl.* **1987**, *99*, 297.
- (24) Boons, G.-J. *Contemp. Org. Synth.* **1996**, *3*, 173.
- (25) Khan, S. H.; O'Neill, R. A. *Modern Methods in Carbohydrate Synthesis*; Harwood Academic: Amsterdam, 1996.
- (26) Fields, S.; Winter, G. *Nature* **1981**, *290*, 213.
- (27) Taylor, C. M. *Tetrahedron* **1998**, *54*, 11317.
- (28) Paulsen, H. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 823.
- (29) Meldal, M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 710.
- (30) Meldal, M.; Bock, K. *Glycoconj. J.* **1994**, *11*, 59.
- (31) Garg, H. G.; von dem Bruch, K.; Kunz, H. *Adv. Carbohydr. Chem. Biochem.* **1994**, *50*, 277.
- (32) Meldal, M.; St Hilaire, P. M. *Curr. Opin. Chem. Biol.* **1997**, *1*, 552.
- (33) Ansfield, S. T.; Lansbury, P. T. *J. Org. Chem.* **1990**, *55*, 5560.
- (34) Cohen-Ansfield, S. T.; Lansbury, P. T. *J. Am. Chem. Soc.* **1993**, *115*, 10531.
- (35) Roberge, J. Y.; Beebe, X.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1998**, *120*, 3915.
- (36) Stowell, C. P.; Lee, Y. C. *Adv. Carbohydr. Chem. Biochem.* **1980**, *37*, 225.
- (37) Aplin, J. D.; Wriston, J. C. *CRC Crit. Rev. Biochem.* **1981**, *10*, 259.
- (38) Marcaurelle, L. A.; Bertozzi, C. R. *Chem. Eur. J.* **1999**, *5*, 1384.
- (39) Lee, Y. C.; Stowell, C. P.; Krantz, M. J. *Biochemistry* **1976**, *15*, 3956.
- (40) Gray, G. R. *Arch. Biochem. Biophys.* **1974**, *163*, 426.
- (41) Bernstein, M. A.; Hall, L. D. *Carbohydr. Res.* **1980**, *78*, C1.
- (42) Zhang, J.; Yergey, A.; Kowalak, J.; Kovac, P. *Tetrahedron* **1998**, *54*, 11783.
- (43) Ragupathi, G.; Koganty, R. R.; Qiu, D. X.; Lloyd, K. O.; Livingston, P. O. *Glycoconj. J.* **1998**, *15*, 217.
- (44) McBroom, C. R.; Samanen, C. H.; Goldstein, I. J. *Methods. Enzymol.* **1972**, *28*, 212.
- (45) Buss, D. H.; Goldstein, I. J. *J. Chem. Soc. C* **1968**, 1457.
- (46) Quéard, C.; Bourgerie, S.; NormandSdiqui, N.; Mayer, R.; Strecker, G.; Midoux, P.; Roche, A. C.; Monsigny, M. *Bioconj. Chem.* **1998**, *9*, 268.
- (47) Lemieux, R. U.; Bundle, D. R.; Baker, D. A. *J. Am. Chem. Soc.* **1975**, *97*, 4076.
- (48) Arakatsu, Y.; Ashwell, G.; Kabat, E. A. *J. Immunol.* **1966**, *97*, 858.
- (49) Ashwell, G. *Methods Enzymol.* **1972**, *28*, 219.
- (50) Lönngren, J.; Goldstein, I. J.; Niederhuber, J. E. *Arch. Biochem. Biophys.* **1976**, *175*, 661.
- (51) Rude, E.; Westphal, O.; Hurwitz, E.; Fuchs, S.; Sela, M. *Immunochemistry* **1966**, *3*, 137.
- (52) Kamath, V. P.; Diedrich, P.; Hindsgaul, O. *Glycoconj. J.* **1996**, *13*, 315.
- (53) Ticha, M.; Cerny, M.; Trnka, T. *Glycoconj. J.* **1996**, *13*, 681.
- (54) Boratynski, J.; Roy, R. *Glycoconj. J.* **1998**, *15*, 131.
- (55) Wong, C. H.; Schuster, M.; Wang, P.; Sears, P. J. *Am. Chem. Soc.* **1993**, *115*, 5893.
- (56) Witte, K.; Seitz, O.; Wong, C. H. *J. Am. Chem. Soc.* **1998**, *120*, 1979.
- (57) Schuster, M.; Wang, P.; Paulson, J. C.; Wong, C.-H. *J. Am. Chem. Soc.* **1994**, *116*, 1135.
- (58) Granovsky, M.; Biefeldt, T.; Peters, S.; Paulsen, H.; Meldal, M.; Brockhausen, J.; Brockhausen, I. *Eur. J. Biochem.* **1994**, *221*, 1039.
- (59) Paulson, J. C.; Hill, R. L.; Tanabe, T.; Ashwell, G. *J. Biol. Chem.* **1977**, *252*, 8624.
- (60) Witte, K.; Sears, P.; Martin, R.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 2114.
- (61) Takegawa, K.; Tabuchi, M.; Yamaguchi, S.; Kondo, A.; Kato, I.; Iwahara, S. *J. Biol. Chem.* **1995**, *270*, 3094.
- (62) Cross, G. A. M.; Tackle, G. B. *Annu. Rev. Microbiol.* **1993**, *46*, 383.
- (63) Tsuboi, S.; Isogai, Y.; Hada, N.; King, J. K.; Hindsgaul, O.; Fukuda, M. *J. Biol. Chem.* **1996**, *271*, 27213.
- (64) Liu, Y.-L.; Hoops, G.; Coward, J. K. *Bioorg. Med. Chem.* **1994**, *2*, 1133.
- (65) Stanley, P. *Glycobiology* **1992**, *2*, 99.
- (66) Jenkins, N.; Parekh, R. B.; James, D. C. *Nature Biotechnol.* **1996**, *14*, 975.
- (67) Lee, E. U.; Roth, J.; Paulson, J. C. *J. Biol. Chem.* **1989**, *264*, 13848.
- (68) Tkacz, J. S.; Lampen, J. O. *Biochem. Biophys. Res. Commun.* **1975**, *65*, 218.
- (69) Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125.
- (70) Yarema, K. J.; Mahal, L. K.; Bruehl, R. E.; Rodriguez, E. C.; Bertozzi, C. R. *J. Biol. Chem.* **1998**, *273*, 31168.
- (71) Orr, G. A.; Rando, R. R. *Nature* **1978**, *272*, 722.

- (72) Schmidt, C.; Stehling, P.; Schnitzer, J.; Reutter, W.; Horstkorte, R. *J. Biol. Chem.* **1998**, *273*, 19146.
- (73) Rodriguez, E. C.; Winans, K. A.; King, D. S.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1997**, *119*, 9905.
- (74) Lemieux, G. A.; Bertozzi, C. R. *TIBTECH* **1998**, *16*, 506.
- (75) Cervigni, S. E.; Dumy, P.; Mutter, M. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1230.
- (76) Davis, N. J.; Flitsch, S. L. *Tetrahedron Lett.* **1991**, *32*, 6793.
- (77) Wong, S. Y. C.; Guile, G. R.; Dwek, R. A.; Arsequell, G. *Biochem. J.* **1994**, *300*, 843.
- (78) Macindoe, W. M.; van Oijen, A. H.; Boons, G.-J. *J. Chem. Soc., Chem. Commun.* **1998**, 847.
- (79) Davis, B. G.; Lloyd, R. C.; Jones, J. B. *J. Org. Chem.* **1998**, *63*, 9614.
- (80) Speculative aspects of this methodology have been discussed by Winter and co-workers, see Bill, R. M.; Winter, P. C.; McHale, C. M.; Hodges, V. M.; Elder, G. E.; Caley, J.; Flitsch, S. L.; Bicknell, R.; Lappin, T. R. *J. Biochim. Biophys. Acta* **1995**, *1261*, 35 and by Boons and co-workers, see Ref. 78, but details have not been published.
- (81) Wynn, R.; Richards, F. M. *Methods Enzymol.* **1995**, *201*, 351.
- (82) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
- (83) Fukuda, M. H., O. *Molecular Glycobiology*; Oxford University Press: Oxford, 1994.
- (84) Stabile, M. R.; Lai, W. G.; DeSantis, G.; Gold, M.; Jones, J. B.; Mitchinson, C.; Bott, R. R.; Graycar, T. P.; Liu, C.-C. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2501.
- (85) Ellman, G. L.; Courtney, K. D.; Andres, V.; Featherstone, R. M. *Biochem. Pharmacol.* **1961**, *7*, 88.
- (86) Davis, B. G.; Lloyd, R. C.; Jones, J. B. *Submitted for Publication* **1999**.
- (87) Plettner, E.; DeSantis, G.; Stabile, M. R.; Jones, J. B. *J. Am. Chem. Soc.* **1999**, *121*, 4977.
- (88) Jones, J. B.; Desantis, G. *Acc. Chem. Res.* **1999**, *32*, 99.
- (89) Dickman, M.; Lloyd, R. C.; Jones, J. B. *Tetrahedron: Asymm.* **1998**, *9*, 4099.
- (90) Lloyd, R. C.; Davis, B. G.; Jones, J. B. **1999**.
- (91) Matsumoto, K.; Khumtaveeporn, K.; Jones, J. B., Unpublished results.
- (92) Fordham-Skelton, T.; Gatehouse, J. A.; Davis, B. G. *Unpublished Results*.
- (93) Wong, S. Y. C. *Curr. Opin. Chem. Biol.* **1995**, *5*, 599.
- (94) Yarema, K. J.; Bertozzi, C. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 49.
- (95) Michael, A. *Nature Biotechnol.* **1997**, *15*, 1233.
- (96) Hill, T. G.; Wang, P.; Huston, M. E.; Wartchow, C. A.; Oehler, L. M.; Smith, M. B.; Bednarski, M. D.; Callstrom, M. R. *Tetrahedron Lett.* **1991**, *32*, 6823.
- (97) Wang, P.; Hill, T. G.; Wartchow, C. A.; Huston, M. E.; Oehler, L. M.; Smith, M. B.; Bednarski, M. D.; Callstrom, M. R. *J. Am. Chem. Soc.* **1992**, *114*, 378.
- (98) Wartchow, C. A.; Wang, P.; Bednarski, M. D.; Callstrom, M. R. *J. Org. Chem.* **1992**, *60*, 2216.
- (99) Wang, P.; Hill, T. G.; Bednarski, M. D.; Callstrom, M. R. *Tetrahedron Lett.* **1991**, *32*, 6827.
- (100) Baek, W. O.; Vijayalakshmi, M. A. *Biochim. Biophys. Acta* **1997**, *1336*, 394.
- (101) Simanek, E. E.; Huang, D.-H.; Pasternack, L.; Machajewski, T. D.; Seitz, O.; Millar, D. S.; Dyson, H. J.; Wong, C.-H. *J. Am. Chem. Soc.* **1998**, *120*, 11567.
- (102) Eisele, F.; Owen, D. J.; Waldmann, H. *Bioorg. Med. Chem.* **1999**, *7*, 193.
- (103) Tsuda, T.; Nishimura, S.-I. *J. Chem. Soc., Chem. Commun.* **1996**, 2779.
- (104) Ashwell, G.; Harford, J. *Annu. Rev. Biochem.* **1982**, *51*, 531.
- (105) Wadhwa, M. S.; Rice, K. G. *J. Drug Targeting* **1995**, *3*, 111.
- (106) Hassan, N.; Wu, G. Y. *Nature Medicine* **1995**, *1*, 210.
- (107) Wu, G. Y.; Wu, C. H. *Adv. Drug Deliv. Rev.* **1993**, *12*, 159.
- (108) Meijer, D. K. F.; Molema, G.; Moolenaar, F.; deZeeuw, D.; Swart, P. J. *J. Controlled Release* **1996**, *39*, 163.
- (109) Franssen, E. J. F.; Jansen, R. W.; Vaalburg, M.; Meijer, D. K. F. *Biochem. Pharmacol.* **1993**, *45*, 1215.
- (110) Molema, G.; Jansen, R. W.; Visser, J.; Herdewijn, P.; Moolenaar, F.; Meijer, D. K. F. *J. Med. Chem.* **1991**, *34*, 1137.
- (111) Molema, G.; Meijer, D. K. F. *Adv. Drug Deliv. Rev.* **1994**, *14*, 25.
- (112) Simon, P. M.; Goode, P. L.; Mobasser, A.; Zopf, D. *Infect. Immun.* **1997**, *65*, 750.
- (113) Avery, O. T.; Goebel, W. F. *J. Exp. Med.* **1929**, *50*, 533.
- (114) Francis, T.; Tillett, W. S. *J. Exp. Med.* **1930**, *52*, 573.
- (115) Ragupathi, G.; Park, T. K.; Zhang, S.; Kim, I. J.; Graber, L.; Adluri, S.; Lloyd, K. O.; Danishefsky, S. J.; Livingston, P. O. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 125.
- (116) Toyokuni, T.; Singhal, A. K. *Chem. Soc. Rev.* **1995**, 231.
- (117) Koganty, R. R.; Reddish, M. A.; Longenecker, B. M. *Drug Discovery Today* **1996**, *1*, 190.
- (118) Harding, C. V.; Kihlberg, J.; Elofsson, M.; Magnusson, G.; Unanue, E. R. *J. Immunol.* **1993**, *151*, 2419.
- (119) Brady, R. O.; Barton, N. W. In *Lectins and Cancer*; Gabius, H.-J., Gabius, S., Eds.; Springer Publ. Co.: New York, 1991.
- (120) Takakura, Y.; Masuda, S.; Tokuda, H.; Nishikawa, M.; Hashida, M. *Biochem. Pharmacol.* **1994**, *47*, 853.
- (121) Sakurai, K.; Miyazaki, K.; Koder, Y.; Nishimura, H.; Shingu, M.; Inada, Y. *Glycoconj. J.* **1997**, *14*, 723.
- (122) Anzai, J.-I.; Kobayashi, Y.; Nakamura, N. *J. Chem. Soc., Perkin Trans. 2* **1998**, 461.
- (123) Dintzis, R. Z.; Okajima, M.; Middleton, M. H.; Greene, G.; Dintzis, H. M. *J. Immunol.* **1989**, *143*, 1239.

Article Identifier:

1437-2096,E;1999,0,09,1495,1507,ftx,en;T00999ST.pdf