A Conformational Study of Linkage Positions in Oligosaccharides Investigated by 2-D NMR Spectroscopy and Molecular Modeling

Eunsun Yoo Yoon

School of Natural Science, Honam University, Gwangju 506-714, Korea
Received November 12, 2002

The conformation of synthetic oligosaccharide can be elucidated by employing molecular modeling and high-field proton NMR (nuclear magnetic resonance) spectroscopy. Information with respect to the composition and configuration of saccharide residues and the sequence and linkage positions of the oligosaccharide can be obtained by employing a variety of one- and two-dimensional NMR techniques and molecular modeling. These techniques are also useful in establishing the solution conformation of the oligosaccharide moiety. This study is focused on the elucidation of linkage positions of synthetic trisaccharides, Gal(β1-4)Glc(β1-3)Glc, Gal(β1-4)Glc(β1-4)Glc and Gal(β1-4)Glc(β1-6)Glc.

Key Words : Linkage, Oligosaccharide, Conformation, 2-D NMR, Molecular modeling

Introduction

The growing interest in the biological function of oligosaccharides has stimulated the constant search for new methods for analyzing their primary and secondary structures.[1,2] Linkage position among other structural parameters of carbohydrate may be the most sensitive to variance in vibrational freedom, particularly rotation, due to the different steric hindrance and degrees of freedom of motion between sugars.

NMR study combined with molecular modeling has become an increasingly important tool for the chemical and physical characterization of carbohydrates and their derivatives.[3-6] The usefulness of the NMR method follows from the assignment of individual proton resonances to particular hydrogens. The ability to make these assignments has been enormously facilitated by the introduction of 2-D NMR techniques. A variety of 2-D ^{1}H experiments and especially homonuclear chemical shift correlation spectroscopy (COSY), has aided signal assignments in ^{1}H-spectra. The 2-D NOE (nuclear overhauser effect) experiment and the COSY experiment were also used for determining oligosaccharide interglycosidic linkages. The structure of different linkages can be defined in terms of NMR parameters of their structural reporter groups. The relative intensities of the signals in the NMR spectrum can be used as marker for the purity of the compound. The analysis of reducing oligosaccharides showed that the anomeric configuration of the reducing end sugar also exerts its influence on the spectral parameters of residues in its spatial neighborhood, being sometimes even the non-reducing end sugars.

The computer program used herein, modified MM2, is one of the molecular mechanics programs that optimize the atomic coordinates of a molecule to produce a structure at a local minimum on a multidimensional hypersurface of potential energy. It includes potentials for bond stretching, bending, and stretch-bending, 3-fold tortional potentials, Van der Waals interactions, and dipole-dipole interactions. A set of oligosaccharides differing only in linkage position, Gal(β1-4)Glc(β1-3)Glc, Gal(β1-4)Glc(β1-4)Glc and Gal(β1-4)Glc(β1-6)Glc, was prepared by enzymatic synthesis.[7] These sets of compounds would be useful for development of structural analysis techniques geared to resolve linkage type of saccharides. Compounds like these are parts of core structures for glycolipids.[1] In this study, a simple and systematic method is described to obtain the structural informations from the protons of linkage position in the synthetic oligosaccharides.

Experimental Section

Materials. Oligosaccharides were synthesized as described previously.[7] Laminaribiose, gentiobiose, UDP-galactose, α-lactoalbumin and lactose synthase were purchased from Sigma Chemical Co. All other chemicals were reagent grade quality.

NMR. ^{1}H NMR spectra of synthetic trisaccharides were recorded on a Bruker AM-400 equipped with an Aspect 3000 computer. Approximately 700 μg of the sample was used for the NMR studies. Prior to the NMR experiment, the sample was dissolved in D_2O (99.9 atom% from Sigma Chemical Co.), and lyophilized to remove exchangable hydroxyl protons. This procedure was repeated three times, at the last time, the sample was placed in the NMR tube. The sample was dissolved in 99.96 atom% D_2O from MSD Isotopes to give a final concentration of about 31 mM. Both one- and two-dimensional NMR experiments were performed at 298 K. Higher temperature experiments (323 K) were performed as needed, to shift the solvent (HDO) peak away from one of the anomeric resonances. Acetone at 2.225 ppm was used as an internal reference, indirectly referenced to 2,2-dimethyl-2-silapentane-5-sulfonic acid. The homonuclear chemical shift correlated spectrometry (COSY) is obtained

*Corresponding Author. Phone: +82-62-940-5564; Fax: +82-62-940-5206, E-mail: yooeun@honam.ac.kr

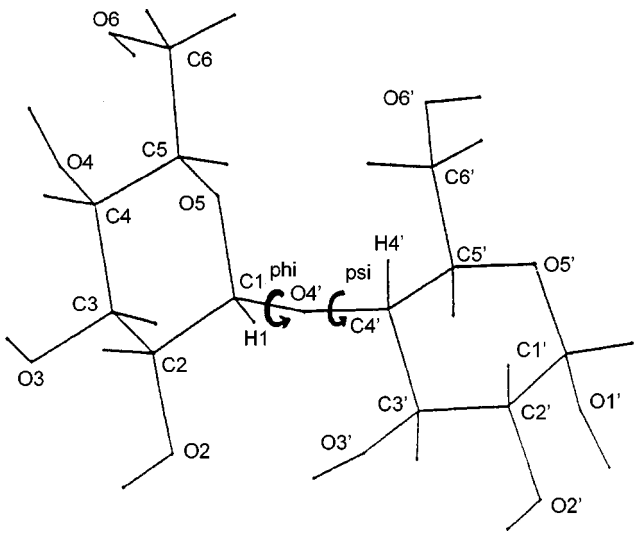

Figure 1. The constitution of disaccharide, Glc(β1-4)Glc, giving atom numbering and the torsional angles phi and psi. Phi and psi are defined by atoms H1-C1-O4'-C4' and H4'-C4'-O4'-C1, respectively.

using 2048 data points in the t_2-domain and a spectral width of 2200 Hz. In the t_1 domain, 1024 data points were acquired and zero-filled to 2048 points before Fourier transformation to give a digital resolution of about 2.2 Hz/point.

Nuclear Overhauser experiments were performed as described by Bax A. *et al.*[8] with a three second presaturation pulse and a one second delay between successive pulses. The 2-D NOESY experiment was performed as described by Kumar A. *et al.*[9] The spectral width was 1200 Hz. 512 data points were used in each value of t_1. The mixing time was 300 msec. The total accumulation time was 6 hours.

Molecular Modeling. The constitution of disaccharide, Glcβ1-4Glc(glucoseβ1-4glucose), giving atom numbering and the tortional angles, phi and psi, are shown in Figure 1. Phi and psi are defined by atoms H1-C1-O4-C4 and H4'-C4'-O4'-C1' of the second and third saccharides of each synthetic trisaccharide, respectively. The definition of a tortional angle follows the IUPAC convention.[10]

Molecular calculations were performed on a DEC MicroVax 3500 or IBM-3090 using SYBYL and Alchemy2000 (Tripos Associates Inc., 1998) and modified MM2 software. Energy contour maps were made with TOPO and SURF programs from the SURFER package (Golden Software, Inc., Golden Co.). The first step in the modified MM2 calculations was determination of the interatomic distances, bond angles and torsional angles in the starting geometry made by SYBYL and Alchemy 2000 programs. The values obtained were used in the different potential function expressions to calculate an initial steric energy, which was simply the sum of various potential energies calculated for all bonds, bond angles, torsional angles, nonbonded pairs of atoms and so forth in the molecule. The modified MM2 program uses a block diagonal Newton-Raphson optimization. Once the optimization had converged, the program printed the final steric energy and optimized geometry. Calculation of minimized energies and optimization of geometries were repeated at each 20° increment of phi and psi torsional angles from -180° to 160°. At each 20° increment of torsion angles, the energy is minimized, providing a value for a point on the energy map.

Results and Discussion

Using lactose synthase, the product trisaccharides were Gal(β1-4)Glc(β1-3)Glc, Gal(β1-4)Glc(β1-4)Glc and Gal (β1-4)Glc(β1-6)Glc. These trisaccharides have the same molecular weight, the only difference being linkage position between the second glucose and the third glucose.

The GalNAc(β1-3)Gal, Gal(α1-4)Gal and Gal(β1-4)Glc disaccharide fragments have very similar conformational behavior in globotetraosyl ceramide and Forssman antigen.[1] It was suggested that the differences observed at the oligosaccharide level are probably due to the cumulative effect of minor differences at the individual disaccharide linkages.[7] The structural details of the synthetic trisaccharides were established by NMR. Significantly, both reducing end anomers of the compounds gave good signal/noise ratios in both the one- (1D) and two-dimensional (2D) spectra. One of the anomers (the α) was only 30-40% of the total concentration and each anomer has its own independent set of resonances. The anomeric configuration and the number

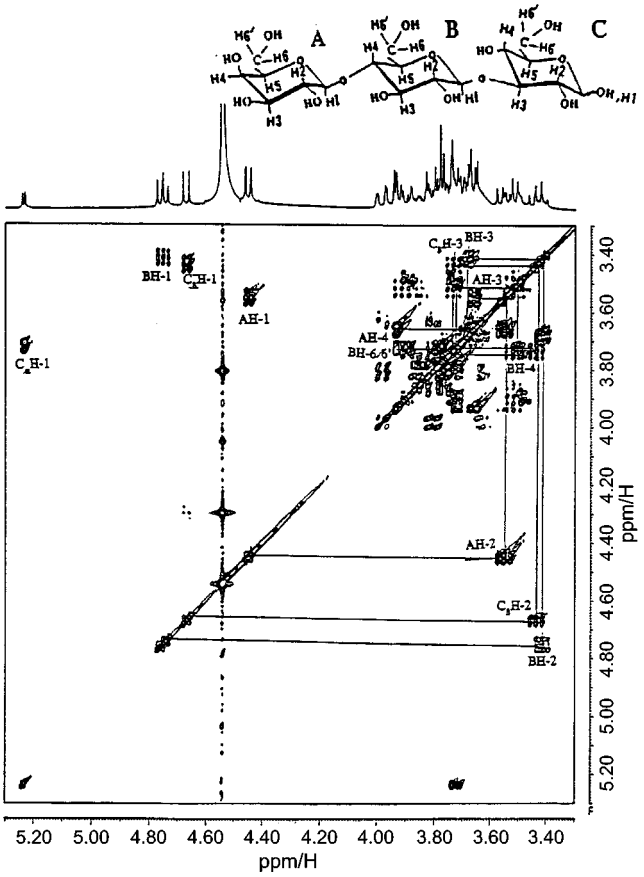

Figure 2. (A) Expansion of the 1-D ^{1}H NMR spectrum and (B) 2-D COSY ^{1}H NMR spectrum of Gal(β1-4)Glc(β1-3)Glc at 323 K.

Table 1. ^{1}H chemical shift (ppm) and vicinal ^{1}H coupling constants ($J_{1,2}$ Hz) of synthetic trisaccharides, Gal(β1-4)Glc(β1-3)Glc and Gal(β1-4)Glc(β1-6)Glc

		Gal(β1-4) Glc(β1-3)Glc	Gal(β1-4) Glc(β1-6)Glc
chemical shift (ppm)	galactose		
	H-1	4.45	4.45
	H-2	3.55	3.55
	H-3	3.75	3.75
	H-4	3.90	3.90
	internal glucose		
	H-1	4.75	4.53
	H-2	3.42	3.39
	H-3	3.71	3.67
	H-4	3.51	3.51
	H-5	3.72	3.86
	H-6/6'	3.93	3.96
	glucose		
	H-1	4.66(β)/5.23(α)	4.65(β)/5.23(α)
	H-2	3.43	3.25
	H-3	3.74	3.50
coupling constant (Hz)	galactose	7.78	7.78
	internal glucose	8.19	7.76
	glucose (β)	8.00	7.96
	glucose (α)	3.72	3.74

of sugar residues were determined by measuring coupling constants and integral intensities of H-1 doublets. The 2-D NOE experiment and the COSY experiment were used for determining oligosaccharide interglycosodic linkages. Both of these approaches to establishing the interglycosidic linkage rely on proton NMR, and are often complicated because of several spectral overlap. Additionally, the effects on which those methods are based, NOE effect across the glycosidic linkage and the $^4J_{HCOCH}$ scalar coupling, are strongly conformation dependent and therefore not always unambiguous.

Figure 2 shows the 2-D COSY spectrum of Gal(β1-4)Glc(β1-3)Glc at 323 K. The 2-D COSY spectrum provides diagnostic evidence on ^{1}H couplings and with expanded spectra, confirms J values. In the expansion of the 1-D spectrum on the top of its 2-D COSY spectrum, the H-1 peak of internal glucose of the trisaccharide apparently resonates as triplelet instead of doublet. However, it is because the two forms of the reducing end glucose in this trisaccharide influences the H-1 of the adjacent internal glucose. Hence, the H-1 of internal glucose that apparently resonates as a triplet, is actually an overlapping doublet. Interestingly, this happens only in the 1-3 and 1-6 linked compounds. From the 1-D NMR spectrum, the synthetic oligosaccharide was confirmed as a trisaccharide. Table 1 was the summary of the chemical shifts of structural reporter group protons of constituent monosaccharides of synthetic Gal(β1-4)Glc(β1-3)Glc and Gal(β1-4)Glc(β1-6)Glc. The anomeric protons were at δ 4.45, 4.75, 4.66 and 5.23 ppm with coupling constants of 7.78 8.19, 8.00 and 3.72 Hz, respectively (Table 1).

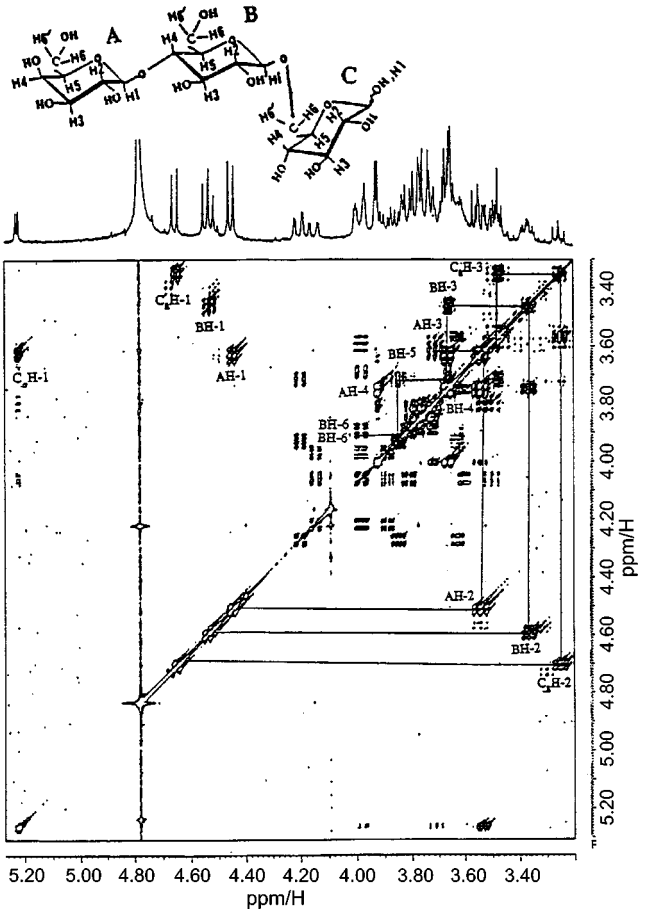

Figure 3. Expansion of the 1-D ^{1}H NMR spectrum (top) and 2-D COSY ^{1}H NMR spectrum of Gal(β1-4)Glc(β1-6)Glc at 298 K.

From coupling constant information,[4,11] peak at δ 4.75 and peak at δ 4.66 ppm are assigned to glucose and peak at δ 4.45 ppm to galactose. Usually, integration rate of protons from monosaccharides is about 0.7-1.0.[11] According to the integration rate, the peak at δ 4.45 ppm was galactose H-1, peak at δ 4.75 ppm was assigned to the penultimate glucose and peak at δ 4.66 ppm and the small α-doublet (δ 5.23 ppm) together were assigned to the reducing end glucose. From the 2-D COSY spectrum, each proton of the internal glucose and of the non-reducing terminal galactose was assigned by cross peaks starting from assignments of the anomeric protons. For Gal(β1-4)Glc(β1-3)Glc, begining at 4.45 ppm with a galactose H-1 resonance, and lowering the contour level somewhat it is possible to find a weak cross peak showing a connectivity to the galactose H-2 resonance at 3.55 ppm. That this connectivity is weak, is a result of a small $J_{1,2}$ coupling constant. From galactose H-2 resonance a connectivity to the galactose H-3 resonance at 3.75 ppm can be found and, likewise, from galactose H-3 to the galactose H-4 resonance at 3.90 ppm. Thus, even in this rather difficult case, it is possible to assign resonances at 4.45, 3.55, 3.75 and 3.90 ppm, to the galactose 1, 2, 3, and 4 protons of non-reducing terminal galactose residue.

By comparing several reference spectra,[8,11,12] the chemical shifts of all protons of Gal(β1-4)Glc(β1-3)Glc (Figure 2)

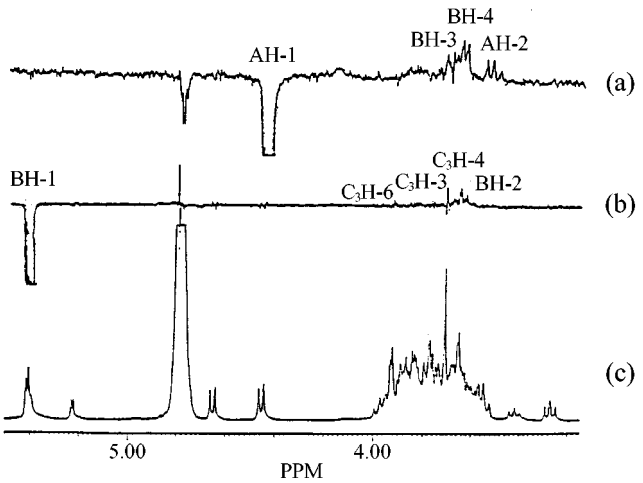

Figure 4. 1-D NOE ^{1}H NMR spectrum of Gal(β1-4)Glc(β1-6)Glc at 298 K, (a) pre-irradiation of AH-1, (b) pre-irradiation of BH-1 and (c) pre-irradiation of CH-1.

were slightly changed except the internal glucose H-4 which shows a much larger shift (3.67 ppm → 3.51 ppm). The chemical shift of the internal glucose H-4 (δ 3.51 ppm) of Gal(β1-4)Glc(β1-3)Glc shifted towards lower field than an earlier reference spectrum of glucose H-4 (δ 3.67 ppm). It shows a glycosylation shift which was the diagnostic signal for the Gal(β1-4)Glc linkage position. Figure 3 is the 2-D COSY spectrum of Gal(β1-4)Glc(β1-6)Glc. The chemical shift of the middle glucose H-4 (δ 3.51 ppm) of Gal(β1-4)Glc(β1-6)Glc also shifted towards lower field than the reference spectrum. Each internal glucose H-4 of Gal(β1-4)Glc(β1-3)Glc and Gal(β1-4)Glc(β1-6)Glc shows a glycosylation shift which was the diagnostic signal for the Gal(β1-4)Glc linkage position.

Figure 4 is the 1-D NOE ^{1}H-NMR spectrum of Gal(β1-4)Glc(β1-6)Glc. In this figure, preirradiation of galactose H-1 (δ 4.45 ppm), generates three NOE difference signals which are the interresidue signals of internal glucose H-4 (δ 3.51 ppm) and of internal glucose H-3 (δ 3.71 ppm) and an intraresidue signal from galactose H-2 (δ 3.42 ppm). The largest NOE signal at δ 3.51 ppm among others indicates the 1-4 linkage between galactose and internal glucose (Table 2). The interresidue signal of glucose H-3 (δ 3.71 ppm) and an intraresidue signal from galactose H-2 (δ 3.42 ppm) appeared because of proximity. The nuclear overhauser effect characterizes conformations by probing H-H distances in the molecule. Since the H-H distances are present in the formula to the 6th power, only protons that are close will affect the response. Hence, 2-D NOESY experiments are required to elucidate the conformation around the interglycosidic bonds. The 2-D NOESY spectrum of Gal(β1-4)Glc(β1-3)Glc is shown in Figure 5. The H-1 peak of galactose (δ 4.45 ppm) is dipole-dipole coupled with H-4 peak of internal glucose (δ 3.64 ppm). This intense signal indicates the 1-4 linkage position between galactose and internal glucose. By NMR study, the synthetic oligosaccharides are proven to trisaccharides which has the specific linkage positions. The synthetic oligosaccharide, Gal(β1-4)Glc(β1-3)Glc and Gal(β1-4)Glc(β1-6)Glc may be interesting acceptor for sialyl α 2-3 transferase enzymes as core structures for neutral sugar residue attached to membrane ceramide. These oligosaccharide structures are a part of ceramides such as lactosyl ceramide, globotriaosyl ceramide, globotetraosyl ceramide, Forssman antigen and asialoGM-1.[1,2,13,14] They are embedded in the lipid bilayer while the sugar is located on the outer surface of the cell membrane. The sugar acts as a specific receptor for several glycoproteins that regulate important physiological functions.

Table 2. Interresidue and Intraresidue NOE contacts of synthetic trisaccharides, Gal(β1-4)Glc(β1-3)Glc and Gal(β1-4)Glc(β1-6)Glc

	Gal(β1-4)Glc(β1-3)Glc	Gal(β1-4)Glc(β1-6)Glc
interresidue contact	AH-1/BH-4 AH-1/BH-3 BH-1/C$_\beta$H-3	AH-1/BH-4 AH-1/BH-3 BH-1/C$_\beta$H-6 BH-1/C$_\beta$H-6'
intraresidue contact	AH-1/AH-2 BH-1/C$_\beta$H-4 BH-1/BH-2	AH-1/AH-2 BH-1/BH-2 BH-1/C$_\beta$H-3 BH-1/C$_\beta$H-4

The computer program used herein, modified MM2, is one of the molecular mechanics programs that optimize the

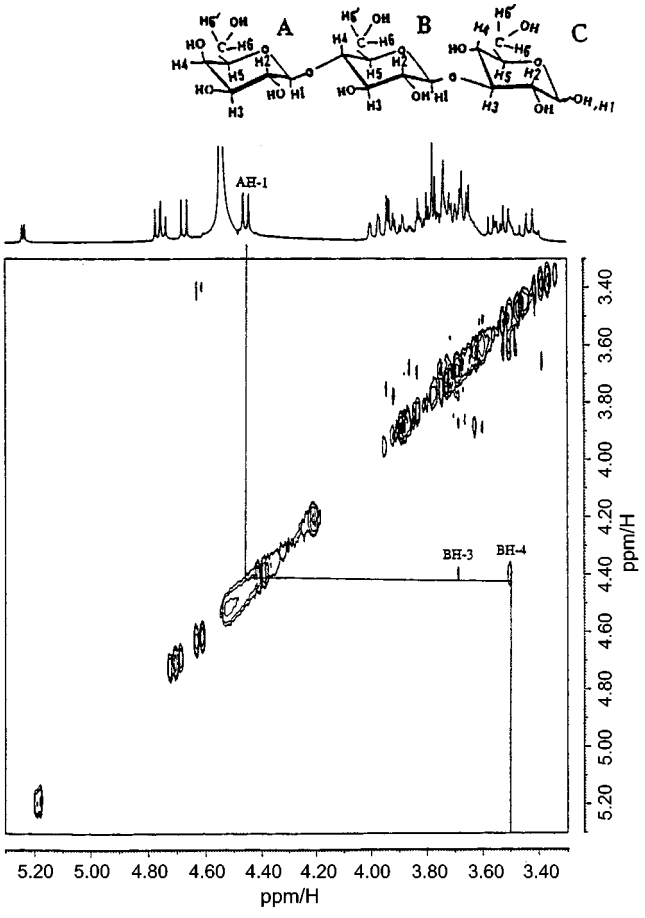

Figure 5. 2-D NOESY ^{1}H NMR spectrum of Gal(β1-4)Glc(β1-3)Glc at 323 K.

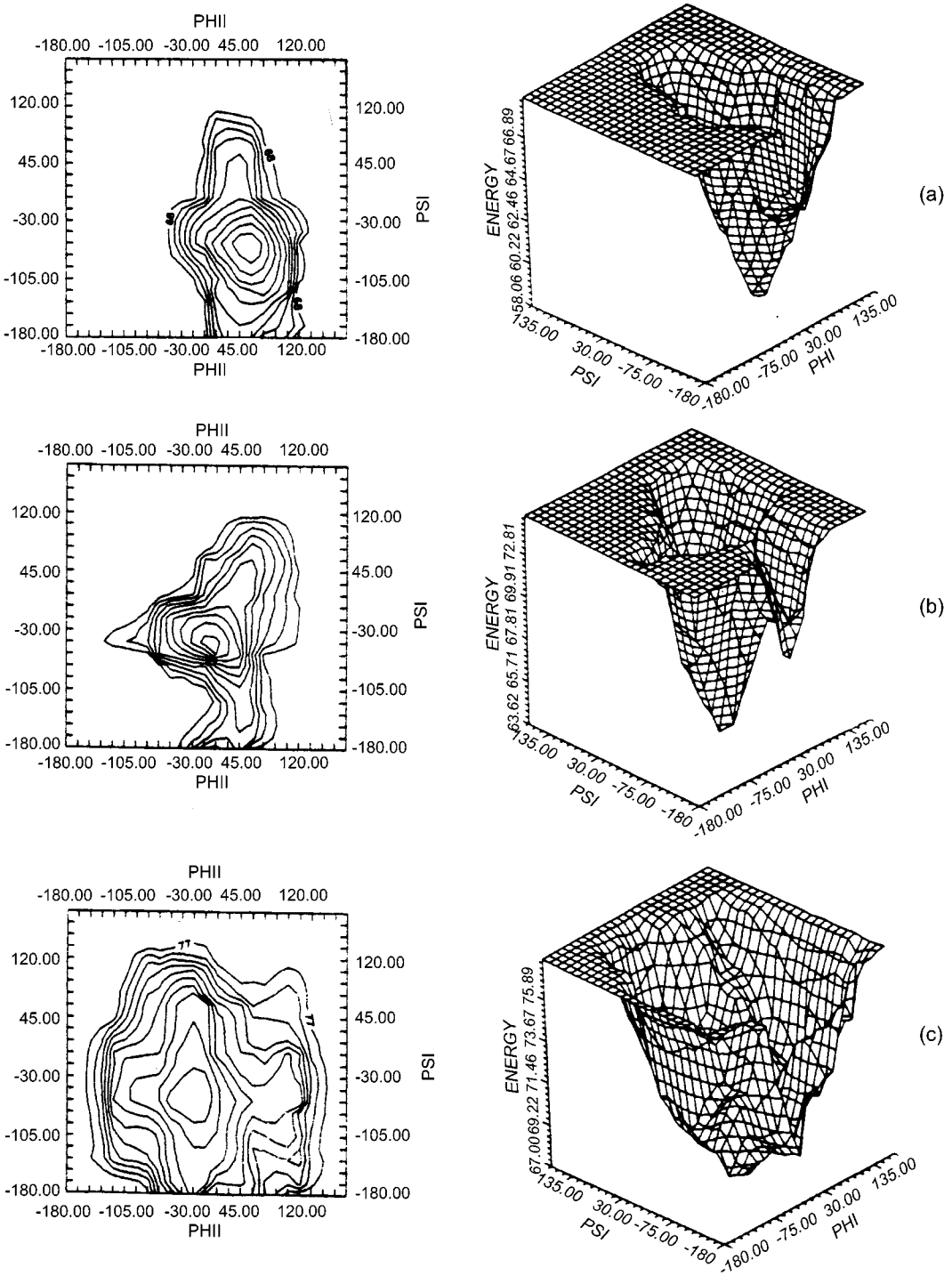

Figure 6. The phi and psi plots of the total energies and energy wells derived from the modified MM2 calculation. The drawings were made with the SURF program for SURFER from Golden software. (a) Gal(β1-4)Glc(β1-3)Glc, (b) Gal(β1-4)Glc(β1-4)Glc and (c) Gal(β1-4)Glc(β1-3)Glc.

atomic coordinates of a molecule to produce a structure at a local minimum on a multidimensional hypersurface of potential energy.[1,15] The MM2 molecular modeling programs was modified for use on an IBM3090 and were also used on DEC Microvax equipment to calculate minimum energy structures and freedom of motion volumes near the minima.[3,16] It includes potentials for bond stretching, bending, and stretch-bending, 3-fold tortional potentials, Van der Waals interactions, and dipole-dipole interactions. The modified MM2 program was chosen for this work as following reasons: first, carbohydrate has a number of possible ring conformers such as $^4C_1$, $^1C_4$, $^1S_5$, etc.[14] The MM2 program can alter the ring geometry to a low energy form for a particular saccharide shape. Secondly, the MM2 version, automatically provides the anomeric effects that are important for sugars. Also, the lone pairs are treated as if

they are atoms because modified MM2 requires lone pairs of electrons on all ether and hydroxyl oxygen atoms to fit the data on alcohols and ethers which are major components of carbohydrate. Finally, this version has "dihedral driver" facility which accepts the initial, final, and increment size values of two tortional angles and energy-minimization at each increment of these torsion angles. It permits conformational analysis of a disaccharide by rotating of glycosidic bond and was further modified to give a rigid dihedral driver option that starts with the same intra-residue geometry at each increment of the driven torsion angles. This avoids the propagation of residue distortions from one conformation to the next. The maps for the α-anomers are on the left and those for β-anomers are on right. Contour lines are graduated in 1 kcal/mol increments above the global minimum.

Molecular modeling of the oligosaccharides supported the relative stability of each glycosidic linkage by examining the degree of rotational freedom around the isomeric linkage. Figure 6 shows 2-D energy maps and 3-D energy wells derived from the modified MM2 calculations on neutral, uncharged molecules which depict degrees of phi-psi rotational freedom. The Gal(β1-4)Glc(β1-3)Glc, the most rigid isomer, generates a volume which I will depict as 1.0, while the Gal(β1-4)Glc(β1-4)Glc generates an intermediate volume of 1.2. The Gal(β1-4)Glc(β1-6)Glc, being the most flexible with its three rotational bonds generates a much larger well (1.8). Comparing the coupling constants (Table 1) of internal glucose between Gal(β1-4)Glc(β1-3)Glc and Gal(β1-4)Glc(β1-6)Glc, the most rigid isomer (Gal(β1-4)Glc(β1-3)Glc) has 7.76 Hz, whereas the most flexible one (Gal(β1-4)Glc(β1-6)Glc) has 8.19 Hz. Calculated linkages from conformations whose energies were based on the modified MM2 parameters gave results in good accord with the conformational properties of oligosaccharides in solution as reflected from NMR data. We are currently studying the dynamic kinetics of the linkage-isomeric oligosaccharides.

The trisaccharides should be linear, not stacked underneath the second or third glc rings, because there is no intra-NOE signals from hydrogen bonding between the second or third glc hydroxyl protons. Therefore synthetic trisaccharides, Gal(β1-4)Glc(β1-3)Glc and Gal(β1-4)Glc(β1-6)Glc, have linear structures, not folded structures. The results derived from these techniques are complementary and are integrated with each other to arrive at the final model. A fine example is the use of interatomic distance information gathered from NMR studies in force field calculation to determine the solution conformation.

Acknowdgement. This work was supported by the Korea Science and Engineering Foundation (R04-2000-00044).

References

1. Rao, V. S. R.; Qasba, P. K.; Balaji, P. V.; Chandrasekaran, R. *Conformation of Carbohydrates*; Harwood Academic Publishers: Amsterdam, Netherlands, 1998; pp 29-42, pp 131-182.
2. Kobata, A. *Acc. Chem. Res.* **1993**, *26*, 319.
3. French, A. D.; Mouhous-Riou, N.; Perez, S. *Carbohyd. Res.* **1993**, *247*, 51.
4. Krishna, N. R.; Choe, B.; Prabhakaran, M.; Ekborg, G. C.; Roden, L.; Harvey, S. C. *J. Biol. Chem.* **1990**, *265*, 18256.
5. Lee, K.; Kim, Y. *Bull. Korean Chem. Soc.* **1996**, *17*, 118.
6. Shim, G.; Lee, S.; Kim, Y. *Bull. Korean Chem. Soc.* **1997**, *18*, 415.
7. Yoon, E. Y.; Laine, R. A. *Glycobiology* **1992**, *2*, 161.
8. Bax, A.; Eagan, W.; Kovac, P. *J. Carbohyd. Chem.* **1984**, *3*, 593.
9. Kumar, A.; Ernst, R. R.; Wuthrich, K. *Biochim. Biophys. Acta* **1980**, *95*, 1.
10. IUPAC Tentative Rules for the Nomenclature of Organic Chemistry, *Eur. J. Biochem.* **1971**, *18*, 151.
11. Gamlan, A.; Ramanowska, E.; Dabrowski, U.; Dabrowski, J. *Biochemistry* **1991**, *30*, 5032.
12. Egge, H.; Dabrowski, J.; Hanfland, P. *Pure Appl. Chem.* **1984**, *56*, 807.
13. Kannagi, R.; Fukushi, Y.; Tachikawa, T.; Noda, A.; Shin, S.; Shigeta, K.; Hiraiwa, K.; Fukuda, N.; Inamoto, T.; Hakomori, S.; Imura, H. *Cancer Res.* **1986**, *46*, 2619.
14. Lowe, J. B.; Stoolman, L. M.; Nair, R. P.; Larsen, R. D.; Berhend, T. L.; Marks, R. M. *Cell* **1990**, *63*, 475.
15. Li, J. H.; Allinger, N. L. *MM2 Program (QCPE 543)*; University of Georgia: 1987.
16. Duben, A. J.; Hricovini, M.; Tvaroska, I. *Carbohyd. Res.* **1993**, *247*, 71.