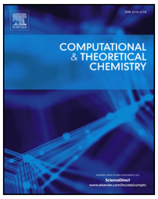
Computational and Theoretical Chemistry 1140 (2018) 24–31

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc



Preferred conformational structures of disaccharides with β-1,4-linked *N*-acetylglucosamine and D-mannose in the gas phase: A tree-step computational approach study

Jianming Gao[a], Dong Chen[a,*], Ran Song[a], Huanlu Xue[a], Tianxiang Wang[a], Bo Liu[a,b,*]

[a] *Physics and Electronics Department, Henan University, Kaifeng 475004, China*
[b] *Lab of Functional Molecules and Materials, School of Physics and Optoelectronic Engineering, Shandong University of Technology, Zibo 255000, China*

ARTICLE INFO

Keywords:
Disaccharides
Tree-step computational approach
IR vibration signature

ABSTRACT

N-glycans are highly crucial to the potential application prospect of glycoprotein. However, its structural diversity and instability severely limit its further research. Here we apply the tree-step computational approach to investigate preferred structures of disaccharides with β-1,4-linked *N*-acetylglucosamine and D-mannose of core pentasaccharide in *N*-glycans. Due to the structural character of disaccharides, hydrogen bonds (H-bonds) are classified according to the upper and lower parts of GlcNAcβ(1,4)GlcNAc and Manβ(1,4)GlcNAc. Almost all building conformers (95 for GlcNAcβ(1,4)GlcNAc and 50 for Manβ(1,4)GlcNAc) are generated depending on the combination of H-bonds in the upper and lower parts, except for one conformer. After establishing initial configurations, theoretical simulations for optimization, preferred structures and frequency calculations are employed. Both energy and IR vibration signature are in good accord with the experimental and theoretical literature. Additionally, five new low-energy structures are found, which have not been reported by using the traditional conformational search.

1. Introduction

Glycoprotein as complex saccharide is formed by covalent attachment of oligosaccharide and polypeptide side-chains. Particularly, oligosaccharide plays an important role in maintaining the biological activities and functions of the substance [1–5]. The core pentasaccharide is common in oligosaccharide in *N*-linked glycoproteins and especially indispensable for *N*-glycosylation of proteins in all eukaryotes [6–9]. As a connection part of the core pentasaccharide, a good understanding of structural information for disaccharides with β-1,4-linked *N*-acetylglucosamine and D-mannose is particularly crucial. In recent years, the conformation of saccharides has been extensively studied [10–17]. α-maltose conformers were investigated by DFT calculations at the B3LYP/6-311++G** level of theory [11]. The hydroxyl group orientations and their effects on the glycosidic dihedral angles were especially presented. NMR spectroscopy and quantum mechanics (QM) calculations [18] were employed to analyze preferred conformation and dynamic behavior of *N*-acetyl and *N*-formyl substituents of 3-amino-3,6-dideoxy-α-D-galactopyranose. The configuration isomer in the gas phase was experimentally determined by the analysis of the infrared (IR) vibration spectra [6,19–22]. The configuration information is primarily assigned to the vibration peak of particular structure in combination with DFT and *ab initio* calculations. With the current rate of increase in computer power, several theoretical techniques have been used to generate starting geometries, such as dihedral angle drive method [14,18,23], molecular dynamics [15,24–27], Monte-Carlo Multiple Minimum (MCMM) conformational searches [12,13,19,20,28,29]. From examination of our previous studies [30–33], it is clear that the MCMM conformational searches describe a result of having sufficient conformations. Unfortunately, the number of initial conformers generated by conformational search is enormous, which requires plenty of time to complete conformational optimization and still hardly identifies the ensemble of low-energy configurations. Especially for the lack of computational power, further quantum chemical calculations even cannot be completed. In addition, some high-energy conformers can optimize to become stable low-energy configurations, which are easily missed by an inadequate conformational search or a very rough selection in the traditional method.

Tree-step computational approach [32–34], which reasonably solves the contradiction of computing cost and the comprehensiveness of low-energy conformers, provides us an effective method to build initial configurations. The strategy generates starting conformers

---

[*] Corresponding authors at: Physics and Electronics Department, Henan University, Kaifeng 475004, China (D. Chen and B. Liu).
E-mail addresses: dongchen@henu.edu.cn (D. Chen), liub@sdut.edu.cn (B. Liu).

https://doi.org/10.1016/j.comptc.2018.07.014
Received 10 May 2018; Received in revised form 19 July 2018; Accepted 22 July 2018

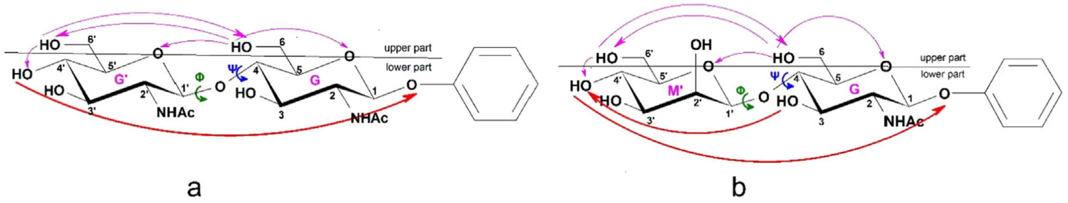


Fig. 1. The structural formulas of GlcNAcβ(1,4)GlcNAc (a) and Manβ(1,4)GlcNAc (b). Here, glycosidic linkage torsion angles (φ and ψ) are used for distinguishing the structures of *cis* and *trans*. φ = OG′-C1′-O′-C4, ψ = C1′-O′-C4-C3. The magenta arrows indicate the potential inter-ring H-bonds. The red arrows refer to the possible orientations of the cooperative hydrogen bonding networks. φ and ψ are the dihedral angles rotating about the glycosidic bond. The black solid line is defined as a split line to separate the upper and lower parts.

depending on the glycosidic linkage (*cis* and *trans*) [35,36], the categories of inter-ring hydrogen bonds (H-bonds) and cooperative H-bonds (clockwise or counterclockwise chain). Our previous works [32–34] suggest that the initial configurations constructed by tree-step computational approach are much smaller in number than those obtained by MCMM conformational search. Moreover, this method is well-suited for identifying as many as possible low-energy conformers in a timesaving and accurate way.

In this work, we determine the preferred structures of GlcNAcβ(1,4)GlcNAc (see Fig. 1a) and Manβ(1,4)GlcNAc (see Fig. 1b) by using the tree-step computational approach. Our present work aims to simplify the building procedure and improve the tree-step method by probing the preferred structures of disaccharides. At first, we build starting structures of GlcNAcβ(1,4)GlcNAc and Manβ(1,4)GlcNAc by the glycosidic linkage, the categories of inter-ring hydrogen bonds and cooperative H-bonds. DFT high-level optimizations and single point MP2 calculations will be performed. A combination of H-bonds in the upper and lower parts is proposed and can be easily used to build complicate carbohydrates. Then, the energy calculations reveal that the derivative conformers will follow the similar trend as the original conformers. In the calculations, the corresponding conformers derived from the conformers with high energy will be ruled out to reduce the computing costs. Finally, the predicted IR vibration signatures are carefully discussed. Following sections provide a detailed description on the conformational determination of GlcNAcβ(1,4)GlcNAc and Manβ(1,4)GlcNAc in the gas phase. The computational details are presented in Section 2. The building trees of initial structures are illustrated in Sections 3.1. The optimized structures and energies are discussed in Section 3.2. The vibrational signatures are analyzed in Section 3.3.

2. Computational details

The stability of carbohydrate structure is mainly attributed to glycosidic linkage, inter-ring H-bond and cooperative intra-ring H-bond. According to the three factors, the process of building starting structures can be divided into three steps. The structural formulas of GlcNAcβ(1,4)GlcNAc and Manβ(1,4)GlcNAc are shown in Fig. 1. As shown in Fig. 1, we connect the two oxygen atoms on the sugar rings (OG′ and OG for GlcNAcβ(1,4)GlcNAc; OM′ and OG for Manβ(1,4)GlcNAc) as a split line. OG′ and OG are labeled as the oxygen atoms located on rings G′ and G of GlcNAcβ(1,4)GlcNAc and OM′ and OG are labeled as the oxygen atoms located on rings M′ and G of Manβ(1,4)GlcNAc, respectively. The structure above the split line is defined as the upper part and that below refers to the lower part. When H-bonds are properly considered in the building procedure, it is possible to construct all low-energy conformers easily and reasonably. Meanwhile, the combination of H-bonds in the upper and lower parts can effectively avoid conformer duplication. In this work, the building tree does not strictly follow the glycosidic linkage, inter-ring H-bond

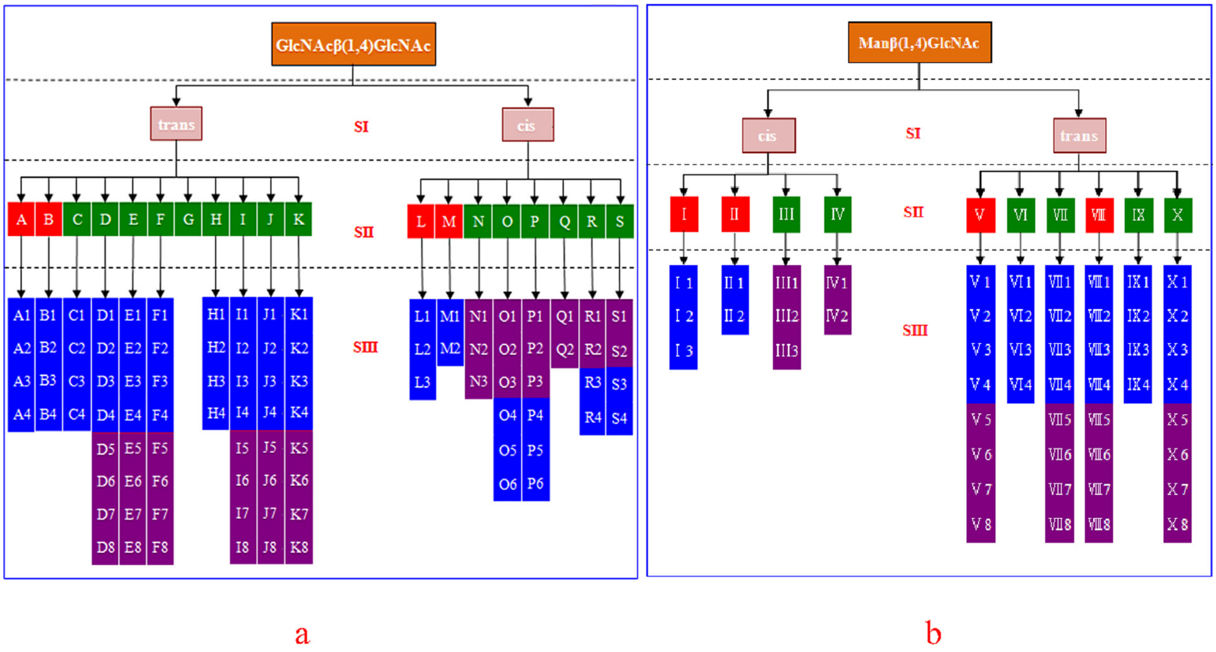


Fig. 2. Building tree of GlcNAcβ(1,4)GlcNAc (a) and Manβ(1,4)GlcNAc (b). Regains SI, SII and SIII reflect tree-step construction. SI considers the constraint of glycosidic bond linkage. SII takes into account H-bonds in the upper and lower parts. SIII refers to the transformation of H-bonds in the upper and lower parts and the counterclockwise and clockwise forms.