Uncertainties in structural determinations of oligosaccharide conformation, using measurements of nuclear Overhauser effects

E. Wrenn Wooten, Christopher J. Edge, Renzo Bazzo, Raymond A. Dwek, and Thomas W. Rademacher*

Glycobiology Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU (Great Britain)

ABSTRACT

A fundamental problem in the determination of molecular structure by n.m.r. spectroscopy is insufficient experimental constraints. This problem is particularly marked for oligosaccharides, where few constraints are available across glycosidic linkages. By calculating distances as a function of dihedral angle, it is shown that, in general, two n.O.e. constraints result in two possible conformations for each glycosidic linkage, one of which can usually be discarded on the basis of model building or energy calculations. Using these calculations, an estimate of the uncertainty in the structure can be obtained.

INTRODUCTION

Determination of the structure of an oligosaccharide, using ^{1}H–^{1}H n.O.e. data, relies upon sequential rather than long-range constraints. Few constraints, however, are generally available across glycosidic linkages[1-4]. The question arises of how precise structural information can be obtained given the few constraints available. As an example, $Man_9GlcNAc_2$ (1) is taken, the structure of which has been determined[5]. Only two n.O.e.s can be measured across each glycosidic linkage in this structure[5] and the question of how well these constraints define the structure is now examined.

1

---
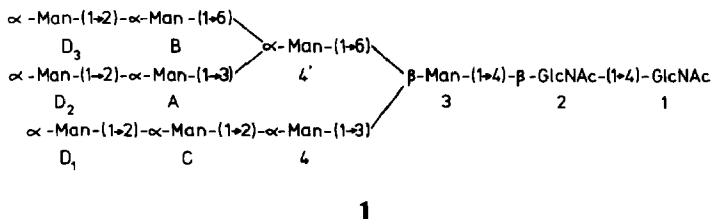
* Author for correspondence.

DISCUSSION

It is necessary to find a relationship for the distance between two protons as a function of the dihedral angles φ and ψ. This can easily be done by using the coordinates of individual atoms. Consider the situation of a linear chain H-1–C-1–O-1–C-X–H-X, with dihedral angles φ H-1–C-1–O-1–C-X and ψ C-1–O-1–C-X–H-X. For simplicity, it is assumed that the atoms in this chain are labelled A–B–C–D–E (A = H-1 *etc.*) and are described in the coordinate system shown in Fig. 1, with C at the origin and all the atoms lying in the x,y plane when φ and ψ are both zero. The interatomic distances and bond angles are assumed to be fixed, and it is assumed that the n.O.e.s can be interpreted in terms of a fixed conformation.
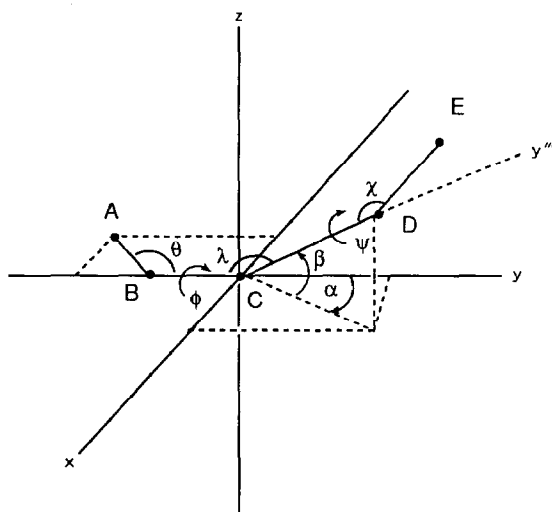

Fig. 1. Coordinate system for calculating the distance between two atoms as a function of dihedral angle. The chain of atoms A–E is shown in a coordinate system x,y,z where C is at the origin. The atoms A–C lie in the x,y plane. The distances AB, BC, CD, and DE are fixed. The angles θ ABC, λ BCD, and χ CDE are also fixed. The dihedral angles are defined as φ ABCD and ψ BCDE. The line y''' is the y-axis in the co-ordinate system in which D is at the origin. The angles α and β are two of the Euler angles which convert the x, y, z system into the x''', y''', z''' system. The angle γ is difficult to visualise and is not depicted.

From Fig. 1, it is easy to obtain the coordinates of A–D as a function of φ and ψ in the coordinate system of C:

$\mathbf{A} = (-AB \sin \theta, AB \cos \theta - CB, 0)$

$\mathbf{B} = (0, -CB, 0)$

$\mathbf{C} = (0, 0, 0)$

$\mathbf{D} = (-CD \sin \lambda \cos \varphi, -CD \cos \lambda, CD \sin \lambda \sin \varphi)$.

The coordinates of E in the coordinate system of D, in which D is at the origin, can also be easily obtained:

$\mathbf{E}''' = (-DE \sin \chi \cos \psi, -DE \cos \chi, DE \sin \chi \sin \psi)$.

The coordinate systems of D and C are related by the Euler angles α–γ:

$\mathbf{E} = \mathbf{R}(\alpha)\mathbf{R}(\beta)\mathbf{R}(\gamma)\mathbf{T}(+CD)\mathbf{E}'''$

with transformation matrices:

$$\mathbf{R}(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{R}(\beta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix}$$

$$\mathbf{R}(\gamma) = \begin{bmatrix} \cos\gamma & 0 & \sin\gamma \\ 0 & 1 & 0 \\ -\sin\gamma & 0 & \cos\gamma \end{bmatrix}$$

$$\mathbf{T}(CD) = \begin{bmatrix} 0 \\ +CD \\ 0 \end{bmatrix}$$

The angles α–γ are defined as those which take the x,y,z coordinate system into the x''',y''',z''' system with positive values corresponding to anticlockwise rotations. Expressions for the Euler angles can be derived as:

$\sin\alpha = \sin\lambda \cos\varphi/(1-\sin^2\lambda \sin^2\varphi)^{\frac{1}{2}}$ $0 \leqslant \alpha \leqslant 2\pi$

$\cos\alpha = -\cos\lambda/(1-\sin^2\lambda \sin^2\varphi)^{\frac{1}{2}}$

$\sin\beta = \sin\lambda \sin\varphi$ $-\pi/2 \leqslant \beta \leqslant \pi/2$

$\cos\beta = (1-\sin^2\lambda \sin^2\varphi)^{\frac{1}{2}}$

$\sin\gamma = -\tan\beta/\tan\lambda$ $0 \leqslant \gamma \leqslant 2\pi$

$\cos\gamma = -\sin\alpha/\sin\lambda.$

TABLE I

n.O.e. data for components of **1**

a-Man-(1→3)-β-Man

Atoms	η^a	r (Å)
4H-1–4H-2	1.00 ±0.1	2.62
4H-1–3H-3	4.00 ±0.4	2.08 ±0.07
4H-1–3H-2	0.40 ±0.04	3.05 ±0.1

a-Man-(1→6)-β-Man

4'H-1–4'H-2	1.00 ±0.1	2.63
4'H-1–3H-6	1.30 ±0.1	2.52 ±0.09
4'H-1–3H-6'	0.70 ±0.07	2.79 ±0.09

a-Man(1→2)-a-Man

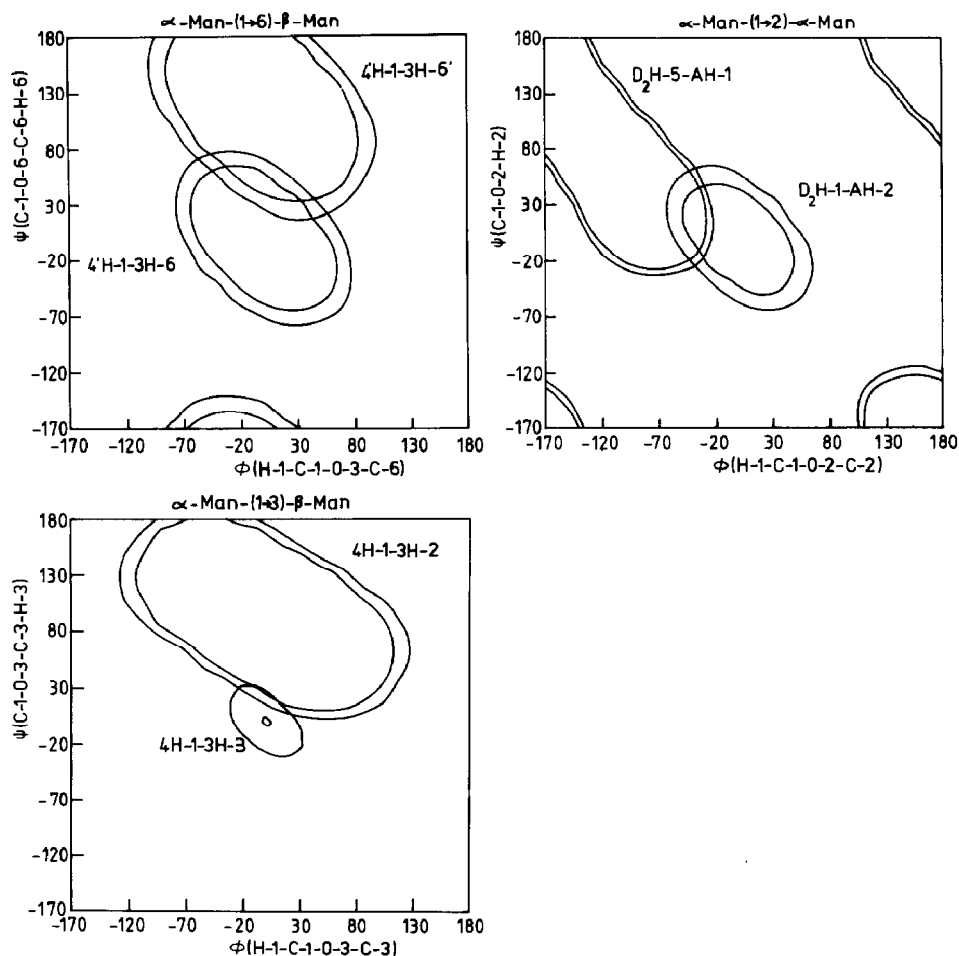|		
D_2H-1–D_2H-2	1.00 ±0.1	2.63
D_2H-1–AH-2	2.08 ±0.2	2.33 ±0.08
D_2H-5–AH-1	1.30 ±0.1	2.53 ±0.1

[a] Relative n.O.e. values.

Fig. 2. Level curves of distances computed from n.O.e. data are plotted as a function of the dihedral angles φ and ψ for three linkages of 1. Each pair of curves represents the error bounds of the distances computed from the n.O.e.s (Table I).

Once the coordinates have been obtained, the distance AE can be calculated:

$$AE = \{(X_A - X_E)^2 + (Y_A - Y_E)^2 + (Z_A - Z_E)^2\}^{\frac{1}{2}}$$

For distances that span > 5 atoms, transformation matrices can be calculated and applied sequentially.

Table I shows the n.O.e.s measured for the a-Man-$(1 \rightarrow 3)$-β-Man, a-Man-$(1 \rightarrow 6)$-β-Man, and a-Man-$(1 \rightarrow 2)$-a-Man linkages of 1 (ref. 5). An uncertainty of 10% has been assumed in the measured integrals and no uncertainty in the calibration distance. Using the above treatment, the distances between the two hydrogen atoms involved in the n.O.e. effect were calculated for each linkage as a function of φ and ψ, with φ and ψ varied in steps of 10°. The bond lengths, bond angles, and dihedral angles used as input were measured from disaccharide structures that were calculated from molecular

orbital (AM1) theory. No uncertainties were assumed in these values. Level curves of the two functions for each linkage at the error bounds of the distances calculated from the n.O.e.s were then plotted on the same graph, using the definitions of φ and ψ shown in Fig. 1.

Two major points emerge from the plots in Fig. 2. The first is that, for two n.O.e. constraints, there are, in general, two solutions, although the a-Man-$(1 \rightarrow 3)$-β-Man unit effectively has only one solution. The second point is that the uncertainty in the calculated φ,ψ values varies markedly between the three linkages.

The first problem, therefore, in obtaining conformation is how to distinguish between the two solutions, assuming that additional n.O.e. constraints cannot be obtained. One method is to use energy calculations. Molecular orbital calculations have been performed in which energy is varied as a function of φ and ψ for the disaccharides[5]. It turns out that only one of the two possible solutions is energetically favorable. It may also be possible to discard structures because of the absence of n.O.e.s that would be expected in a particular conformer.

The second important piece of information which can be obtained from these plots is a measure of the degree of uncertainty in the n.m.r.-derived structure once a given solution has been chosen. For example, the $(1 \rightarrow 2)$ linkage has an extremely well-defined position ($\varphi \pm 10°$, $\psi \pm 10°$) whereas that of the $(1 \rightarrow 3)$ linkage is less well defined ($\varphi \pm 50°$, $\psi \pm 25°$). Such uncertainties in the structure would be especially important when examining receptor–ligand interactions. It is important to emphasise, however, that these are uncertainties only in a fixed calculated structure; conformational averaging as in the $(1 \rightarrow 6)$ linkage[2] will add additional uncertainty.

REFERENCES

1 J. N. Scarsdale, R. K. Yu, and J. H. Prestegaard, *J. Am. Chem. Soc.*, 108 (1986) 6778–6784.
2 S. W. Homans, R. A. Dwek, and T. W. Rademacher, *Biochemistry*, 26 (1987) 6571–6578.
3 K. Bock, T. Frejd, J. Kihlberg, and G. Magnusson, *Carbohydr. Res.*, 176 (1988) 253–270.
4 J. Dabrowski and L. Poppe, *J. Am. Chem. Soc.*, 111 (1989) 1510–1511.
5 S. W. Homans, A. Pastore, R. A. Dwek, and T. W. Rademacher, *Biochemistry*, 26 (1987) 6646–6655.