MONTE CARLO CALCULATIONS OF PREFERRED CONFORMATIONS FOR 1, 2-DIACETAMIDO-GLUCOSE AND 1-METHOXY-2-ACETAMIDO-GALACTOSE

Anthony J. Duben
Chemistry Department
South Dakota School of Mines and Technology
Rapid City, South Dakota 57701-3995
C. Allen Bush
Chemistry Department
Illinois Institute of Technology
Chicago, Illinois 60616

ABSTRACT

The beta turn region of proteins has been suggested to be the likely site for glycosylation in the formation of glycoproteins. In order to search for likely conformations of glycosylated beta turns in N- and O- linked analogs of glycoprotein linkages, the empirical conformational energy method of Momany and coworkers was employed in a series of Monte Carlo calculations using the Metropolis algorithm. The specific compounds studied were the beta anomer of the N-linked 1,2-diacetamido-glucose and the alpha anomer of the O-linked 1-methoxy-2-acetamido-galactose. Conformational density of states' maps for the compounds in question were plotted. They indicated that only restricted sets of angular variables provided energetically favorable conformations. Comparison of the predicted favored conformations provides a critical test of the torsional potential functions used.

INTRODUCTION

When a protein chain is at the surface of a protein, it can assume a beta turn conformation. The beta turn is suspected of being a likely spot for glycosylation (Aubert et al., 1976; Beeley, 1977) since an attached carbohydrate in this location could interact with solvent or with carbohydrates on other cells or proteins.

Two of the most common linkages (Sharon, 1975) are the one between the beta anomer of 2-deoxy-2-N-acetyl-glucosamine and asparagine which involves a C-N bond at the anomeric carbon and the one between the alpha anomer of 2-deoxy-2-acetamido-D-galactopyranose and either serine or threonine which involves a C-O bond.

The peptide chains in many proteins tend to assume preferred conformations. As a consequence, the question is posed whether the glycopeptide linkage assumes a preferred conformation or set of conformations.

In a preliminary study (Bush, in press) of the conformational energy of 1-N-acetyl-1,2-dideoxy-2-acetamido-β-D-glucopyranosyl amine (to be called β-1,2-DAG), the empirical conformational energy method of Momany and coworkers (Momany, et al., 1974A, 1974B) was applied in order to determine the lowest energy conformation of this compound. The calculated results were compared with those obtained from circular dichroism and NMR experiments on these compounds.

In the work reported here, the search for preferred conformations was based on Boltzmann statistics instead of using the deliberate search method of the previous study. Specifically, a Monte Carlo calculation using the Metropolis algorithm (Metropolis, et al., 1953) was applied using the empirical conformational energy method to calculate energies. The results of this calculation include probabilities of finding conformations within a given range of geometric parameters. Low energy confromations are identified as well. If a sufficient number of conformations are generated, the set of conformations retained by the Metropolis algorithm should have relative probabilities corresponding to the Boltzmann distribution.

β-1,2-DAG, the analog of the N-linked glycoproteins, is reexamined in this work. In addition, the O-methyl glycoside of 2-N-acetyl-galactosamine (GalNAc-OMe) was studied as the analog of the O-linked glycoproteins.

COMPUTATIONAL METHODS

A. Empirical Energy Method

The empirical conformational energy method has been discussed at length in the original papers (Momany, et al., 1974A, 1974B).

In this method, four contributions to the energy are identified. They are the electrostatic, a Lennard-Jones nonbonded, a hydrogen bond, and a torsional contribution. The electrostatic, Lennard-Jones, and hydrogen bond contributions are calculated on the basis of atom pairs. These contributions depend only on the interatomic distances between the atoms and the chemical identity of the atoms. CNDO charges are used to calculate electrostatic interactions. One calculates *either* a Lennard-Jones nonbonded interaction using a 6-12 potential function *or* a hydrogen bond interaction using a 10-12 potential function after identifying the chemical types of the atoms involved. A torsional energy contribution is added for every dihedral angle that changes in the calculation.

The conformation of the sugar ring itself was fixed in the $^4C_1$ conformation. Atom coordinates of the starting conformations for

both compounds are listed in Table 1. The numbering system is shown on the accompanying Figs. 1 and 2.

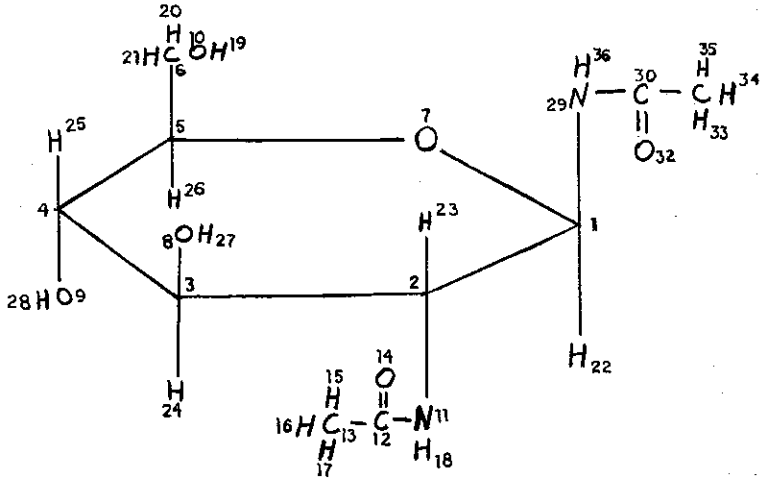

Figure 1. Structural Formula for β-1,2-DAG.

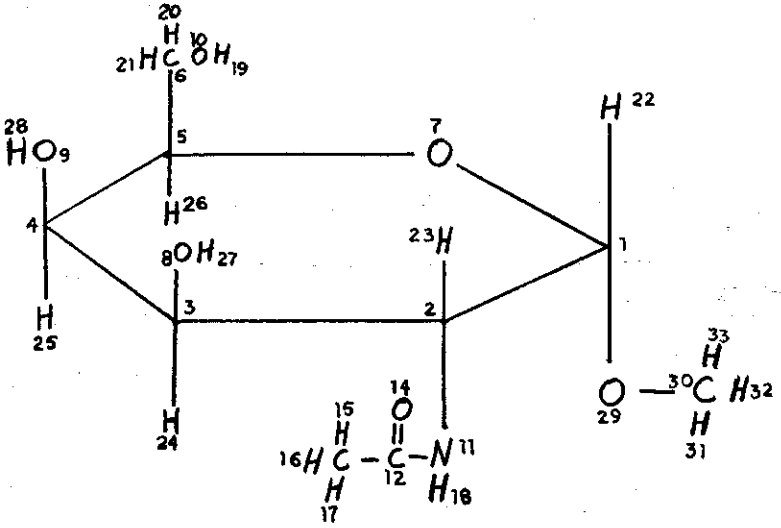

Figure 2. Structural Formula for GalNAc-OMe.

Two dihedral angles were allowed to change in each compound. For the N-linked β-1,2-DAG, the dihedral angles are identified as

Tau_1: C(carbonyl) - N(amide) - C_1(ring) - C_2(ring)
Tau_2: C(carbonyl) - N(amide) - C_2(ring) - C_1(ring)
For the O-linked GalNAc-OMe, the dihedral angles are identified as:
Tau_1: C(methyl) - O - C_2(ring) - C_3(ring)
Tau_2: C(carbonyl) - N(amide) - C_2(ring) - C_1(ring)

TABLE 1
Sample Conformation of β-1,2-DAG

Atom	X (Angstroms)	Y (Angstroms)	Z (Angstroms)
1	0.0	0.0	0.0
2	1.523	0.0	0.0
3	2.056	1.425	0.0
4	1.442	2.224	−1.142
5	−0.079	2.113	−1.108
6	−0.744	2.890	−2.224
7	−0.472	0.732	−1.136
8	3.484	1.402	0.0
9	1.942	3.559	−1.140
11	2.108	−0.765	1.090
10	−0.434	4.282	−2.159
12	1.873	−0.464	2.367
13	2.555	−1.356	3.358
14	1.157	0.466	2.729
15	1.869	−1.805	3.931
16	3.164	−0.810	3.934
17	3.089	−2.048	2.872
18	2.701	−1.536	0.860
19	−0.893	4.753	−2.912
20	−1.733	2.765	−2.148
21	−0.427	2.519	−3.097
22	−0.333	0.471	0.817
23	1.754	−0.442	−0.867
24	1.731	1.902	0.817
25	1.673	1.790	−2.014
26	−0.419	2.472	−0.239
27	3.825	2.342	0.0
28	1.527	4.063	−1.897
29	−0.512	−1.362	0.0

30	−0.230	−2.219	−0.979
31	−0.840	−3.579	−0.819
32	0.472	−1.943	−1.949
33	−1.836	−3.495	−0.786
34	−0.512	−3.993	0.030
35	−0.579	−4.155	−1.594
36	−1.091	−1.648	0.764

TABLE 2
Sample Conformation of GalNAc-OMe

Atom	X (Angstroms)	Y (Angstroms)	Z (Angstroms)
1	0.0	0.0	0.0
2	1.523	0.0	0.0
3	2.056	1.425	0.0
4	1.442	2.224	−1.142
5	−0.079	2.113	−1.108
6	−0.744	2.890	−2.224
7	−0.472	0.732	−1.136
8	3.484	1.402	0.0
9	1.759	1.622	−2.396
10	−0.434	4.282	−2.159
11	2.108	−0.765	1.091
12	3.426	−0.874	1.246
13	3.854	−1.699	2.421
14	4.249	−0.350	0.498
15	4.401	−1.137	3.041
16	4.401	−2.473	2.103
17	3.046	−2.037	2.903
18	1.490	−1.217	1.734
19	−0.893	4.753	−2.912
20	−1.733	2.765	−2.148
21	−0.427	2.519	−3.097
22	−0.326	−0.946	0.0
23	1.754	−0.442	−0.867
24	1.731	1.902	0.817
25	1.807	3.154	−1.121
26	−0.419	2.471	−0.239
27	3.825	2.343	0.0

28	2.753	1.602	−2.505
29	−0.502	0.668	1.157
30	−1.996	0.723	1.252
31	−2.363	−0.207	1.274
32	−2.363	1.207	0.458
33	−2.259	1.205	2.087

B. Monte Carlo Calculations

The application of the Monte Carlo method to the problem at hand is discussed briefly in this section. For a more extended discussion, the reviews of Wood (1968) and Valleau and Whittington (1977) should be consulted. Although these references discuss the method in the context of the simulation of liquids, the general principles discussed are applicable to the problem of biopolymer conformations as well as to the statistical thermodynamics of liquids.

Instead of explicitly incorporating both the value of a property of a state and its probability in a calculation of the average value of the property, the average is obtained as a sum over discrete states of unit weight. The states are generated as links in a Markov chain in which the probability of going to another state is used to select the states incorporated into the chain.

From the current conformation, a new conformation is generated by randomly selecting which angle was to be twisted and randomly choosing the angle of the twist. A full 360° range was chosen in order to preserve ergodicity. The energy of the new conformation was calculated, and the difference between the new energy and the old energy was found. If the energy of the new conformation was less than that of the old one, the new conformation was accepted without question as the next element in the Markov chain. If the new energy was higher than the old energy, a random number was generated and compared against the Boltzmann factor using the conformational energy difference. If the random number was smaller than the Boltzmann factor, the conformation was accepted. If not, it was rejected and the next link in the Markov chain was taken to be a repetition of the old conformation. In all of the calculations, the temperature was chosen to be 25°C.

The conformations in the Markov chain were used to generate probability density maps for the angles used in the calculation. Both single angle density plots and two angle density maps were generated. Although the two angle density plots resemble Ramachandran maps, they are not the same. The probability of entries

result from summing over all of the variables *not* being plotted. Ramachandran maps are not unique; probability maps are.

The vibrational heat capacity per torsion is calculated as a check on thermodynamic convergence of the calculation. Although a vibrational heat capacity less than R, the equipartition value, is not a sufficient condition on convergence, it is a necessary condition. This quantity can be calculated in the course of the calculation by using

$$C_v = \frac{[\,(E^2)_{ave} - (E_{ave})^2\,]}{R\,T^2} \qquad (1)$$

RESULTS

A. β-1,2-DAG

15,000 trials were used for β-1,2-DAG. The acceptance rate for new conformations was approximately 26%. 15,000 trials were more than sufficient for convergence. The long Markov chain was divided into two shorter chains of 7500 conformations. Each of them had the same distribution of conformations as the longer chain.

In the previous study (Bush, in press), the conformational energy map showed two major isolated regions of low energy. In order to satisfy the ergodicity requirement, both theoretically and practically, the full 360° range was taken so that both regions of low energy could be accessed. The double well conformational energy map is reflected in the probability maps for single angles as well as in the map for both angles (Figs. 3-5).

The two angle map (Fig. 5) has four major regions in which acceptable conformations may be found. The two angle map represents the coupled behavior of the two angles in contrast to the single angle graphs. The regions of acceptable conformations are separated by spaces in which the conformational energies are so high that the probability of accepting conformations in those regions is negligible.

The angles Tau_1 and Tau_2 have been discussed in previous papers (Bush and Duben, 1978; Bush, et al., 1980) as being "cis" or "trans." These descriptors specifically refer to the relationship between the hydrogen on the sugar ring carbon and the amide group hydrogen. "Cis" corresponds to an angle of −60° and "trans" to an angle of 120°.

Both cis and trans angles are possible for either dihedral angle. Tau_1 is more flexible than Tau_2. The most probable angle for Tau_1 is approximately 180°. An angle of 180° for Tau_1 would make the amide N-H bond parallel to the ring C_1-C_2 bond.

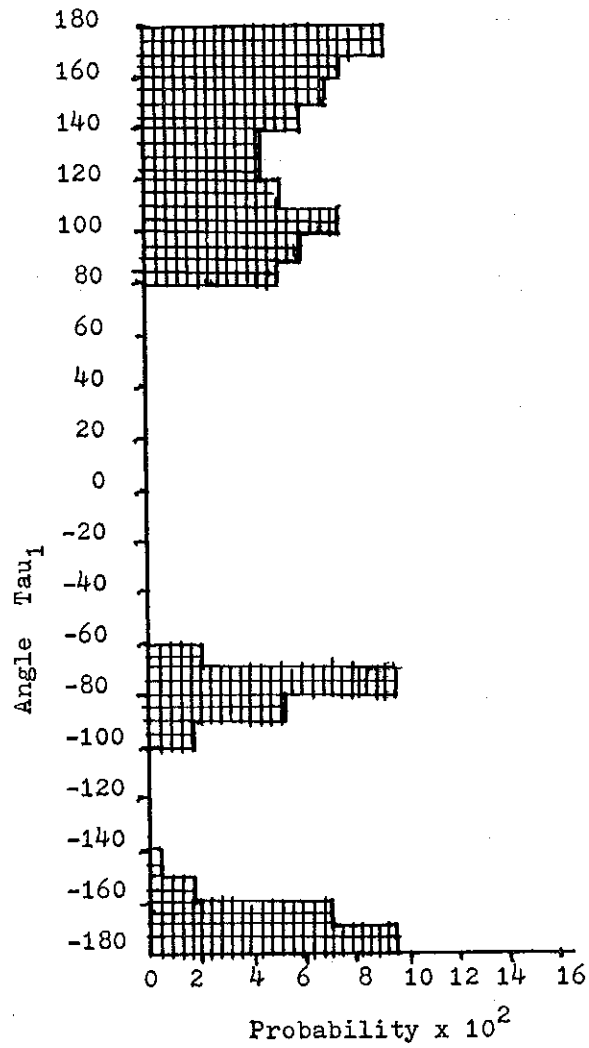

Figure 3. Probability Graph for Tau_1 in β-1,2-DAG.

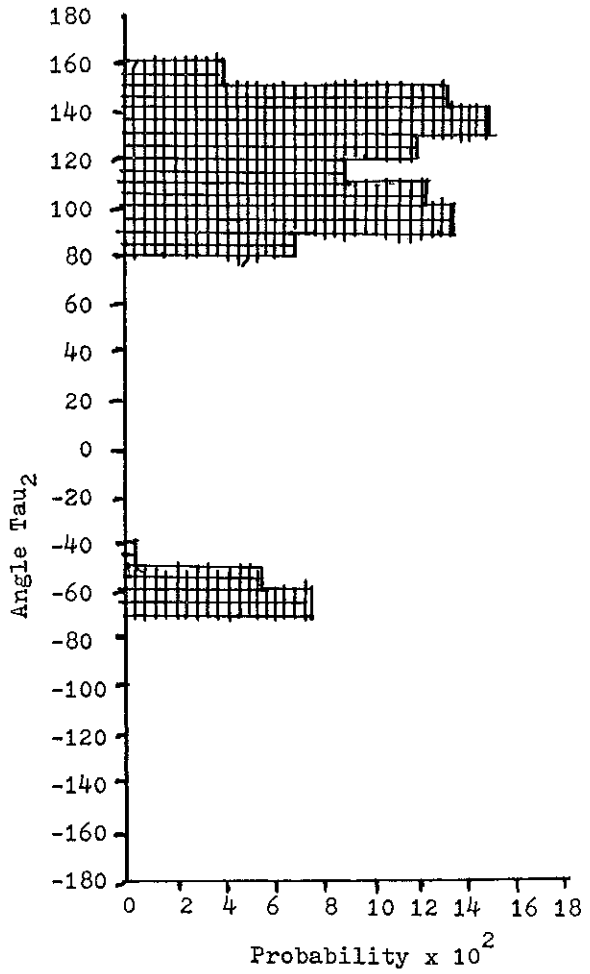

Figure 4. Probability Graph for Tau_2 in β-1,2-DAG.

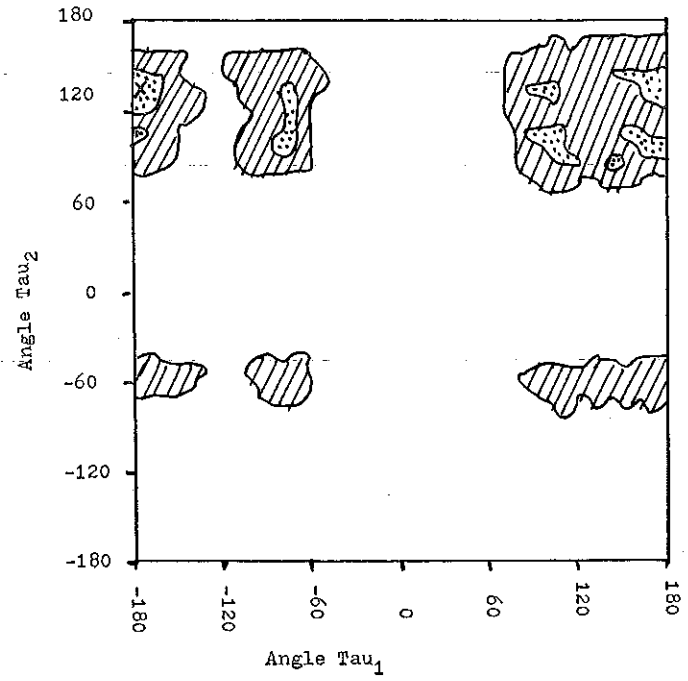

Figure 5. Probability Map for Tau_1 and Tau_2 in β-1,2-DAG. Cross-Lined Areas are Regions of Probability Between 0 and 1 x 10^{-2}; Dotted Areas are Regions of Probability Greater than 1 x 10^{-2}. X Marks the Section of Maximum Probability, 1.8 x 10^{-2}.

Tau_2 is more restricted than Tau_1. It is more probable that the amide hydrogen would be closer to a trans orientation for Tau_2 than for Tau_1.

For both angles the probability of assuming a pure "trans" orientation of 120° is reduced. This effect is probably due to non-bonded repulsions between the hydrogen on the ring carbon and the carbonyl oxygen.

In Fig. 5, the regions in which acceptable conformations have been found are labeled with Roman numerals I through IV. The lowest energy conformation was found in region I. The values of the dihedral angles characterizing the lowest energy conformations in each of the regions are listed in Table 3. Total probabilities for each region are given in the same table.

TABLE 3
Low Energy Conformations and Probabilities for β-1,2-DAG

Region (fig. 5)	Tau_1	Tau_2	Energy Above Overall Minimum (Joules)	Probability for Region
I	102	99	.....	.74
II	−78	99	450	.18
III	180	−60	1360	.06
IV	−77	−60	1550	.02

The heat capacity per degree of freedom was calculated to be 0.25R using eq. (1). This value is consistent with a convergent calculation.

B. GalNAc-OMe

The alpha methyl glycoside of 2-acetamido galactose is the simplest example of an O-linked species. 7500 trials were used for GalNAc-OMe. 18.2% of new conformations were accepted for inclusion in the Markov chain. This percentage is comparable to the area of the two angle map (Fig. 8) occupied by acceptable conformations.

The probabilities for the single angles Tau_1 and Tau_2 are presented in Figs. 6 and 7. Considerable steric hindrance is exhibited for the methoxy substituent at the anomeric carbon. The overall significance of steric hindrance is much higher in this compound than in β-1,2-DAG. Only a very small set of angles provide low energy conformations.

The terms "cis" and "trans" are still meaningful for angle Tau_2. They are not appropriate for Tau_1. $Tau_1 = O°$ corresponds to the O-CH_3 group eclipsing the ring C_1-C_2 bond. An angle of −120° has the O-CH_3 bond eclipsing the C_1-H bond. The O-CH_3 bond is farthest from the C_1-H bond when $Tau_1 = 60°$ by putting the O-CH_3 directly under the sugar ring.

No Tau_2 angles were found to have been cis. Tau_1 assumes angles with a negative sign. In the narrow range −160° to −170°, 30% of the Tau_1 angles were found.

The minimum energy conformation was found to have a Tau_1 value of −167° and a Tau_2 angle of 76°. The Tau_2 angle is more staggered than trans. The orientation of Tau_1 is intermediate between a staggered conformation in which the O-CH_3 bond bi-

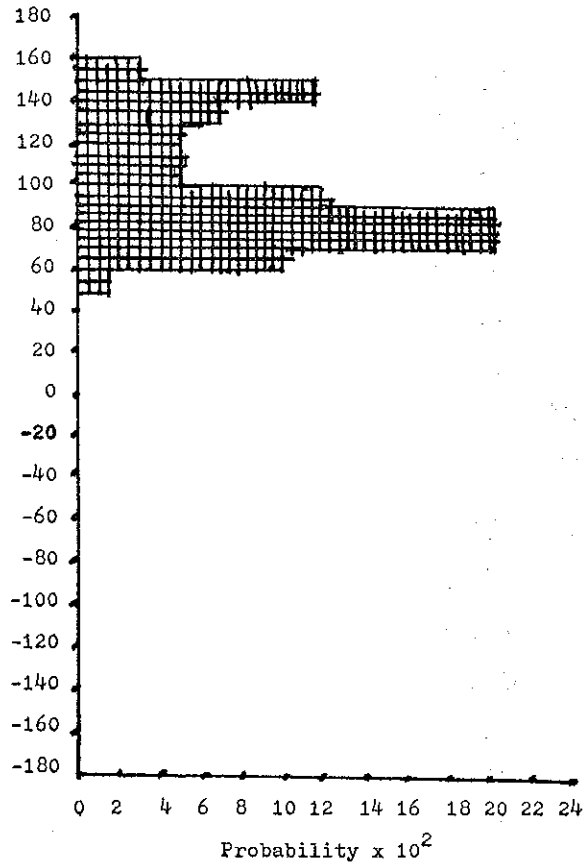

Figure 6. Probability Graph for Tau_1 in GalNAc-OMe.

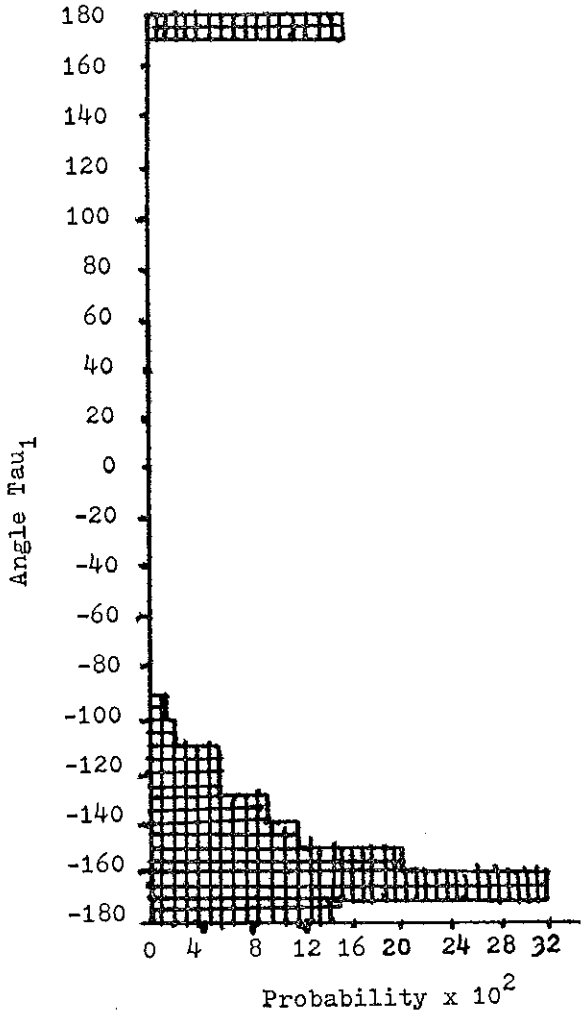

Figure 7. Probability Graph for Tau_2 in GalNAc-OMe.

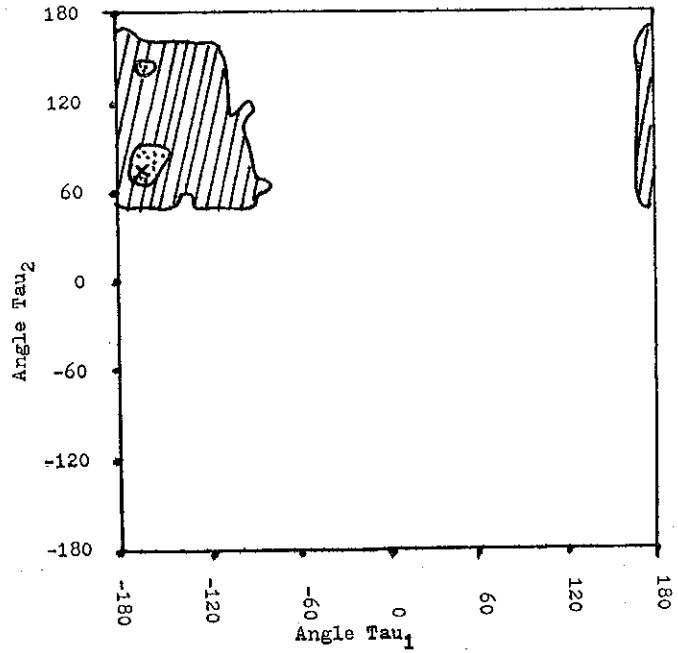

Figure 8. Probability Map for Tau_1 and Tau_2 in GalNAc-OMe. Cross-Lined Areas are Regions of Probability Between 0 and 4×10^{-2}; Dotted Areas are Regions of Probability Greater than 4×10^{-2}. X Marks the Section of Maximum Probability, 8.2×10^{-2}.

sects the angle between the anomeric carbon and the acetal oxygen and the conformation in which the O-CH_3 eclipses the C_1-H bond.

The vibrational heat capacity per degree of freedom was found to have been 0.4 R. This is consistent with a convergent calculation.

DISCUSSION

In discussing these results, there are two main concerns: the empirical conformational energy method employed and the application of the Monte Carlo method.

A. Empirical Conformational Energy Method

Angle Tau_1 in GalNAc-OMe is strongly oriented to values between $-160°$ and $-170°$. This angular range is consistent with the

expectation that O-methyl glycosides should exhibit an exo-anomeric effect (Jeffrey, et al., 1974) with the methyl groups taking a gauche orientation with respect to the ring oxygen. The most likely angle in this calculation was found to have been approximately 15°-20° more extended than a strictly gauche (staggered) conformation. The steric interactions present in GalNAc-OMe because of the substitution of a large acetamido group at C_2 may have introduced some complications into the torsional potential. However, the experimental evidence quoted by Jeffrey, et al. (1974) suggests that the potential function used for the O-methyl glycoside may have over-emphasized the non-bonded interactions producing staggered conformations. The experimental angles are much closer to 180° (using the present angle definitions), the angle associated with an unmodified exo-anomeric effect.

β-1,2-DAG is a more complicated problem than GalNAc-OMe. In a recently obtained X-ray structure of β-1,2-DAG, the dihedral angles Tau_1 and Tau_2 were found to have been 122.2° and 116.6°, respectively. These angles are nearly trans (Bush, Blumberg, Brown).

Based on the exo-anomeric effect, Tau_1 would be expected to take angles of −60° or 180°. Both of these angles are predicted in these calculations. The disagreement between the predicted angles and the experimental value of 120° indicates that an improved torsional potential function for the acetamido group attached to the anomeric carbon is required. Since the empirical potential energy method had been originally parameterized to reproduce amino acid and peptide conformations, its lack of ability to handle sugars, especially N-linked sugars with an N-acetal anomeric linkage, is not surprising.

Calculations in which O-acetyl substitution at the anomeric carbon had been modelled using small molecules (Jeffrey and Yates, 1980) showed that a non-staggered orientation of the acetyl group is a secondary minimum and that it may be the only minimum in a complete pyranose molecule. If the O-acetyl substituent is an appropriate analogy for the N-acetamido group, then the need for an improved torsional potential may be more clearly seen.

B. Application of the Monte Carlo Method

The number of conformations used in these calculations is much smaller than the number of conformations usually considered in a liquid modelling problem—300,000 to 500,000 (Jorgenson and Ibrahim, 1981). In spite of the many points of similarity between the two problems, they differ significantly. The primary difference is the much flatter potential surface of the liquids. The potential energy surface for a simple sugar has a number of barriers separating isolated regions in which energetically acceptable conformations may be found. The problem is that the calculation could become "trapped" in one of these acceptable regions and find it difficult to move to other points on the potential energy surface. Few new acceptable conformations will be generated, and the acceptance rate will be low as a result. This problem would become even more serious in modelling a simple glycopeptide in which more degrees of freedom are involved and the opportunities for steric hindrance are multiplied.

LITERATURE CITED

Aubert, J. P., G. Biserte, and M. N. Louchex-Lefebvre. 1976. Arch. Biochem. Biophys. **175,** 410.

Beeley, J. G. 1977. Biochem. Biophys. Res. Commun. **76,** 1051.

Bush, C. A. and A. J. Duben. 1978. J. Amer. Chem. Soc. **100,** 4987.

Bush, C. A., A. J. Duben, and S. Ralapati. 1980. Biochem. **19,** 501.

Bush, C. A., Biopolymers, in press.

Bush, C. A., K. Blumberg, and J. N. Brown, to be published.

Jeffrey, G. A., J. A. Pople, and L. Radom. 1974. Carbohyd. Res. **38,** 81.

Jeffrey, G. A. and J. H. Yates. 1980. Carbohyd. Res. **79,** 155.

Jorgenson, W. L. and M. Ibrahim. 1981. J. Amer. Chem. Soc. **103,** 3976.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. J. Chem. Phys. **21,** 1087.

Momany, F. A., L. M. Carruthers, R. F. McGuire, and H. A. Scheraga. 1974A, J. Phys. Chem. **78,** 1595.

Momany, F. A., R. F. McGuire, A. W. Burgess, and H. A. Scheraga. 1974B, J. Phys. Chem. **78,** 2361.

Sharon, N. 1975. **Complex Carbohydrates: Their Chemistry, Biosynthesis and Functions.** Addison-Wesley, Reading, Mass.

Valleau, J. P. and S. G. Whittington. 1977. **Statistical Mechanics. Part A: Equilibrium Techniques.** (B. J. Berne, ed.) Plenum Press, New York.

Wood, W. W. 1968. **Physics of Simple Liquids** (H. V. N. Temperly, G. S. Rushbrooke, eds.) North Holland Pub. Co. Amsterdam, Chap. 5.