Carbohydrate Research 344 (2009) 1575–1581

Contents lists available at ScienceDirect

Carbohydrate Research

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

Metadynamics modelling of the solvent effect on primary hydroxyl rotamer equilibria in hexopyranosides

Vojtěch Spiwok [a,b,*], Igor Tvaroška [b]

[a] Department of Biochemistry and Microbiology, Institute of Chemical Technology in Prague, Technická 3, 166 28 Prague 6, Czech Republic
[b] Department of Structure and Function of Saccharides, Institute of Chemistry, Centre for Glycomics, Slovak Academy of Sciences, Dúbravska cesta 9, 84538 Bratislava, Slovakia

ARTICLE INFO

Article history:
Received 25 February 2009
Received in revised form 29 April 2009
Accepted 13 May 2009
Available online 23 May 2009

Dedicated to Professor Dr. Hans Kamerling on the occasion of his 65th birthday

Keywords:
Saccharide conformation
Gauche effect
C-5–C-6 rotamers
Free energy
Metadynamics
GLYCAM

ABSTRACT

Accurate modelling of rotamer equilibria for the primary hydroxyl groups of monosaccharides continues to be a great challenge of computational glycochemistry. The metadynamics technique was applied to study the conformational free energy surfaces of methyl α-D-glucopyranoside and methyl α-D-galactopyranoside, employing the GLYCAM06 force field. For both molecules, seven to eight conformational free-energy minima, differing in the ω (O-5–C-5–C-6–O-6) and χ (C-3–C-4–O-4–HO-4) dihedral angles, were identified in vacuum or in a water environment. The calculated rotamer equilibrium of the primary hydroxyl group is significantly different in vacuum than in water. The major effect of a water environment is the destabilisation of a hydrogen bond between O-4–HO-4 and O-6–HO-6 groups. It was possible to calculate the free-energy differences of individual rotamers with an accuracy of better than 2 kJ/mol. The calculated *gg*, *gt* and *tg* rotamer populations in water are in close agreement with experimental measurements, and therefore support the theoretical background of metadynamics.

1. Introduction

Carbohydrate recognition by proteins and other molecular partners is essential to many biological processes such as interactions of immune system components, cell–cell adhesions, metastasis and pathogen entry into a cell. Numerous carbohydrates and carbohydrate mimics are currently being developed to therapeutically intervene in these processes. Analogously, artificial carbohydrate-recognising systems offer great potential as carbohydrate sensors or therapeutics. Moreover, cleavage of highly abundant carbohydrates is nowadays widely studied as a potential source of future fuels. An understanding of carbohydrate conformation and dynamics is essential for future development of carbohydrate-based biotechnologies.

An important feature of oligosaccharides is that they usually do not contain secondary structure elements typical for most proteins, nucleic acids or polysaccharides with repeating sequences. Instead, oligosaccharides are typically present in solution in multi-conformational families. This complicates experimental determination of their three-dimensional structures using X-ray crystallography and NMR spectroscopy, yet, on the other hand, it provides a great opportunity for the application of molecular modelling techniques. Numerous successful applications of molecular dynamics simulation for modelling carbohydrate conformation have been presented in the last several decades.[1]

The major disadvantage of molecular dynamics simulation is that this method is relatively computationally expensive. In principle, it is possible to calculate conformational equilibria from probabilities of individual conformations during a long molecular dynamics simulation. However, time scales accessible by molecular dynamics simulations are often not long enough to allow the modelled carbohydrate to explore all populated conformations. Transitions between these families are often too rare to be observed in a reasonable time scale. Even if the carbohydrate can change between conformational families, their populations usually cannot be measured within an acceptable degree of accuracy. Numerous molecular modelling techniques have been developed to improve sampling of molecular dynamics simulation by use of an artificial degree of freedom (e.g., free-energy perturbation), bias potential or force (e.g., umbrella sampling or metadynamics), parallel tempering or other principles. Recently introduced metadynamics is a rapidly developing technique that improves sampling and allows for the calculation of a low-dimensional free-energy surface.[2,3]

The first step of the application of metadynamics is a selection of few (in this study two) collective variables. These are parameters

* Corresponding author. Tel.: +420 220 44 30 28; fax: +420 220 44 51 67.
E-mail addresses: chemspiw@savba.sk, spiwokv@vscht.cz (V. Spiwok), chemitsa@savba.sk (I. Tvaroška).

0008-6215/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2009.05.019

describing the most important slow degrees of freedom in the studied system. The repertoire of collective variables applied in metadynamics includes distances, angles, coordination numbers, collective motions, crystallographic parameters and others (see Ref. 3). Then the system is simulated by a standard molecular dynamics simulation with an additional history-dependent bias potential. The bias potential is defined as a sum of Gaussian hills located in the space of collective variables:

$$V_{\text{bias}} = \sum_{t'<t} \prod_{j} w_{t'} \exp\left[\frac{-(s_j(x) - s_j^{t'})^2}{2\delta s_j^2}\right] \quad (1)$$

where w is height of a hill and δs are its widths. During the simulation, every G-th step (in this study $G = 1000$ steps, 1 ps) a Gaussian hill is added to the standard molecular mechanical potential. These hills cumulate until they flood certain free-energy minimum and allow the system to overcome the free-energy barrier and to explore much wider range of configurations. Moreover, a final bias potential approximates a low-dimensional free-energy surface of the studied system, that is, for certain values of collective variables it is possible to calculate equilibrium probability of corresponding states (e.g., conformations). Metadynamics has been successfully applied in numerous fields, such as the modelling of conformational equilibria,[4] chemical reactivity[5] or phase transitions.[6]

In spite of recent progress in the modelling of saccharides, a simple description of the conformational equilibria of hydroxyl and primary hydroxyl groups is a challenging issue. Primary hydroxyl groups, such as the C-5–C-6–O-6–HO-6 moiety in hexopyranoses (Fig. 1), show an interesting conformational behaviour, which was inaccurately modelled by early generations of carbohydrate force fields. Experimental crystal structures of carbohydrate–protein complexes and carbohydrates showed that the C-5–C-6 bond can exist in three canonical conformers, namely in *gauche–gauche* (*gg*), *gauche–trans* (*gt*) and *trans–gauche* (*tg*). Populations of individual rotamers in solution can be estimated using NMR spectroscopy from measured homonuclear ($^3J_{\text{H-5–H-6}}$) or heteronuclear ($^3J_{\text{C-4-H-6}}$) vicinal coupling constants. For methyl α-D-glucopyranoside, the experimentally predicted percentage populations for *gt*, *tg* and *gg* rotamers are 38%, 5% and 57%, respectively.[7] For methyl α-D-galactopyranoside the corresponding populations are 47%, 39% and 14%.[7]

Potential energies of these rotamers can be calculated for static structures in vacuum using molecular mechanics, ab initio, or density functional theory (DFT) methods.[8,9] However, their energetics in vacuum and at 0 K is poorly related to the situation in solvent and at biologically relevant temperatures, even if a high level of theory is used. For example, the *trans–gauche* (*tg*) conformation of methyl α-D-glucopyranoside was predicted to be energetically favourable in vacuum, when, in fact, it was almost absent in a water environment. In order to realistically model populations of C-5–C-6 rotamers it was necessary to include the effect of solvent.

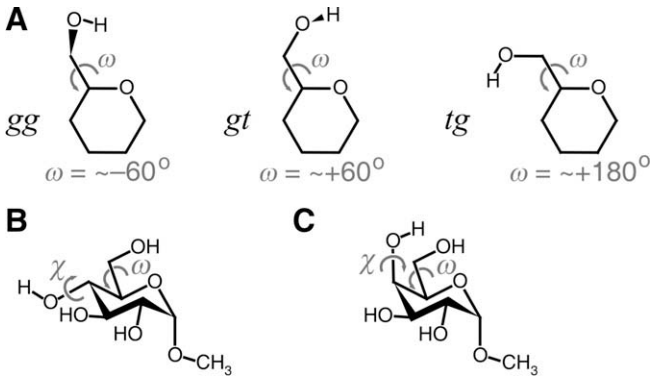

Figure 1. (A) Schematic representation of *gauche–trans* (*gt*), *trans–gauche* (*tg*) and *gauche–gauche* (*gg*) conformations, (B) methyl α-D-glucopuranoside (α-D-Glc*p*-OMe, **1**), (C) methyl α-D-galactopyranoside (α-D-Gal*p*-OMe, **2**).

In the context of hexopyranoses, the propensity of ω torsional angle (O-5–C-5–C-6–O-6) to adopt the *gauche* conformation is often being explained by the *gauche* effect,[10] that is, the rotamers with two oxygens in *gauche* orientation tend to be more stable. However, this effect can be compensated by hydrogen bonding between O-6–HO-6 and O-4–HO-4 moieties. Both these phenomena, the *gauche* effect and hydrogen bonding, can be accurately addressed by DFT and ab initio techniques in the gas phase[8,9] and in the development of empirical force fields. However, the situation becomes more complicated when a solvent comes into play. The first attempt in finding a solution to the problem of solvent effects was the application of an implicit solvation model. Energies calculated at the DFT level of theory, together with the self-consistent interaction field solvent model, were in sound agreement with experimental populations.[8]

Alternatively, it is possible to perform a very long molecular dynamics simulation in an explicitly modelled solvent. The resulting trajectory can be analysed to calculate populations of individual rotamers. This approach was applied by Kirschner and Woods.[11] In a 50-ns MD simulation of methyl α-D-glucopyranoside, they observed 15 transitions between *gauche–trans*, *trans–gauche* and *gauche–gauche* conformations. By analysis of the resulting trajectory they found that the *gauche–gauche* conformation was the most populated (54%), followed by *gauche–trans* (40%), and the *trans–gauche* conformation was least populated (6%), all in good agreement with experimental results. A correlative simulation for methyl α-D-galactopyranoside led to *gt*, *tg* and *gg* rotamer populations of 64%, 28% and 8%, respectively.

In this study we apply the metadynamics method to simulate rotamer equilibria of a primary hydroxyl group (Fig. 1) in methyl α-D-glucopyranoside (α-D-Glc*p*-OMe, **1**) and methyl α-D-galactopyranoside (α-D-Gal*p*-OMe, **2**) in vacuum and an explicit water environment. The free-energy surface for **1** and **2** was calculated as a function of two variables. The first collective variable was the ω torsional angle (O-5–C-5–C-6–O-6). The χ torsional angle (C-3–C-4–O-4–HO-4) was chosen as the second collective variable because it controls clockwise versus counter-clockwise orientations of secondary hydroxyl groups that are especially important in vacuum.

2. Models and methods

Metadynamics calculates the free-energy surface in the space of a few (typically two) parameters (collective variables), and their choice is the first step in the application of metadynamics. In order to quantitatively study torsional equilibrium of primary hydroxyl rotamers, it was essential to choose the torsional angle ω (O-5–C-5–C-6–O-6, Fig. 1) as the first collective variable. It is useful to use a bias potential that improves sampling of all potentially slow degrees of freedom within the system. Important degrees of freedom that are not biased by the metadynamics potential may cause poor convergence and hysteresis in calculated free-energy surfaces. The second important degree of freedom of the studied system is rotation of the secondary hydroxyl groups. Especially in vacuum, these hydroxyl groups form a hydrogen-bond network with either a clockwise (*c*) or a counter-clockwise (*cc*) orientation. These orientations influence the populations of primary hydroxyl rotamers. Therefore, the torsional angle χ (C-3–C-4–O-4–HO-4) was selected as the second collective variable.

The duration of each metadynamics run was initially set to 10 ns. However, after approximately 8 ns of the metadynamics run of α-D-Glc*p*-OMe (**1**) in water, the molecule started to explore