Conformational Studies of Oligosaccharides and Glycopeptides: Complementarity of NMR, X-ray Crystallography, and Molecular Modelling

Mark R. Wormald,*,† Andrei J. Petrescu,†,‡ Ya-Lan Pao,† Ann Glithero,†,§ Tim Elliott,§,# and Raymond A. Dwek†

Oxford Glycobiology Institute, Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, United Kingdom, Institute of Biochemistry of the Romanian Academy, Splaiul Independentei 296, 77700 Bucharest 17, Romania, and Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom

Received September 20, 2001

Contents

I. Introduction

The term "glycobiology", introduced in 1988, focused attention on the role of oligosaccharides in the context of the proteins to which they were attached.[1] Inspection of the protein databases suggests that as many as 70% of proteins have potential *N*-glycosylation sites (Asn−X−Ser: X not proline) and *O*-glycosylation may be even more ubiquitous.

Glycans attached to proteins can have a wide variety of roles; there is no single unifying function for glycosylation.[2,3] Some of the functions of glycans depend on the bulk of the glycan rather than its detailed three-dimensional structure. Simply because of their large size and hydrophilicity glycans can alter the behavior of proteins, making them more soluble, protecting them from proteolysis, covering antigenic sites, altering the orientation of the protein on a membrane, etc. The presence of glycans can also alter the structural properties of a protein. The presence of individual glycans can alter the conformation of peptides. It also appears to reduce the backbone flexibility, although not the conformation, of folded proteins leading to increased protein stability.[4] Multiple glycosylation, as found in heavily *O*-glycosylated mucin like domains, results in the whole chain becoming more rigid and extended.[5] To explore these functions, it is only necessary to know the overall size, shape, and charge of the glycan rather than the detailed structure. Another major function of protein-linked glycans is to provide additional recognition epitopes for protein receptors.[6] Such recognition events are involved in a wide range of processes including protein folding and trafficking, host defense, and inflammation. These recognition phenomena depend on the precise three-dimensional shape of the glycan.

This review will concentrate on the available techniques for determining the conformation and dynamics of oligosaccharides and glycoproteins, discuss their advantages and limitations, and show how the complementary data from NMR spectroscopy, X-ray crystallography, and molecular modeling can be used to address these problems.

II. Methods for Determining Oligosaccharide Conformations

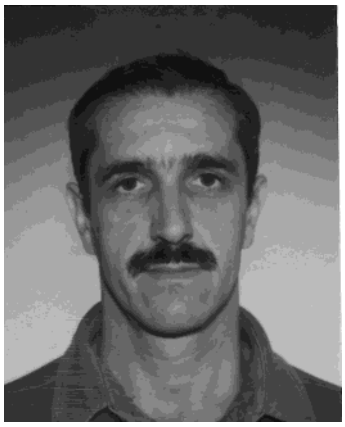
Monosaccharides are usually assumed to have a rigid ring structure. This is a good approximation for hexapyranoses with up to one axial group, which includes the majority of residues found in biological oligosaccharides. Thus, determination of the conformation of an oligosaccharide structures reduces to characterizing the glycosidic linkages between the rigid monosaccharide units. In practice, this involves determining either two or three torsion angles for each glycosidic linkage (Figure 1). Although this sounds simple, it is difficult to do in practice for two main reasons.

The first problem in characterizing the conformation of a glycosidic linkage is that the majority of

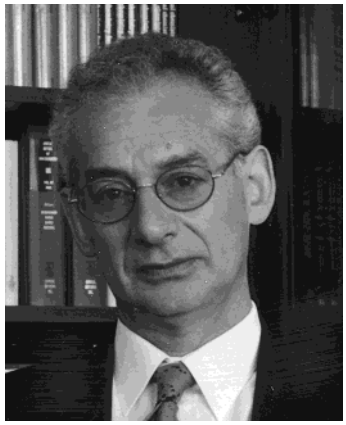
* To whom correspondence should be addressed. Phone/fax: +44 (0)1865 275738, E-mail: mark@glycob.ox.ac.uk.
† Oxford Glycobiology Institute.
‡ Institute of Biochemistry of the Romanian Academy.
§ John Radcliffe Hospital.
# Current address: School of Medicine, University of Southampton, Biomedical Sciences Building, Bassett Crescent East, Southampton, SO16 7PX, United Kingdom.

10.1021/cr990368i CCC: $39.75
Published on Web 01/29/2002

Dr. Mark Wormald is a University Research Lecturer in the Department of Biochemistry, Oxford University, and is a Fellow in Biochemistry and Chemistry at Corpus Christi College, Oxford. He graduated with a degree in chemistry from Oxford University in 1985, followed by a D.Phil. in the Inorganic Chemistry Laboratory (doing a mixture of physics, biophysics, and bioinorganic chemistry) with Prof. R. J. P. Williams, FRS. He moved to the then Oxford Oligosaccharide Group in 1989 to work with Prof R. A. Dwek, FRS, on the conformations of oligosaccharides by NMR spectroscopy. From 1990 to 1993, he held a Junior Research Fellowship at Corpus Christi College, Oxford. He is currently running the structural glycobiology group in the Oxford Glycobiology Institute, his main interests being in the conformations and dynamics of oligosaccharides, glycopeptides, and glycoproteins.



Dr. Andrei-J. Petrescu is a graduate in biophysics from the University of Bucharest. This was followed by postdoctoral positions in the Department of Biochemistry, University of Oxford, UK, and at the Centre de Etudes Atomique, Saclay, France. He is a member of the Oxford Glycobiology Institute and head of the Structural Biochemistry Group of the Biochemistry Institute of the Romanian Academy. He has been involved in developing physical methods for the study of unfolded states of proteins. He determined the structures of glucosylated oligomannose glycans specific to the early glycosylation stages in the ER and has contributed to the characterization of the glycan recognition elements of glycoproteins by lectin-like chaperones during glycoprotein folding. Currently, he is working on the development of a database of structural information on glycoproteins.



Professor Raymond Dwek is Professor of Glycobiology, Director of the Glycobiology Institute, and Head of the Department of Biochemistry, Oxford University. He is a Professorial Fellow at Exeter College, Oxford. He obtained his B.Sc. (1963) and M.Sc. (1964) degree at Manchester University and his D.Phil. (1966) degree in Oxford. He founded the Glycobiology Institute at Oxford University in 1991 and has received several awards for his work on glycobiology, including the 7th Wellcome Trust Award for Research in Biochemistry Related to Medicine and the First Scientific Leadership Award, Hepatitis B Foundation, Philadelphia, PA. He is a member of the European Molecular Biology Organization and Fellow of the Royal Society. In 1996, Professor Dwek was awarded a Doctoris Honoris Causa by the Katholieke Universiteit, Leuven, Belgium, for his research contributions to NMR, antibodies, and glycobiology. In 2000, he was awarded the National Romanian Order for merit with rank of Commander for his major contribution to the Romanian–British cooperation in biochemistry and molecular biology, and this year has been awarded the Doctor Philosophiae Honoris Causa from the Ben-Gurion University of the Negev, Israel, for his pioneering work in glycobiology.

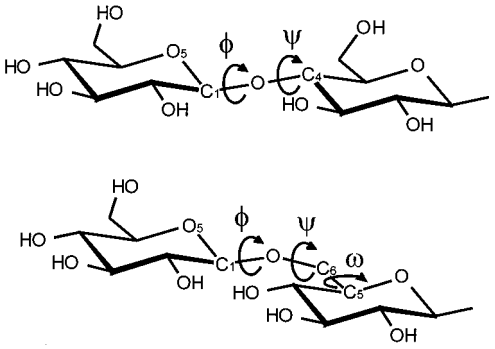


Figure 1. Schematic diagrams of a 1−4 linkage (upper) and 1−6 linkage (lower) showing the torsion angles that need to be determined to characterize the linkage conformation.

linkages are not rigid but flexible. This can include large scale vibrations of a single well-defined linkage conformation and/or transitions between two or more distinct conformations. The degree of flexibility differs from linkage to linkage, six-linkages usually being the most flexible. This variable flexibility of oligosaccharides makes linkage conformational analysis more difficult,[7] complete characterization of a given glycosidic linkage requiring knowledge of (a) the number of conformers adopted by the linkage, (b) the time spent in each conformer, and (c) the flexibility of each conformer. The problem becomes even more complex if the different linkages in an oligosaccharide do not behave independently but conformational transitions in different linkages are correlated.

The second problem is that the standard experimental techniques that are used to provide atomic level structural information about biological macromolecules, X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, have considerable limitations when applied to oligosaccharides. X-ray crystallography does not work well on highly flexible systems and NMR only provides time-aver-

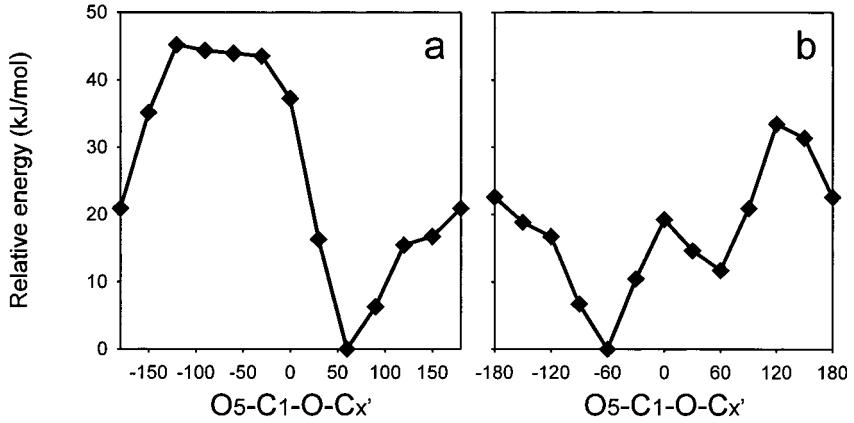


Figure 2. Plots of relative energy versus ϕ torsion angle for the α- and β-anomers of a model compound calculated using a restricted Hartree−Fock model.[10] (a) α-anomer (b) β-anomer.

aged conformational data. This has led to the widespread use of molecular modeling in determining linkage conformations; however, theoretical calculations are limited by the accuracy of the theory used.

In the past few years, more systematic use has been made of the crystallographic information that is available, more accurate and extensive data are being obtained from NMR experiments, and the theoretical models used to calculate oligosaccharides conformations have improved. The results of these techniques are complementary, and the combination of such results is giving a much clearer picture of the conformational properties of oligosaccharides.

The nomenclature used in this paper for these torsion angles is for a 1-x linkage $\phi = O_5-C_1-O-C_{x'}$ and $\psi = C_1-O-C_{x'}-C_{x-1'}$ and for a 1−6 linkage $\omega = O-C_{6'}-C_{5'}-C_{4'}$ (Figure 1).

A. Molecular Modelling

To predict the conformation of a molecule, it is necessary to be able to calculate the total energy of the molecule as a function of the conformation. It is then possible to determine the conformation with the lowest energy, the most stable conformation, and also the conformations that can be adopted at any given temperature. By applying Newton's Laws of motion to the individual atoms, it is also possible to simulate the dynamic behavior of the molecule. A molecular force field is simply an equation with terms for calculating the energies of all the different interatomic interactions (bond stretching, bond bending, bond rotation, electrostatic, hydrogen bonding, etc.), the total energy of the molecule in a given conformation being given by a summation of all the interactions. Many force fields have been developed for a range of molecules. To use such calculations for oligosaccharides, it is necessary to modify these force fields to include the interactions that are important in determining the conformations of glycosidic linkages.

The ϕ torsion angle of a glycosidic linkage is determined largely by the *exo-anomeric effect*.[8,9] This is a stereoelectronic effect involving the lone pairs on the linkage oxygen. Figure 2 shows the calculated effect of altering the ϕ angle on the energy of a model compound for both α and β anomers,[10] and this energy term needs to be included in the force field

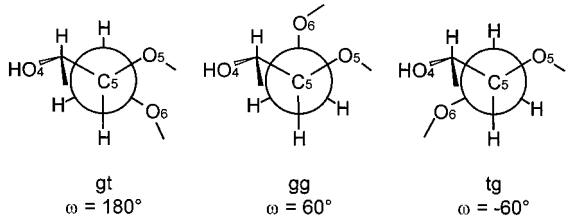


Figure 3. Schematic representation of the three sterically allowed rotamers for the ω torsion angle in mannose, looking along the C_5-C_6 bond. The tg rotamer is not observed in solution.

used. The ψ torsion angle is determined largely by steric interactions and hydrogen bonding between residues and with the solvent. The ω torsion angle for 1−6 linkages can adopt three staggered rotamers based on steric interactions, referred to as gauche−trans (gt, g referring to the O5−C5−C6−O6 and t to the C4−C5−C6−O6 arrangements), gauche−gauche (gg), and trans-gauche (tg) (Figure 3). For six-substituted saccharides with an equatorial 4-hydroxyl (such as mannose and glucose) the tg rotamer ($\omega = -60°$) is not stable,[11] referred to as the *gauche effect*.[12] The instability of the tg rotamer is proposed to result from repulsive interactions, particularly syn-periplanar repulsion between the O4 and O6 oxygen atoms. However, this interaction is only dominant when water is present to disrupt any favorable intramolecular hydrogen bonding.[13] Thus, simulations of oligosaccharide should ideally be done in the presence of water molecules, or additional energy terms need to be added to the force field to model the gauche effect.

Protocols for modeling oligosaccharide structures are well developed and have been recently reviewed[14,15] and so will not be discussed any further here.

B. Crystallographic Studies

X-ray crystallographic analysis has the potential advantage over NMR and theoretical methods that it can provide a complete oligosaccharide conformation from experimental data. The major limitation of this technique is that it requires regular crystals.

Few underivatized oligosaccharides crystallize. This is probably due to the inherent flexibility of most

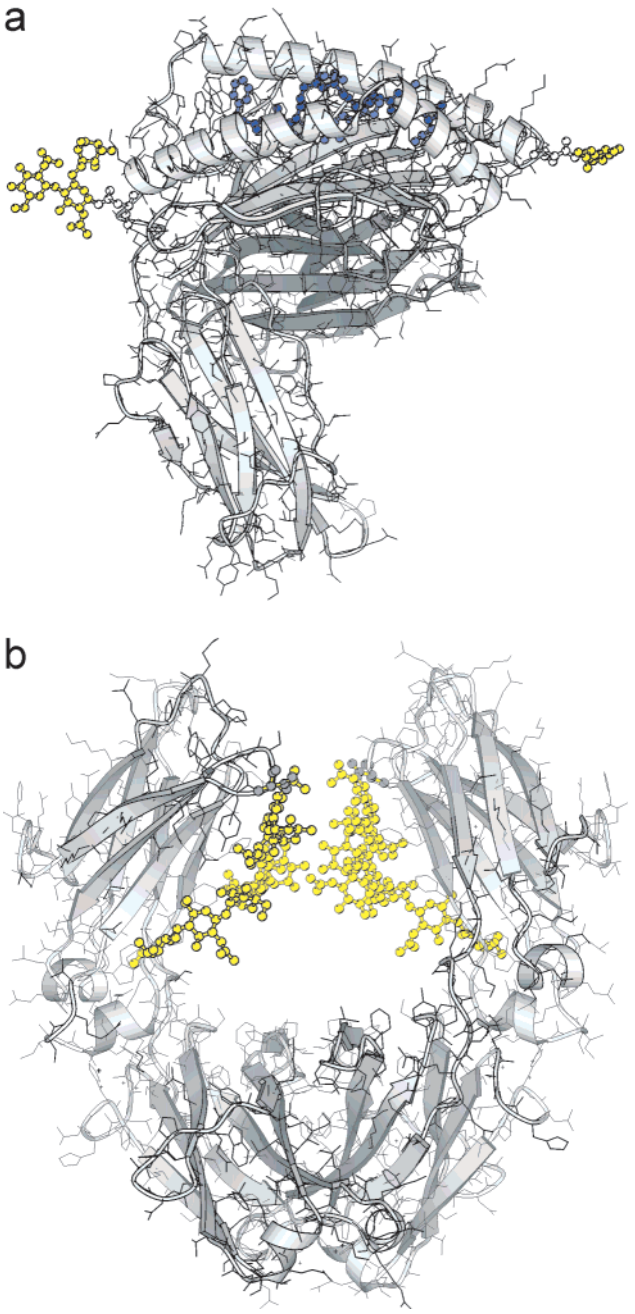


Figure 4. X-ray structures of two glycoproteins. Figures were prepared using the Molscript program.[76] (a) The crystal structure[77] of mouse MHC Class I H-2K^{bm1}. The glycans (yellow) point away from the protein surface. Density for the glycans can be seen at low resolution, but only a few residues are visible after crystallographic refinement[17] due to the flexibility of the glycans in the crystal. (b) The crystal structure[18] of IgG Fc. The glycans (yellow) are immobilized by interactions with the surface of the protein and so are well resolved in the electron density map.

oligosaccharides, flexible structures being less likely to pack in a uniform way to give crystals. X-ray crystallography has been used to determine the structure of very small, relatively rigid glycans, such as the Lewis-X trisaccharide[16] where stacking of the fucose and galactose rings reduces the flexibility of the glycosidic linkages.

It is much easier to produce crystals in which glycans are covalently attached to proteins or bound to proteins, the formation of regular crystals being dominated by the protein interactions. However, in most cases part or all of the glycan is not observed in the high-resolution electron density map. For instance, in MHC class I crystals electron density for the *N*-linked glycans can be observed at lower resolution but only a few residues of the glycan can be seen after crystallographic refinement[17] (Figure 4a). Again this is likely to be due to inherent flexibility, the glycans adopting different conformations in different unit cells and hence leading to averaged electron density. In a very few cases, such as IgG Fc[18,19] (Figure 4b), well-resolved structures of an entire glycan can be obtained. The whole glycan can be seen in cases where the glycan is immobilized either by interactions with the surface of the protein, as is the case with IgG Fc, or by crystal contacts. These structures may be distorted by being immobilized and so may not be representative of the normal solution structure of the glycan.

The lack of X-ray structures for complete glycans has been used to demonstrate the flexibility of most oligosaccharides. Recently, we have adopted a more systematic approach to using the large amount of X-ray data that is available.[20] While there are few intact glycan structures available, there are a large number of incomplete structures in the crystallographic databases in which glycosidic linkages can be observed. Any one structure will not give much insight into the solution conformation of particular glycosidic linkage; however, a statistical analysis of the all the examples of that linkage will give a give an indication of the range of conformations that the linkage can adopt. This is equivalent to generating experimental "Ramachandran-type" plots for glycosidic linkages. Tables 1 and 2 and Figures 5−7 give the current results of this statistical analysis (updated from Petrescu et al.[20]) for linkages found in *N*- and *O*-linked glycans.

X-ray data on linkage conformations can be used in a number of ways. The range of observed conformers can be used to validate the results of molecular models, comparing the observed and predicted behavior (for example, see ref 15). Distorted or unusual linkage conformations can be identified by comparison of experimental results on a specific glycan with the average data for that linkage (for example, see ref 20). In addition, the average linkage conformations can be used to quickly construct approximate models of specific glycans for modeling glycoproteins (for example, see ref 21).

C. NMR Studies

NMR has been a major technique for the identification[22] and the determination of the covalent structure of oligosaccharides, being the only available technique that can determine both the anomericities and linkages of a novel glycan.[23] Advances in this area have been recently reviewed.[24] In addition to providing information about the covalent structure, NMR has been the main experimental technique to provide conformational information for oligosaccharides (summarized in Table 3).

The major source of conformational information comes from the nuclear Overhauser effect (NOE or

Table 1. Public Domain Crystal Structures at a Resolution Better than 3.0 Å Containing Oligosaccharides with Glycosidic Linkages Found in *N*- or *O*-Linked Glycans[a]

type of structure	no. of crystal structures	no. of glycan structures	no. of linkages between undistorted saccharides[b]	no. of incorrect linkages or linkages between distorted saccharides
unmodified oligosaccharides	9 {*9*}	10 {*10*}	11 {*11*}	0 {*0*}
glycoproteins with *N*-glycans	228 {*110*}	503 {*208*}	1091 {*441*}	292 {*134*}
glycoproteins with *O*-glycans	2 {*2*}	2 {*2*}	2 {*2*}	0 {*0*}
proteins with glycan ligands	29 {*23*}	74 {*64*}	200 {*185*}	1 {*0*}

[a] Values in {italics} are the values reported in 1999. [b] Six linkage structures do not meet the criteria for inclusion in Table 2, which are one example each of Araα1−2Man, Fucα1−4GlcNAc, Galα1−4Gal, Gal*β*1−4Man, Glcα1−3Glc, and Manα1−4Man.

Table 2. Average Torsion Angles and Standard Deviations for all Distinct Conformers[a] of Glycosidic Linkages Found in Either *N*- or *O*-Linked Glycans

glycosidic linkage	total number of structures	average linkage torsion angles for distinct conformers			conformer population
		ϕ	ψ	ω	
Linkages for which there are at least 10 examples from at least five different crystal structures:					
Fuc α1−3 GlcNAc[b]	46	−71.9 ± 8.0	−98.7 ± 7.1		40
Fuc α1−6 GlcNAc	34	−75.7 ± 13.7	204.9 ± 24.0	63.4 ± 12.3	24
Gal *β*1−4 GlcNAc[c]	44	−71.4 ± 10.9	132.2 ± 7.4		28
GlcNAc *β*1−4 GlcNAc	398	−75.9 ± 11.6	119.0 ± 15.4		376
GlcNAc *β*1−2 Man	70	−80.1 ± 12.6	−97.6 ± 22.3		53
		58.3 ± 9.4	−87.2 ± 15.2		8
Man *β*1−4 GlcNAc	240	−86.5 ± 11.6	110.7 ± 19.4		197
Man α1−2 Man	94	63.3 ± 5.5	−179.6 ± 4.2		26
		70.9 ± 12.3	−106.2 ± 15.1		59
Man α1−3 Man	163	71.5 ± 8.8	−120.6 ± 16.8		130
Man α1−6 Man	145	59.4 ± 7.5	94.0 ± 17.5	188.5 ± 12.3	29
		67.0 ± 10.5	178.5 ± 13.7	186.0 ± 12.8	29
		64.7 ± 10.4	181.6 ± 10.0	59.7 ± 14.0	38
Xyl *β*1−2 Man	20	−86.8 ± 11.5	−106.1 ± 8.5		19
Others (for which there are examples from at least two different crystal structures):					
Fuc α1−2 Gal	8	−97.8 ± 23.7	103.9 ± 27.7		8
Gal *β*1−3 GlcNAc	12	−74.3 ± 10.0	−131.5 ± 18.3		12
GlcNAc *β*1−4 Man	2	−170.0 ± 10.7	94.7 ± 6.1		2
NeuAc α2−3 Gal	14	68.7 ± 13.6	−125.1 ± 15.5		14
NeuAc α2−6 Gal	8	144.3 ± 2.5	188.6 ± 1.9	51.3 ± 4.8	3
		44.0 ± 13.0	159.4 ± 14.1	186.6 ± 13.5	3
		148.7	130.4	158.5	1
		294.5	122.2	30.1	1

[a] See ref 20 for discussion. For definitions of ϕ, ψ, and ω see Figure 1. [b] Includes both core and outer arm fucose linkages. [c] Includes four structures with sulfated galactose at the 3- or 4- position. These all have conformations within a distinct conformer region.

ROE). The NOE between two nuclei depends on the distance between them ($\propto r^{-6}$). This enables the distance between two protons across a glycosidic linkage to be determined (assuming a rigid structure, see below). A given distance between two protons is frequently consistent with a range of conformations, usually displayed as allowed regions on a ϕ/ψ map (Figure 8) and termed a *conformational constraint*. If enough of these constraints are available, then a single structure consistent with all these constraints can be obtained.

There are three problems with using NOEs to determine the conformation of oligosaccharides. The first problem is that frequently there are not enough constraints available to define the conformation of a linkage.[25] One NOE is consistent with a large region of conformational space (Figure 8b). A second NOE will reduce this considerably (Figure 8c), but there are still usually two or more regions that are consistent with two NOEs. Three or more NOEs may be needed to unambiguously define a conformation.

The second problem is that the NOE is very short range (up to about 4 Å) and so is usually only observed between nuclei within a monosaccharide or across a glycosidic linkage, no long range information is available. It may be possible to define the conformation of one glycosidic linkage reasonably accurately, if enough constraints are available. However, determination of the overall structure of a large oligosaccharide relies on adding together the local conformations of the individual linkages. Any uncertainties or errors in the local structures will then add cumulatively when trying to determine the large scale structure, unless long range as well as sequential NOEs are observed.

The third and biggest problem is that the NOEs are measured on a 50 ms to 1 s time scale. This does not matter if the molecule is rigid, but if it is flexible or converts between several conformations then an average NOE will be measured. These cannot even be easily interpreted in terms of an average conformation because of the r^{-6} dependence of the NOE.

The first of these problems, the uncertainty in conformation from too few NOEs, can be addressed to some degree by obtaining additional conformational constraints. These include using the absence

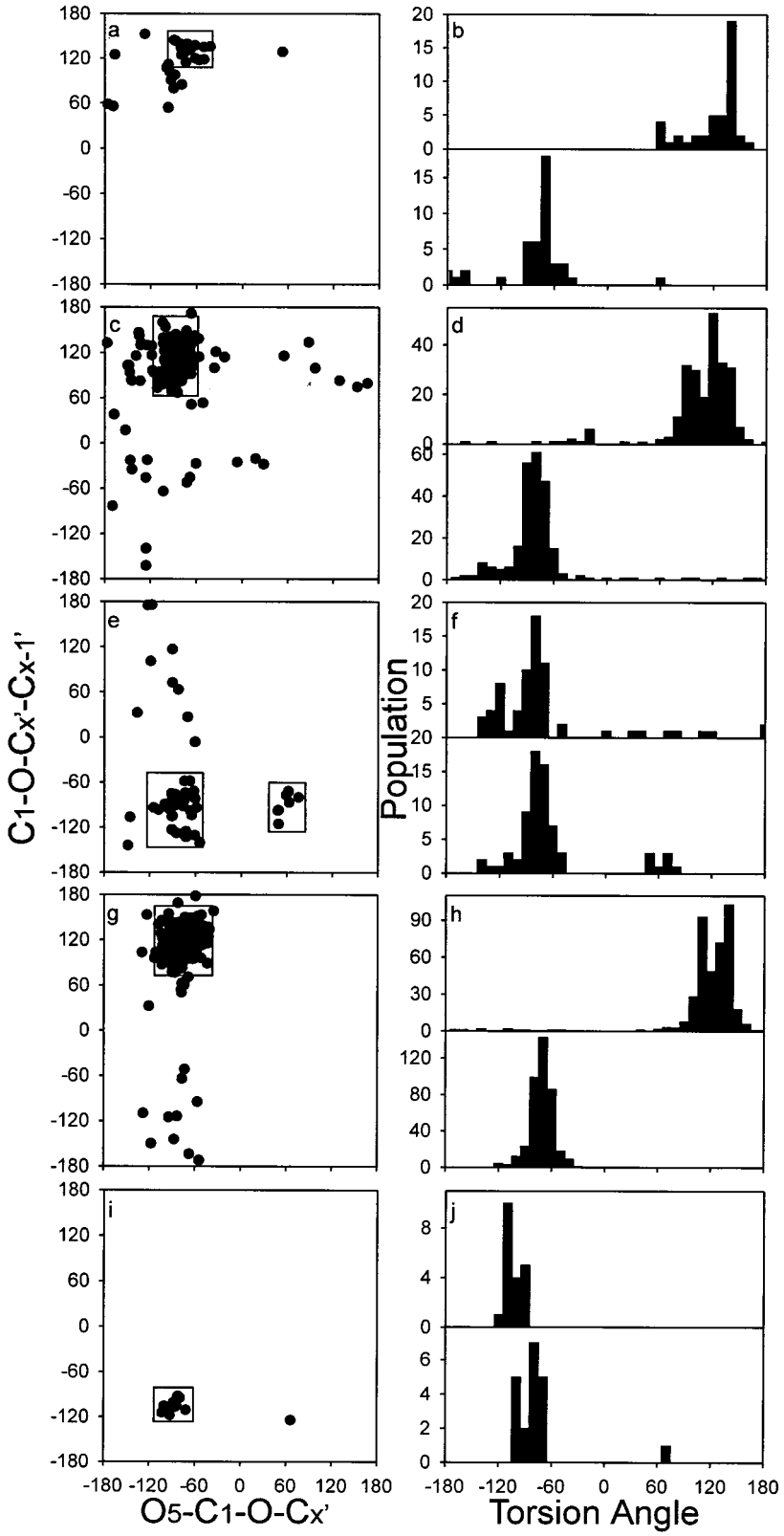


Figure 5. ϕ/ψ torsion angle plots and population histograms for β-glycosidic linkages for which there are at least 10 examples from at least five different crystal structures, updated from.[20] Panels (a), (c), (e), (g), and (i) are plots of $O_5-C_1-O-C_{x'}$ versus $C_1-O-C_{x'}-C_{x-1'}$ for a β1-x linkage. The boxed regions show the areas identified as distinct conformers (see Table 2). Panels (b), (d), (f), (h), and (j) are plots of histogram population (using a 10 degree window) versus torsion angle, the lower panel of each giving the $O_5-C_1-O-C_{x'}$ histogram and the upper panel giving the $C_1-O-C_{x'}-C_{x-1'}$ histogram. (a) and (b) Galβ1−4GlcNAc linkage. This also includes four structures with sulfated Gal at either the 3- or 4-position. These all fall within the boxed region. (c) and (d) Manβ1−4GlcNAc linkage; (e) and (f) GlcNAcβ1−2Man linkage; (g) and (h) GlcNAcβ1−4GlcNAc linkage; (i) and (j) Xylβ1−2Man linkage.

of NOEs to place lower limits on internuclear distances.[26] This allows regions of ϕ/ψ conformational space to be excluded because those conformations would give rise to NOEs which are not seen. Scalar

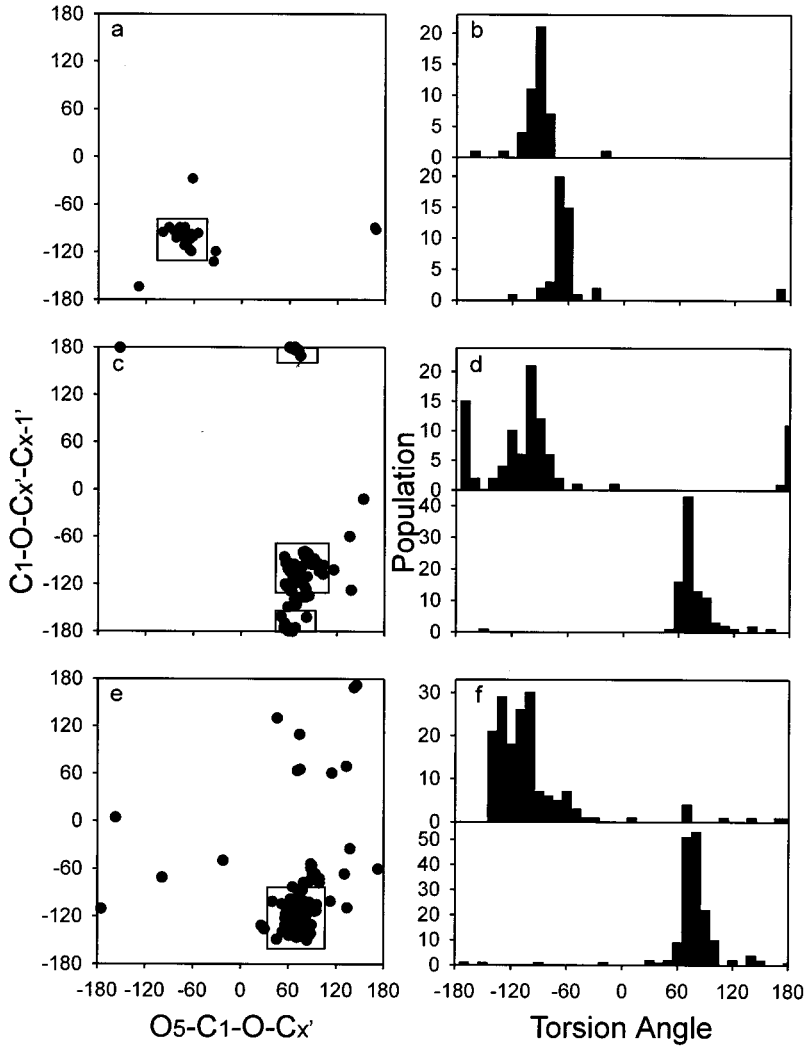


Figure 6. ϕ/ψ torsion angle plots and population histograms for α-glycosidic linkages (except α1−6, see Figure 7) for which there are at least 10 examples from at least five different crystal structures, updated from.[20] Panels (a), (c), and (e) are plots of $O_5-C_1-O-C_{x'}$ versus $C_1-O-C_{x'}-C_{x-1'}$ for a α1-x linkage. The boxed regions show the areas identified as distinct conformers (see Table 2). Panels (b), (d), and (f) are plots of histogram population (using a 10 degree window) versus torsion angle, the lower panel of each giving the $O_5-C_1-O-C_{x'}$ histogram and the upper panel giving the $C_1-O-C_{x'}-C_{x-1'}$ histogram. (a) and (b) Fucα1−3GlcNAc linkage; (c) and (d) Manα1−2Man linkage; (e) and (f) Manα1−3Man linkage.

coupling constants (J values) can be used to determine torsion angles,[27,28] although there may be several torsion angles that give the same *J* value. Determination of ϕ and ψ requires the measurement of three-bond carbon−proton coupling constants ($^3J_{CH}$), which can be very difficult if insufficient sample is available to give a sufficient carbon signal. Three-bond proton−proton couplings ($^3J_{HH}$) have been used to determine ω in 1−6 linkages.[29] More recently residual dipolar couplings have been used to give the relative orientations of CH bonds in a saccharide.[30,31] This requires partially orienting the saccharide by using bicelle or phage solutions.

The second problem, that the only information available is very short range, is also a problem with using scalar couplings. However, residual dipolar couplings overcome this to some degree because the accuracy with which the relative angle of two CH bonds to the magnetic field can be measured is independent of the distance between the two units.

The third problem, that only average structural constraints are measured, is a problem with all data obtained by NMR. The different NMR parameters are averaged on different time scales, and all these time scales are long compared to the internal motion of the saccharide.

Thus, NMR can be used to measure a large number of parameters that are sensitive to the conformation of the molecule. For rigid molecules, it is relatively straightforward to interpret these in terms of a single structure. However, for flexible molecules this is not possible because all the parameters measured are average values.

D. Oligosaccharide Flexibility

Evidence for oligosaccharide flexibility comes from the fact that in most cases any attempt to determine a unique structure for a given glycan from experimental data does not produce a meaningful result. X-ray crystallographic analyses of the same glycan

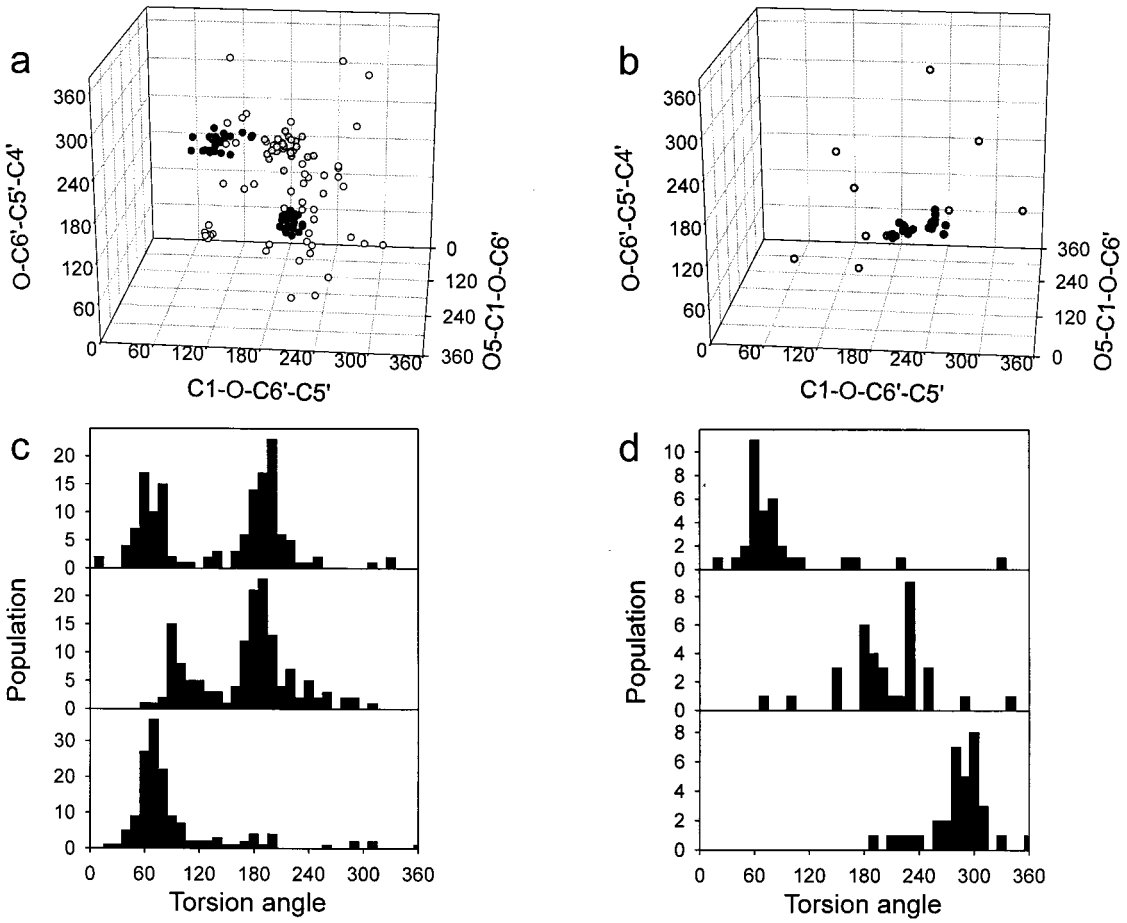


Figure 7. $\phi/\psi/\omega$ torsion angle plots and population histograms for α1−6 glycosidic linkages for which there are at least 10 examples from at least five different crystal structures, updated from ref 20. Panels (a) and (b) are plots of $O_5-C_1-O-C_{6'}$ versus $C_1-O-C_{6'}-C_{5'}$ versus $O-C_{6'}-C_{5'}-C_{4'}$. Filled circles − structures belonging to distinct conformers (Table 2), open circles − other structures. Panels (c) and (d) are plots of histogram population (using a 10 degree window) versus torsion angle, the lower panel of each giving the $O_5-C_1-O-C_{6'}$ histogram, the middle panel giving the $C_1-O-C_{6'}-C_{5'}$ histogram and the upper panel giving the $O-C_{6'}-C_{5'}-C_{4'}$ histogram. (a) and (c) Manα1−6Man linkage; (b) and (d) Fucα1−6GlcNAc linkage. The O5−C1−O−C6′ axis in panel (b) is reversed relative to panel (a) to enable easy comparison between the linkage conformations for D-Man and L-Fuc.

Table 3. Conformational Information Available from NMR

NMR parameter	structural information (all information is averaged)
Nuclear Overhauser effect (NOE or ROE)	distance between two nuclei (up to 4 Å)
one and three bond scalar coupling constants (J)	bond torsion angle
residual dipolar coupling	distance between two nuclei and the angle between the inter nuclear vector and the applied field
relaxation rates	mobility (bulk tumbling and internal motion)
temperature variation of amide resonances	degree of hydrogen bonding
$^1H/^2H$ solvent exchange rate	accessibility of an exchangeable proton to the solvent

structure in different glycoprotein crystals give different results, and NMR conformational data is frequently inconsistent with a single conformation. Molecular dynamics simulations, of course, provide a detailed prediction of the dynamic behavior, as well as the average conformer(s), of a glycan. However, obtaining experimental information about the degree of flexibility is much more difficult. The X-ray ϕ/ψ plots give a range of conformations that a linkage can adopt, but this can only be taken as a maximum range; the range adopted in a specific glycan may be smaller than this. NMR conformational analysis, based on NOEs and scalar couplings, of a single glycan will give the minimum range of conformations necessary to fit the experimental data, but will not give the time scales of any conformational changes. NMR relaxation measurements can be used to study glycan dynamics, giving an estimate of the flexibility on the ns time scale.[32,33] A combined approach is to use interproton distances determined by simulation and experimental NOE intensities to calculate the dynamic behavior of specific linkages in an oligosaccharide.[34]

III. Conformation of $Man_9GlcNAc_2$ and $Glc_3Man_9GlcNAc_2$

The oligomannose oligosaccharide $Man_9GlcNAc_2$ (Figure 9) is the least processed *N*-glycan found on mature glycoproteins. The glucosylated glycans, $Glc_{1-3}Man_9GlcNAc_2$, are not normally found on mature proteins and are involved in a number of

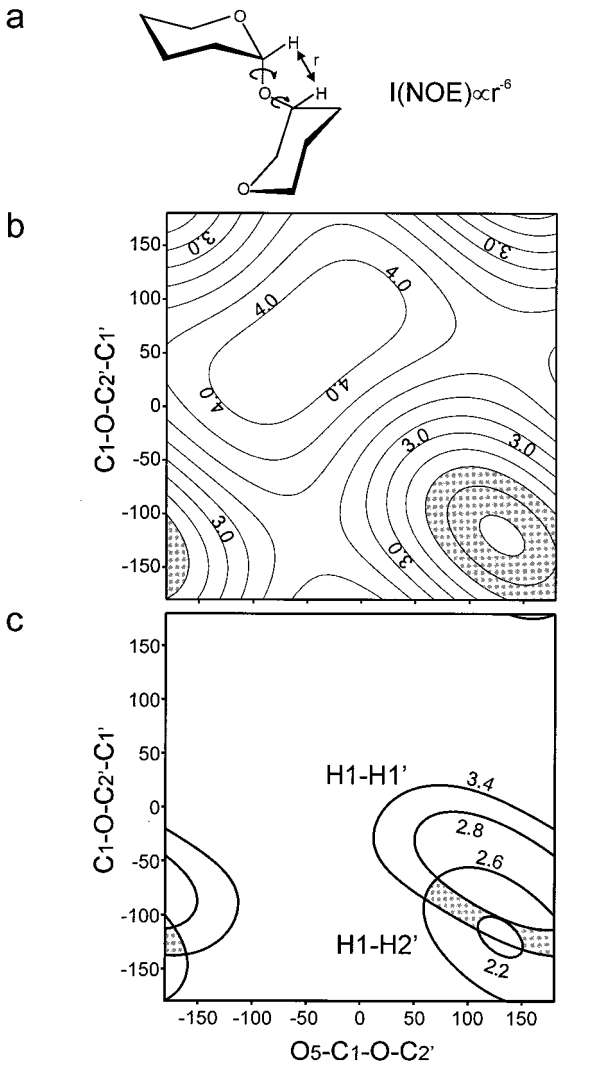


Figure 8. The use of glycosidic conformational constraints derived from NOEs. (a) Schematic drawing of an α1−2 linked disaccharide, showing the H1 and H2′ protons. (b) A contour plot showing the distance between the two protons highlighted in (a) as a function of the ϕ ($O_5-C_1-O-C_{2'}$) and ψ ($C_1-O-C_{2'}-C_{1'}$) torsion angles. The region of conformational space (i.e., the set of ϕ/ψ angles) consistent with an H1−H2′ interproton distance of 2.2 to 2.6 Å is shaded. (c) The same plot as in panel (b) showing the regions of conformational space consistent with both an H1−H2′ distance of 2.2 to 2.6 Å and an H1−H1′ distance of 2.8 to 3.4 Å.

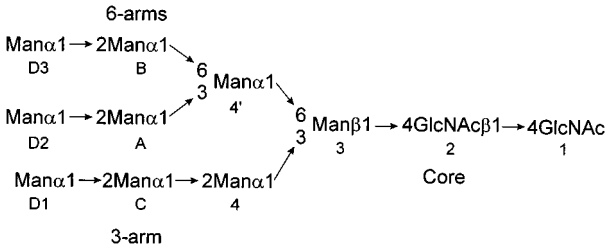


Figure 9. A schematic diagram of $Man_9GlcNAc_2$ showing the glycosidic linkages.

important steps during glycoprotein biosynthesis and folding, including initial *N*-glycosylation of the nascent polypeptide and the chaperone-dependent folding of the glycoprotein in the endoplasmic reticulum.[35]

Oligomannose oligosaccharides, and their fragments, have been the subject of numerous experimental and theoretical investigations of their conformational properties both as free glycans in solution using a wide variety of techniques[29,31,36−40] and when attached to a protein using NMR.[41] The structure of the Glc_3 cap found on glucosylated glycans has been determined in DMSO[42] and the structure of $Glc_3Man_7GlcNAc_2$ has been determined in water,[43] both by NMR.

We will start by considering the conformational analysis of some of the separate linkages present in these structures before going on to what we know about the structures of the molecules as a whole.

A. Manα1−2Man Linkages

Typical results of a molecular dynamics simulation for a Manα1−2Man linkage in $Man_9GlcNAc_2$[39] are shown in Figure 10a. This shows the linkage to exist in two flexible conformations, both with similar ϕ values (∼80°) in line with the exo-anomeric effect and with different ψ values (−60° and −160°), which interconvert relatively slowly. The X-ray ϕ/ψ plot (Figure 10b) also shows two distinct conformations with ϕ values of ∼65° and different ψ angles (−100° and −180°). It is interesting to note that the ϕ value obtained from the crystal structure of the Manα1−3Manα-OMe disaccharide (ϕ/ψ = 115°/−102°) is significantly different from the average values obtained, suggesting that crystal contacts in the disaccharide have a significant effect on the conformation. The NMR ϕ/ψ plot for a typical Manα1−2Manα linkage[39] is shown in Figure 10c. In this case, there are four NOEs observed across the glycosidic linkage. There are more than sufficient constraints to completely determine the conformation if the linkage adopted a single conformation. In fact, the NMR constraints are not consistent with any single conformation, and thus the linkage must either be highly flexible or adopt more than one conformation. A range of conformations with $\phi \sim 60°$ and ψ in the range −80° to −180° is the smallest range of conformers that would satisfy the constraints. Thus, all three independent analyses are consistent with this linkage adopting two interconverting conformations (Figure 10d).

B. Manα1−3Man Linkages

The Manα1−3Man linkages in $Man_9GlcNAc_2$ show different conformational properties. As for the Manα1−2Man linkages, the NMR conformational constraints are not consistent with a single rigid structure but could be satisfied by a range of conformations with $\phi \sim 60°$ and ψ between −60° and −180°. The X-ray data (Figure 6e,f) suggest that there is a single, fairly flexible conformation with $\phi \sim 70°$ and $\psi \sim -120°$. The molecular dynamics results on $Man_9GlcNAc_2$ suggest that the linkage adopts two conformers, with ϕ/ψ of 75°/−130° and 110°/−170°, which interconvert very rapidly. NMR and modeling studies on a small oligomannose fragment, Manα1−3(Manα1−6)Manα-OMe, give similar results.[40]

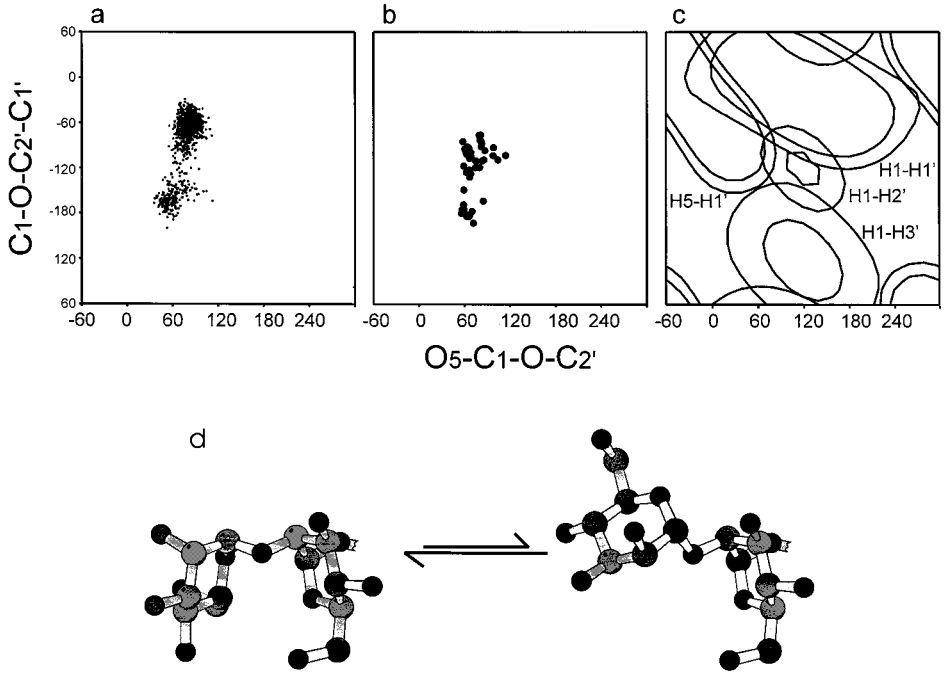


Figure 10. Conformational study of a Manα1−2Manα glycosidic linkage, the D1-C linkage in $Man_9GlcNAc_2$. Panels (a) to (c) show ϕ versus ψ plots. (a) Results of an unrestrained molecular dynamics simulation of $Man_9GlcNAc_2$39. Each dot represents a snapshot of the structure at one pico-second intervals. (b) The crystallographic structures for the discrete conformers (as in Figure 6c, except that the statistically nonsignificant structures have been removed). (c) Distance constraints obtained from NOE data.[39] Four interproton distance constraints can be obtained from the solution NMR data, which would be sufficient to define completely a glycosidic linkage that adopted a single conformation. Each set of contours on the ϕ/ψ plot shows the region of conformational space consistent with a single observed NOE. (d) Molecular models of the two distinct conformations for the Manα1−2Manα linkage, left (ϕ/ψ = +70°/−105°), right (ϕ/ψ = +65°/−170°). These are in dynamic equilibrium, the equilibrium lying slightly more to the left. Each distinct conformer is not static but shows limited oscillations for both ϕ and ψ.

C. Manα1–6Man Linkages

The Manα1−6Man linkage presents rather more of a problem. The X-ray $\phi/\psi/\omega$ plot (Figure 7a,c) gives a constant ϕ value of ∼ 60°, but two values for both the ψ and ω torsion angles, 90° and 180° for the former and 60° and 180° for the latter. $\omega = -60°$ is not expected because of the gauche effect. The ψ/ω combination of 90°/60° is only observed in six structures, but the other three ψ/ω combinations are all seen with similar frequency (Table 2). The molecular dynamics simulations for the Manα1−6Man linkages of $Man_9GlcNAc_2$ show a much more restricted set of conformers, with the $\psi = 90°$ conformer only transiently observed for one linkage (the B-4′ linkage, Figure 11a) and not for the other (the 4′−3 linkage). Both $\omega = 60°$ and 180° conformers are observed for the B-4′ linkage (Figure 11a), but only the $\omega = 60°$ conformer for the 4′−3 linkage. In this case, there are too few NMR constraints to be able to define the structure even if it is rigid. The NMR ϕ/ψ plot (Figure 11b) shows two regions consistent with both NOEs. The conformer populations for the ω torsion angle have been determined by measuring scalar coupling constants.[29] This suggests that both linkages adopt both the $\omega = 60°$ and 180° conformers, and both have a much higher population in the $\omega = 60°$ conformer, also consistent with the NOE data.[39] NMR and modeling studies on a small oligomannose fragment, Manα1−3(Manα1−6)Manα-OMe, suggest that the six linkage adopts two major conformations with $\phi/\psi/\omega$ of 64°/180°/60° and 64°/180°/180°,[40] and measurement of residual dipolar couplings constants has indicated significant motion.[31]

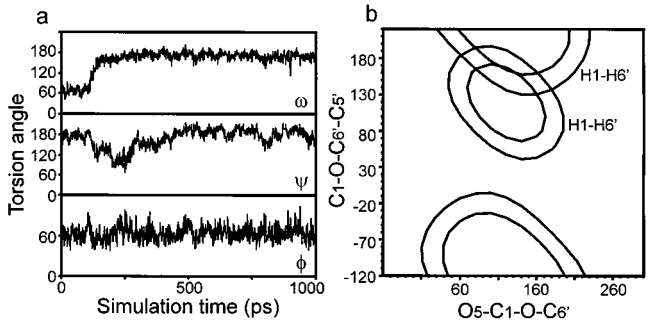


Figure 11. Conformational study of a Manα1−6Man linkage, the B-4′ linkage in $Man_9GlcNAc_2$.[39] (a) Results of an unrestrained molecular dynamics simulation of $Man_9GlcNAc_2$. Torsion angle is plotted versus simulation time. (b) ϕ versus ψ plot showing the distance constraints obtained from NOE data. Two interproton distance constraints can be obtained from the solution NMR data, which is not sufficient to define completely a glycosidic linkage that adopted a single conformation.

Thus, the X-ray data suggest that the Manα1−6Man linkage adopts three major and one minor discrete conformations. NMR and molecular modeling of a small trisaccharide in solution indicates that the linkage only adopts two of these major conformations. For $Man_9GlcNAc_2$, molecular modeling suggests that both linkages adopt one major conformation that are different; the NOE analysis is not conclusive because of too few constraints, and the scalar coupling data suggests that both linkages

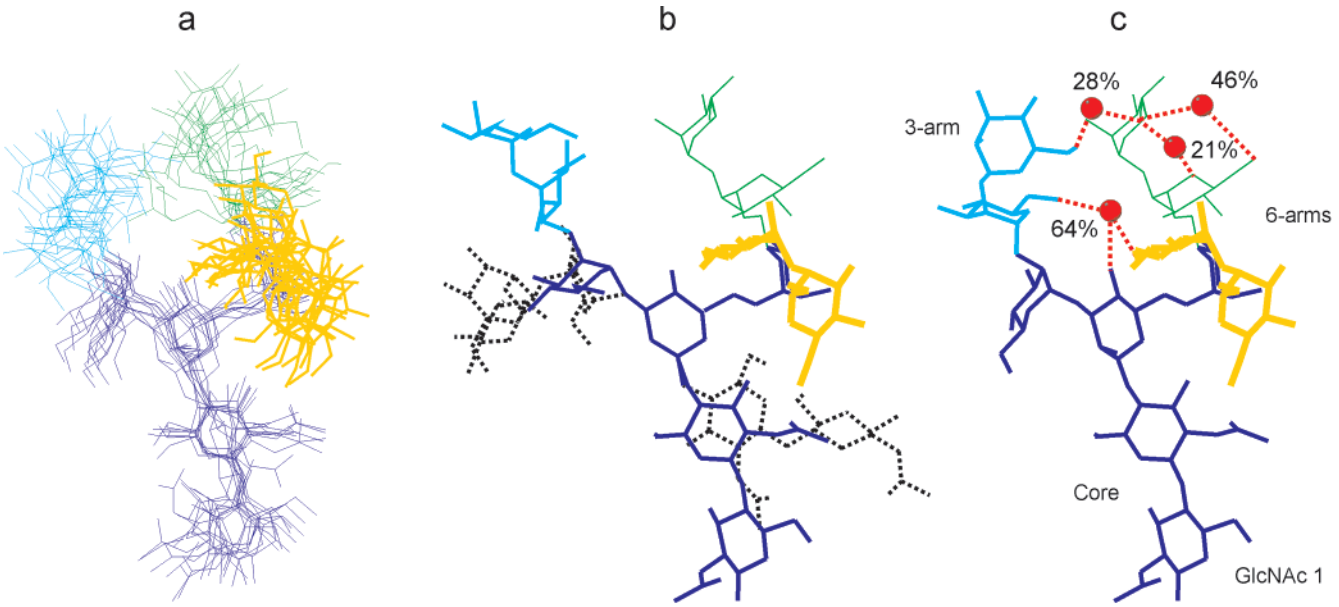


Figure 12. Molecular dynamics results on $Man_9GlcNAc_2$.[39] (a) Overlay of individual snapshots taken at 100 picosecond intervals. The glycan arms are color coded (dark blue: $Man_3GlcNAc_2$ core; light blue: 3-arm; green and yellow: two 6-arms). The overall topology of the molecule is well preserved given the high degree of flexibility for each individual linkage. (b) Two structures of $Man_9GlcNAc_2$ showing the range of conformations that could be adopted by the 3-arm and the core on the basis of the flexibility of the individual linkages. (c) Single snapshot of $Man_9GlcNAc_2$ showing the positions of water molecules (red spheres) involved in interarm hydrogen bonding, together with their occupancies based on the molecular dynamics results.

convert between two ω conformers. The apparent discrepancy between the molecular dynamics results and the experimental data could be due to inadequacies of the theoretical model or it could be due to the simulation time being too short to adequately sample the whole of the available conformational space.

D. Overall Structure of $Man_9GlcNAc_2$

The overlay of the structures produced by the molecular dynamics simulation of $Man_9GlcNAc_2$ is shown in Figure 12a. Each linkage is very flexible, but the overall shape is remarkably well conserved. Figure 12b shows two extreme structures that have been generated by adjusting the ϕ/ψ values for the linkages in the chitobiose core and the three-arm (the six-arms have not been altered) within the range of values for each linkage observed during the simulation. These models give equally good fits to the X-ray ϕ/ψ plots and to the experimental NMR data for $Man_9GlcNAc_2$ as the structures shown in Figure 12a. The molecular dynamics simulations suggest that the overall structure is much better defined than that suggested by a model based on independent behavior of the linkages, that the rotation around the linkage bonds is correlated. There are currently no experimental data to test this theoretical prediction, because the techniques available only provide short range conformational constraints.

If this prediction is correct, that the overall topology is less flexible than the linkages,[39] the explanation probably lies with the solvent. The simulation shows the presence of a number of water-mediated hydrogen bonds between the arms of the oligosaccharide (Figure 12c), which would hold these arms closer together. The overall shape of the glycan may also be constrained by the surrounding solvent shell, providing a well-defined "hole" within which the molecule can move.

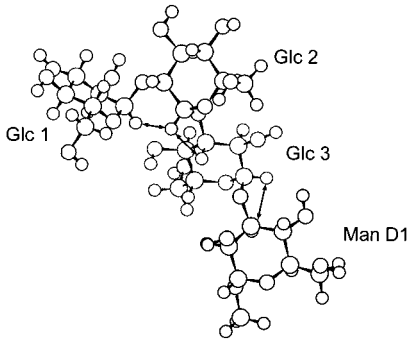


Figure 13. Molecular model of the Glc_3Man unit determined by NMR studies on $Glc_3Man_7GlcNAc_2$.[43] The arrows show all the observed interresidue NOEs.

E. Structure of $Glc_3Man_9GlcNAc_2$

The triglucosyl cap of $Glc_3Man_9GlcNAc_2$ shows distinctively different conformational properties from those of the flexible oligomannose part of the structure. Only one NOE is seen across each of the glycosidic linkages of the cap[43] as compared to three or four for the oligomannose linkages. This absence of NOEs is very unusual for a glycan. There are very few linkage conformations that result in only one proton on each side of the linkage being close enough to give an NOE, and this indicates that each linkage adopts a single rigid conformation. The solution structure of the cap has been solved by a combination of NMR and molecular modeling (Figure 13). A rigid structure has advantages for the recognition events in the endoplasmic reticulum that involve the triglucosyl cap, minimizing the loss of entropy on binding.

IV. Linking Glycans to Peptides

The study of small glycopeptides presents even more problems than that of oligosaccharides. The peptide components tend to be even more flexible than the glycan components, with far more degrees of conformational freedom. Each peptide linkage has two torsion angles (similar to the glycosidic linkage) and there are additional torsion angles associated with the side chains. This flexibility makes X-ray crystallography impossible unless the glycopeptide is immobilized by some means. NMR results on small peptides are less conclusive than for glycans because averaging of NMR parameters occurs over a wider range of conformations. NMR can still provide considerable conformational information, but it is much more difficult to interpret this in terms of detailed conformers. Molecular modeling is also more difficult for peptides than oligosaccharides, because the much greater degree of freedom means that it takes much longer to adequately explore the conformational space. In practice, instead of using NMR and molecular modeling as independent techniques to determine conformation and then comparing the results (as for glycans above), NMR results are used as constraints during the molecular modeling. The models produced should be viewed more as an exploration of the experimental data than as full theoretical predictions.

A. Conformation of the Glycan-peptide Linkage and the Effect on the Glycan Conformation

The conformational behavior of the linkage between an *N*-glycan and its asparagine side chain has been well characterized by NMR studies of glycopeptides in solution[44,45] and by a statistical survey of crystallographic data.[46] These show the linkage between the glycan and the Asn to be relatively rigid, planar, and extended. The rest of Asn side chain, rotation around the Cα−Cβ and Cβ−Cγ bonds, shows the same conformational preferences as unsubstituted Asn residues.

The large majority of studies have shown that attaching glycans to either small peptides or proteins has little effect on the glycan conformation. For example, the conformation of oligomannose glycans has been studied attached to the IgM tailpiece[44] and to domain 1 of CD2.[41] In both cases, the same pattern of NMR constraints is observed as for the free glycan (discussed above). Thus, the presence of the peptide does not appear to significantly alter the average glycan structure, although it does appear to reduce the flexibility of the core GlcNAc residues in CD2.[41] Similarly, NMR studies on isotopically ^{13}C,^{15}N-enriched α-chain in the intact hCG dimer have shown that the glycans appear to extend into the solvent having similar conformational properties to the free oligosaccharides,[47] while NMR studies on the free α-subunit have shown extensive contacts between the first two GlcNAc residues of the glycan at Asn 78 and the peptide hydrophobic core,[48,49] leading to a reduction in mobility of the core glycan residues.

B. Effect of the Glycan on the Peptide Conformation

There have been many studies of the conformational effects of N-glycosylation on the peptide in the vicinity of the glycosylation site.[44,50−55] The general conclusions from these studies are that *N*-glycosylation does not induce any permanent secondary structure in unstructured peptides but does alter the conformational preferences of the peptide backbone, leading to a higher probability of more compact conformations. These effects only appear to involve the first few residues of the glycan and are probably mediated by steric and hydrophobic/hydrophilic interactions between the core glycan residues and the neighboring amino acid side chains. For example, calcitonin is a membrane spanning peptide with an *N*-glycosylation site near the *N*-terminus. The structures of the aglycosyl peptide, the peptide glycosylated with a single GlcNAc residue, and the peptide glycosylated with $Man_6GlcNAc_2$ have been determined in SDS micelles by NMR spectroscopy. The overall conformation of the peptide in these micelles was not affected by the glycosylation, whereas amide exchange experiments in water suggest that glycosylation reduces fluctuations in the membrane spanning helix.[56]

Multiple *O*-glycosylation can have a very pronounced effect on peptide conformation. The extensive *O*-glycosylation of mucins leads to peptide chain extension and increased rigidity of the backbone.[57,58] This observation is confirmed by another study that showed a significant decrease in end-to-end distance and radius of gyration upon removal of carbohydrate from the mucin-like glycopeptide by stochastic dynamic simulations.[59] Similar results are also obtained for multiply glycosylated model peptides.[60] Multiple *O*-glycosylation can also result in the formation of tertiary structure. CD and NMR conformational studies of a model 30 residue collagen-based peptide have shown that multiple *O*-glycosylation is required for the formation of a stable triple helix.[61]

As with *N*-glycosylation, single site *O*-glycosylation frequently has little or no effect on the conformation of peptides, although it may alter the conformational preferences slightly.[62,63] For example, *O*-glycosylation of drosocin, a 19-amino acid peptide, does not effect the conformational properties of the peptide in water but does have small effects on some of the folded conformers observed in TFE/water.[64] The effect of *O*-glycosylation on peptide conformation also appears to depend on whether glycosylation occurs at a serine or threonine residue.[65]

In some cases, single site *O*-glycosylation appears to have a larger conformational effect. *O*-Glycosylation of a 24-residue peptide epitope (residue 308−331, containing the immunodominant tip of the V3 loop) from the HIV-$1III_B$ isolate, has been shown to stabilize a reverse turn and rigidify the backbone conformations proximal to the glycosylation site.[66] In another study, a 23-residue peptide derived from the C-terminus of a leucine zipper domain was shown by CD spectroscopy to adopt a beta-sheet conformation in water, whereas the singly *O*-glycosylated peptide assumes a helical structure.[67]

C. Conformational Studies on Glycopeptides Recognized by MHC

The conformational study of peptides involved in MHC recognition provides a good example of how X-ray crystallography, NMR spectroscopy, and molecular modeling can be combined to address a specific biological problem, in this case the affinity of modified peptides for MHC.

The immune response to foreign antigens is triggered by T lymphocytes recognizing complexes of foreign peptides bound to major histocompatibility complex (MHC) molecules on the cell surface. MHC class I molecules present small peptides generated in the cytosol. Modification of such peptides by glycosylation has been shown to affect both peptide binding to MHC class I molecules and recognition of the peptide/MHC complex by T lymphocytes.[68] A significant population of peptides presented by MHC class I in vivo are glycosylated, mostly by cytosolic *O*-β-GlcNAc.[69]

Peptides bind to MHC class I in a deep, narrow groove on the protein surface. In most cases, glycosylation results in either unchanged or reduced binding to MHC class I. However, we have found a case where peptide glycosylation increases peptide binding to the murine MHC class I molecule H-2D^b.[70] In an in vitro binding assay, the analogue peptide K2 (FAPGSYPAL, Figure 14) based on an antigenic peptide from *Sendai virus nucleoprotein* (residues 324−332, denoted wild type in Figure 14) showed no detectable binding to H-2D^b. This is presumably due to replacement of the anchor residue Asn-5 with Ser. Addition of an *O*-linked GlcNAc to Ser-5 (to give the glycopeptide K2G, Figure 14) partially restored binding to H-2D^b, although K2G still binds with low affinity (100-fold lower than for the wild-type peptide).

It was initially thought that the restoration of binding might be due to the glycan occupying the anchor residue pocket in the MHC binding site. Crystallographic studies[71] show that the wild-type peptide, K2 and K3G (with a nonanchor residue modified, Figure 14) bind to MHC class I in fully extended backbone conformations (K3G shown in Figure 15a). K2G also binds in a chiefly extended conformation, although the peptide backbone has to

		Binding to H-2D^b
Wild type (wt)	F A P G N* Y P A L*	✓
K2	F A P G S Y P A L	×
K2G	F A P G S(GlcNAc) Y P A L	(✓)
K3G	F A P S(GlcNAc) N Y P A L	✓

Figure 14. Amino acid sequences of *Sendai Virus Nucleoprotein* (SEV-9, residue 324−332, wild type) and analogues K2, K2G, and K3G. * − anchor residue for binding to H-2D^b.

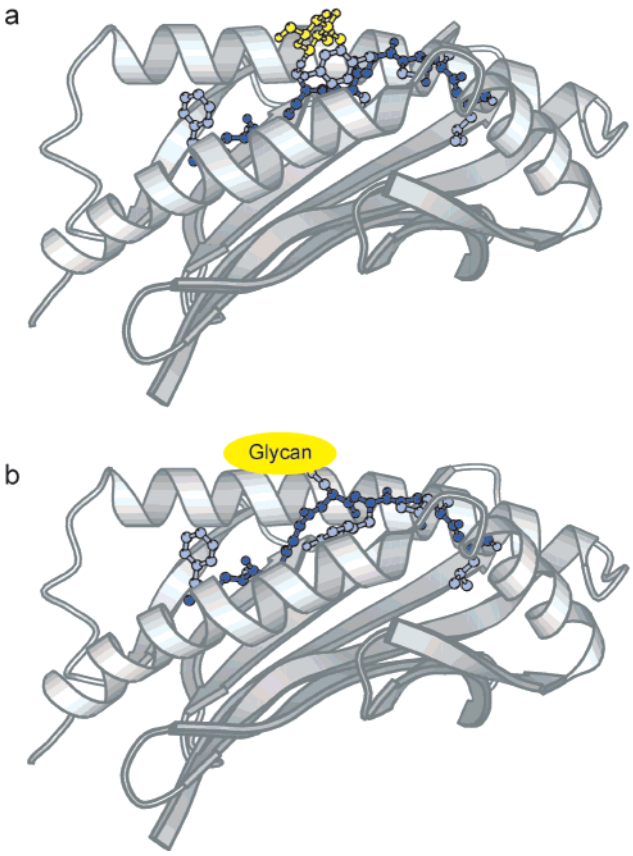


Figure 15. Crystal structures of MHC/glycopeptide complexes.[71] MHC − gray, peptide backbone − dark blue, peptide side chains − light blue, glycan − yellow. (a) Structure of the H-2D^b/K3G complex. The glycan is shown in one of its crystallographic positions (it is also found flipped by 180° while retaining the stacking with Tyr-6, with equal occupancy to the conformer shown). (b) Structure of the H-2D^b/K2G complex. The glycan is seen as a diffuse crown of electron density in the crystal structure pointing away from the binding site, indicated by the ellipse.

bulge somewhat around residue five to allow the saccharide to project out of the binding groove (Figure 15b). The GlcNAc residue can be clearly seen in the K3G/MHC complex, stacking against Tyr-6, whereas only blurred electron density is seen for the GlcNAc in the K2G/MHC complex. The peptide backbone of K2G does not fit as well in the groove of MHC as does K2 because of the steric bulk of the GlcNAc residue and the GlcNAc of K2G is mobile in the crystal showing no interactions with the protein, and yet K2G binds with much higher affinity to MHC than does K2.

K2 and K2G in solution have very similar NMR parameters, typical of random coil structures.[72] In both K2 and K2G, Pro-7 is found in both the cis and trans forms in similar ratios, as expected for a proline preceded by an aromatic residue,[73] while Pro-3 only adopts the trans form. However, there are significant differences between the NMR results for K2 and K2G. Interside chain NOEs are observed for K2G involving Phe-1, Pro-3, Tyr-6, Pro-7, and the GlcNAc residue (Figure 16), indicating that these residues are clustered together for at least a significant part of the time. The solvent exchange of NH protons can be monitored by presaturating the water resonance

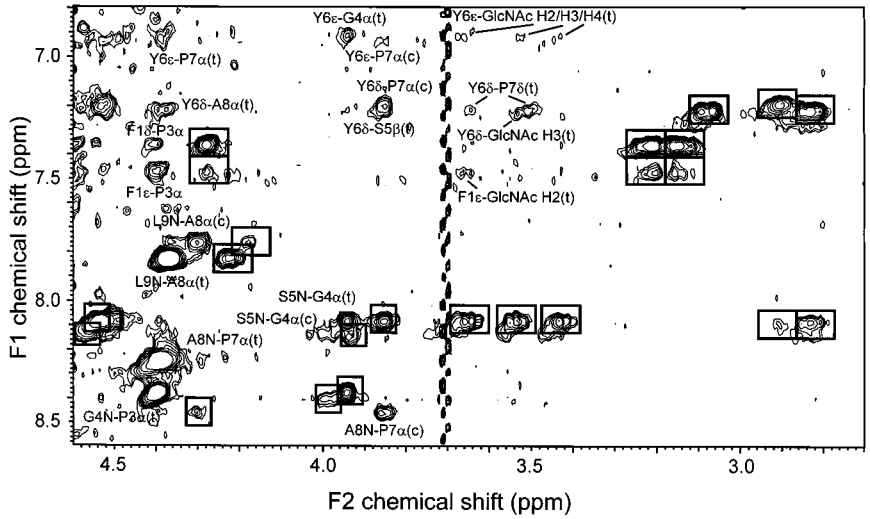


Figure 16. Region of the ROESY spectrum of K2G, mixing time 300 ms. The interresidue side-chain NOEs are marked. The ring protons of the *O*-GlcNAc residue are labeled as H1, H2, etc. (t) − resonance from the Pro-7 trans isoform, (c) − resonance from the Pro-7 cis isoform. Boxes indicate cross-peaks present in the TOCSY spectrum.

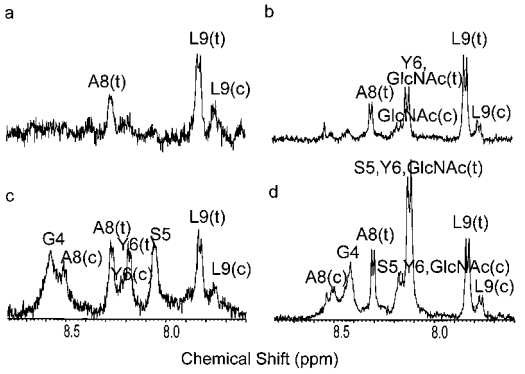


Figure 17. Amide proton region of the 1D NMR spectra of K2 and K2G, pH = 6.5, showing the effects of solvent presaturation. (t) − resonance from the Pro-7 trans isoform, (c) − resonance from the Pro-7 cis isoform. The assignment of Tyr-6 in (b) was confirmed by a COSY spectrum. (a) K2 with 1.0 s solvent presaturation. (b) K2G with 1.0 s solvent presaturation. (c) K2 without solvent presaturation. (d) K2G without solvent presaturation.

and observing the decreased intensity of the NH resonances, the smaller the decrease the slower the exchange with solvent. The Leu-9 backbone NH is protected from solvent exchange relative to the majority of backbone amide protons in both K2 and K2G, the GlcNAc NH is protected from solvent exchange in K2G, and the Tyr-6 backbone NH is protected from solvent exchange in K2G but not in K2 (Figure 17). This indicates that the peptide backbone in the region of residues five and six is less accessible to solvent and thus probably less flexible.

The pattern of interside chain NOEs, with Phe-1 giving an NOE to GlcNAc H2 and Tyr-6 giving NOEs to GlcNAc H3 and H6, places Phe-1 on one side of the GlcNAc ring and Tyr-6 on the other, all three residues being on the same side of the peptide chain. This hydrophobic environment for the GlcNAc is consistent with the slow solvent exchange rate of its NH proton and the hydrophobic clustering of the Phe, Tyr, and GlcNAc together with the reduced flexibility of the backbone would also explain the reduced solvent exchange of Tyr-6 NH.

The restrained simulated annealing results for K2 and K2G show very different behavior for the amino acid side chains. In K2, there is a random distribution of the side chains relative to the peptide backbone, whereas in K2G (Figure 18c) the hydrophobic residues cluster around the GlcNAc. The stacking between the GlcNAc and Tyr-6 is similar to those seen in the crystal structure of K3G (Figure 18a). Intermolecular hydrophobic interactions between monosaccharides and aromatic amino acid side-chains have been found to be important in glycan recognition by proteins.[74,75]

The interactions between Phe-1, GlcNAc, and Tyr-6 in K2G may provide the explanation for the restored binding of K2 to MHC on glycosylation. Although the peptide backbone remains quite flexible, the clustering of these side chains would greatly reduce the configurational entropy of the peptide in solution, and so there would be a much smaller loss of configurational entropy on complex formation. Indeed, these three residues all are found on the same side of the peptide chain pointing out of the binding groove in the K2G/MHC complex (Figure 18b), in a similar arrangement to that given by the restrained simulated annealing results (Figure 18c).

V. Conclusions

The combination of experimental and theoretical techniques gives considerable insight into oligosaccharide conformation. It is now relatively straightforward to characterize the conformational properties of individual glycosidic linkages. Different types of linkages show different types of conformational behavior, one or more conformers, and different degrees of flexibility. A single type of linkage can show different behavior in different environments, the conformational freedom being further reduced by the surrounding residues. However, the problem of determining the overall shape of a large oligosaccharide remains, with molecular modeling suggesting that the flexibility of the overall shape is less than would

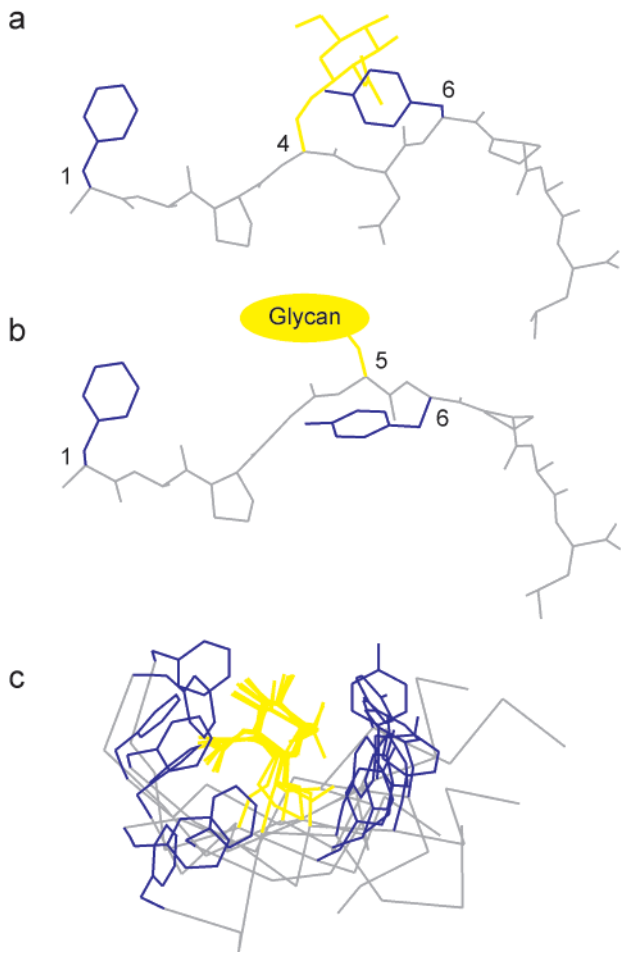


Figure 18. Conformational analysis of MHC binding glycopeptides. Crystallographic results from the glycopeptide/MHC complex[71] and NMR results from solution. (a) K3G peptide from the crystal structure of the K3G/MHC class I complex. The backbone is almost fully extended with residues 5 and 9 pointing down into the MHC binding grove. Phe-1 and Tyr-6 (blue) point out of the grove, as does the GlcNAc residue (yellow). (b) K2G peptide from the crystal structure of the K2G:MHC class I complex. The backbone bulges out of the MHC binding grove at position 5 (as compared to K3G). Phe-1, Tyr-6, and the GlcNAc residue all point out of the binding grove. (c) Overlay of the six best structures of K2G (Pro-7 trans isoform) generated by restrained simulated annealing. Only the backbone trace and residues 1, 5, 6, and the GlcNAc are shown. The backbone is more compact than in the complex. The backbone and side-chains are flexible, but residues Phe-1, Tyr-6, and the GlcNAc are positioned on the same side of the peptide chain.

be predicted from the flexibilities of all the separate linkages. This cooperative internal flexibility may well depend on the properties of the solvent.

The conformational study of glycopeptides is more difficult and generally leads to a less precise structural interpretation. In general, the addition of a glycan to a peptide does not alter the glycan conformation. However, it does reduce the flexibility of the peptide backbone in the vicinity of the glycosylation site and can result in some conformations of the peptide backbone becoming more favorable. This has implications not only for the solution properties of glycopeptides but also for the early stages of the unassisted folding of glycoproteins.

VI. References

(1) Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Annu. Rev. Biochem.* **1988**, *27*, 785.
(2) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. *Essentials of Glycobiology*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1999.
(3) Dwek, R. A. *Chemical Rev.* **1996**, *96*, 683.
(4) Wormald, M. R.; Dwek, R. A. *Structure* **1999**, *7*, R155.
(5) Van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. *Crit. Rev. Biochem. Molec. Biol.* **1998**, *33*, 151.
(6) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637.
(7) Xu, Q.; Gitti, R.; Bush, C. A. *Glycobiology* **1996**, *6*, 281.
(8) Lemieux, R. U.; Koto, S.; Voisin, D. In *Anomeric Effect: Origin and Consequences*; Szarek, W. A., Horton, D., Eds.; American Chemical Society: Washington, DC, 1979; Vol. 87.
(9) Tvaroska, I.; Bleha, T. *Adv. Carbohydr. Chem. Biochem.* **1989**, *28*, 45.
(10) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
(11) Marchessault, R. H.; Perez, S. *Biopolymers* **1979**, *18*, 2369.
(12) Wolfe, S. *Acc. Chem. Res.* **1972**, *5*, 102.
(13) Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541.
(14) Woods, R. J. *Glycoconjugate J.* **1998**, *15*, 209.
(15) Imberty, A.; Perez, S. *Chem. Rev.* **2000**, *100*, 4567.
(16) Perez, S.; Mouhousriou, N.; Nifantev, N. E.; Tsvetkov, Y. E.; Bachet, B.; Imberty, A. *Glycobiology* **1996**, *6*, 537.
(17) Rudd, P. M.; Wormald, M. R.; Stanfield, R. L.; Huang, M.; Mattsson, N.; Speir, J. A.; DiGennaro, J. A.; Fetrow, J. S.; Dwek, R. A.; Wilson, I. A. *J. Mol. Biol.* **1999**, *293*, 351.
(18) Deisenhofer, J. *Biochemistry* **1981**, *20*, 2361.
(19) Saphire, E. O.; Parren, P. W.; Pantophlet, R.; Zwick, M. B.; Morris, G. M.; Rudd, P. M.; Dwek, R. A.; Stanfield, R. L.; Burton, D. R.; Wilson, I. A. *Science* **2001**, *293*, 1155.
(20) Petrescu, A. J.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **1999**, *9*, 343.
(21) Rudd, P. M.; Wormald, M. R.; Harvey, D. J.; Devasahayam, M.; McAlister, M. S. B.; Brown, M. H.; Davis, S. J.; Barclay, A. N.; Dwek, R. A. *Glycobiology* **1999**, *9*, 443.
(22) Vliegenthart, J. F. G.; Dorland, L.; van Halbeek, H. *Adv. Carbohydr. Chem. Biochem.* **1983**, *22*, 209.
(23) Dwek, R. A.; Edge, C. J.; Harvey, D. J.; Wormald, M. R.; Parekh, R. B. *Annu. Rev. Biochem.* **1993**, *32*, 65.
(24) Duus, J. O.; Gotfredsen, C. H.; Bock, K. *Chem. Rev.* **2000**, *100*, 4589.
(25) Wooten, E. W.; Edge, C. J.; Bazzo, R.; Dwek, R. A.; Rademacher, T. W. *Carbohydr. Res.* **1990**, *203*, 13.
(26) Wormald, M. R.; Edge, C. J. *Carbohydr. Res.* **1993**, *246*, 337.
(27) Tvaroska, I.; Taravel, F. R. *Adv. Carbohydr. Chem. Biochem.* **1995**, *34*, 15.
(28) Milton, M. J.; Harris, R.; Probert, M. A.; Field, R. A.; Homans, S. W. *Glycobiology* **1998**, *8*, 147.
(29) Wooten, E. W.; Bazzo, R.; Edge, C. J.; Zamze, S.; Dwek, R. A.; Rademacher, T. W. *Eur. Biophys. J.* **1990**, *18*, 139.
(30) Almond, A.; Bunkenborg, J.; Franch, T.; Gotfredsen, C. H.; Duus, J. O. *J. Am. Chem. Soc.* **2001**, *123*, 4792.
(31) Tian, F.; Al-Hashimi, H. M.; Craighead, J. L.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 485.
(32) Maler, L.; Widmalm, G.; Kowalewski, J. *J. Phys. Chem.* **1996**, *100*, 17103.
(33) Rundolf, T.; Venable, R. M.; Pastor, R. W.; Kowalewski, J.; Widmalm, G. *J. Am. Chem. Soc.* **1999**, *121*, 11847.
(34) Lommerse, J. P.; Kroon-Batenburg, L. M.; Kamerling, J. P.; Vliegenthart, J. F. *Biochemistry* **1995**, *34*, 8196.
(35) Ware, F. E.; Vassilakos, A.; Peterson, P. A.; Jackson, M. R.; Lehrman, M. A.; Williams, D. B. *J. Biol. Chem.* **1995**, *270*, 4697.
(36) Homans, S. W.; Pastore, A.; Dwek, R. A.; Rademacher, T. W. *Biochemistry* **1987**, *26*, 6649.
(37) Imberty, A.; Perez, S.; Hricovini, M.; Shah, R. N.; Carver, J. P. *Int. J. Biol. Macromolecules* **1993**, *15*, 17.
(38) Balaji, P. V.; Qasba, P. K.; Rao, V. S. *Glycobiology* **1994**, *4*, 497.
(39) Woods, R. J.; Pathiaseril, A.; Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Eur. J. Biochem.* **1998**, *258*, 372.
(40) Sayers, E. W.; Prestegard, J. H. *Biophys. J.* **2000**, *79*, 3313.
(41) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269*, 1273.
(42) Alvarado, E.; Nukada, T.; Ogawa, T.; Ballou, C. E. *Biochemistry* **1991**, *30*, 881.
(43) Petrescu, A. J.; Butters, T. D.; Reinkensmeier, G.; Petrescu, S.; Platt, F. M.; Dwek, R. A.; Wormald, M. R. *EMBO J.* **1997**, *16*, 4302.
(44) Wormald, M. R.; Wooten, E. W.; Bazzo, R.; Edge, C. J.; Feinstein, A.; Rademacher, T. W.; Dwek, R. A. *Eur. J. Biochem.* **1991**, *198*, 131.
(45) Davis, J. T.; Hirani, S.; Bartlett, C.; Reid, B. R. *J. Biol. Chem.* **1994**, *269*, 3331.

(46) Imberty, A.; Perez, S. *Protein Eng.* **1995**, *8*, 699.
(47) Weller, C. T.; Lustbader, J.; Seshadri, K.; Brown, J. M.; Chadwick, C. A.; Kolthoff, C. E.; Ramnarain, S.; Pollak, S.; Canfield, R.; Homans, S. W. *Biochemistry* **1996**, *35*, 8815.
(48) De Beer, T.; Van Zuylen, C. W. E. M.; Leeflang, B. R.; Hard, K.; Boelens, R.; Kaptein, R.; Kamerling, J. P.; Vliegenthart, J. F. G. *Eur. J. Biochem.* **1996**, *35*, 229.
(49) van Zuylen, C. W.; de Beer, T.; Leeflang, B. R.; Boelens, R.; Kaptein, R.; Kamerling, J. P.; Vliegenthart, J. F. *Biochemistry* **1998**, *37*, 1933.
(50) Davis, J. T.; Hirani, S.; Bartlett, C.; Reid, B. R. *J. Biol. Chem.* **1994**, *269*, 3331.
(51) Conrad, S. F.; Byeon, I. J.; DiGeorge, A. M.; Lairmore, M. D.; Tsai, M. D.; Kaumaya, P. T. *Biomed. Pept., Proteins Nucleic Acids* **1995**, *1*, 83.
(52) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 97.
(53) Lommerse, J. P.; Kroon Batenburg, L. M.; Kroon, J.; Kamerling, J. P.; Vliegenthart, J. F. *J. Biomol. NMR* **1995**, *6*, 79.
(54) Live, D. H.; Kumar, R. A.; Beebe, X.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12759.
(55) Carotenuto, A.; D'Ursi, A. M.; Nardi, E.; Papini, A. M.; Rovero, P. *J. Med. Chem.* **2001**, *44*, 2378.
(56) Hashimoto, Y.; Toma, K.; Nishikido, J.; Yamamoto, K.; Haneda, K.; Inazu, T.; Valentine, K. G.; Opella, S. J. *Biochemistry* **1999**, *38*, 8377.
(57) Gerken, T. A.; Butenhof, K. J.; Shogren, R. *Biochemistry* **1989**, *28*, 5536.
(58) Shogren, R.; Gerken, T. A.; Jentoft, N. *Biochemistry* **1989**, *28*, 5525.
(59) Butenhof, K. J.; Gerken, T. A. *Biochemistry* **1993**, *32*, 2650.
(60) Biondi, L.; Filira, F.; Gobbo, M.; Pavin, E.; Rocchi, R. *J. Pept. Sci.* **1998**, *4*, 58.
(61) Bann, J. G.; Peyton, D. H.; Bachinger, H. P. *FEBS Lett.* **2000**, *473*, 237.
(62) Wilce, J. A.; Otvos, L.; Craik, D. J. *Biomed. Pept., Proteins Nucleic Acids* **1996**, *2*, 59.
(63) Kirnarsky, L.; Prakash, O.; Vogen, S. M.; Nomoto, M.; Hollingsworth, M. A.; Sherman, S. *Biochemistry* **2000**, *39*, 12076.
(64) McManus, A. M.; Otvos, L.; Hoffmann, R.; Craik, D. J. *Biochemistry* **1999**, *38*, 705.
(65) Naganagowda, G. A.; Gururaja, T. L.; Satyanarayana, J.; Levine, M. J. *J. Pept. Res.* **1999**, *54*, 290.
(66) Huang, X. L.; Barchi, J. J.; Lung, E. D. T.; Roller, P. P.; Nara, P. L.; Muschik, J.; Garrity, R. R. *Biochemistry* **1997**, *36*, 10846.
(67) Narasimhamurthy, S.; Naganagowda, G. A.; Janagani, S.; Gururaja, T. L.; Levine, M. J. *J. Biomol. Struct. Dyn.* **2000**, *18*, 145.
(68) Elliott, T. *Sci. Med.* **1998**, *5*, 44.
(69) Haurum, J. S.; Hoier, I. B.; Arsequell, G.; Neisig, A.; Valencia, G.; Zeuthen, J.; Neefjes, J.; Elliott, T. *J. Exp. Med.* **1999**, *190*, 145.
(70) Haurum, J. S.; Tan, L.; Arsequell, G.; Frodsham, P.; Lellouch, A. C.; Moss, P. A. H.; Dwek, R. A.; McMichael, A. J.; Elliott, T. *Eur. J. Immunol.* **1995**, *25*, 3270.
(71) Glithero, A.; Tormo, J.; Haurum, J. S.; Arsequell, G.; Valencia, G.; Edwards, J.; Springer, S.; Townsend, A.; Pao, Y. L.; Wormald, M. R.; Dwek, R. A.; Jones, E. Y.; Elliott, T. *Immunity* **1999**, *10*, 63.
(72) Pao, Y.-L.; Glithero, A.; Arsequell, G.; Valencia, G.; Elliott, T.; Dwek, R. A.; Wormald, M. R. **2001**, in preparation.
(73) Wuthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1984**, *180*, 715.
(74) Dwek, R. A.; Leatherbarrow, R. J. *Period. Biol.* **1983**, *85*, 21.
(75) Weis, W. I.; Drickamer, K. *Annu. Rev. Biochem.* **1996**, *35*, 441.
(76) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946.
(77) Rudolph, M. G.; Speir, J. A.; Brunmark, A.; Mattsson, N.; Jackson, M. R.; Peterson, P. A.; Teyton, L.; Wilson, I. A. *Immunity* **2001**, *14*, 231.

CR990368I