Modeling Carbohydrate Conformations from NMR Data: Maximum Entropy Rotameric Distribution about the C5–C6 Bond in Gentiobiose

Leszek Poppe*

Contribution from the Complex Carbohydrate Research Center and Department of Biochemistry, The University of Georgia, 220 Riverbend Road, Athens, Georgia 30602-4712

Received March 16, 1993®

Abstract: The conformation about the C5–C6 bond in β-gentiobiose, (β-D-Glc*p*-(1→6)-β-D-Glc*p*), has been analyzed in terms of the maximum entropy probability density distribution of the dihedral angle Ω. This analysis used the information carried by proton–proton and carbon–proton vicinal coupling constants and proton–proton cross-relaxation rates. Two major conformations characterized by Ω = –68.9° ± 6.3° and Ω = 79.0° ± 3.4° were found with populations of 0.34 ± 0.06 and 0.66 ± 0.06, respectively.

Introduction

Nuclear magnetic resonance spectroscopy is a powerful technique for determining the solution conformations of molecules.[1] However, the existence of multiple conformations in flexible molecules such as oligosaccharides results in an averaging of NMR parameters.[2] In addition, there are usually an insufficient number of constraints with which to construct a 3D model in the absence of minimum-energy calculations.[3] The modeling of glycosidic linkages by simple "distance-mapping" procedures[4] gives only qualitative assignment of the probabilities to certain conformational states, even with the use of extensive experimental data based on the hydroxyl and amido proton resonances.[5]

The torsional flexibility about the exocyclic C5–C6 bond is believed to be an important conformational feature in oligosaccharides.[6] Typically, only three staggered states (see Figure 1) are considered, and their corresponding populations are then calculated from the vicinal coupling constants, $^3J_{H5H6R,S}$, and the Karplus equation.[6,7] Alternatively the average coupling constants can be calculated, using different force-field calculations to derive a continuous distribution of the Ω angle, and then compared with the experimentally obtained values.[6g,h] However, these approaches have limitations. For example, the nonconservative assumption of three discrete conformational states can lead to physically unrealistic results, and the results of force-field calculations that utilize only two coupling constants are of questionable value.

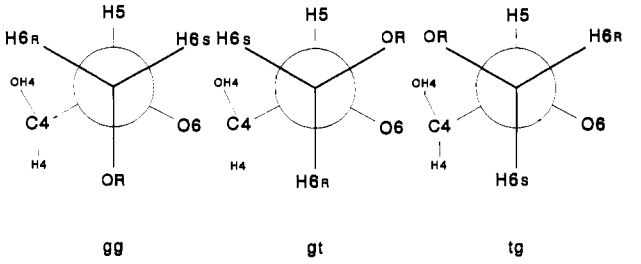


Figure 1. Newman projections of three major rotameric states about the C5–C6 bond in gentiobiose.

To overcome such difficulties, methods have been developed to use NMR data to determine the continuous probability distributions of conformations. For example, NOE data has been used to determine the conformation of the *N*-glycosyl bond in nucleosides in terms of the Gaussian distribution of the torsional angle.[8] Recently, Džakula et al. have exploited this approach and developed CUPID (continuous probability distribution of rotamers from NMR data) to derive a continuous probability distribution of rotamers from vicinal coupling constants and NOE effects.[9] A different approach has been to apply the distribution $p(\Omega)$, derived from the maximum-entropy principle,[10] to conformational problems.[11] For example, the internal rotations in biphenyl derivatives have been estimated from direct coupling constants obtained by liquid crystal NMR measurements.[11a] We now describe the use of the maximum-entropy probability density distribution to describe the conformation about the C5–C6 bond in β-gentiobiose (β-D-Glc*p*-(1→6)-β-D-Glc*p*). An explicit form of the probability distribution on conformational space Ω was calculated using the principle of maximum entropy and combined with geometric information from both vicinal coupling constants and NOE effects.

Unlike direct couplings, the interpretation of NOE effects in the liquid state is not straightforward, as it depends on both the

* The address for correspondence: Dr. Leszek Poppe, Complex Carbohydrate Research Center, The University of Georgia, 220 Riverbend Road, Athens, GA 30602-4712. Telephone: (706) 542-4414. Telefax: (706) 542-4412.

® Abstract published in *Advance ACS Abstracts*, August 15, 1993.

(1) (a) Noggle, J. H.; Schirmer, R. E. *The Nuclear Overhauser Effect*; Academic Press: New York, 1971. (b) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH Publishers: New York, 1989.

(2) Jardetzky, O. *Biochim. Biophys. Acta* **1980**, *621*, 227–232.

(3) Carver, J. P. *Curr. Opin. Struct. Biol.* **1991**, *1*, 716–720.

(4) (a) Poppe, L.; Lieth, C. W.; Dabrowski, J. *J. Am. Chem. Soc.* **1990**, *112*, 7762–7771. (b) Wooten, E. W.; Edge, C. J.; Bazzo, R.; Dwek, R. A.; Rademacher, T. W. *Carbohydr. Res.* **1990**, *203*, 13–17.

(5) Dabrowski, J.; Poppe, L. *J. Am. Chem. Soc.* **1989**, *111*, 629–630.

(6) (a) Gagnaire, D.; Horton, D.; Taravel, F. *Carbohydr. Res.* **1973**, *27*, 363–372. (b) Streefkerk, D. G.; de Bie, M. J. A.; Vliegenthart, J. F. G. *Tetrahedron* **1973**, *29*, 833–844. (c) Marchessault, R. H.; Perez, S. *Biopolymers* **1979**, *18*, 500–521. (d) Bock, K.; Arnap, J.; Lonngren, J. *Eur. J. Biochem.* **1982**, *129*, 171–178. (e) Brisson, J. R.; Carver, J. P. *Biochemistry* **1983**, *22*, 3680–3686. (f) Ohrui, H.; Nishida, Y.; Watanabe, M.; Hori, H.; Meguro, H. *Tetrahedron Lett.* **1985**, *26*, 3251–3254. (g) Homans, S. W.; Dwek, R. A.; Boyd, J.; Mahmoudian, J. B.; Richards, W. G.; Rademacher, T. W. *Biochemistry* **1986**, *25*, 6342–6350. (h) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6676–6683. (i) de Vries, N. K.; Buck, H. M. *Carbohydr. Res.* **1987**, *165*, 1–16. (j) Kroon-Batenburg, L. M. J.; Kroon, J. *Biopolymers* **1990**, *29*, 1243–1248.

(7) Haasnoot, C. A. G.; de Leeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783–2792.

(8) Schirmer, R. E.; Davis, J. P.; Noggle, J. H.; Hart, P. A. *J. Am. Chem. Soc.* **1972**, *94*, 2561–2572.

(9) (a) Džakula, Z.; Westler, W. M.; Edison, A. S.; Markley, J. L. *J. Am. Chem. Soc.* **1992**, *114*, 6195–6199. (b) Džakula, Z.; Edison, A. S.; Westler, W. M.; Markley, J. L. *J. Am. Chem. Soc.* **1992**, *114*, 6200–6207.

(10) Jaynes, E. T. In *The Maximum Entropy Formalism*; Massachusetts Institute of Technology: Cambridge, MA, 1979; pp 15–118.

(11) (a) Catelano, D.; Di Bari, L.; Veracini, C. A.; Shilstone, G. N.; Zannoni, C. *J. Chem. Phys.* **1991**, *94*, 3928–3935. (b) Berardi, R.; Spinozzi, F.; Zannoni, C. *J. Chem. Soc., Faraday Trans.* **1992**, *88* (13), 1863–1873.

0002-7863/93/1515-8421$04.00/0

overall and the internal motions of the molecule.[12] For intramolecular motions occurring on the time scale equal to, or shorter than, that of the overall molecular tumbling, the 'NOE constraints' do not have a simple geometrical interpretation, which limits their utility in building 3D structures. This problem will be addressed in the present study.

Methods

Maximum-Entropy Method. The ensemble average of an observable, O_m, which in this study will be a scalar coupling constant and the inverse sixth power of the internuclear distance, can be expressed by

$$\langle O_m \rangle = \int_0^{2\pi} O_m(\Omega) p(\Omega)\, d\Omega \tag{1}$$

where $p(\Omega)$ is the probability density distribution of the Ω angle. Among all possible distributions, the least biased one is the maximum-entropy probability distribution, $p^{ME}(\Omega)$, having the form[10]

$$p^{ME}(\Omega) = \frac{1}{Z} \exp(-\sum_{m=1}^{N} \lambda_m O_m(\Omega)) \tag{2}$$

where Z is a partition function

$$Z = \int_0^{2\pi} \exp(-\sum_{m=1}^{N} \lambda_m O_m(\Omega)) \tag{3}$$

and N is the total number of all constraints. The Lagrange multipliers, λ_m, can be found either by solving the nonlinear set of equations (1) or by minimizing the function F[13]

$$F = \ln Z + \sum_{m=1}^{N} \lambda_m \langle O_m \rangle_{exp} \tag{4}$$

which, for a consistent and linearly independent set of constraints, is a concave function in the entire $\{\lambda_m\}$ space.[13a] Function F is also an upper bound to the entropy of the distribution, $S(p)$.[13a]

In reality, our observables contain experimental errors, and models describing the functional dependencies $O_m(\Omega)$ may not be accurate. As a consequence, the constraints, especially if their number is large, may become inconsistent.[14] For these cases function F does not have a minimum[13a] and thus $P^{ME}(\Omega)$ cannot reproduce the data. However, maximum-entropy distributions which reproduce experimental data to within a specified accuracy can be determined.[14] For example, consistent sets of constraints corresponding to the minimum of F can be derived by randomly scanning NMR observables within certain ranges corresponding to experimental error.[11b] However, for this study, a different computational strategy was chosen: instead of minimizing function F, we bounded the feasible solutions by minimizing the quantity[14]

$$\sum_{m=1}^{N} w_m(\langle O_m \rangle - \langle O_m \rangle^{\#})^2 = Q_{min} \tag{5}$$

where w_m are weights chosen to be $\langle O_m \rangle^{\#-2}$. $\langle O_m \rangle$ are computed averages from eqs 1 and 2, and $\langle O_m \rangle^{\#}$ are "improved" experimental observables which can differ from the measured $\langle O_m \rangle_{exp}$ by no more than the estimated error bound, δ_m. For computational reasons they are expressed in the following form:

$$\langle O_m \rangle^{\#} = \langle O_m \rangle_{exp} + \delta_m \sin \gamma_m \tag{6}$$

where $\sin \gamma_m$ is an error-scaling factor. Since both λ_m and γ_m were optimized during minimization, Q_{min} should practically converge to zero ($<10^{-10}$). In these calculations all $\langle O_m \rangle^{\#}$ values within $\pm\delta_m$ were assumed to be equally probable, so the final probability distribution is calculated as an average over a sufficiently large number of solutions, M

$$P(\Omega) = \frac{1}{M} \sum_{k=1}^{M} P_k^{ME}(\Omega) \tag{7}$$

where $P_k^{ME}(\Omega)$ are distributions corresponding to the different solutions sets, $\{\lambda_m,\gamma_m\}_k$, of eq 5. Alternatively one can test individual $P_k^{ME}(\Omega)$ distributions against additional information, like, for example, temperature dependence of scalar coupling constants and NOE effects.

Constraint Equations. The Ω angle is defined by the planes containing the O6–C6–C5–O5 atoms. Ω is zero when O6 eclipses O5. Looking along the C6–C5 bond toward C5, anticlockwise rotation of the O6–C6 bond creates a positive Ω angle. The stereospecific assignment of prochiral protons H6R and H6S in β-glucose has been obtained by a selective deuteration technique.[6f] For the reducing glucosyl unit in β-gentiobiose this assignment is unequivocal, on the basis of homonuclear and heteronuclear coupling constants, and is the same as that in β-glucose.[6f] The functional dependencies, $^3J(\Omega)$, for ^1H–^1H and ^{13}C–^1H coupling constants were adopted from refs 7 15, respectively. The NOE constraint equations were obtained by nonlinear least-squares fitting of internuclear distances, derived from an energy-minimized X-ray structure, to the parametric form $r^{-6}(\Omega) = [\alpha + \beta \cos(\Omega + \gamma)]^{-3}$. In the calculations, the magnitudes of the NOE constraints and J couplings were balanced by multiplying the former by the factor 10^3.

Computational Details. All calculations were performed with 386-MATLAB software[16] running on an IBM 386 PC. For minimization of the F function in eq 4, the "fminu" procedure based on the quasi-Newton BFGS algorithm[17] was utilized. The quadratic form in eq 5 was minimized together with the normalization condition, $Z = 1$ in eq 3, using the "leastsq" procedure, which is based on the Levenberg–Marquardt algorithm.[17] Every calculation was restarted from different initial conditions, $\{\lambda_{m0},\gamma_{m0}\}$, where all λ_{m0} were set to zero and γ_{m0} were randomly picked up from the interval (–90°,90°). For each method of calculation (methods A and B in Table III), about sixty different minimizations were performed, rejecting all results (ca. 10%) with $Q_{min} > 10^{-8}$.

NMR Measurements and Evaluation. The NMR measurements were obtained using a Bruker AM 500 spectrometer. Gentiobiose (5 mg) was dissolved in 0.4 mL of $D_2O/(CD_3)_2CO$ (4:1, v/v).

The homonuclear coupling constants were measured by fitting the spectral lines to Lorentzian functions, using Felix software (Hare Research, Inc.) running on an IRIS 4D/220 GTX computer. The heteronuclear coupling constants $^3J_{C4H6R}$ and $^3J_{C4H6S}$ were obtained by the selective, proton-detected HSQC[18] experiment,[19] using a 180° selective proton DANTE[20] pulse and carbon chemical shift selective filter,[21] which was systematically incremented.[22] The same couplings were measured independently by the proton-flip 2D J-resolved ^{13}C technique,[23] with selective 180° DANTE pulses for protons.[24] In both experiments the spectral resolution was 0.2 Hz/pt.

The NOE measurements were obtained by selective 1D NOESY.[25] For selective excitation, the composite 270° DANTE pulse[26] ($90°_x$ – $180°_{x,y,-x,-y}$), of duration 20–50 ms depending on the desired selectivity, was used. The cross-relaxation rates for the probability distribution

(12) (a) Werbelow, L. G. *J. Am. Chem. Soc.* **1974**, *96*, 4747–4749. (b) Tropp, J. *J. Chem. Phys.* **1980**, *72*, 6036–6043. (c) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546–4570. (d) Olejniczak, E. T.; Dobson, C. M.; Karplus, M.; Levy, R. J. *J. Am. Chem. Soc.* **1984**, *106*, 1923–1930. (e) LeMaster, D. M.; Kay, L. E.; Brünger, A. T.; Prestegard, J. H. *FEBS Lett.* **1988**, *236*, 71–76. (f) Farmer, B. T.; Macura, S.; Brown, L. R. *J. Magn. Reson.* **1988**, *80*, 1–22. (g) Kessler, H.; Griesinger, Ch.; Lautz, J.; Müller, A.; van Gunsteren, W. F.; Berendsen, H. J. C. *J. Am. Chem. Soc.* **1988**, *110*, 3393–3296. (h) Koning, T. M. G.; Boelens, R.; Kaptein, R. *J. Magn. Reson.* **1990**, *90*, 111–123. (i) Duben, A. J.; Hutton, W. C. *J. Am. Chem. Soc.* **1990**, *112*, 5917–5924. (j) Withka, J. M.; Swaminathan, S.; Beveridge, D. L.; Bolton, P. H. *J. Am. Chem. Soc.* **1991**, *113*, 5041–5049. (k) Brüschweiler, R.; Roux, B.; Blackledge, M.; Griesinger, C.; Karplus, M.; Ernst, R. R. *J. Am. Chem. Soc.* **1992**, *114*, 2289–2302. (l) Poppe, L.; van Halbeek, H. *J. Am. Chem. Soc.* **1992**, *114*, 1092–1094.

(13) (a) Alhassid, Y.; Agmon, N.; Levine, R. D. *Chem. Phys. Lett.* **1978**, *53*, 22–26. (b) Agmon, N.; Alhassid, Y.; Levine, R. D. *J. Comput. Phys.* **1979**, *30*, 250–258.

(14) Manz, J. *J. Chem. Phys.* **1976**, *14*, 385–391.

(15) Spoormaker, T.; De Bie, M. J. A. *Recl. Trav. Chim. Phys-Bas* **1978**, *97*, 81–90.

(16) (a) *386-MATLAB*; The MathWorks, Inc.: Natick, MA, 1990. (b) *Optimization Toolbox*; The MathWorks, Inc.: Natick, MA, 1990.

(17) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*; Cambridge University Press: Cambridge, England, 1988.

(18) Bodenhausen, G.; Ruben, D. *Chem. Phys. Lett.* **1980**, *69*, 185–188.

(19) Poppe, L.; van Halbeek, H. *J. Magn. Reson.* **1991**, *92*, 636–641.

(20) Morris, G. A.; Freeman, R. *J. Magn. Reson.* **1978**, *29*, 433–462.

(21) Delsuc, M. A.; Guittet, E.; Lallemand, J. Y. *Magn. Reson. Chem.* **1985**, *23*, 213–215.

(22) Otting, G. *J. Magn. Reson.* **1990**, *86*, 496–508.

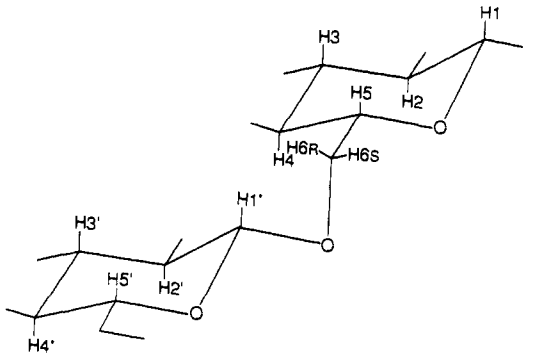
(23) Bax, A.; Freeman, R. *J. Am. Chem. Soc.* **1982**, *104*, 1099–1100.

(24) Morat, C.; Taravel, F. R.; Vignon, M. R. *Magn. Reson. Chem.* **1988**, *26*, 264–270.

(25) Kessler, H.; Mronga, S.; Gemmecker, G. *Magn. Reson. Chem.* **1991**, *29*, 527–557.

(26) Sklenar, V.; Feigon, J. *J. Am. Chem. Soc.* **1990**, *112*, 5663–5666.

Chart I



calculations were calculated using the full relaxation matrix approach.[27] The spin system was reduced to six spins: H6*R*, H6*S*, H5, H4(3), H1, H1′ (see Chart I). The strongly coupled, almost degenerate H4 and H3 protons, can then be treated as a single spin. Selective excitation was applied to the H6*R*, H6*S*, H5, and H1 proton resonances. The 1D NOESY data were collected for ten mixing times within 0.2–2 s at 263 and 301 K. All spectra were multiplied before FT transformation with a Gauss–Lorentz window function[28] (LB = –0.7, GB = 0.04) and zero filled to give a final resolution of 0.3 Hz/pt. All experimental intensities were corrected for differential relaxation losses during the selective pulses. The relaxation matrix elements were optimized by fitting simulated peak intensites to all experimentally available intensities using the 'leastsq" procedure within the 386-MATLAB software. For standard error estimates[29] the calculations were repeated for over sixty different sets of data points obtained by random variations of the experimental intensities, I (the total number was 160). The bias was either generated by normal distribution with variance reaching $0.2I$ or by uniform distribution within the interval $\pm 0.3I$. Signal intensities from repeated experiments, on average, differed by no more than 10%. In all other cases cross-relaxation rates were calculated by linearizing multiexponential build-up curves of NOE effects[30] using the formula

$$\sigma_{\parallel} = \lim_{\tau_m \to 0} \frac{I_{ij}}{I_{ii}\tau_m}$$

where I_{ij} and I_{ii} are NOE and selectively excited signal intensities, respectively, at a mixing time τ_m.

Results

Synthetic Data. Before the analysis of experimental data, we first generated synthetic sets of the eight average NMR constraints, $\langle r^{-6}_{H5H6R}\rangle, \langle r^{-6}_{H5H6S}\rangle, \langle r^{-6}_{H4H6R}\rangle, \langle r^{-6}_{H4H6S}\rangle, \langle {}^3J_{H5H6R}\rangle, \langle {}^3J_{H5H6S}\rangle, \langle {}^3J_{C4H6R}\rangle, \langle {}^3J_{C4H6S}\rangle$, obtained from the different Gaussian distributions (solid lines in Figure 2). The maximum-entropy density distributions, obtained by minimization of function F in eq 4, are also plotted in Figure 2 as dotted lines. Essentially the same results were also obtained by solving the least-squares problem (eq 5). A number of different simulations have shown that with eight, and in some cases even five, constraints the maximum-entropy distributions adequately reproduced up to three peak Gaussian distributions. Interestingly, the efforts of fitting synthetic data to the distribution of the Gaussian form and using eqs 1 and 5 always failed unless the initial guesses of the fitted parameters were very close to the true values.

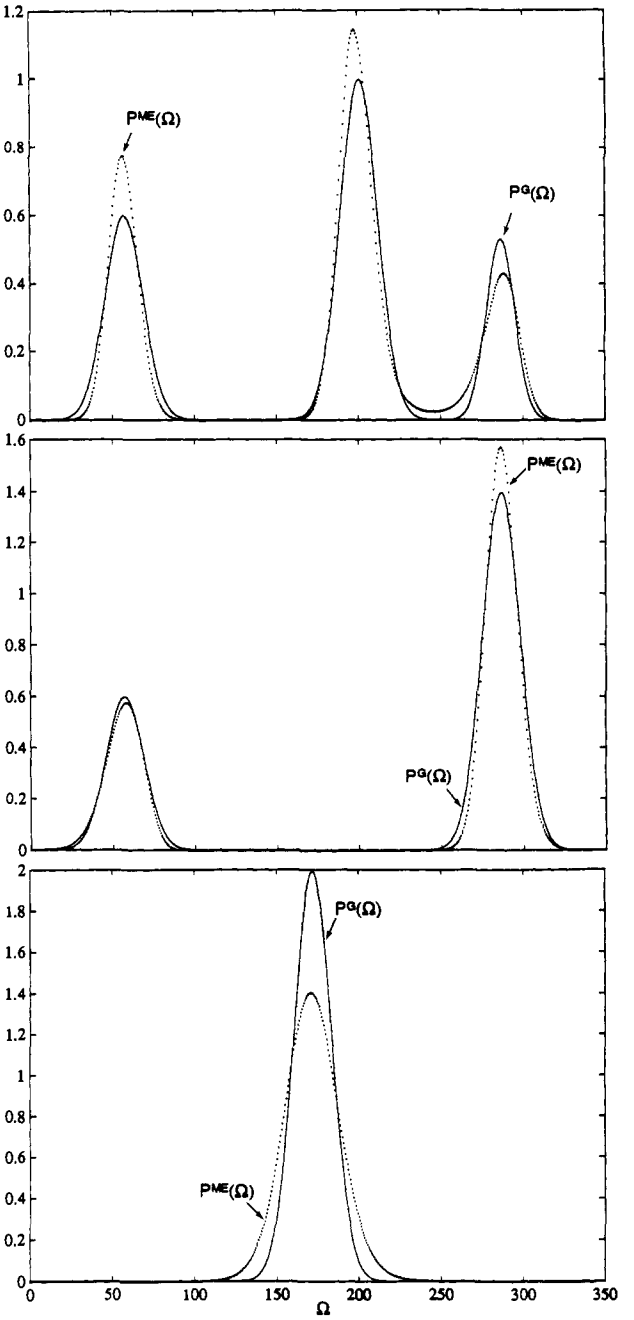


Figure 2. Maximum-entropy probability density distributions $P^{ME}(\Omega)$ (dotted) obtained after minimization of eq 4 with $\langle O_m\rangle$ values simulated from the corresponding Gaussian distribution $P^{G}(\Omega)$ (solid).

NMR Constraints. Since the experimental constraints are noisy, the error bounds for each constraint were determined before solving eq 5.

The functional dependence of vicinal coupling constants on the Ω angle is given in a form of the Karplus equation.[31] For proton–proton couplings in carbohydrates this equation has been parametrized by Haasnoot et al.[7] The root mean square error was reported to be <0.5 Hz. We adopted this value as an error bound even though we can measure these couplings with greater precision (compare Table I). However using a larger error bound takes into account the uncertainty of parameters in the Karplus equation.

The heteronuclear coupling constants, ${}^3J_{C4H6R,S}$, were measured by a selective 1D HSQC experiment[19] (Figure 3) and independently from a proton-flip 2D J-resolved ${}^{13}C$ spectrum.[23] For the

(27) (a) Massefski, W.; Bolton, R. *J. Magn. Reson.* **1985**, *65*, 526–530. (b) Olejniczak, E. T.; Gampe, R. T.; Fesik, S. W. *J. Magn. Reson.* **1986**, *65*, 28–41. (c) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1988**, *79*, 493–512. (d) Post, C. B.; Meadows, R. P.; Gorenstein, D. G. *J. Am. Chem. Soc.* **1990**, *112*, 6796–6803.

(28) Weiss, G. H.; Ferretti, J. A.; Byrd, R. A. *J. Magn. Reson.* **1987**, *71*, 97–105.

(29) Alper, J. S.; Gelb, R. I. *J. Phys. Chem.* **1990**, *94*, 4747–4751.

(30) (a) Eaton, H. L.; Andersen, N. H. *J. Magn. Reson.* **1987**, *74*, 212–225. (b) Baleja, J. D.; Moult, J.; Sykes, B. D. *J. Magn. Reson.* **1990**, *87*, 375–384. (c) Majumdar, A.; Hosur, R. V. *Biochem. Biophys. Res. Commun.* **1989**, *159*, 886–892.

(31) Karplus, M. *J. Chem. Phys.* **1957**, *30*, 11–15.

(32) Lucas, N. *J. Mol. Phys.* **1971**, *22*, 147–154.

Table I. Vicinal Coupling Constants in Hz from H6*R* and H6*S* Protons in Gentiobiose

atom pair	coupling constant	experimental error	error bound[a]
H5–H6*R*	6.0	0.1	0.5
H5–H6*S*	2.0	0.1	0.5
C4–H6*R*	1.0	0.5	1.0
C4–H6*S*	3.0	0.5	1.0

[a] This error bound was used for the calculations of maximum-entropy distribution.

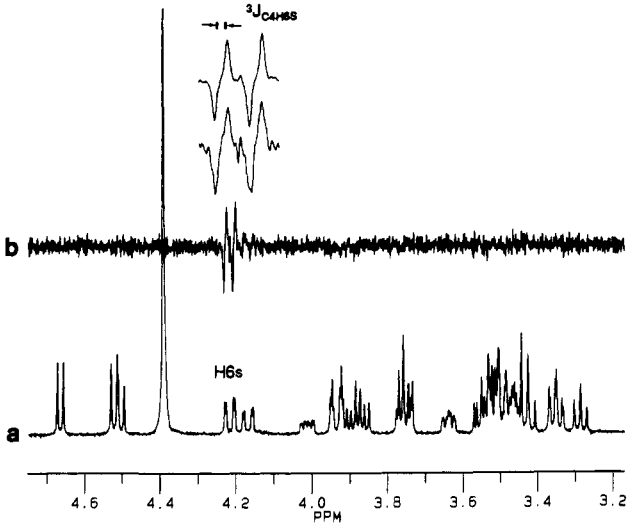


Figure 3. Example of the measurement of heteronuclear coupling constant: (a) ^{1}H ground spectrum of gentiobiose at 500 MHz, 1 scan; (b) 1D selective, in both proton and carbon domains, HSQC experiment, 3200 scans; (c, inset) expansions of experimental (bottom) and theoretical (top) proton multiplets. The last was obtained by convolution of spectrum a with a function $\mathrm{sgn}(\nu - \nu_0)\delta(|\nu - \nu_0| - 0.5(^3J_{C4H6S}))$, where ν_0 is the chemical shift, in Hz, of the H6*S* resonance.

Karplus equation we adopted the parametrization of Spoormaker and De Bie,[15] which gives satisfactory agreement for a number of measured ^{13}C–^{1}H coupling constants.[24] The precision of the experimental data as well as model parameters is lower than for homonuclear couplings, so we chose the error bound of 1 Hz (Table I).

1D NOESY measurements for H6*R*,*S*–H5,H4 proton pairs were performed at 265 and 301 K, where $\omega\tau_0 \approx 3$ and ≈ 0.5, respectively.[33] Typical spectra are shown in Figure 4a and b. The relaxation rates were obtained using the full relaxation matrix approach,[27] and experimental errors (see Table II) were estimated from Monte–Carlo analysis,[29] as described in the Methods section. Since this relaxation analysis does not include cross-correlation effects, which might introduce systematic errors, we examined whether it is justified to ignore these effects. The simulated NOE build-up curves with (dotted lines) and without (solid lines) cross-correlation effects[35] are shown in Figure 5, where it is apparent that differences between both types of curves are much smaller than experimental error for $\omega\tau_0 = 0.5$ and practically vanish when $\omega\tau_0 = 3$.

(33) The $\omega\tau_0$ values were obtained from H1α–H2 cross-relaxation rates.
(34) Similar values for the correlation times, τ_i, have been obtained from $T_1(^{13}C)$ measurements for the CH_2OH groups of the glucose residues in lactose derivatives, as reported by: Ejchart, A.; Dabrowski, J. *Magn. Reson. Chem.* **1992**, *30*, 115–124.
(35) (a) Werbelow, L. G.; Grant, D. M. *Advances in Magnetic Resonance*; Waugh, J. S., Ed.; Academic Press: New York, 1977; Vol. 9, pp 189–299. There is a typographical error in Table VIIa, where all Δ_{jlk} should be δ_{jlk}. (b) Bull, T. E. *J. Magn. Reson.* **1987**, *72*, 397–413. (c) Dalvit, C.; Bodenhausen, G. *Advances in Magnetic Resonance*; Warren, W. S., Ed.; Academic Press: San Diego, CA, 1990; Vol. 14, pp 1–33.
(36) Topping, J. *Errors of Observation and Their Treatment*; Champan and Hall Ltd.: London, 1955.

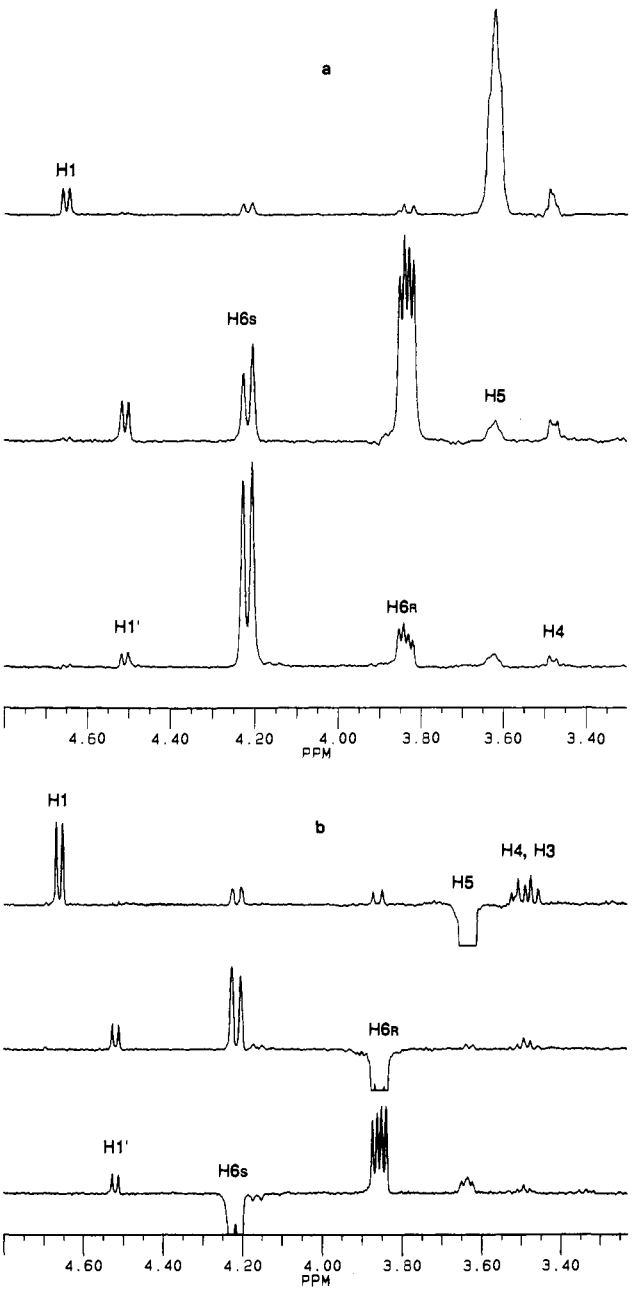


Figure 4. 1D NOESY spectra of gentiobiose in aqueous solution: (a) at $\omega\tau_0 \approx 3$ and mixing time 260 ms; (b) at $\omega\tau_0 \approx 0.5$ and mixing time 330 ms.

Table II. Cross-Relaxation Rates and Average 'NOE Distances' for Maximum-Entropy Probability Distribution Calculations, Obtained at 265 and 301 K from 1D NOESY Spectra at 500 MHz

	$\omega\tau_0 \approx 0.5$[a]		$\omega\tau_0 \approx 3$	
proton pair	σ_m/s^{-1}	$\langle r_m^{-6}\rangle \times 10^3$[b]/Å^{-6}	σ_m/s^{-1}	$\langle r_m^{-6}\rangle \times 10^3$/Å^{-6}
H6*R*–H6*S*	0.393 ± 0.052[c]		−0.736 ± 0.07	
H6*R*–H5	0.033 ± 0.004	2.64 ± 0.51[d]	−0.117 ± 0.010	5.0 ± 0.72
H6*S*–H5	0.054 ± 0.010	4.16 ± 1.00	−0.183 ± 0.012	7.82 ± 1.05
H6*R*–H4	0.039 ± 0.006	3.12 ± 0.67	−0.115 ± 0.011	4.91 ± 0.74
H6S–H4	0.021 ± 0.005	1.68 ± 0.47	−0.068 ± 0.007	2.90 ± 0.45

[a] These values were obtained from the H1α–H2 cross-relaxation rate. [b] Calculated from eq 8, where $r_0 = 1.78 \pm 0.02$ Å. [c] Standard errors obtained from Monte–Carlo simulations.[29] [d] Standard errors calculated from the error propagation laws.[36]

As a constraint, the quantity required is the so-called 'NOE distance', $\langle r_{ij}^{-6}\rangle$, between the H_i–H_j proton pair, where $\langle\ \rangle$ denotes the ensemble average value. If the internal motions are slower

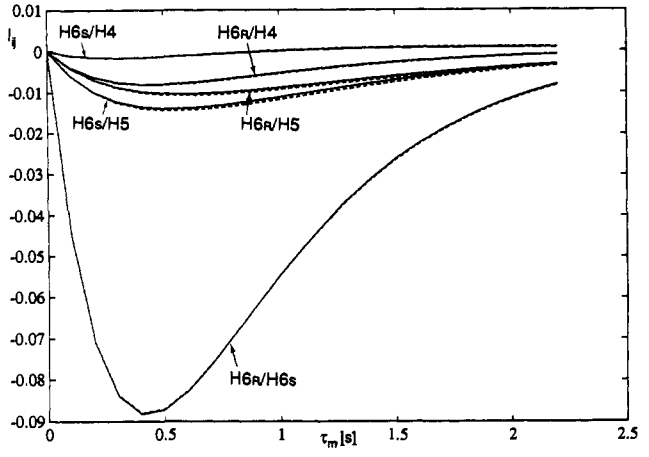


Figure 5. Theoretical NOE build-up curves for the H6*R*H6*S*H5H4 spin system at $\omega\tau_o = 0.5$ without (solid lines) and with (dashed lines) cross-correlation effects. For the calculations of spectral densities we assumed slow ($\sigma < k < \tau_o^{-1}$) interconversion between equally populated *gg* and *gt* conformations.

than the overall tumbling ($\tau_i > \tau_o$, where τ_i and τ_o are correlation times for internal and overall motion, respectively), the following equation is commonly used:

$$\langle r_{kl}^{-6} \rangle = r_o^{-6} \frac{\sigma_{kl}^{\parallel}}{\sigma_o^{\parallel}} \tag{8}$$

where r_o is some conformationally independent internuclear distance and $\sigma_o^{\parallel}$ is the laboratory frame cross-relaxation rate. In this study H6*R*–H6*S* was chosen for a reference interaction because $\sigma_o^{\parallel}$ is measured with higher precision for this proton pair and because vibrational averaging[32] has a small effect on the NMR-derived distance, which is 1.78 ± 0.02 Å. The apparent average values, $\langle r_{ij}^{-6} \rangle$, obtained from eq 8 are listed in Table II. At $\omega\tau_o \approx 3$ these values are in strong variance with those obtained at $\omega\tau_o \approx 0.5$ (Table II). These differences, however, can be well explained in terms of the fast, 10^9–10^{10} s^{-1}, conformational transitions about the C5–C6 bond. The τ_i values were estimated from the temperature dependence, from 261 up to 330 K, of the H6*R*–H6*S* cross-relaxation rate, relative to those of the cross-relaxation rates for proton pairs having fixed positions within the pyranose ring (H1α–H2, H1–H3, H1–H5, H1β–H2 in Chart I). The apparent H6*R*–H6*S* distance, as obtained from eq 8, has been found to be longer (by ca. 0.1 Å) at low temperatures than at ambient temperatures and above, where its value is close to 1.8 Å. This observation strongly suggests that, at 261–269 K, $\tau_i < \tau_o$ and, at ambient temperature, $\tau_i > \tau_o$. As a consequence we should expect that only the data in the second column in Table II adequately represent the ensemble average values, $\langle r_{ij}^{-6} \rangle$ (see also Discussion).

Calculations of Maximum-Entropy Distributions. The maximum-entropy distributions, $P_k^{ME}(\Omega)$, were computed according to the procedure outlined in the Methods using coupling constants and 'NOE distances' measured at 301 K. The average parameters and their standard deviations obtained from simulations are listed in Table III (column A). The average probability density has been calculated from eq 7 and is shown in Figure 6 as a solid line. The calculations were repeated for the same set of coupling constants, but 'NOE constraints' were expressed in the new form:

$$\langle r_m^{-6} \rangle^{\#} = \xi(\sigma_m + \delta_m \sin \gamma_m) \tag{9}$$

where δ_m stands for the error bound for cross-relaxation rates, set to standard errors given in Table II. The new parameter ξ is common for all distance constraints and together with λ_m and γ_m was optimized during the course of the calculations. The differences between ξ factors for different proton pairs, caused by the anisotropic motion, should be small enough to diffuse during the course of calculations into the uncertainties of σ_m values. The results from both methods (Table III and Figure 6) are very similar. For the NOE data obtained at low temperatures (the last column in Table II) the calculations did not converge. As will be discussed shortly, application of eq 9 in this case is not well justified.

Table III. Results from Three Different Calculations of Maximum-Entropy Distributions

parameter[a]	calculation A[b]	calculation B[c]
λ_1	0.031 ± 0.070	−0.004 ± 0.042
λ_2	0.892 ± 0.206	0.933 ± 0.170
λ_3	0.484 ± 0.160	0.525 ± 0.125
λ_4	−0.319 ± 0.189	−0.455 ± 0.177
λ_5	0.445 ± 0.209	0.699 ± 0.250
λ_6	−0.433 ± 0.312	−0.617 ± 0.193
λ_7	−0.550 ± 0.227	−0.369 ± 0.117
λ_8	0.527 ± 0.317	0.609 ± 0.241
$^3J_{H5H6R}$ (Hz)	6.26 ± 0.21	6.23 ± 0.29
$^3J_{H5H6S}$ (Hz)	2.42 ± 0.06	2.41 ± 0.12
$^3J_{C4H5R}$ (Hz)	1.30 ± 0.04	1.26 ± 0.08
$^3J_{C4H6S}$ (Hz)	3.47 ± 0.10	3.50 ± 0.14
$\langle r_{H6RH5}^{-6} \rangle \times 10^3$ (Å^{-6})	2.48 ± 0.25	2.30 ± 0.16
$\langle r_{H6SH5}^{-6} \rangle \times 10^3$ (Å^{-6})	3.52 ± 0.11	3.50 ± 0.10
$\langle r_{H6RH4}^{-6} \rangle \times 10^3$ (Å^{-6})	2.49 ± 0.06	2.47 ± 0.11
$\langle r_{H6SH4}^{-6} \rangle \times 10^3$ (Å^{-6})	1.50 ± 0.22	1.58 ± 0.13
p_{gt}	0.66 ± 0.06	0.69 ± 0.04
p_{gg}	0.34 ± 0.06	0.31 ± 0.04
Ω_{gt} (deg)	79.0 ± 3.4	80.0 ± 2.5
Ω_{gg} (deg)	−68.9 ± 6.3	−65.01 ± 4.2
ξ		68.1 ± 4.8
S^d	0.819 ± 0.100	0.721 ± 0.108

[a] These are average values with their standard errors obtained from 60 different calculations. [b] Calculation using 'NOE distances' obtained at $\omega\tau_o \approx 0.5$ and eq 6. [c] Calculation using cross-relaxation rates obtained at $\omega\tau_o \approx 0.5$ and eq 9. [d] S is an upper bound for entropy calculated from eq 4 and can be compared with $S = \ln 2\pi \approx 1.84$ for the unconstrained case.

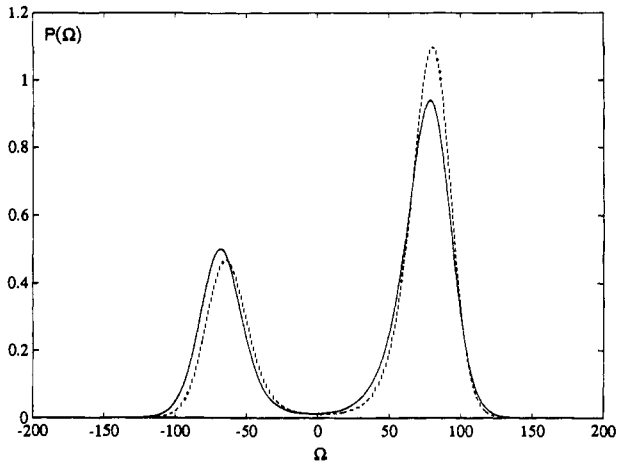


Figure 6. Average probability density distribution obtained from calculation A (solid line) and B (dashed line) from Table III.

Discussion

The distribution which is shown in Figure 6 is the least biased in the sense that no information was introduced that is not warranted by the data. On the other hand, this distribution might not necessarily reproduce the physical reality either because there is not enough data, or existing data is too uncertain, or both. New insights, however, can be gained by more careful analysis of the componet distributions, $P_k^{ME}(\Omega)$. All simulations showed two-peak distributions, where *gt* conformations prevailed with the maxima shifted away from the staggered position significantly more than in the *gg* conformers. As exemplified in Figure 7, the dispersion of the distributions is quite large and in the absence

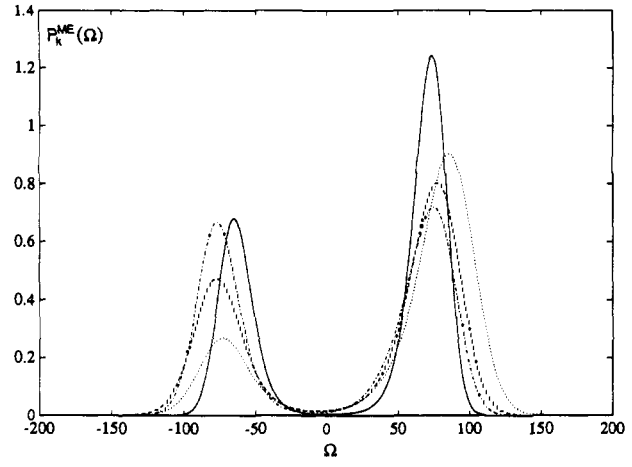


Figure 7. Some component distributions, $P_k^{ME}(\Omega)$, obtained from calculation A in Table III: (solid line) 0.36, –62°, 74°, 0.41, 0.562; (dashed line) 0.36, –78°, 77°, 0.46, 0.991; (dotted line) 0.23, –73°, 86°, 0.54, 0.947; (dashed-dotted line) 0.43, –78°, 77°, 0.48, 0.928. The numbers represent p_{gg}, Ω_{gg}, Ω_{gt}, root mean square error, F value (eq 4), respectively.

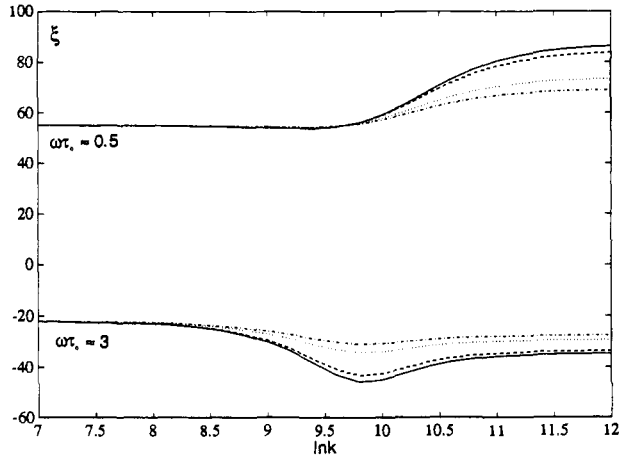


Figure 8. ξ values in Å^{-6}s (eq 9) calculated from the two-state, jump model with $p_{-62°} = 0.36$ and $\Omega_{gt} = 74°$ for H5–H6$_R$ (solid line), H5–H6$_S$ (dashed line) H4–H6$_R$ (dotted line), and H4–H6$_S$ (dashed-dotted line). $k = k_1 + k_{-1} = \tau_i^{-1}$ in s^{-1}, where k_1 and k_{-1} denote forward and reverse rate constants.

of any additional information there are no reasons to distinguish any of the $P_k^{ME}(\Omega)$ distributions. In this study however, two extra observations were not incorporated into the probability density calculations. First, the double selective 1D TOCSY–ROESY experiment for gentiobiose in H_2O solution has shown that the H6S–OH4 NOE interaction is clearly stronger than the H6R–OH4 interaction.[37] From the model (see Figure 1) this result is in excellent agreement with the distributions weighted toward the gt conformations. The second observation concerns relative temperature dependence of the cross-relaxation rates and needs slightly more elaboration. In order to rationalize the increase of the apparent average 'NOE distances' for H5,4–H6R,S pairs (compare the second and the fourth columns in Table II), we simulated the corresponding cross-relaxation rates in eq 8 using a two-state, jump model[12b] with p_{gt}:p_{gg} = 0.65:0.35. From these calculations and the geometry of the glucose residue it follows that, for $\Delta\Omega = \Omega_{gt} - \Omega_{gg} > 135°$, if $\tau_i < \tau_o$, the apparent $\langle r_{ij}^{-6}\rangle$ values are decreased relative to the $\tau_i > \tau_o$ case (keeping in mind that H6R–H6S is the reference interaction). For $\Delta\Omega < 135°$ the effect is completely opposite. Since the $\langle r_{ij}^{-6}\rangle$ values in Table II are larger at $\omega\tau_o \approx 3$, one should, accordingly, regard the distributions with $\Delta\Omega > 135°$ as less probable than those where $\Delta\Omega \lesssim 135°$ (compare distributions in Figure 7).

The differences between statistical averages and apparent averages obtained from eq 8 are suppressed at $\omega\tau_o \approx 0.5$ due to the overall molecular tumbling which acts like a bandpass filter for the internal motions.[12k] This can also be envisioned by calculating the ξ coefficients from eq 9 for the H6R,S–H5,4 atom pairs as a function of τ_i and τ_o (Figure 8). Figure 8 demonstrates that within the expected range of τ_i one can not assume a common ξ factor for all interactions at low temperature.

Concluding Remarks. This study presents a new approach to quantitative evaluation of the vicinal coupling constants and NOE effects in terms of multiple molecular conformations. It utilizes the maximum-entropy probability density distribution, which, as was shown for synthetic data, can adequately reproduce fairly complex Gaussian distributions. Moreover, the convergence of the fitting procedure (eq 5) is significantly better when using the maximum-entropy probability distribution instead of the Gussian form. On the other hand, however, the maximum-entropy distribution is built up from empirical relationships, which can be inaccurate. This certainly introduces a bias to the method, and with the large data sets, it might be preferable to apply variations to the parametric equations as well.

The computations started from the flat distribution where every angle Ω was equally probable. For many molecular systems one can derive the prior distribution, $m(\Omega)$, for example, on a purely stereochemical basis. Then, before solving the constraint problem (eq 5), $m(\Omega)$ can be combined with $p(\Omega)$ from eq 2, according to the Bayes rule.[10]

The approach presented here might be particularly useful for studying the side-chain conformations in carbohydrates such as exocyclic, amido, and hydroxyl groups, where the initial guess about rotameric states cannot be clearly defined. For example, the assumption about staggered rotameric states in gentiobiose gives, from $^3J_{H5H6R}$, $^3J_{H5H6S}$, and the Karplus equation, an inconsistent result ($p_{gg} = 0.56, p_{gt} = 0.57, p_{tg} = -0.13$). Extension of the method to more than one dimension is straightforward and might give interesting results if applied to the study of conformations of various glycosidic linkages.

Acknowledgment. This research is supported by National Institutes of Health Grants P41-RR-05351, P01-Al-27135, and S10-RR-04720. The author thanks Dr. Herman van Halbeek for general support and useful discussions and Malcom O'Neill and Willie York for their helpful comments in the preparation of the manuscript.

(37) Poppe, L.; van Halbeek, H. *J. Magn. Reson.* **1992**, *96*, 185–190.