Glycoconjugate Journal (1998) **15:** 107–113

Three-dimensional structure of a glycosphingolipid having a novel carbohydrate linkage, Galβ1-4(Fucα1-3)Glcβ1-3Galβ, determined by theoretical calculations

Jun Kawakami[1]*, Yasushi Kawakami[2,3], Kazuo Nakamura[3], Hisako Kojima[3], Shoei Ito[1] and Yoichi Tamai[3]

[1] *Department of Chemistry, Faculty of Science, Hirosaki University, Hirosaki, Aomori 036, Japan*
[2] *Department of Environmental Health, Azabu University, Sagamihara, Kanagawa 229, Japan*
[3] *Department of Biochemistry, Kitasato University School of Medicine, Sagamihara, Kanagawa 228, Japan*

The novel glycosphingolipid, SEGLx (Galβ1-4(Fucα1-3)Glcβ1-3GalβCer), which was identified by us (Kawakami Y, *et al.* (1993) *J Biochem* 114: 677–83), shows a characteristic spectrum on ^{1}H-NMR analysis, in which the anomeric proton resonances of a reducing end galactose and a glucose are split. To elucidate the structural characteristics of SEGLx, we determined its three-dimensional (3D) structure by means of computer simulation, involving such techniques as molecular mechanics (MM2), the semiempirical molecular orbital method (AM1), molecular dynamics (Amber), and computer 3D modelling. With the hypothesis that all OH group(s) of a ceramide participate in intramolecular hydrogen bonds, two kinds of stable conformers, horizontal and right-angled ones, were formed, depending on the ceramide species. The present findings suggest that the chemical species of both the long chain base and fatty acid moieties, mainly the occurrence of OH group(s), affect the chemical shifts of the anomeric proton resonances not only of the reducing terminal galactose but also the penultimate glucose through the formation of intramolecular hydrogen bonds. Computer simulation through theoretical calculation and 3D modelling was shown to be the best means of confirming the results obtained by experimental analysis.

Key words: **conformation, glycosphingolipid, theoretical calculation**

Abbreviations: **Cer, ceramide; Fuc, fucose; Gal, galactose; Glc, glucose; Hex, hexose; NMR, nuclear magnetic resonance; ROESY, rotating frame Overhauser effect spectroscopy; SEGLx, Galβ1-4(Fucα1-3)Glcβ1-3Galβ1-Cer (a glycosphingolipid from *S. erinacei*)**

Introduction

Glycosphingolipids are localized on cell surface membranes, their carbohydrate moieties protruding from them [1, 2], and are thought to be involved in various recognition processes, *eg* as blood group, transplantation and tumour antigens, or as receptors for toxins, viruses and bacteria. In view of these important roles played by the glycolipid headgroups, it is considered that an understanding of the structure and dynamics of the carbohydrate moiety will provide valuable insight into the functional significance of specific glycolipids on the membrane surface, at the molecular level. We recently isolated a fucosylated glycosphingolipid with a novel carbohydrate structure (SEGLx, Galβ1-4(Fucα1-3)Glcβ1-3Galβ1Cer) from the parasite, *Spirometra erinacei* [3]. This glycolipid is characterized by: (i) the occurrence of a penultimate glucose molecule attached to the reducing end galactose through a β1-3 linkage; and (ii) the presence of a fucose attached to a glucose through an α1-3 linkage. In addition, ^{1}H-NMR analysis revealed that the ceramide moiety affects the carbohydrate structure, suggesting a unique three-dimensional (3D) structure for this glycolipid. Conformational studies on carbohydrate antigens such as glycosylated proteins, blood group antigens and glycosphingolipids have shown that individual glycoconjugate molecules exhibit characteristic 3D structures [4]. Considering the novel chemical structure of SEGLx and its possible functional significance in parasitism, elucidation of the 3D structure of this glycolipid is of great importance.

In the present study, we investigated the conformation of SEGLx to determine the 3D structure of this molecule by means of computer simulation. First, the carbohydrate part of SEGLx was calculated by means of molecular mechanics (MM2) [5, 6] according to the 3^n algorithm [7], and then

* To whom correspondence should be addressed. Tel: + 81-172-39-3570; Fax: + 81-172-39-3561.

0282–0080

the highest population conformers in the Boltzmann distribution were optimized by means of the semiempirical molecular orbital method (AM1) [8] and molecular dynamics (Amber) [9, 10], including the effects of solvation and temperature. Second, using the carbohydrate part determined by means of theoretical calculations and the ceramide part based on the results of X-ray single crystal analysis of a galactosylceramide [11], the intramolecular hydrogen bond between the carbohydrate part and the ceramide part in SEGLx was determined by computer 3D modelling.

Materials and methods

The isolation and structural determination of SEGLx, and determination of its configuration by means of ^{1}H-NMR were previously reported [3].

Theoretical calculations (MM2, AM1 and Amber) were performed as follows. On a Macintosh Centris 660AV computer, the Chem3D Plus 3.0 energy minimization program was used to perform MM2 force field calculations of the energies associated with several potential conformations of the sugar part of SEGLx. First, three glycopyranoses (β-D-Gal, β-D-Glc and α-L-Fuc) were optimized. Next, 4Glcβ1-3Galβ1, 4(Fucα1-3)Glcβ1-3Galβ1 and Gal1β1-4(Fucα1-3)Glcβ1-3Galβ1 were calculated by means of MM2 according to the 3^n algorithm, which is applicable to torsional energy surfaces of flexible molecules. For example, assuming that two glycopyranoses are optimized, 4Glcβ1-3Galβ1 will contain two main rotating single bonds (φ1 and φ2), each of which is expected to have three staggered rotamers (Figure 1). This means that 9 ($= 3^2$) conformations must be optimized. The population of the ith conformation, pi, was calculated using equation (1), based on the assumption that the distribution of conformations obeys the Boltzmann distribution,

$$\mathrm{p}i = gi\exp(-\mathrm{SE}i/\mathrm{RT})/\Sigma gi\exp(-\mathrm{SE}i/\mathrm{RT}) \qquad (1)$$

where SEi is the calculated conformational energy ($=$ steric energy) for the ith conformation. The highest population conformer in the Boltzmann distribution was optimized by means of AM1, based on the geometries calculated by means of MM2. AM1 calculations with MOPAC ver. 6.0 were performed on an ACOS-930 computer at the Hirosaki University Center for Computer and Communications. The molecule was set at a minimum energy position of 0 K for MM2 and AM1. The geometry calculated by means of AM1 was then optimized by means of Amber, including the effects of solvation and temperature. The calculations were carried out using the Amber calculations in DISCOVER (Biosym Technologies Inc., CA, USA) on a COMTEC 4D computer at the Hirosaki University Center for Computer and Communications. In the calculation of the electrostatic force, the dielectric constant was set at 46.6 for DMSO in order to approximate the electrostatic screening effect of the surrounding medium *in vacuo*. Energy minimization was achieved using the steepest descent method for the initial minimization and the conjugated gradient method for the final minimization. On ^{1}H-NMR analysis, the energy of a molecule in solution

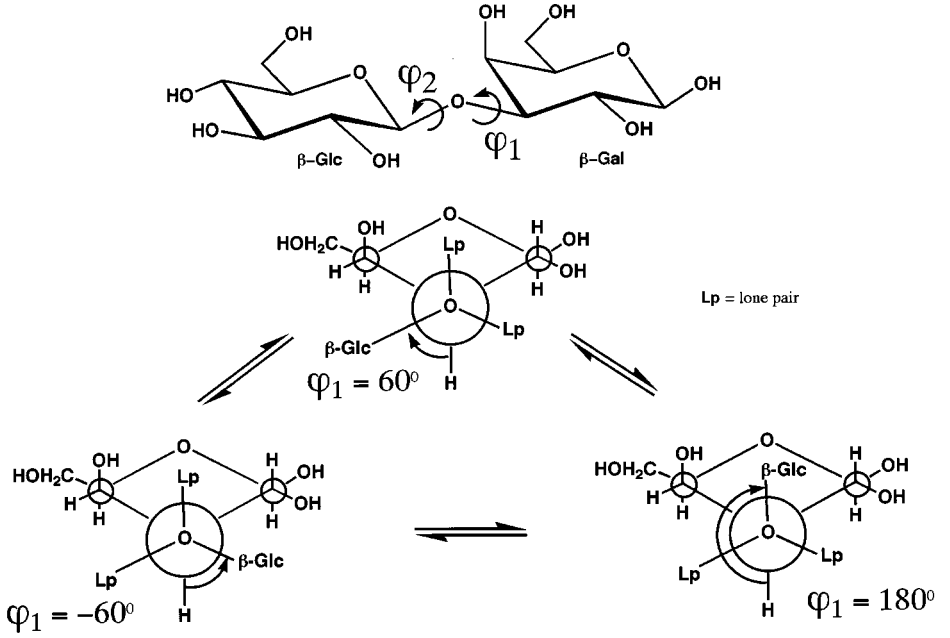

Figure 1. The Newman projection formula for 4Glcβ1-3Galβ1.

dynamically fluctuates at room temperature. In order to determine the stability of the molecule at room temperature, the time course of the molecular motion under thermal fluctuation must be determined. Thus, Amber calculations were carried out with a step size of 1fs and at a constant temperature of 298 K. The molecule was equilibrated for a period of 0.1 ps on simulation. Data analysis was performed for 300 ps. With the hypothesis that all OH group(s) participate in intramolecular hydrogen bonds, SEGLx, as a whole molecule, was investigated by means of computer 3D modelling using the carbohydrate chain geometry determined by theoretical calculations, and the ceramides based on the results of X-ray single crystal analysis of a galactosylceramide [11].

Results and discussion

Conformational analysis of the carbohydrate of SEGLx

SEGLx contains several single bonds per molecule, for each of which three staggered rotameters are expected. We performed MM2 force field calculations to determine the most stable conformation of the carbohydrate part of SEGLx. MM2 is well suited for calculating a carbohydrate chain,

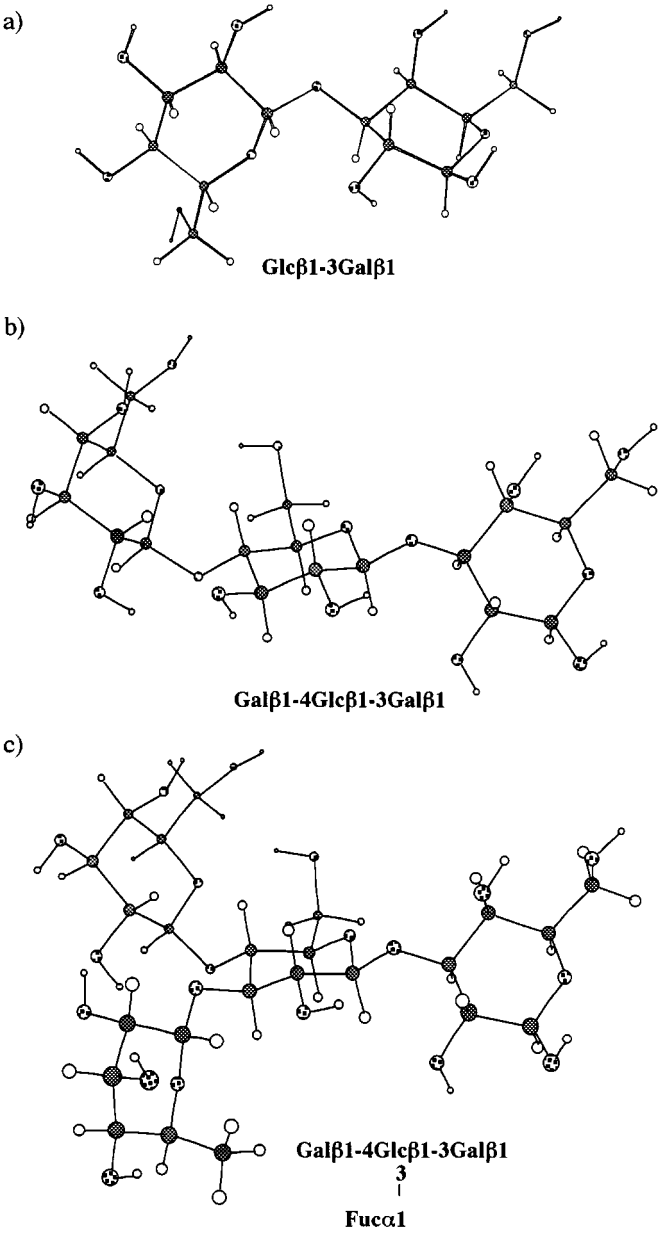

Figure 2. Optimized geometries, which exhibit the lowest steric energies and the highest Boltzman distribution, calculated by means of MM2. a) Glcβ1-3Galβ1, b) Galβ1-4Glcβ1-3Galβ1, and c) Galβ1-4(Fucα1-3)Glcβ1-3Galβ1.

since it reproduces important interactions in the carbohydrate chain, *ie* the anomeric and exoanomeric effects. The optimized geometries, which exhibit the lowest steric energy and the highest Boltzmann distribution, are shown in Figure 2. The population of these conformers, based on the assumption that the distribution of conformations obeys the Boltzmann distribution, was more than 70% in total. These structures showed exoanomer effects. AM1 was used for the highest population conformer of Galβ1-4(Fucα1-3)Glcβ1-3Galβ1. This method is based on quantum mechanics and its confidence level is accepted in general. The optimized geometries obtained by means of AM1 are shown in Figure 3. Two optimized conformers calculated by means of MM2 and AM1 were found to have the same geometry, being low in steric hindrance and stable in structure.

In theoretical calculations it is still questionable whether fairly large systems including solvent molecules around the solute molecules can be calculated, in principle, by a computer simulation method, and indeed, such simulation is still limited because of the limited capability of computers. For the calculation of large molecules such as carbohydrates, therefore, the effects of solvent molecules should be treated with some approximations. In most cases, such effects have been partly included by changing the dielectric constant (ε) as an electrostatic screening effect of the surrounding medium [12–14]. We attempted to calculate the carbohydrate part of SEGLx, including the effects of solvation ($\varepsilon = 46.6$) and temperature (298 K), based on the geometry of the lowest energy conformer in a vacuum obtained by means of AM1. The optimized geometries obtained by means of Amber are shown in Figure 4. The influence of solvation and temperature on the geometry of the sugar chains was observed: the directions of hydroxy groups and the carbon skeleton, and the steric correlation in carbohydrate molecules were changed. The calculated geometry was consistent with the interresidual ROESY spectrum obtained in our previous study [3]. This carbohydrate geometry of SEGLx was considered to be the most stable conformation. The results obtained with the Amber force field are known to fit the structural, energetic and dynamic properties well in calculations of the carbohydrate [15].

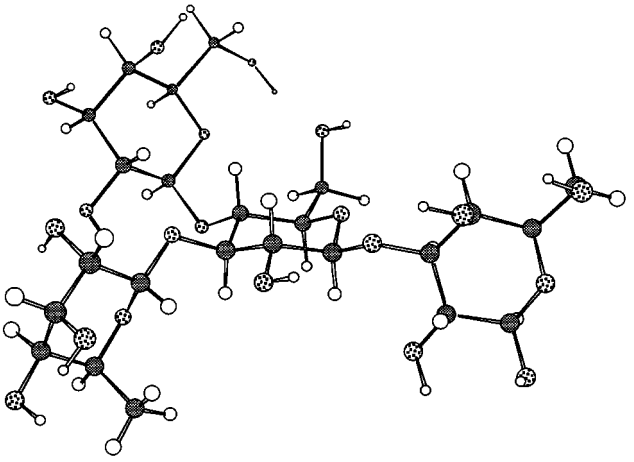

Figure 4. Optimized geometry of the carbohydrate part of SEGLx, calculated by means of Amber at $\varepsilon = 46.6$ and 298 K.

The 3D structure of SEGLx

The ceramides of SEGLx were found to consist of four molecular species (I, II, II and IV), as shown in Figure 5 [3]. ^{1}H-NMR analysis of SEGLx in DMSO revealed that both the long chain base and fatty acid moieties affect the chemical shifts of the anomeric proton resonances of not only the reducing terminal galacose but also the penultimate glucose (Table 1); the same findings were also obtained for lyso-SEGLx on which the amide-linked fatty acids were liberated [3]. These findings imply that the splitting of anomeric protons observed on ^{1}H-NMR analysis was due to the intramolecular hydrogen bond between the carbohydrate part and the ceramide part. To determine the 3D structure

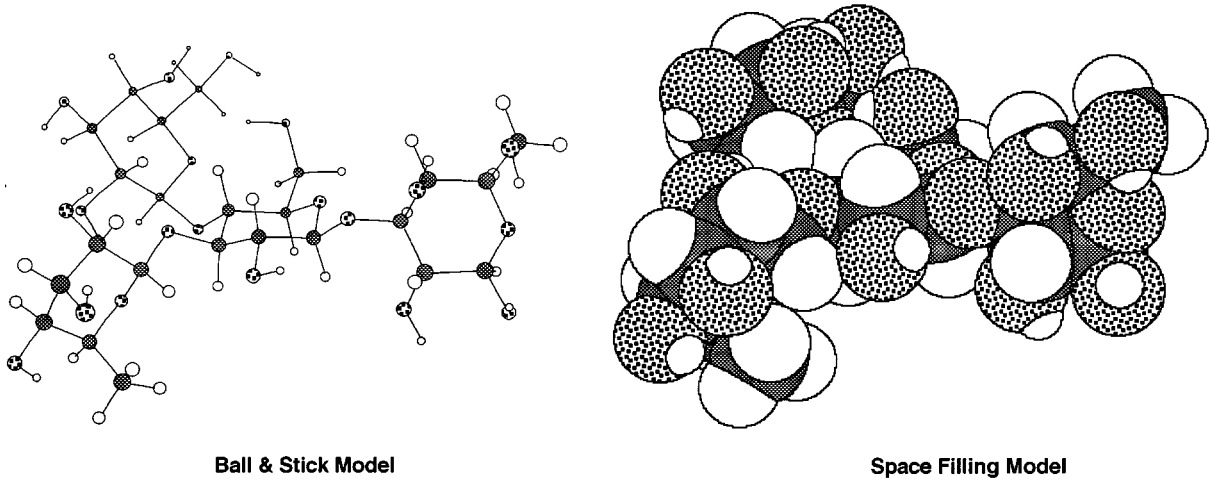

Figure 3. Optimized geometry of the carbohydrate part of SEGLx, calculated by means of AM1.

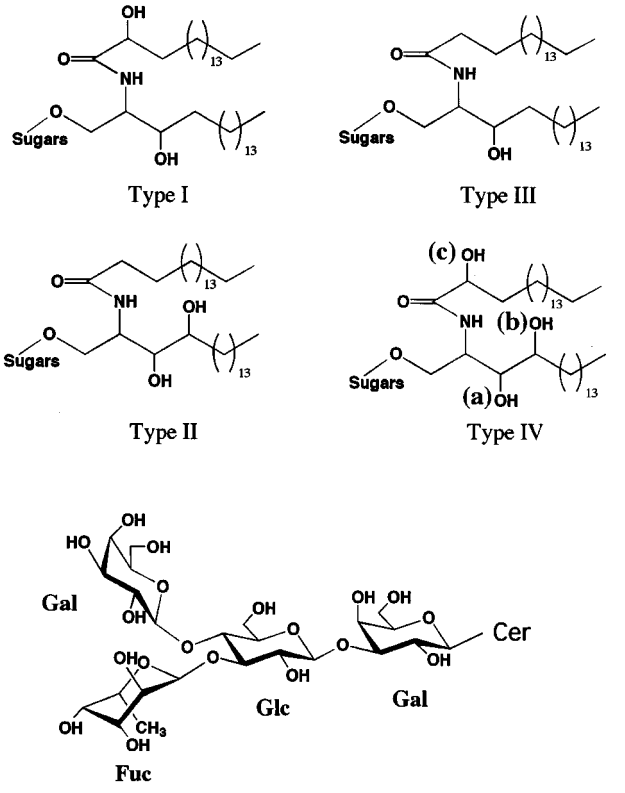

Figure 5. Four ceramide types of SEGLx. The ceramide parts of types I, II, III and IV consist of sphinganine and 2-hydroxystearic acid, 4-D-hydroxysphinganine (phytosphingosine) and stearic acid, sphinganine and stearic acid, and phytosphingosine and 2-hydroxystearic acid, respectively.

Table 1. Chemical shifts of anomeric protons of the reducing end galacose and the penultimate glucose in DMSO $- d_6$

	Gal		*Glc*	
δ/ppm	4.16	4.18	4.49	4.52

The values were calculated from the spectrum presented in [3]

of SEGLx, we performed computer 3D modelling: the carbohydrate part was determined by the theoretical calculations shown above and the ceramide part was based on the results of X-ray single crystal analysis of a galactosylceramide [11]. Using type IV SEGLx having three hydroxy groups in the ceramide part (see Figure 5), we determined the correlation of the anomeric proton of the carbohydrate part with the hydroxy group of the ceramide part. Two conformations, one almost horizontal (3D model I) and the other right-angled (3D model II), were obtained, depending on the ceramide species of SEGLx (Figure 6). Based on the concept that the length of the hydrogen bond between H and OH is about 2.0 Å, in the 3D model I (Figure 6A), two hydrogen bonds could be formed between the anomeric proton of the reducing end galacose and OH (a), and the anomeric proton of the penultimate glucose and OH(a), respectively. It has been reported that glycosphingolipids are able to form a hydrogen bond between the anomeric proton of a galactose and OH (c) of a fatty acid in the case of galactosylceramide [16]. However, interaction between the anomeric proton of the carbohydrate part and OH (a) of the long chain base has not been reported so far. In SEGLx, a stable conformer having two hydrogen bonds was formed. This is probably because a penultimate carbohydrate attaches to a reducing end carbohydrate through a β1-3 linkage, and not a β1-4 linkage which is generally present in glycosphingolipids in nature. Thus, the 3D model I can explain the ^{1}H-NMR spectrum of SEGLx [3], where the ceramide part affects the chemical shifts of the anomeric proton resonances of not only the reducing terminal galactose but also the penultimate glucose. If OH (b) or OH (c) in the ceramide part formed a hydrogen bond with the anomeric proton of the penultimate glucose, a bulky conformation exhibiting steric hindrance would have been formed.

On the other hand, in 3D model II (Figure 6B), two or three hydrogen bonds are present, the stable conformation being formed, *ie* the right-angled structure. One of the two hydrogen bonds is between the anomeric proton of the reducing end galactose and OH (a) and (b), and the other is between the anomeric proton and OH (a) and (c); three two hydrogen bonds are present between the anomeric proton of the reducing end galactose and OH (a), (b) and (c). Thus, the presence of OH (b) and/or (c) causes the formation of two or three additional hydrogen bonds, the right angled conformation being inevitably formed.

Conclusion

The ceramides of SEGLx consist of four molecular species (see Figure 5). The ^{1}H-NMR in DMSO revealed that the chemical species of both the long chain base and fatty acid moieties affect the chemical shifts of the anomeric proton resonances of not only the reducing terminal galactose but also the penultimate glucose, resulting in splitting of the anomeric protons due to the intramolecular hydrogen bond between the carbohydrate part and the ceramide part. To elucidate the structural characteristics of SEGLx, we determined the most favourable conformations that were consistent with the ^{1}H-NMR spectrum, performing molecular simulation of this molecule by means of MM2, AM1, Amber and computer 3D modelling. Two kinds of stable conformers, one almost horizontal (3D model I) and the other right-angled (3D model II), were obtained: SEGLx having only OH (a), type III, formed 3D model I, and SEGLx having OH (a) and OH (b) or OH (c), type II or I, and all three, OH (a), (b) and (c), type IV, formed 3D model II. These 3D structures can be formed only with the presence of a β1-3 linkage between the reducing end galactose and the penultimate glucose, which characterizes the structure of

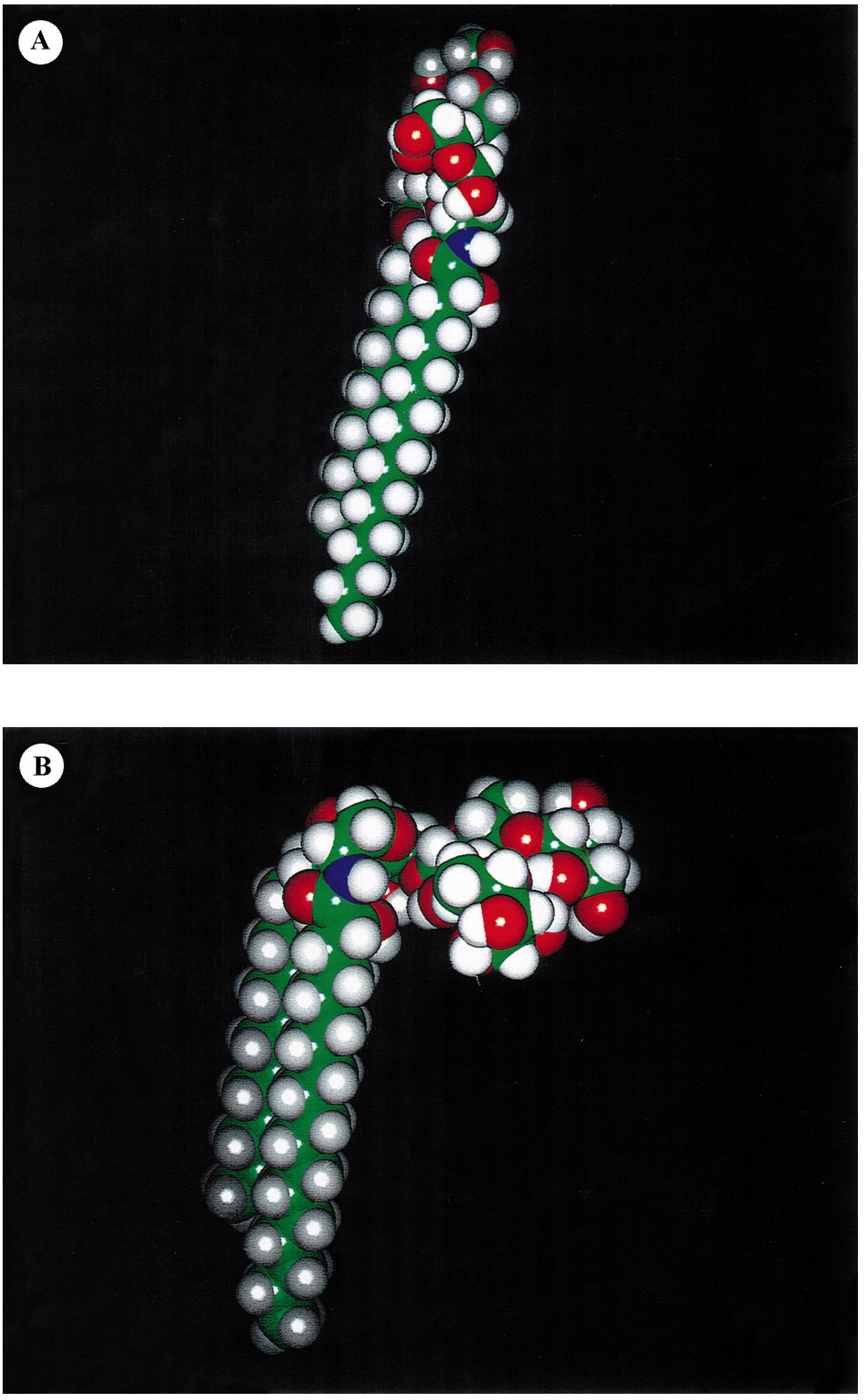

Figure 6. The geometry of SEGLx determined by computer 3D modelling, A, model I; B, model II.

SEGLx. The characteristics of the ^{1}H-NMR spectrum of SEGLx may be explained by the two possible structures.

Furthermore, we wish to emphasize that the molecular simulation performed in this study will be an effective means of determining stable conformations. Elucidation of the 3D structures of glycosphingolipids may lead to an understanding of their functional roles in nature.

References

1 Karlsson KA (1989) *Ann Rev Biochem* **58**: 309–50.
2 Hakomori S (1990) *J Biol Chem* **265**: 18713–16.
3 Kawakami Y, Nakamura K, Kojima H, Suzuki M, Inagaki F, Suzuki A, Sonoki S, Uchida A, Murata Y, Tamai Y (1993) *J Biochem* **114**: 677–83.
4 Imberty A, Mikros E, Koca J, Mollicone R, Oriol R, Perez S (1995) *Glycoconjugate J* **12**: 331–49.
5 Burkert U, Allinger NL (1982) *Molecular Mechanics*. Washington, DC: The American Chemical Society.
6 Clark T (1985) *A Handbook of Computational Chemistry*. New York: John Wiley & Sons, Inc.
7 Osawa E, Goto H, Oishi T, Ohtsuka Y, Chuman T (1989) *Pure Appl Chem* **61**: 597–600.
8 Dewar MJS, Zoebisch EG, Healy EF, Stewart JJP (1985) *J Am Chem Soc* **107**: 3902–9.
9 Weiner SJ, Kollman PA, Case DA, Singh VS, Chio C, Alagon GS, Profeta S, Weiner P (1984) *J Am Chem Soc* **106**: 765–84.
10 Weiner P, Kollman PA, Nguyen DT, Case DA (1986) *J Comp Chem* **7**: 230–52.
11 Nyholm P, Pascher I, Sundell S (1990) *Chem Phys Lipids* **52**: 1–10.
12 Hardy BJ, Sarko A (1993) *J Comp Chem* **14**: 848–57.
13 Dowd MK, Reilly PJ, French AD (1992) *J Comp Chem* **13**: 102–14.
14 Jeffery GA, Taylor R (1980) *J Comp Chem* **1**: 99–109.
15 Yamada H (1997) *J Syn Org Chem* **55**: 29–41.
16 Dabrowski J, Egge H, Hanfland P (1980) *Chem Phys Lipids* **26**: 187–96.

Received 6 November 1996, revised 18 April 1997, accepted 21 May 1997