Modulation of protein structure and function by asparagine-linked glycosylation

Sarah E O'Connor and Barbara Imperiali

In eukaryotic cells, many enzymes are devoted to the construction of the complex glycan structures that decorate secreted and cell-surface proteins. Recent studies have begun to elucidate the effects of asparagine-linked glycosylation on protein folding and on the structure and function of mature glycoproteins.

Address: Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

Chemistry & Biology October 1996, **3**:803–812

 ISSN 1074-5521

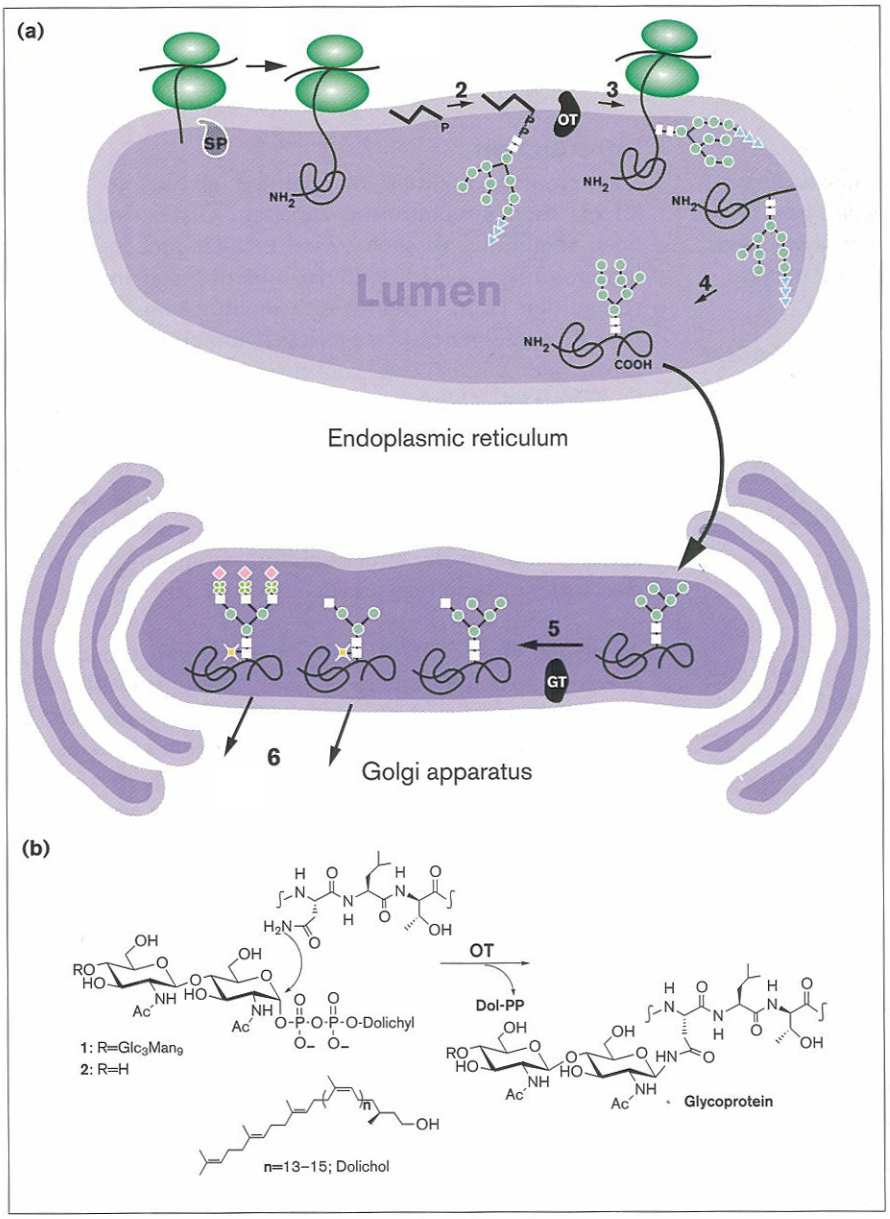
Introduction

The covalent addition of complex glycans to proteins affects their structure and function. Glycoproteins have been implicated in such varied cellular processes as the immune response, intracellular targeting and intercellular recognition [1,2]. Carbohydrate modifications of proteins fall into three general categories: *N*-linked modification of asparagine within the consensus sequence -Asn-Xaa-Ser/Thr- [3,4], *O*-linked modification of serine or threonine [5,6], and glycosylphosphatidyl inositol derivatization of the carboxy-terminal carboxyl group [7]. Here we examine how *N*-linked glycosylation affects the conformation of nascent and folded proteins, and we present some of the current views on the structural implications of this critical co-translational protein-processing event.

Asparagine-linked (*N*-linked) glycosylation occurs as a polypeptide is biosynthesized from the cognate mRNA on membrane-associated ribosomes (Fig. 1). Proteins that are destined to undergo *N*-linked glycosylation are translated with an amino-terminal signal peptide sequence that is removed by proteolysis as the nascent peptide emerges into the lumen of the rough endoplasmic reticulum (ER) [8]. Recent studies have shown that ~14 residues of the newly synthesized peptide must clear the luminal surface of the ER membrane before oligosaccharyl transferase (OT)-mediated glycosylation can occur [9]. This observation implies that the active site of the membrane-bound OT resides in the soluble domain of the enzyme that is exposed to the ER lumen, distal to the membrane surface. Although local secondary structure may influence the propensity for the glycosylation reaction to occur [10], it is generally accepted that global protein structure is not involved in the recognition events that lead to the modification reaction. The initial glycosylation reaction involves transfer of a conserved tetradecasaccharide ($GlcNAc_2Man_9Glc_3$) from the corresponding dolichyl-pyrophosphate-linked donor. The protein-bound glycan structure is then further processed and derivatized through the collective action of a variety of carbohydrate-specific glycohydrolase and glycosyl transferase enzymes [8].

Although the significance of glycosylation is not fully understood in all situations, some of the structural effects of glycosylation are now emerging [11]. Without co-translational glycosylation, immature proteins may misfold, aggregate, and be degraded before leaving the ER. Essentially, glycosylation influences the conformational dynamics of nascent polypeptides, thus influencing the folding process. Glycosylation also has a key role in maintaining

Figure 1



Co-translational protein glycosylation. **(a)** Cellular location of the steps in co-translational glycosylation. 1. Polypeptide synthesized on membrane-associated ribosomes contains a signal peptide sequence that directs it to translocate across the ER membrane. The signal peptide is removed from the protein as it emerges into the lumen of the ER by signal peptidase (SP). 2. Biosynthesis of dolichol-linked oligosaccharide, the donor for glycosylation. ❑ = *N*-acetylglucosamine (GlcNAc), ○ = mannose, Δ = glucose, ◇ = sialic acid, ⋈ = fucose, ⊗ = galactose. 3. Protein glycosylation catalyzed by oligosaccharyl transferase (OT). 4. Glucose trimming, followed by transfer to the Golgi apparatus. 5. Modification of the glycan moiety by glycohydrolase and glycosyl transferase (GT) enzymes. 6. Secretion at various stages of carbohydrate elaboration yields heterogenous glycoprotein products. **(b)** Asparagine-linked glycosylation. OT catalyzes the transfer of saccharide from the dolichyl-pyrophosphate-linked donor to the polypeptide as it emerges into the ER lumen.

the structure and stability of the final folded protein through long range hydrogen-bonding and hydrophobic interactions between the oligosaccharide and the protein. The precise location of the glycosylation site is unimportant in many cases, suggesting that the oligosaccharide has a global effect on structure that does not involve specific interactions [12].

Studies of glycopeptides and glycoproteins

In this review, we describe the analyses of *N*-linked glycopeptides and glycoproteins that have been explicitly carried out in aqueous media, and we highlight the applications of several different spectroscopic techniques in these investigations. While studies in less polar organic media have provided an alternative perspective on conformational effects of protein glycosylation, it is principally the studies in aqueous media that can be extrapolated to biological phenomena.

A variety of techniques have been used for the analysis of glycopeptide and glycoprotein structure. Nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopies, molecular-modeling predictions, fluorescence-energy transfer (FET) and site-directed mutagenesis studies have provided a structural picture of polypeptides, and have been used to investigate the effect of oligosaccharide

modification on the structure or stability of a peptide or protein. X-ray crystallography is not easily applied to glycoprotein systems, because crystallization is hampered by the heterogeneity of glycan structures and the conformational mobility of the saccharide branches on the surface of the protein.

The conformational behavior of small peptides and glycopeptides provides insight into how *N*-linked glycosylation affects protein secondary structure. Smaller peptide systems are simpler to interpret than large intact proteins; the structure of a glycosylated peptide can be directly compared to that of its nonglycosylated counterpart, providing specific information on the conformational consequences of the modification. Moreover, the size and flexibility of peptides may mimic the transient and fluxional secondary structure of partially folded, nascent polypeptides *in vivo*. Recent developments in the chemical and enzymatic synthesis of glycopeptides have provided sufficient amounts of homogeneous material for use in a variety of spectroscopic studies [13].

Unlike studies with model peptides, the complexity of glycoproteins hinders detailed structural analyses by spectroscopic techniques. In addition to spectroscopic analysis, however, conformational changes induced by glycosylation of proteins can be assessed by investigating biological functions such as antibody–epitope recognition or catalytic activity. Calorimetric evaluation of protein thermal stability can provide an indication of the global influence of glycosylation on the native protein structure [14].

The preparation of carbohydrate-free analogs of glycoproteins may be accomplished by several different methods. Site-directed mutagenesis of the essential consensus glycosylation sequence, -Asn-Xaa-Ser/Thr-, or expression of the protein in the presence of the glycosylation inhibitor tunicamycin, can afford an aglycosyl derivative. These methods produce nonglycosylated proteins that may fold either correctly or incorrectly from nonglycosylated folding intermediates. Alternatively, the carbohydrates may be removed from a secreted, folded glycoprotein by enzymatic deglycosylation. This process probes the effect on protein conformation of removing the saccharide from a glycoprotein that is already folded. Removal of more than one carbohydrate attachment site is often required to disrupt the folding process, indicating that saccharides work in concert to generate a functional protein conformation [12]. Interestingly, the glycosylation process itself may be cooperative. For example, immunoglobulins generally have several glycosylation sites and are heterogeneously modified; cells grown in the presence of tunicamycin, however, produce only nonglycosylated or fully glycosylated structures, suggesting that the initial glycosylation instigates subsequent glycosylation at other sites [15].

Effects on peptide secondary structure

Surface peptides from hemagglutinin

Peptides representing the sequence of two β-turn surface loops from the HA_1 chain of hemagglutinin have been examined using time-resolved FET techniques [16]. The folded conformation of the glycopeptides in the corresponding protein is known from the X-ray crystal structure of hemagglutinin [17]. Specific conformational features in fluxional peptides can be assessed by FET studies, through the measurement of specific interfluorophore distances. The high sensitivity (μM sample concentrations) of fluorescence measurements, and the fact that they are compatible with the use of aqueous media at neutral pH, makes this method valuable for the study of the conformational effects of protein glycosylation.

The FET study examined the sequences 19–26 (Ala-Val-Pro-Asn-Gly-Thr-Leu-Val) and 282–288 (Ile-Thr-Pro-Asn-Gly-Ser-Ile) of the HA_1 chain. These sequences are from surface loops that form β-turns in the final folded protein [17]. The β-turn is of considerable interest, since this motif has been previously identified as a common feature of glycosylation sites [18]. This motif has also been implicated as a structural nucleation element in protein folding [19]. The native hemagglutinin sequences were modified to allow for the incorporation of two fluorophores, the dansyl and the indole moieties, on opposite sides of the glycosylated region. The R_0 (the interfluorophore distance for 50 % energy transfer) of this pair of fluorophores is 17–19 Å, making them appropriate to use with these sequences, as the interprobe distance in the peptide can vary from <9–30 Å depending upon conformation. The glycopeptides were prepared from the corresponding peptides by oligosaccharyl-transferase-catalyzed glycosylation using dolichylpyrophosphoryl-*N*,*N'*-diacetylchitobiose as the lipid-linked carbohydrate donor. In time-resolved measurements, the donor fluorescence decays were analyzed for the effect of glycosylation on energy transfer, providing information on the distributions of interprobe distances and thus revealing discrete information on the peptide conformation.

When examined in aqueous solution, the average interfluorophore distance dropped significantly on glycosylation. The distance changed from 14.8 Å to 9.5 Å for the peptide corresponding to A19–A26 and from 12.4 Å to 7.7 Å for the peptide corresponding to A282–A288. This result is consistent with a model indicating that glycosylation causes the two peptides to adopt more compact, folded conformations, such as β-turns, upon glycosylation.

T cell epitopes: conformation and antigenicity

Glycosylation sites are frequently found within the regions of viral glycoproteins that are recognized as T-cell epitopes, and it is now established that glycosylation may strongly influence protein conformation and antigenicity. This effect

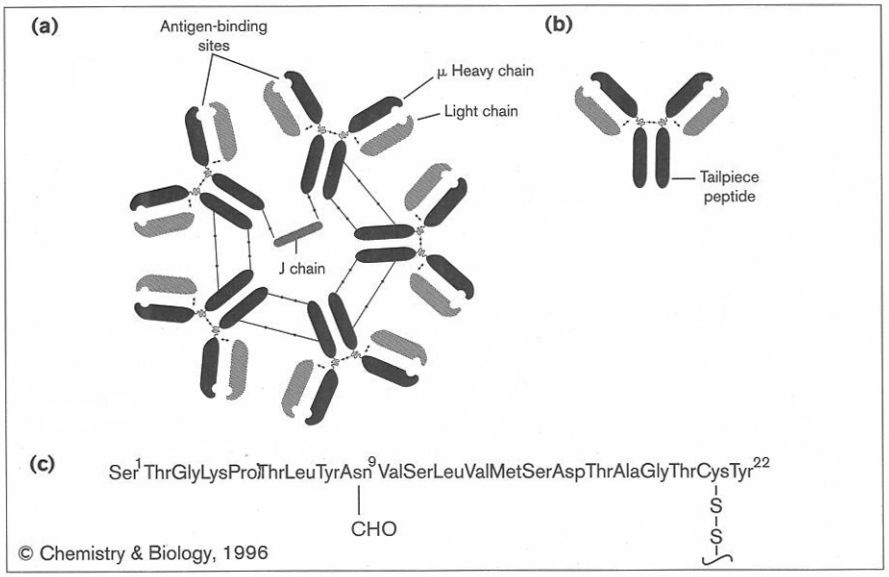
of glycosylation is dependent on the position of the carbohydrate in the protein [20]. Glycosylation induces discrete changes in the secondary structure of peptides representing two epitopes in the rabies virus glycoprotein (VF13, amino acids 29–41: Val-Val-Glu-Asp-Glu-Gly-Cys-Thr-Asn-Leu-Ser-Gly-Phe and GM12, amino acids 312–323: Gly-Lys-Ala-Tyr-Thr-Ile-Phe-Asn-Lys-Thr-Leu-Met), as observed by CD and Fourier transform infrared spectroscopy [21]. The glycopeptides for this study were prepared by chemical synthesis using protected, glycosylated asparagine building blocks that are suitable for incorporation into glycopeptides via solid-phase synthesis methods. Glycosylation with one or two *N*-acetyl glucosamines or one glucose residue disrupted helical (α or 3_{10}) secondary structure and seemed to induce the formation of a type I (III) β-turn. The most dramatic effects were observed upon addition of a single, simple carbohydrate; the effect was not dependent on the conformation of the C-2 acetamido group on the sugar that is directly attached to the oligopeptide. Later studies with both the rabies virus glycoprotein [22] and influenza virus hemagglutinin [23] showed that the T cell stimulatory activity and sensitivity to degradation in human serum could be modified by the site-specific incorporation of a carbohydrate group into the antigenic peptides. Carbohydrate-induced modulation of peptide secondary structure has also been demonstrated with the peptide T fragment of HIV gp120, which inhibits binding of the virus to the CD4 receptor [24]. The ability to manipulate these peptide functions has significant implications for the design of T-cell agonists and peptide-based vaccines [20,23].

The tailpiece peptide from serum IgM

The 22-residue peptide from the carboxy-terminal domains of human serum IgM, containing a single conserved glycosylation site at Asn563, has been investigated by NMR [25]. Glycosylation of the tailpiece peptide may influence the formation of intersubunit disulfide bridges. These form between pairs of Cys575 residues within heavy chains of the five Ig subunits in the intact pentameric immunoglobulin M (IgM), and between these residues on two of the IgM chains and the J chain (Fig. 2). For the NMR studies, the glycopeptide was obtained by limited proteolysis of IgM from the serum of a patient with Waldenstrom's macroglobulinemia, which contains unusually high levels of IgM. The structure of the glycopeptide was compared to that of the corresponding nonglycosylated species.

Although the presence of peptide had little effect on the conformation and dynamics of the oligosaccharide, addition of the carbohydrate moiety clearly caused a decrease in the conformational mobility of the polypeptide backbone (relative to the isolated peptide) in the regions proximal to the glycosylation site. This reduced mobility may favor formation of disulfide bonds. The tailpiece glycopeptide showed distinct conformational tendencies in different regions of the sequence; NOEs typical of an extended conformation were observed for residues 6–9 and NOEs indicative of a more compact structure were noted for residues 9–13. In contrast, residues 1–6 and 13–22 lacked any strong NOEs, suggesting a highly fluxional ensemble of conformations. A detailed analysis of the *N*-glycosidic bond revealed that this bond is both rigid and planar, suggesting that the decreased flexibility of the glycopeptide may be a direct consequence of the reduction in conformational space accessible to this portion of the modified structure. Although several different glycoforms ($Man_nGlcNAc_2$; n = 6,7,8 or 9) of the tailpiece

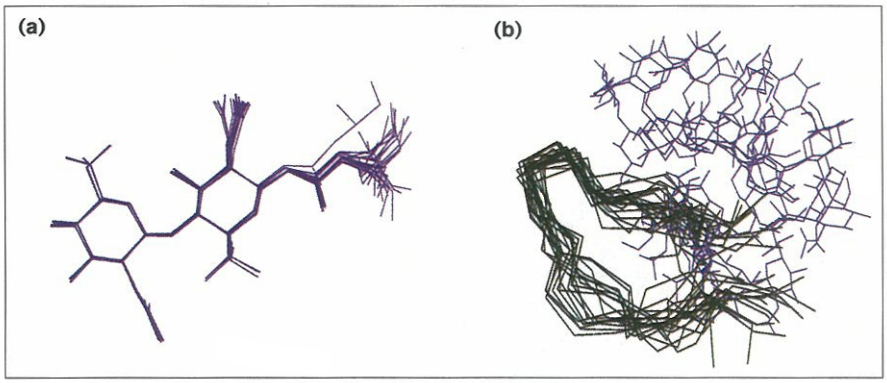
Figure 2



Core carbohydrate moieties affect the conformational mobility of the tailpiece peptide from serum IgM. **(a)** Diagrammatic representation of IgM pentamer. Disulfide bonds between the light and heavy chains are denoted by dots. **(b)** IgM monomer indicating location of tailpiece peptide. **(c)** Sequence of the tailpiece peptide (residues 555–576 of the heavy chain) showing the glycosylation site and the position of the cysteine residue involved in the intersubunit disulfide. Glycosylation causes a decrease in the conformational mobility of the polypeptide backbone, which may favor the formation of disulfide bonds.

Figure 3

^{1}H-NMR analysis of the glycosylated extracellular loop peptide from nAChR. **(a)** Superimposition of the chitobiose unit and the asparagine residue. **(b)** Superimposition of α-carbon traces for the nAChR loop polypeptide with the carbohydrate residues shown in purple. Disorder of the carbohydrate moiety relative to the polypeptide backbone suggests the absence of specific sugar–peptide interactions. Reprinted with permission from [31].



glycopeptide were present in the analysis, only a single set of peptide peaks was observed in the proton NMR spectra. This observation indicates that the peptide conformation is insensitive to changes in the outer antennae of the oligosaccharide moiety and that the effects on local peptide mobility are principally influenced by the core carbohydrate moieties.

Hen ovomucoid glycopeptide

^{1}H NMR has also been used to compare the conformations of a peptide and glycopeptide from hen ovomucoid [26]. The peptide sequence analyzed was Ser-Ile-Glu-Phe-Gly-Thr-Asn-Ile-Ser-Lys and the corresponding glycopeptide was modified with the core pentasaccharide $Man_3GlcNAc_2$. The homogeneous glycopeptide was prepared from the hen ovomucoid B_α-2 mixture by limited proteolysis and sequential treatment with several specific glycosidases [27]. In this example, NMR analysis showed that attachment of the pentasaccharide core did not have a dramatic effect on the average conformation of the polypeptide backbone, but it did appear that the core GlcNAc-1 C2 side chain conformation in the glycopeptide was more rigid in the free pentasaccharide. This conformational effect was proposed to be due to the adjacent *N*-glycosidic linkage. These studies did not reveal the presence of any specific hydrogen-bonding interactions between the carbohydrate and peptide components of the conjugate. Unfortunately, although secondary structure prediction analysis has suggested that this sequence from ovomucoid may be involved in a β-turn [28], the ovomucoid domain from which this peptide is taken has not been characterized crystallographically. It is thus not yet possible to assess whether the apparent lack of a discrete conformational change on glycosylation is expected proximal to the glycosylation site in the native protein structure.

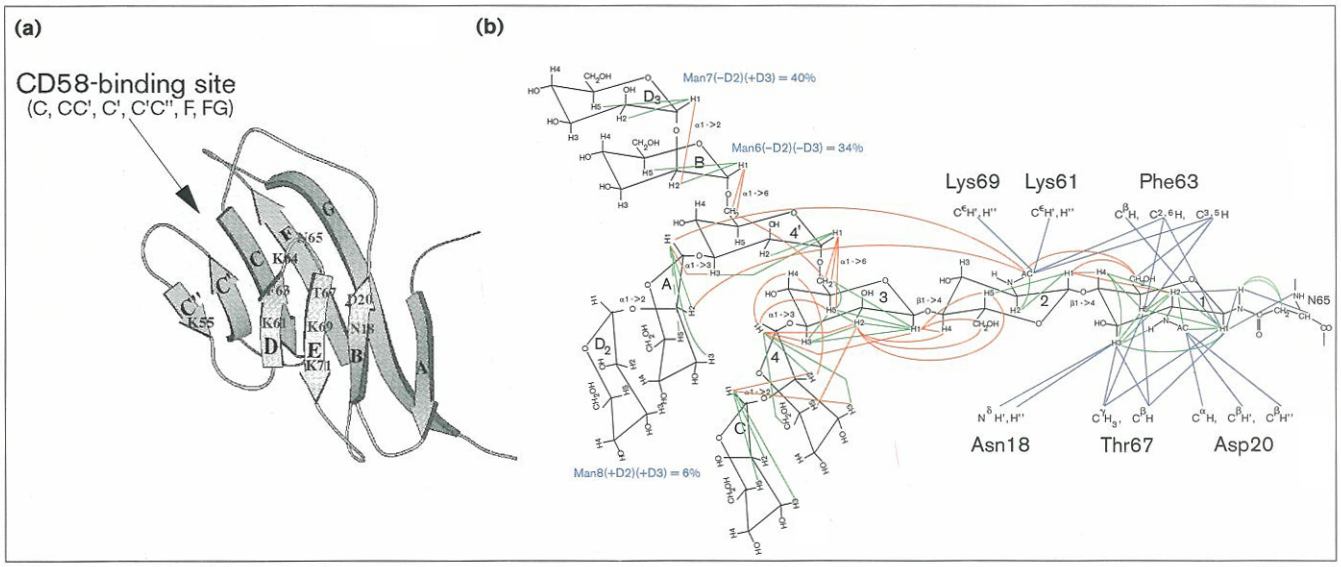
Glycosylation facilitates protein folding

The *cis/trans* proline amide equilibrium in an nAChR peptide

Site directed mutagenesis studies have suggested that specific, highly conserved glycosylation sites may be critical for subunit folding and oligomerization of multimeric proteins. The subunits of the pentameric nicotinic acetylcholine receptor (nAChR) [29] contain a highly conserved 15-residue loop peptide with the consensus sequence Cys-Xaa-Xaa-Xaa-Val-Xaa-Xaa-Phe-Pro-Phe-Asp-Xaa-Gln-Asn-Cys-Thr/Ser, including a disulfide bridge between the two cysteines. This polypeptide forms part of the soluble, extracellular domain of the protein and is thought to be involved in intersubunit interactions that are important for the assembly of intact $\alpha_2\beta\gamma\delta$ complexes. The loop sequence in the mature α-1 subunit of the neuromuscular nAChR from *Torpedo californica* comprises residues 128–143, including an asparagine-linked glycosylation site (-Asn-Cys-Thr-, that includes Cys142), and a proline (residue 136) at the remote end of the loop formed by the disulfide. Site-directed mutagenesis studies have revealed that the glycosylation site is essential for the proper assembly and stability of the receptor complex [30], but the mechanism whereby glycosylation influences the structure of the complex is unknown. The size of this polypeptide made it amenable to a detailed structural analysis, providing insight into the structure of the loop and the importance of the conserved glycosylation site.

NMR studies of the conformation of the glycosylated and nonglycosylated loop peptide (Ac-Tyr-Cys-Glu-Ile-Ile-Val-Thr-His-Phe-Pro-Phe-Asp-Gln-Gln-Asn-Cys-Thr-NH_2) of the nAChR α-subunit revealed that glycosylation had a significant influence on the conformational dynamics of the system [31]. The glycopeptide was prepared chemically and was derivatized with a chitobiosyl disaccharide. After glycosylation, the dithiol/disulfide equilibrium was shifted in favor of the oxidized species, suggesting that the modification brings the termini of the loop into closer proximity, as in the mature protein. Furthermore, in the oxidized peptide, glycosylation significantly altered the *cis–trans* proline equilibrium (relative to the unglycosylated species), favoring the *trans* isomer. It is interesting to note that the folding of the homopentameric α-7 variant of this subunit, which lacks the glycosylation site corresponding to Asn142

Figure 4



Glycosylation of CD2 maintains a functional CD58-binding site by stabilizing a cluster of surface-exposed lysine residues. **(a)** Ribbon diagram of a single representative structure of the adhesion domain of human CD2 ($hsCD2_{105}$), which is composed of two β-sheets containing three (D, E, and B) and five (C'', C', C, F, and G) antiparallel β-strands. **(b)** Summary of the NOE data obtained for the high mannose *N*-glycan of $hsCD2_{105}$. Reprinted with permission from [46].

but still contains the Phe135-Pro136 sequence, is uniquely dependent on the peptidyl proline isomerase cyclophilin [32]. These conformational studies suggest that the cyclophilin-independence of α-1 folding may be a consequence of glycosylation. Since both disulfide-bond formation and *cis–trans* proline equilibration are among the slower steps in the protein folding process, it becomes clear that *N*-linked glycosylation has the potential to significantly alter both the course of protein folding and the final folded protein structure.

^{1}H-NMR analysis of the glycopeptide did not reveal any specific hydrogen-bonding interactions between the carbohydrate and the polypeptide (Fig. 3). These studies suggest that the conformational effects of the oligosaccharide may arise from either steric perturbation or a subtle modulation of the local water environment surrounding the polypeptide.

Multiple glycosylation sites in CD4 and gp120

The extracellular portion of the T-cell marker CD4 serves as the binding site for the HIV envelope glycoprotein gp120. Glycosylation sites in human CD4 are located at Asn271 and Asn300; although glycosylation is necessary for the efficient folding and secretion of CD4 [33], elimination of the saccharide moieties after folding does not affect gp120 binding [34]. Gp120 contains over 20 glycosylation sites, with carbohydrate comprising >50% of the molecular weight of the native protein. Completely nonglycosylated gp120, produced by site-directed mutagenesis of the glycosylation sites, fails to bind to CD4 [35], and mutation of three glycosylation sites leads to reduced CD4 binding efficiency [36]. A single glycosylation site mutation in the V1 loop of gp120 affects the overall conformation of the protein, but not viral infectivity [37]. If gp120 is deglycosylated after secretion, however, it binds to CD4. It is interesting to note that fully glycosylated gp120 cannot bind to CD4 immediately after synthesis; time to fold into a mature conformation is required. It is possible that non-native disulfide bonds may form in nonglycosylated gp120, resulting in a nonfunctional conformation [35]. The complete inhibition of CD4 binding by mutation of two conserved cysteine residues (425 and 452) in gp120 or by treatment with reducing agent substantiates this hypothesis [36]. NMR studies of the gp120 V3 loop in solution revealed a structure that is, on the whole, disordered, but contains two β-turns, one of which is located at the glycosylation site at Asn6 [38]. This illustrates once again that the β-turn appears to be a common feature of glycosylation sites.

Alteration of refolding rates in yeast invertase glycoforms

The availability of homogeneously glycosylated invertase from yeast has made it possible to study the role of the saccharide in the native protein [39]. The nonglycosylated protein folds more slowly than glycosylated forms and has an increased tendency to aggregate [40]. Attempts to suppress aggregation by the addition of the chaperone GroEL to the renaturation process surprisingly prevented renaturation. This effect is thought to be due to the binding of

the nonglycosylated species to the molecular chaperone [41], and it has been suggested that the hydrophilic saccharide moieties may prevent certain deleterious interactions with GroEL by directing the chaperone to other regions of the unfolded polypeptide. The interactions of immature glycoproteins and molecular chaperones appear to work in concert to ensure the quality control of exported protein products.

Glycosylation can stabilize fully folded proteins

Stabilization of unfavorable electrostatic interactions in CD2

The T-cell receptor CD2 is a member of the immunoglobulin superfamily. This protein binds to the antigen-presenting-cell receptor CD58 (for a recent review see [42]). The binding patches of both proteins are highly charged, indicating that adhesion primarily depends on electrostatic interactions. Human CD2 must be glycosylated at Asn65 to bind to CD58 [43], although deglycosylated mutants leave the ER and are expressed on the cell surface, indicating that correct folding occurs.

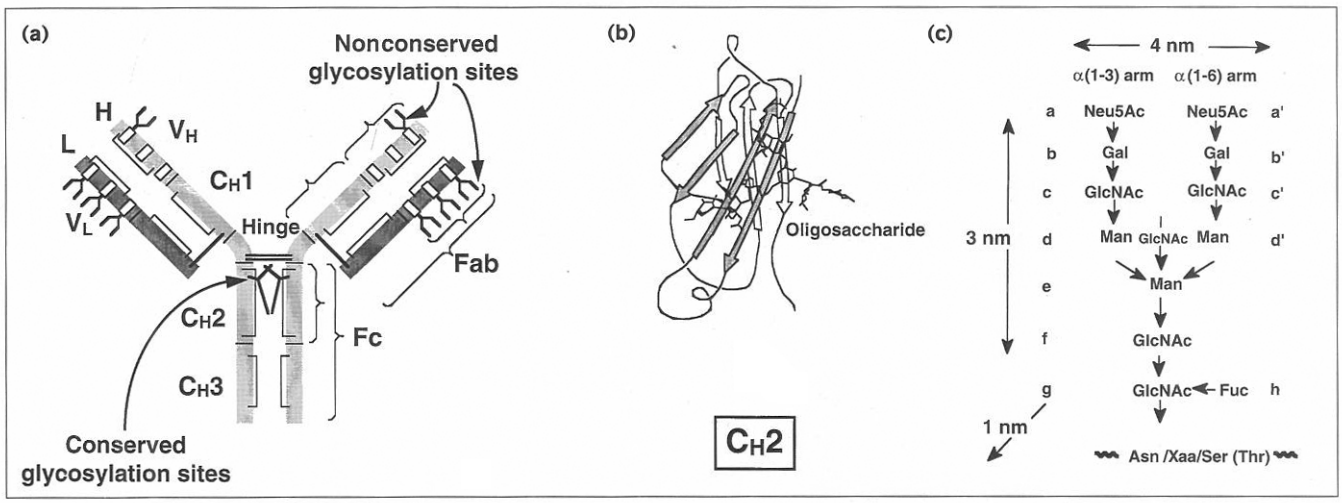
X-ray crystallographic studies have revealed that the carbohydrate on CD2 ($GlcNAc_2Man_{5-9}$) fills the cavity between two β-sheet structures that are formed by the B, E, and D β-strands and the C, F and G β-strands of the CD2 structure [44,45] (Fig. 4a). This placement of the carbohydrate positions it away from the CD58-binding site; therefore, the saccharide is not directly involved in the interaction between CD2 and CD58. NMR studies have shown instead that the saccharide maintains a functional conformation of the CD58-binding site on CD2 by stabilizing a positively charged cluster of five surface-exposed lysine residues (Lys55, Lys61, Lys64, Lys69, and Lys71) [46] (Fig. 4b). Mutagenesis studies provide complementary and independent evidence for the destabilizing effect of the 'lysine cluster' in the nonglycosylated CD2 molecule. Substitution of one of these lysine residues (Lys61) with a negatively charged glutamate stabilizes the surface cationic site and eliminates the glycan requirement [46]. Rat CD2, which has a glutamine substituted in place of Lys61, does not need to be glycosylated to bind to its ligand, CD48. This example emphasizes the essential, but indirect role that carbohydrate modifications have in protein structure and function. The protein/carbohydrate interactions in CD2 result in a net stabilization of the protein surface distal to the glycosylation site, presenting a stable binding site to the cognate receptor, CD58.

The NMR analysis also provided information on the structure of the high mannose triantennary branched saccharide [46]. NOE contacts between the Man4 and GlcNAc2 were observed, indicating that this branch is folded into the Man(3)GlcNAc(2)GlcNAc(1) core when one of the terminal mannoses is trimmed. Broadened NMR line widths also indicate that this branch is less mobile than the other two.

Stabilization of IgG and rheumatoid arthritis

Immunoglobulin G (IgG) contains a conserved glycosylation site in a turn segment between two β-sheets in the C_H2 subunit at Asn297 (Fig. 5). This site seems to be required for numerous biological functions. For example, IgG cannot bind to complement factor C1q if it is deglycosylated [47]. An ensemble of complex biantennary saccharides are present

Figure 5



Glycosylation of the conserved site in the C_H2 domain of the IgG heavy chain is essential for biological function. **(a)** Schematic representation of IgG showing domain structure and location of glycosylation sites. **(b)** Ribbon diagram of C_H2 domain showing location of conserved glycosylation site. (Note that the size of a single IgG domain and a fully extended complex oligosaccharide are similar.) **(c)** Biantennary complex oligosaccharide chain linked to Asn297. Reprinted with permission from [47].

at Asn297, with structural differences arising from oligosaccharide core modification (e.g., with fucose), outer arm galactosylation and sialylation (Fig. 5c) [48].

The carbohydrates are sequestered between the two C_H2 domains of the protein, and numerous protein-carbohydrate interactions are evident in the X-ray crystal structure [49]. Although the flexible α(1–6) antenna can assume two conformations, interaction of the galactoses with Phe243, Pro246 and Thr260 stabilizes a single conformation. Molecular modeling calculations with a 27-residue peptide containing Asn297 showed strong interactions between a terminal galactose and Asn297, Tyr296 and Arg301 [50]. Presumably these interactions stabilize the conformational integrity of the protein. It has been proposed that there is a lateral movement of domains in the hinge region in the absence of sugars. Moreover, the removal of saccharides results in a 60-fold increase in the rate of C_H2 cleavage by trypsin [47].

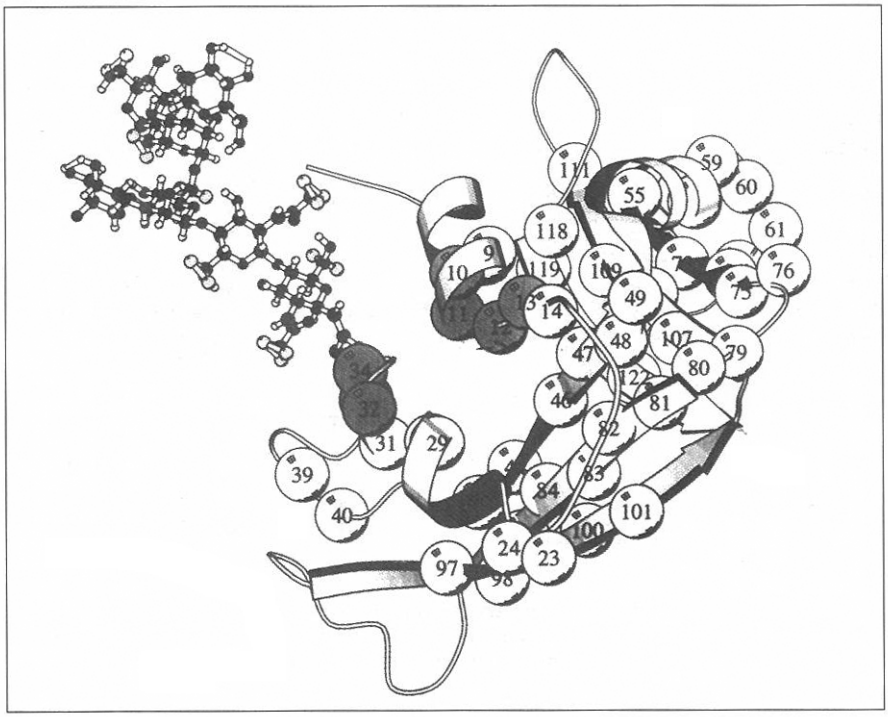
If the terminal galactoses are removed, these protein/carbohydrate interactions are no longer possible, leading to increased saccharide mobility. Although the overall tertiary structures of glycosylated and nonglycosylated C_H2 domains are identical, altered local conformation, particularly at the Gln295 and Thr299 residues, of the agalactosyl protein have been demonstrated in NMR studies [47]. Patients with the autoimmune disorder rheumatoid arthritis (RA) have significantly higher populations of IgG proteins that lack the galactose modification. It has been proposed that increased saccharide mobility in the IgG proteins from these individuals may lead to exposure of the protein surface, causing aggregation or revealing new epitopic sites. Therefore the structural observations described above may explain the symptoms of RA that result from an autoimmune response. In fact, the glycosylation state of the IgG molecules has been used as a diagnostic marker for RA [47].

The dynamic stability of RNase B

RNase is responsible for the hydrolysis of 3',5'-phosphodiester linkages of ribonucleic acids. Native RNase consists of a mixture of nonglycosylated RNase A and RNAse B, a collection of glycoforms substituted on Asn 34 with $Man_xGlcNAc_2$ (x = 5,6,7,8,9). RNase B, modified with truncated carbohydrate structures (x = 0–5) can be prepared by the selective hydrolysis of native RNase B with mannosidase enzymes. RNase activity decreases as larger oligosaccharides are present [51] .

RNase A and the RNase B glycoforms have identical tertiary structures [52]. The saccharide, however, may increase the overall dynamic stability of the protein, as shown by reduced amide proton/deuterium exchange rates for residues both adjacent to (residues 29–35) and remote from (residues 57–61 and 75–76) Asn34 [53]. It has been proposed that the carbohydrate shields the amides by hydrogen bonding contacts or solvent exclusion. The structure of RNase B highlighting the amide protons with perturbed exchange rates is illustrated in Figure 6.

Figure 6



Glycosylation increases the overall dynamic stability of RNase B. A schematic representation of RNase B is shown, highlighting the residues (identified by circles) with amide protons that demonstrate reduced hydrogen–deuterium exchange rates in the glycoprotein. Filled circles represent residues with amide protons that show exchange rates that are dependent on the glycoform (RNaseB, RNase Man-1 or RNase Man-5). Reprinted with permission from [54].

Exchange rates for six residues are dependent on the RNase B glycoform [51]. These rates decrease as the saccharide chains lengthen from x = 0 to x = 5; exchange rates show no further change with the addition of more mannose units. Molecular modeling of RNase reveals a cationic binding site, capable of forming a salt bridge between the enzyme and substrate, that would be occluded by carbohydrate in the glycosylated protein. Thus, the binding between enzyme and substrate deteriorates as saccharides are added. However, since the Man5 isoform covers almost 80 % of the Man9 surface area, the activity reaches a plateau from x = 5 to x = 9.

Conclusion

N-linked glycosylation is one of the most complex of enzyme-catalyzed protein modification reactions. Many enzymes are devoted to the construction and elaboration of the complex glycan structures that are found decorating eukaryotic proteins. It is therefore not surprising that this class of modifications is implicated in a wide variety of structural and functional roles. *N*-linked glycosylation is unusual since it occurs concomitant with polypeptide synthesis. This timing means that the modification can influence early events in protein folding. The structural and functional analyses of glycoproteins emphasize the fact that, because of the large relative size of the oligosaccharide moieties, this modification can mediate significant effects not only locally, on the adjacent polypeptide sequence, but also at remote sites distant from the modified asparagine residue.

References

1. Dwek, R.A. (1996). Glycobiology – toward understanding the function of sugars. *Chem. Rev.* **96**, 683–720.
2. Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* **3**, 97–130.
3. Imperiali, B. & Hendrickson, T.L. (1995). Asparagine-linked glycosylation: specificity and function of oligosaccharyl transferase. *Bioorg. Med. Chem.* **3**, 1565–1578.
4. Silberstein, S. & Gilmore, R. (1996). Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *FASEB J.* **10**, 849–858.
5. Jentoft, N. (1990). Why are proteins *O*-glycosylated? *Trends Biochem. Sci.* **15**, 291–294.
6. Wilson, I.B.H., Gavel, Y. & von Heijne, G. (1991). Amino acid distributions around *O*-linked glycosylation sites. *Biochem. J.* **275**, 529–534.
7. Englund, P.T. (1993). The structure and biosynthesis of glycosyl phosphatidylinositol anchors. *Annu. Rev. Biochem.* **62**, 121–138.
8. Hubbard, S.C. & Ivatt, R.J. (1981). Synthesis and processing of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **50**, 555–583.
9. Nilsson, I. & von Heijne, G. (1993). Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* **268**, 5798–5801.
10. Imperiali, B., Spencer, J.R. & Struthers, M.D. (1994). Structural and functional characterization of a constrained Asx–turn motif. *J. Am. Chem. Soc.* **116**, 8424–8425.
11. Wyss, D.F. & Wagner, G. (1996). The structural role of sugars in glycoproteins. *Curr. Opin. Biotechnol.* **7**, 409–416.
12. Helenius, A. (1994). How *N*-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. *Mol. Biol. Cell* **5**, 253–265.
13. Bill, R.M. & Flitsch, S.L. (1996). Chemical and biological approaches to glycoprotein synthesis. *Chemistry & Biology* **3**, 145–149.
14. Wang, C., Eufemi, M., Turano, C. & Giartosio, A. (1996). Influence of the carbohydrate moiety on the stability of glycoproteins. *Biochemistry* **35**, 7299–7307.
15. Cushley, W., Singer, H.H. & Williamson, A.R. (1983). Cooperative nature of *N*-glycosylation of proteins at multiple sites: evidence from studies with tunicamycin. *Biosci. Rep.* **3**, 331–336.
16. Imperiali, B. & Rickert, K.W. (1995). Conformational implications of asparagine-linked glycosylation. *Proc. Natl. Acad. Sci. USA* **92**, 97–101.
17. Wilson, I.A., Skehel, J.J. & Wiley, D.C. (1981). Structure of the hemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**, 366–373.
18. Beintema, J.J. (1986). Do asparagine linked carbohydrate chains in glycoproteins have a preference for β-bends? *Biosci. Rep.* **6**, 709–714.
19. Oka, M., Montelione, G.T. & Scheraga, H.A. (1984). Chain-folding initiation structures ribonuclease A. *J. Am. Chem. Soc.* **106**, 7959–7969.
20. Otvos, L., *et al.*, & Ertl, H.C. (1994). Glycosylation of synthetic T-helper cell epitopic peptides influences their specific antigenic potency and conformation in a sugar location-specific manner. *Biochim. Biophys. Acta* **1224**, 68–76.
21. Otvos, L., Thurin, J., Kollat, E., Urge, L., Mantsch, H.M. & Hollosi, M. (1991). Glycosylation of synthetic peptides breaks helices. *Int. J. Pept. Protein Res.* **38**, 476–482.
22. Otvos, L., *et al.*, & Ertl, H. (1995). Comparison of the effects of amino acid substitutions and beta-N- vs. alpha-O-glycosylation on the T-cell stimulatory activity and conformation of an epitope on the rabies virus glycoprotein. *Biochim. Biophys. Acta* **1267**, 55-64.
23. Jackson, D.C., Drummer, H.E., Urge, L. & Otvos, L. (1994). Glycosylation of a synthetic peptide representing a T-cell determinant of influenza virus hemagglutinin results in loss of recognition by CD4(+) T-cell clones. *Virology* **199**, 422–430.
24. Urge, L., Gorbics, L. & Otvos, L. (1992). Chemical glycosylation of peptide T at natural and artificial glycosylation sites stabilizes or rearranges the dominant reverse turn structure. *Biochem. Biophys. Res. Commun.* **184**, 1125–1132.
25. Wormald, M.R., *et al.*, & Dwek, R.A. (1991). The conformational effects of *N*-glycosylation on the tailpiece from serum IgM. *Eur. J. Biochem.* **198**, 131–139.
26. Davis, J.T., Hirani, S., Bartlett, C. & Reid, B.R. (1994). ^{1}H-NMR studies on an Asn-linked glycopeptide. *J. Biol. Chem.*, **269**, 3331–3338.
27. Yet, M.G., Chin, C.C.Q. & Wold, F. (1988). The covalent structure of individual *N*-linked glycopeptides from ovomucoid and asialofetuin. *J. Biol. Chem.* **263**, 111–117.
28. Beeley, J.G. (1977). Peptide chain conformation and the glycosylation of glycoproteins. *Biochem. Biophys. Res. Commun.* **76**, 1051–1055.
29. Stroud, R.M., McCarthy, M.P. & Shuster, M. (1990). Nicotinic acetylcholine receptor superfamily of ligand-gated ion channels. *Biochemistry* **29**, 11009–11023.
30. Blount, P. & Merlie, J.P. (1990). Mutational analysis of muscle nicotinic acetylcholine receptor subunit assembly. *J. Cell. Biol.* **111**, 2613–2622.
31. Rickert, K.W. & Imperiali, B. (1995). Analysis of the conserved glycosylation site in the nicotinic acetylcholine receptor: potential roles in complex assembly. *Chemistry & Biology* **2**, 751–759.
32. Helekar, S.A., Char, D., Neff, S. & Patrick, J. (1994). Prolyl isomerase requirement for the expression of functional homo-oligomeric ligand-gated ion channels. *Neuron* **12**, 179–189.
33. Tifft, C.J., Proia, R.L. & Camerini-Otero, R.D. (1993). The folding and cell surface expression of CD4 requires glycosylation. *J. Biol. Chem.* **267**, 3268–3273.
34. Arthos, J., *et al.*, & Sweet, R.W. (1989). Identification of the residues on human CD4 critical for the binding of HIV. *Cell* **57**, 467–481.
35. Li, Y., Luo, L., Rasooi, N. & Yong Kang, C. (1993). Glycosylation is necessary for the correct folding of human immunodeficiency virus gp120 and in CD4 binding. *J. Virol.* **67**, 584–588.
36. Dirckx, L., Lindemann, D., Ette, R., Manzoni, C., Moritz, D. & Mous, J. (1990). Mutation of conserved *N*-glycosylation sites around the CD4 binding site of human immunodeficiency virus type 1 gp120 affects viral infectivity. *Virus Res.* **18**, 9–20.
37. Gram, G.J., *et al.*, & Hanson, J.E. (1994). Identification of an *N*-linked glycan in the V1 loop of HIV-1 gp120 influencing neutralization by anti-V3 antibodies and soluble CD4. *Arch. Virol.* **139**, 253–261.
38. Vrankon, W.F., Budesinsky, M., Fant, F., Boulez, K. & Borremans, F.A.M. (1995). The complete consensus V3 loop peptide of the envelope protein gp120 of HIV-1 shows pronounced helical character

in solution. *FEBS Lett.* **374**, 117–121.
39. Kern, G., Schulke, N., Schmid, F.X. & Jaenicke, R. (1992). Stability, quaternary structure, and folding of internal, external, and core glycosylated invertase from yeast. *Protein Sci.* **1**, 120–131.
40. Kern, G., Kern, D., Jaenicke, R. & Seckler, R. (1993). Kinetics of folding and association of differently glycosylated variants of invertase from *Saccharomyces cerevisiae*. *Protein Sci.* **2**, 1862–1868.
41. Kern, G., Schmidt, M., Buchner, J. & Jaenicke, R. (1992). Glycosylation inhibits the interaction of invertase with the chaperone GroEL. *FEBS Lett.* **305**, 203–205.
42. Davis, S.J. & van der Merwe, P.A. (1996). The structure and ligand interactions of CD2: implications for T-cell function. *Immunol. Today* **17**, 177–187.
43. Davis, S.J., *et al.*, & van der Merwe, P.A. (1995). Ligand binding by the immunoglobulin superfamily recognition molecule CD2 is glycosylation independent. *J. Biol. Chem.* **270**, 369–375.
44. Bodian, D.L., Jones, E.Y., Harlos, K., Stuart, D.I. & Davis, S.J. (1994). Crystal structure of the extracellular region of the human cell adhesion molecule CD2 at 2.5 Å resolution. *Structure* **2**, 755–766.
45. Jones, E.Y., Davis, S.J., Williams, A.F., Harlos, K. & Stuart, D.I. (1992). Crystal structure at 2.8 Å resolution of a soluble form of the cell adhesion molecule CD2. *Nature* **360**, 232–239.
46. Wyss, D.F., *et al.*, & Wagner, G. (1995). Conformation and function of the *N*-linked glycan in the adhesion domain of human CD2. *Science* **269**, 1273–1278.
47. Dwek, R.A., Lellouch, A.C. & Wormald, M.R. (1995). Glycobiology – the function of sugar in the IgG molecule. *J. Anat.* **187**, 279–292.
48. Rudd, P.M., Leatherbarrow, R.J., Rademacher, T.W. & Dwek, R.A. (1991). Diversification of the IgG molecule by oligosaccharides. *Molecular Immunol.* **28**, 1369–1378.
49. Deisenhofer, J. (1981). Crystallographic refinement and atomic models of human Fc fragment and its complex with fragment B of protein A from *Staphylococcus aureus* at 2.9 and 2.8 Å resolution. *Biochemistry* **20**, 2361–2370.
50. Stuikeprill, R. & Meyer, B. (1990). A new force field program for the calculation of glycopeptides and its application to a heptacosa-peptide-decasaccharide of immunoglobin-G1. *Eur. J. Biochem.* **194**, 903–919.
51. Rudd, P.M., *et al.*, & Dwek, R.A. (1994). Glycoforms modify the dynamic stability and functional activity of an enzyme. *Biochemistry* **33**, 17–22.
52. Williams, R.L., Greene, S.M. & McPherson, A. (1987). The crystal structure of ribonuclease B at 2.5 Å resolution. *J. Biol. Chem.* **262**, 16020–16031.
53. Joao, H.C. & Dwek, R.A. (1993). Effects of glycosylation on protein structure and dynamics in ribonuclease B and some of its individual glycoforms. *Eur. J. Biochem.* **218**, 239–244.
54. Joao, H.C., Scragg, I.G. & Dwek, R.A. (1992). Effect of protein glycosylation on amide proton exchange rate in RNase B. *FEBS Lett.* **307**, 343–346.