Modelling the Effect of Solvents on Carbohydrates

Vojtěch Spiwok[1,2]* and Alfred D. French[3]

[1]*Department of Biochemistry and Microbiology, Institute of Chemical Technology Prague, Technická 3, Prague 6, 166 28, Czech Republic*

[2]*Department of Structure and Function of Saccharides, Institute of Chemistry, Center for Glycomics, Slovak Academy of Sciences, Dúbravská cesta 9, 84538 Bratislava, Slovak Republic*

[3]*Southern Regional Research Center, US Department of Agriculture, 1100 Robert E. Lee Blvd., New Orleans, LA 70124, United States*

Abstract: Carbohydrates are polar molecules and their conformational and anomeric equilibrium can be strongly influenced by solvents. This review provides examples of studies addressing different issues of glycochemistry, such as anomeric equilibrium, conformational changes in rings, modelling of inter-residue linkages or complex carbohydrates and formation of carbohydrate-protein complexes. All these phenomena provide benchmark systems for testing of different solvent models.

Keywords: Carbohydrate, solvation, conformation, continuum solvent model.

INTRODUCTION

The importance of carbohydrates has been recently recognized after decades during which nucleic acids and proteins dominated the field of molecular biology. Development of novel carbohydrate-based drugs, biotechnologies and glycomics techniques has led to new tasks for molecular modellers and bioinformatics experts [1]. Major challenges of carbohydrate modelling (also matching the major topics addressed by articles in this special issue) are highlighted in Fig. (**1**) and include predictions of anomer equilibrium of reducing sugars, conformational modelling of ring systems, conformational modelling of linkages between monosaccharide units in oligo- and polysaccharides, conformational modelling of complex carbohydrates and glycoconjugates and recognition of carbohydrates by their biological binding partners, mostly by proteins. The effects of solvent and their accurate modelling play a crucial role in all these topics. This is especially true for highly polar saccharide molecules in highly polar water environment.

CARBOHYDRATES IN THE GAS PHASE

Water is the environment that significantly influences conformational behaviour and dynamics of molecules and stability of their assemblies, but taking it into account by computational modelling can add considerable complexity. Because application of molecular modelling techniques is relatively straightforward and computationally efficient on isolated molecules in vacuum, that was the environment of choice in many pioneering computational studies of carbohydrates. That resulted in a fundamental disconnections because most carbohydrates were observed only in condensed phases. Relatively recently, however, rigorous experimental data on behaviour of carbohydrates in vacuum have been obtained, owing to the development of advanced double resonance spectroscopy techniques in a molecular beam environment. This completes the circle. The simple vacuum models can be compared with the more advanced studies that consider solvent water, and the results from established experimental techniques for solutions and crystals can be compared with those from the new experiments on vacuum phase carbohydrates.

In these sophisticated experimental studies, the molecule/complex must be first transferred into a molecular beam by means of supersonic jet expansion. The molecule is then irradiated by the first tunable monochromatic IR or UV laser (in IR-UV or UV-UV spectroscopy, respectively). If the first laser is tuned to the absorbing wavelength of the molecule, its ground state population is perturbed. The second UV laser is then used to photo-ionize those molecules/complexes that did not absorb at the wavelength of the first laser. This second laser is tuned to ionize a certain conformation of the molecule (or a configuration of the molecular assembly). Finally, the ionized molecule/complex is analysed by a mass spectrometer. If the wavelength of the first laser is resonant with the IR or UV transition, the object cannot be ionized by the second laser and the mass spectrometer signal is thus reduced. These methods provide high resolution and mass- and conformationally-resolved UV and IR spectra of biomolecules and their assemblies. Photochemistry can be elegantly studied by a delay between irradiations by the lasers. The computational interpretation of spectra is an advanced field reviewed for example by DeVries and Hobza [2].

Carbohydrates (together with other groups of biomolecules) have been studied by IR-UV spectroscopy by the group of John P. Simons at the University of Oxford [3-8]. The first subjects of their research were arylglycosides, studied to probe orientations of hydroxyl and hydroxymethyl groups in bare and singly- or multiply-hydrated monosaccharides [3-8]. The presence of an aromatic moiety, attached either covalently [3-6] or noncovalently [7,8], is necessary for photo-ionization. Hydroxyl groups of monosaccharides were observed to be oriented in a highly cooperative pattern, i.e. they exist either in clockwise or counter-clockwise orientations, depending on monosaccharide and degree of solvation (Fig. **1C**). For example, the hydroxyls of a bare molecule of phenyl-β-D-glucopyranoside [4] were experimentally observed only in the counter clockwise orientation. This was true regardless of which of the three staggered orientations of its hydroxymethyl moiety (*gauche-gauche*, *gauche-trans* and *trans-gauche*, Fig. **1D**) existed. Addition of a single water molecule stabilized the clockwise *gauche-gauche* form (with water between O-4–H and O-6–H) and counterclockwise *gauche-trans* orientation (with water either between O-6–H and O-5 or between O-2–H and the aglycon). Similar changes in conformational preferences were observed for other monosaccharides [3-8]. The recent study of Cocinero and coworkers reviews the previous results on hydrated derivatives of α-D-Man*p*, β-D-Glc*p* and β-D-Gal*p* with new results on β-D-Man*p*, α-D-Glc*p* and α-D-Gal*p* [6]. In general, a water molecule or molecules tend to bind to the hydroxymethyl moiety and bridge to neighbouring hydroxyl groups or O-5 atom. The hydrophilic edge of a monosaccharide is more attractive for water molecules than the hydrophobic face. These studies have shown that the fine details of monosaccharide conformations can be significantly influenced even by a single molecule of water.

*Address correspondence to this author at the Department of Biochemistry and Microbiology, Institute of Chemical Technology Prague, Technická 3, Prague 6, 166 28, Czech Republic; Tel: +420 220 443 028; Fax: +420 220 445 167;
E-mail: spiwokv@vscht.cz

1570-193X/11 $58.00+.00

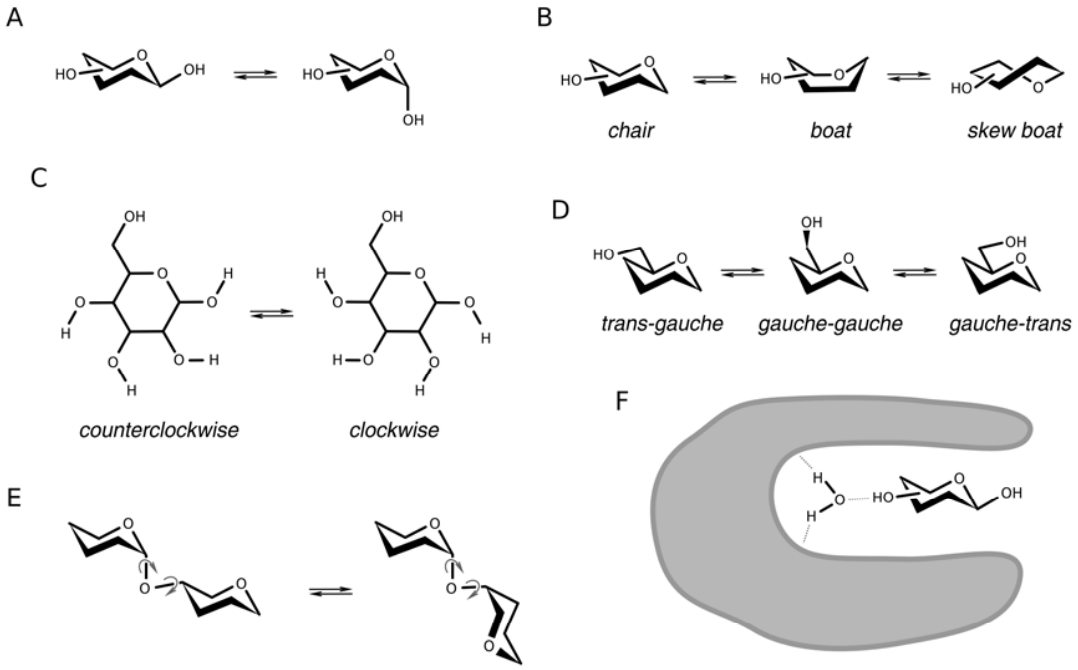

Fig. (1). Schematic representation of the phenomena addressed in this review. **A**, anomeric equilibrium. **B**, ring conformations. **C**, clockwise/counterclockwise orientation of hydroxyl groups. **D**, rotamer equilibrium of the hydroxymethyl group. **E**, conformational behaviour of oligosaccharides. **F**, water-mediated carbohydrate-protein interaction.

CARBOHYDRATES IN SOLVENT

While multi-photon ionisation spectroscopic studies on isolated carbohydrates and their computational interpretation is a fine surgery of today's molecular sciences, computational modelling of solvated carbohydrates is rather a molecular butchery. The solvent can be modelled using implicit (continuum-based) or explicit (atom-resolved) models. The effect of solvent molecules can be dissected into three major factors. First, water is highly polar and therefore it reduces the strength of electrostatic interactions compared to the situation in vacuum. Numerous implicit solvent models, such as self-consistent reaction field (SCRF) [9], polarizable continuum model (PCM) [10] or conductor-like screening model (COSMO) [11] have been developed to model the screening of electrostatic interactions by highly polar solvent molecules. Second, water (or any other solvent) is viscous and its viscosity influences the dynamics of the solute. That effect can be modelled implicitly by adding a friction term to dynamics of atoms of the solute. Finally, there are several examples where an individual water molecule or a cluster of waters might bind into a certain position and thus stabilize a certain state of the solute. The explicit function of water molecules, together with electrostatic screening and solvent viscosity, can be modelled by explicit (i.e. atom-resolved) solvent models. Different water parametrisations such as single-point charge (SPC) [12], transferable intermolecular potential (TIP) [13] and others are frequently used, mostly in molecular dynamics simulations. Implicit solvent models are widely used in computationally expensive quantum chemical studies, which are usually based on calculations of single-point energies and energies of geometry optimized structures (i.e. non-dynamical calculations). Energetics of conformational change are simply evaluated by comparing energies of the molecule in both conformations. Such calculations can be performed either in vacuum or in implicitly modelled solvents. On the other hand explicit solvents are preferred in time-resolved molecular simulations, i.e., molecular dynamics. The reason for this segregation comes from the fact that addition of each explicit water molecule is associated with additional degrees of freedom. In principle it is possible to compare energies of two conformations of a molecule, both solvated by hundreds of explicitly modelled waters. However, this is similar to an attempt to weigh a milligram quantity by weighing yourself on a personal balance with and without the sample in your hand and subtracting the values. Random energy differences associated with the solvation shell might be significantly higher than those associated with the conformational change in the solute.

The design of implicit and explicit solvent models is a challenging task. Parametrization of explicit solvents, similar to other molecular mechanics functions, can employ different strategies. Some force fields and explicit solvents were designed to achieve maximum accuracy in fine atomic details whereas others were designed to reproduce bulk properties. As a result, a comparison of potential energy calculated by molecular mechanics with a quantum chemistry reference does not necessarily reflect the quality of the molecular mechanics potential. This is because some force fields and explicit solvents show a good performance in prediction of solvation free energies, modelling of hydrophobic effect or simulations of large conformational changes, but they fail in energetics of fine structural details. The consequence is that some force fields were designed solely for simulations in solvent. Moreover, different force fields are suitable for different explicit or implicit solvent models.

To conclude this part, implicit continuum solvents, which do not increase the complexity of the studied system, are suitable for computationally expensive quantum chemistry calculations where we want to obtain maximum information from a static structure. Explicit solvents are on the other hand popular in biomolecular simulations where thousands of snapshots of the system are being evaluated along the time axis. Both these models – implicit as well as explicit – have been applied to solve the above listed key issues of glycochemistry.

ANOMERIC EQUILIBRIUM

The existence of an anomeric carbon and the equilibrium between the two anomers on carbohydrate reducing termini (Fig. **1A**) is a spice of glycochemistry. The pioneering article of Jeffrey, Pople and Radom from 1972 [14] compared *ab initio* energies of

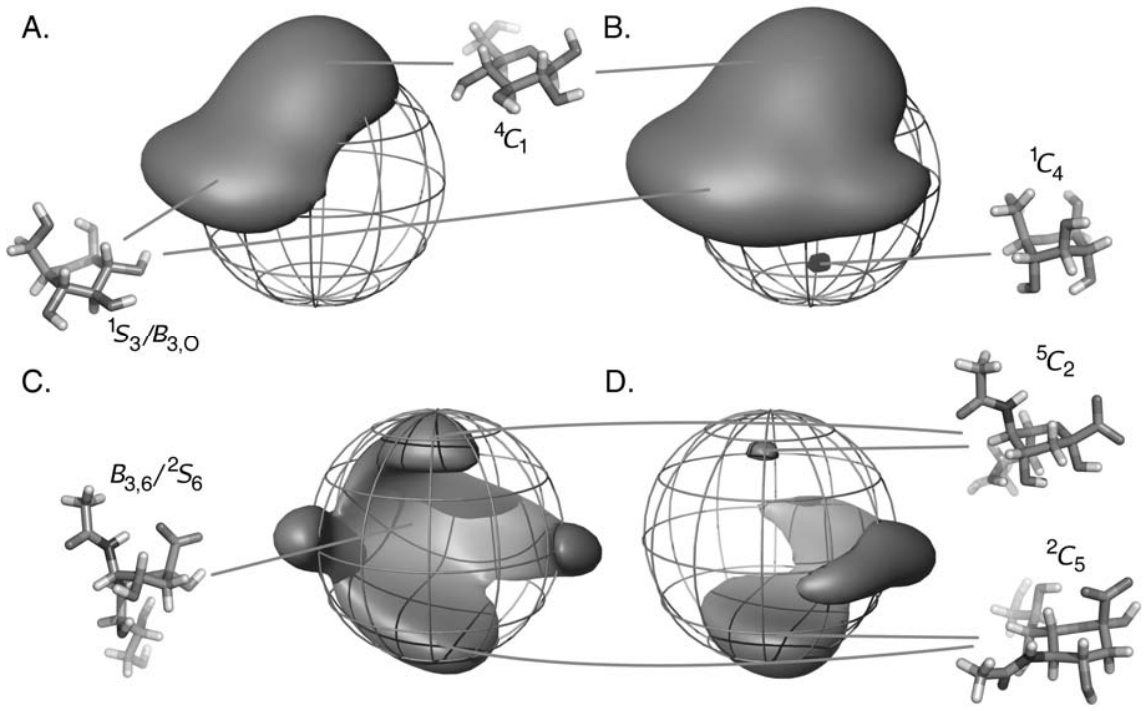

Fig. (2). The effect of water on the conformational free energy surface of monosaccharides. Energy isosurfaces are plotted in Cremer-Pople spherical coordinates with chair structures located at both poles. Boat and skew forms are represented on the equator. **A**, β-D-glucopyranose in vacuum. **B**, β-D-glucopyranose in water. Free energy is depicted as isosurfaces at 50 kJ/mol, relative to $^4C_1$. **C**, α-*N*-acetylneuraminic acid in vacuum. **D**, α-*N*-acetylneuraminic acid in water. Free energy is depicted as isosurfaces at 30 kJ/mol, relative to $^2C_5$. The effect of water on β-D-glucopyranose is small and rather qualitative, however, for α-*N*-acetylneuraminic acid the effect is significantly higher due to formation of an ionic hydrogen bond in the $B_{3,6}/^2S_6$ conformation.

methanediol and the results were discussed in relation to the anomerism of α- and β-D-glucopyranose. This study was nominated to be among the 66 most influential theoretical chemistry papers of the 20th century in an end-of-the century issue of Theoretical Chemistry Accounts [15], owing to the fact that it is a pioneering work on scanning potential energy surfaces.

The anomeric equilibrium of pyranose sugars was studied by Tvaroška and Carver using 2-hydroxytetrahydropyran as a model [16] (following an earlier study on 2-methoxytetrahydropyran [17]). This example shows an important effect of the solvent. *Ab initio* energies have shown that axial α-anomer is more favoured in vacuum by 5.4 kJ/mol over the equatorial β-anomer (both in *gauche-trans* conformations of the hydroxyl group). When free energies were taken into the account (introduced by a harmonic oscillator model), the situation changes only quantitatively, not qualitatively (from 5.4 to 2.1 kJ/mol). This result is in good agreement with NMR data measured in nonpolar solvents. However, when a polar water environment is mimicked by the Poisson-Boltzmann equation, the conformational preference changes so that the β-anomer is preferred by 0.2 kJ/mol. This is in good agreement with experimental results measured in water (0.4 to 2.5 kJ/mol). This pioneering study on a minimal sugar model (2-hydroxy tetrahydropyran) was further extended to the whole picture of anomerism in glucopyranose by a detailed mechanistic study of the mutarotation process [18].

RING CONFORMATION

Another important issue of carbohydrate modelling is ring conformation, especially in pyranose sugars (Fig. **1B**). Saturated six-membered rings can exist in different conformations, namely in two different chair, six boat and six skew-boat conformers. Transition states and rings incorporating double bonds can be described by half-chair and envelope conformers as well. Steric factors usually restrict a given molecule to only one or a few forms, and non-chair structures will often not correspond to any of the characteristic forms. The majority of pyranoses dominantly favour the chair conformation with a maximum number of equatorial substituents, however, there are many important exceptions to this rule such as conformationally distorted substrates in active sites of glycosidases [19], mechanically stressed polysaccharides [20] or L-iduronic acid residues in sulphated glucosaminoglycans [21]. A well-established method for analysis of ring conformation was developed by Cremer and Pople in 1975 [22]. The conformation of a saturated six-membered ring can be described by three coordinates q_x, q_y and q_z. Stable conformations are located on a sphere where chair structures are located on both poles ($q_x \approx q_y \approx 0$) and boat/skew-boat configurations are placed on the equator ($q_z \approx 0$). This Cartesian formulation of Cremer and Pople coordinates is computationally convenient, but puckering parameters are usually expressed in terms of spherical polar coordinates ϕ, θ and Q. The Q represents deviation of the ring atoms from a mean plane, and the θ value indicates the type of ring (chair, boat or skew). The ϕ value indicates which of boat or skew shape is involved (see Fig. **2**).

Several quantum chemistry and molecular mechanics studies aimed at the elucidation of conformational equilibria between these forms have been reported. For example Barrows and co-workers [23] calculated energies of $^4C_1$ and $^1C_4$ structures of β-D-glucopyranose in vacuum, each represented by a pair of O–6 conformers, for different levels of theory (Hartree-Fock, Møller-Plesset and coupled cluster). In vacuum the $^4C_1$ is favoured by 33 to 38 kJ/mol, expressed as $\Delta G_{298\ K}$, relative to $^1C_4$. Inclusion of the solvent by SM4 model resulted in values of 54 to 79 kJ/mol in water and 42 to 54 kJ/mol in *n*-hexadecane. This study demonstrated that each canonical free energy minimum (in this case each ring conformer) is represented by a number of conformers differing in addi-

tional degrees of freedom such as orientation of hydroxymethyl and hydroxyl groups or interaction with water. These additional degrees of freedom cause the potential energy surface of β-D-glucopyranose to be rough and conformational sampling probably represents a more important problem than the level of theory. That said, the various levels of quantum theory that were employed in vacuum calculations resulted in a range of more than 140 kJ/mol for the difference in energy between the two chair forms. In some cases the $^1C_4$ was favoured in vacuum by 70 kJ/mol, and in others, including the largest calculation, the $^4C_1$ was favoured by as much as 70 kJ/mol.

A potential solution to the problems of additional degrees of freedom and entropic effects has been provided by Biarnés and coworkers [24]. They simulated the dynamics of β-D-glucopyranose using Car-Parrinello molecular dynamics (CPMD) method in combination with metadynamics. Metadynamics allows for modelling of slow processes (e.g. slow conformational changes) and simultaneously it estimates the conformational free energy surface. Conformational sampling in the space of Cremer-Pople puckering coordinates was being enhanced in this study and a conformational free energy surface was calculated in vacuum. The free energy difference associated with the transition from $^4C_1$ to $B_{3,O}$ boat was calculated as 11 kJ/mol. Multiple boat/skew-boat local free energy minima were observed, but the $^1C_4$ chair and the effects of water were not studied.

The effect of water was independently studied by our group (together with testing of different force fields) [25] and by Hansen and Hünenberger (together with their new free energy modelling methodology) [26]. Both studies applied a free energy modelling method (our group used metadynamics and Hansen and Hünenberger used their innovative local elevation umbrella sampling) to calculate conformational free energy surfaces of β-D-glucopyranose in vacuum and in water. Comparisons of those free energy surfaces showed some specific effects of the water environment (for example increased stabilization of a boat form relative to the chair in water), but these differences were not dramatic. Fig. (**2AB**) compares conformational free energy surfaces of β-D-glucopyranose, one calculated in vacuum and one in water, both calculated with the GLYCAM06 forcefield (data taken from [25]). Free energies are visualized as isosurfaces at a certain free energy value. For an isosurface at 50 kJ/mol this means that free energy values of conformations located inside the isosurface "blob" are not higher than 50 kJ/mol relative to the global energy minimum. The $^4C_1$ chair is located in the coordinate system on the north pole of the sphere, the inverted $^1C_4$ chair is located on the south pole and boat/skew-boat conformations are located on the equator. It shows that the effect of water on this molecule is relatively small. The $^4C_1$ chair is the global free energy minimum in both environments, and also pathways of conformational changes are identical. In the other words, the free energy surfaces in both environments differ quantitatively, rather than qualitatively. Similar effects were observed for three different carbohydrate-tuned force fields [25].

The situation is likely to be different in *N*-acetylneuraminic acid. We have studied the ionized form of this molecule in vacuum and in water [27]. Free energy surfaces were calculated again using metadynamics in GLYCAM06 molecular mechanics potential. In contrast to β-D-glucopyranose, conformational free energy surfaces of α-*N*-acetylneuraminic acid in vacuum and in water were strikingly different (Fig. **2CD**, data from [27]). The major difference is the high stability of boat/skew-boat conformations in vacuum is due to formation of an ionic hydrogen bond between the amide N–H bond and carboxylic acid moiety. This conformation is disfavoured in water solution as a result of screening of this interaction. Here we have two monosaccharides with strikingly different effects of water on conformation of their rings. On one hand, β-D-glucopyranose shows relatively low effect of water on ring conformation. On the other hand, the water environment changes the conformational free energy surface of *N*-acetylneuraminic acid dramatically; not only quantitatively but also qualitatively, i.e. in the order of free energy minima and in pathways of conformational changes. These findings must be further verified experimentally, nevertheless, it illustrates the issue of solvent-dependence of monosaccharide conformations.

HYDROXYMETHYL ROTAMER EQUILIBRIA

Another traditional issue of conformational modelling of carbohydrates is the equilibrium of the orientations of the hydroxymethyl (-CH_2-OH) group (Fig. **1D**). This moiety can adopt either *gauche-gauche* (*gg*), *gauche-trans* (*gt*) and *trans-gauche* (*tg*) conformations, relative to the ring atoms. For the example in β-D-glucopyranose they can be defined by the value of O-5–C-5–C-6–O-6 torsion angle (ω) with equilibrium values –60, +60 and +180 deg, respectively. There are several reasons why this system became a popular model in computational glycochemistry. Mainly, this equilibrium is relevant for modelling of conformations of oligo/polysaccharides with 1→6 linkage. Moreover, populations of these conformations are large enough that they can be determined using NMR from homo-or heteronuclear spin-spin couplings. Also, inter-conversion among these three forms is relatively fast (in contrast to for example ring equilibria) and populations can be sampled by a standard (if lengthy) molecular dynamics simulation. These two reasons make rotamer equilibria of a hydroxymethyl moiety a suitable benchmark for evaluation of newly developed carbohydrate empirical potentials.

The first correct predictions of conformational equilibria for structures varying in the ω-torsion angle by quantum chemistry methods showed that accuracy mostly depends on including a solvent treatment rather than on the level of quantum mechanics theory. The first solution to the problem of solvent was introduced by Tvaroška and Carver [28]. They were successful in prediction of the correct rotamer equilibria by application of free energy corrections together with a continuum solvent model (namely the self-consistent reaction field). Their approach can be illustrated for example on methyl α-D-glucopyranoside. That compound was geometry optimized in different conformational families at the B3LYP/6-31G** level of theory and then their energies were calculated at the B3LYP/6-311++G** level. The energetically most favourable was the *tg* conformation with the counter-clockwise orientation of hydroxyl groups. However, when free energies (calculated from vibration frequencies) were taken into account the counter-clockwise *gg* conformation had the minimum energy. This preference was furthered by use of an implicit solvent model. The resulting free energies agreed well with conformational equilibria determined by NMR, not only for α-D-glucopyranoside, but also for other molecules, especially those with the *galacto* configuration.

Instead of comparing energies obtained in implicitly modelled solvents, populations of different conformations can be calculated in an explicitly modelled solvent by a lengthy molecular dynamics simulation. This approach was applied by Kirschner and Woods [29]. They studied methyl α-D-glucopyranoside and methyl α-D-galactopyranoside using the GLYCAM03 molecular mechanics potential. In vacuum they obtained a good agreement between GLYCAM and quantum chemistry. However, there was poor correlation between the GLYCAM potential energies calculated for variations of the ω-torsion in vacuum and the experimental data on the corresponding equilibrium in water. Therefore they performed a long molecular dynamics simulation in explicitly modelled water and calculated the relative populations of individual rotamers. During 50 ns simulations in explicitly modeled TIP3P water they observed multiple conformational transitions. Populations of ω-torsion rotamers were in good agreement with experimental data. Moreover, this study revealed that a correct modelling of the conformational equilibrium can be achieved when 1–4 interactions (i.e. interactions separated by three covalent bonds) are not scaled (the

scaling of 1–4 interactions is a standard practice on AMBER calculations). This can be explained by the fact that scaling of 1–4 interactions unbalances interaction of O-6 with O-5 (via a 1–4 interaction) and with O-4 (via a 1–5 interaction). We have studied the same system [30] using the metadynamics method, which has shown excellent performance in modelling ring equilibria [25,27]. Instead of using metadynamics to speed up ring puckering, we focused on conformations of the hydroxymethyl moiety. The free energy surface was calculated by a 10 ns metadynamics run with GLYCAM energies. In vacuum we observed a strong cooperativity of hydroxyl groups. A preference for clockwise/counterclockwise hydroxyl orientations was linked to the preference for *gt*, *tg* and *gg* conformations. Free energy differences between clockwise/counterclockwise conformations are higher in vacuum than among different hydroxymethyl rotamers. As expected, the equilibrium of *gt*, *tg* and *gg* conformations calculated in vacuum was in poor agreement with experimental data measured in water. Addition of explicitly modelled water decreased the cooperativity of hydroxyl groups as well as barriers between clockwise/counterclockwise orientations. Also, the predicted conformational equilibrium of *gt*, *tg* and *gg* conformations calculated in water is in remarkably good agreement with experiment. This study further extended the study of Kirschner and Woods [29] by adding the resolution to hydroxyl orientations and by providing statistics.

INTER-RESIDUE LINKAGE

The next step from monosaccharides to oligosaccharides and glycoconjugates requires accurate modelling of glycosidic linkages (Fig. **1E**). As most monosaccharides are relatively rigid, in most cases only two degrees of freedom per single monosaccharide building block predominantly determine the conformation of the chain. These degrees of freedom are torsions describing the rotation around the first and the second bond in the C–O–C linkage. This is similar to the situation in proteins where two Ramachandran angles per single amino acid residue determine the conformation of the backbone. For this reason, corresponding torsions are often also referred to as "carbohydrate Ramachandran angles" (O-5–C-1–O-1–C'-n and C-1–O-1–C'-n–C'-(n-1) are ϕ and ψ IUPAC standards for aldopyranoses [31], but, H–C–O–C and C–O–C–H ϕ_H and ψ_H angles are also frequently used). Exceptions are for example oligosaccharides where one residue is linked to a hydroxymethyl moiety of the second residue. For this kind of linkage, also the third torsional angle (i.e. C–O–CH_2–C) is necessary to determine the chain conformation. Even a simple disaccharide when placed to the gas phase tends to form inter-residue hydrogen bonds beside previously discussed clockwise or counterclockwise hydrogen bonding networks (however, usage of the term "hydrogen bond" for vicinal diols is a question of debate [32,33]). On the other hand, these hydrogen bonds are likely to be screened when in water. The fact that hydroxyl groups of disaccharide residues are "stickier" in the gas phase can change the preferred Ramachandran dihedral angles of the linkage.

Cellulose and related cellooligosaccharides are intensively studied as future sources of green energy. Cellobiose – the disaccharide fragment of cellulose – is likely to show an interesting effect of water on its conformation. French and Johnson have performed a thorough conformational analysis of cellobiose linkage at the quantum mechanics level [34]. They identified the structure of $\phi \sim +60$ deg and $\psi \sim -120$ deg as the global energy minimum in vacuum, which is in agreement with computational as well as experimental studies [35,36]. This structure is characterized by O-2'–H ... O-3 and O-6'–H ... O-6 hydrogen bonds forming a wider inter-residue hydrogen bonding network. However, a recent parallel tempering study of cellobiose in explicitly modelled water revealed that a structure showing similar values of Ramachandran angles is present but it is not the most populated one [37]. Instead, the structure of $\phi = -75$ deg and $\psi = -120$ deg was identified as a global free energy minimum. This study uses the method called parallel tempering (often also referred to as the replica exchange method; however, this name is more general and includes techniques where other principles are used to improve sampling). It is possible to make conformational and other sampling faster by running the simulation at elevated temperatures, however, conformational preferences might become significantly different from those at the temperature of interest. In parallel tempering the system is simulated in parallel replicas differing in temperature. During the simulation, a pair of replicas can occasionally exchange. This means that Cartesian coordinates and velocities of atoms of a colder replica are exchanged for a warmer one and *vice versa*, velocities are rescaled to new temperatures and the simulation continues. These exchanges are random and their probability depends on the potential energy difference between replicas. For example, if the simulation starts from some energetically unfavourable conformation of cellobiose, the molecule tends to exchange for high-temperature replicas and climbs towards higher temperatures. Then, since it reaches some energetically favourable conformation (owing to better sampling at high temperature) the molecule tends to exchange for colder replicas and returns to biological temperatures. The criterion of exchanges ensures that trajectory snapshots recorded at a certain temperature represent realistic conformational preferences of the molecule at the given temperature. These two studies have shown a significant difference between conformational behaviour of a cellulose building block in vacuum and in water. This is especially intriguing because cellulose interests us in various environments including different crystal forms, solubilized cellulose when undergoing biological degradation as well as a molecule in gas phase during a technological pyrolysis process.

Another interesting disaccharide in the context of solvation is α,α-trehalose, often used as a protective agent because of its strong interaction with water. A pair of studies indicated that its conformational preferences are significantly different in vacuum and in water, namely that the structure of Ramachandran angles $\phi \sim \phi' \sim +180$ deg is the vacuum energy minimum, but the solution minimum is at $\phi \sim \phi' \sim +60$ deg [38,39]. The later conformation was also observed in crystal structures of anhydrous trehalose and in mono-hydrate [40] as well as all other molecules with similar linkages. However, a quantum chemistry study by French and co-workers revealed that the 180 deg vacuum preference was a result of inaccuracy in the molecular mechanics potential rather than an effect of the absence of a solvent [40]. Comparison of quantum mechanical calculation with and without an implicit solvent model confirmed that the solvent effect is very low [40].

Michelle M. Kuttel has studied Ramachandran plots of numerous disaccharides by advanced free energy modelling techniques (umbrella sampling with weighed histogram analysis method) [41-43]. For example globobiose (α-D-Gal*p*-(1→4)-β-D-Gal*p*) [43] was studied in vacuum, in explicitly modelled water and also in two simple implicit models (one modelled using a high dielectric constant and the second modelled by a distance-dependent dielectric constant). Free energy surfaces calculated in all environments were very similar. In the major free energy well, there are two shallow sub-minima. In vacuum, the global minimum is located at $\phi_H = 5$ deg, $\psi_H = 35$ deg. This conformation is stabilized by a hydrogen bond between O-2 of the first residue and the O-6–H of the second residue. The secondary minimum ($\phi_H = -40$ deg, $\psi_H = -5$ deg) is higher in free energy by 6.7 kJ/mol. On the other hand, the conformation at $\phi_H = -40$ deg, $\psi_H = -10$ deg was modelled as the global free energy minimum in explicit water. This conformation is stabilized by a water molecule between O-6 of the first residue and O-2 of the second residue. The conformation at $\phi_H = 20$ deg, $\psi_H = 30$ deg was modelled as a local minimum higher by 4.6 kJ/mol in free energy. Interestingly, a similar result was obtained in both implicit water models, despite the fact that these models were very simple.

The water environment modelled by a simple distance-dependent dielectric constant provides the free energy surface in a good qualitative and quantitative agreement with the results from explicit water.

Variation of dielectric constant has been also intensively studied for use in predicting conformations in crystals when using just the isolated disaccharide. French and co-workers have calculated Ramachandran potential energy maps of per-substituted β-maltose [44], sucrose [45], maltose [45] and laminarabiose [45]. These studies employed an MM3 force field for exo-cyclic groups and HF/6-31G* for the backbone atoms. Interestingly, numerous crystal structures were located in higher-energy areas of Ramachandran potential energy surfaces (> 2 kcal/mol) when the default, vacuum dielectric constant wes used. However, much better agreements with crystal structures were obtained for potential surfaces calculated at an elevated relative dielectric constant (7.5 instead of 1.5 recommended for vacuum MM3 studies). A similar effect has been observed independently for all studied disaccharides. These results clearly show that screening of electrostatic interactions plays an important role not only in water solution but in any condensed phase and that accurate electrostatics modelling is necessary to study carbohydrate conformations.

SOLVENT EFFECT ON COMPLEX CARBOHYDRATES

In previous sections we discussed several examples of the solvent effect and its impact on saccharides. Here we highlight one interesting idea employed in conformational modelling of complex carbohydrates by the "carbohydrate hydroxyl groups represented by extended atoms" (CHEAT) empirical potential [46]. Hydroxyl groups can act as donors or acceptors of hydrogen bonding. However, many of such potential hydrogen bonds are screened in water solution. A conformational search or simulation of a complex carbohydrate in vacuum using classical force field would likely lead to structures with artificially enhanced hydrogen bonding, relative to the situation in water. The CHEAT force field attacks this problem by replacing each hydroxyl group by a single particle in a semi-coarse-grained fashion. The hydroxyl particle represents the polar character of the group but its hydrogen bonding capacity is reduced. The effect of water is therefore modelled by this special character of hydroxyl groups. The CHEAT force field has been applied in the original article on modelling of a trisaccharidic structural variant of LewisX antigen [46]. The systematic search of possible conformers followed by geometry optimizations of low energy conformers lead to a prediction of structure in good agreement with experiment (NMR) or molecular dynamics studies in explicitly modelled water and significantly better than *in vacuo* simulations. This result is challenging but testing on a larger set of complex carbohydrates is necessary. Nevertheless, in the opinion of one of authors (V.S.) this 'noncovalent interaction coarse graining' seems to be a fruitful strategy and deserves further development, not only in molecular simulations, but also for example in carbohydrate-protein docking.

MOLECULAR RECOGNITION

Prediction of carbohydrate-protein interactions is considered to be more difficult in comparison to interactions of protein with general drug-like molecules. This is due to the fact that these interactions are relatively weak, binding sites are shallow, saccharides are pseudo-symmetric (multiple hydroxyl groups are present in a similar orientation), an important role is played by CH/π interactions and there are other issues. Another important complication is the fact that carbohydrate-protein interactions are often water-mediated, i.e. that direct hydrogen bonds between the protein and the carbohydrate are replaced by hydrogen bonds from protein to water to carbohydrate (Fig. **1F**). An interesting example is recognition of galactose-containing saccharides and glycosides by heat-labile enterotoxin. Minke and coworkers [47] used this system as a benchmark of carbohydrate-protein docking. The experimental structure of this toxin revealed numerous water molecules in the binding site involved in recognition of the saccharide. Such water molecules are usually removed in standard protein-ligand docking protocols. Minke and co-workers showed that inclusion of water molecules in the receptor site might positively influence the results of docking. For example, when lactose was docked into a dehydrated binding site, the docking resulted in inaccurate binding poses (RMSD > 3Å). However, when the same molecule is docked into the binding site with four water molecules, it leads to a generally correct binding pose (RMSD < 1Å). Some ligands have shown a similar trend, whereas the situation was more complicated in for example *p*-nitrophenyl-β-D-galactopyranoside. Existence of water molecules in a binding site is one of the important factors that must be taken into the account when the binding site of the target is being prepared for docking. Other recently recognised target heterogeneities include a conformational heterogenity, different protonation states and histidine tautomers. The role of water in protein-ligand recognition is another example of phenomena recognised early and studied on carbohydrates as model systems.

CONCLUSION

This review presented examples of studies showing how the water environment influences carbohydrates and how these effects can be modelled. This topic deserves to be compared with other classes of biomolecules. For example, structures of nucleic acids are determined by inter-base hydrogen bonds and stacking interactions. Nucleic acids are also highly negatively charged and the interaction with solvent and counter-ions significantly influences the structure as well as the outcome of molecular modelling. It is therefore necessary to focus on hydrogen bonds, stacking interactions and electrostatics treatment in molecular modelling studies of nucleic acids. Protein structure and folding is driven by a hydrophobic collapse and by formation of hydrogen-bonded secondary structure elements. Hydrophobicity and secondary structure propensities must be carefully addressed in protein simulations. Carbohydrates (usually) do not fold like proteins and they (usually) do not show high net charges like nucleic acids. In our opinion, compared to well known protein- and nucleic-acid-specific issues, the topics addressed in this review are carbohydrate-specific issues that must be accurately addressed to successfully model structure and dynamics of carbohydrates and carbohydrate-protein assemblies.

ACKNOWLEDGEMENT

Support from the Czech Ministry of Education, Youth and Sports (MSM6046137305) is acknowledged. We also thank the reviewers for their fruitful comments and help.

REFERENCES

[1] DeMarco, M.L.; Woods, R.J. Structural glycobiology: a game of snakes and ladders. *Glycobiology* **2008**, *18*(6), 426-440.

[2] de Vries, M.S.; Hobza, P. Gas-phase spectroscopy of biomolecular building blocks. *Annu. Rev. Phys. Chem.* **2007**, *58*, 585-612.

[3] Hünig, I.; Painter, A.J.; Jockusch, R.A.; Carçabal, P.; Marzluff, E.M.; Snoek, L.C.; Gamblin, D.P.; Davis, B.G.; Simons, J.P. Adding water to sugar: a spectroscopic and computational study of α- and β-phenylxyloside in the gas phase. *Phys. Chem. Chem. Phys.* **2005**, *7*(12), 2474-2480.

[4] Carçabal, P.; Jockusch, R.A.; Hünig, I.; Snoek, L.C.; Kroemer, R.T.; Davis, B.G.; Gamblin, D.P.; Compagnon, I.; Oomens, J.; Simons, J.P. Hydrogen bonding and cooperativity in isolated and hydrated sugars: mannose, galactose, glucose, and lactose. *J. Am. Chem. Soc.* **2005**, *127*(32), 11414-11425.

[5] Carçabal, P.; Patsias, T.; Hünig, I.; Liu, B.; Kaposta, C.; Snoek, L.C.; Gamblin, D.P.; Davis, B.G.; Simons, J.P. Spectral signatures and structural motifs in isolated and hydrated monosaccharides: phenyl α- and β-L-fucopyranoside. *Phys. Chem. Chem. Phys.* **2006**, *8*(1), 129-136.

[6] Cocinero, E.J.; Stanca-Kaposta, E.C.; Scanlan, E.M.; Gamblin, D.P.; Davis, B.G.; Simons, J.P. Conformational choice and selectivity in singly and multiply hydrated monosaccharides in the gas phase. *Chem. Eur. J.* **2008**, *14*(29), 8947-8955.

[7] Screen, J.; Stanca-Kaposta, E.C.; Gamblin, D.P.; Liu, B.; Macleod, N.A.; Snoek, L.C.; Davis, B.G.; Simons, J.P. IR-spectral signatures of aromatic-sugar complexes: probing carbohydrate-protein interactions. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*(20), 3644-3648.

[8] Stanca-Kaposta, E.C.; Gamblin, D.P.; Screen, J.; Liu, B.; Snoek, L.C.; Davis, B.G.; Simons, J.P. Carbohydrate molecular recognition: a spectroscopic investigation of carbohydrate-aromatic interactions. *Phys. Chem. Chem. Phys.* **2007**, *9*(32), 4444-4451.

[9] Tapia, O.; Goscinski, O. Self-consistent reaction field theory of solvent effects. *Mol. Phys.* **1975**, *29*(6), 1653-1661.

[10] Miertus, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilization of *ab initio* molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117-129.

[11] Klamt, A.; Schüürmann, G.J. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans.* **1993**, *2*, 799-805.

[12] Berendsen, H.J.C.; Postma, J.P.M.; van Gunsteren, W.; Hermans, J. *Interaction models for water in relation to protein hydration*. In *Intermolecular Forces*; Pullman, B. Ed.; Reidel: Dordrecht, **1981**, pp. 331-342.

[13] Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935.

[14] Jeffrey, G.A.; Pople, J.A.; Radom, L. The application of *ab initio* molecular orbital theory to the anomeric effect. A comparison of theoretical predictions and experimental data on conformations and bond lengths in some pyranoses and methyl pyranosides. *Carbohydr. Res.* **1972**, *25*, 117-131,

[15] Cramer, C.J. Sugar anomerism - a short and sweet digression. *Theor. Chem. Acc.* **2000**, *103*, 308-310.

[16] Tvaroška, I.; Carver, J.P. The anomeric and exo-anomeric effects of a hydroxyl group and the stereochemistry of the hemiacetal linkage. *Carbohydr. Res.* **1998**, *309*(1), 1-9.

[17] Tvaroška, I.; Kožár, T. Theoretical studies on the conformation of saccharides. 3. Conformational properties of the glycosidic linkage in solution and their relation to the anomeric and exoanomeric effects. *J. Am. Chem. Soc.* **1980**, *102*(23), 6929-6936.

[18] Silva, A.M.; da Silva, E.C.; da Silva, C.O. A theoretical study of glucose mutarotation in aqueous solution. *Carbohydr. Res.* **2006**, *341*(8), 1029-1040.

[19] Davies, G.J.; Ducros, V.M.-A.; Varrot, A.; Zechel, D.L. Mapping the conformational itinerary of β-glycosidases by X-Ray crystallography. *Biochem. Soc. Trans.* **2003**, *31*, 523-527.

[20] Marszalek, P.E.; Oberhauser, A.F.; Pang, Y.P.; Fernandez, J.M. Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring. *Nature* **1998**, *396*, 661-664.

[21] Jin, L.; Hricovíni, M.; Deakin, J.A.; Lyon, M.; Uhrín, D. Residual dipolar coupling investigation of a heparin tetrasaccharide confirms the limited effect of flexibility of the iduronic acid on the molecular shape of heparin. *Glycobiology* **2009**, *19*(11), 1185-1196.

[22] Cremer, D.; Pople, J. A general definition of ring puckering coordinates. *J. Am. Chem. Soc.* **1975**, *97*(6), 1354-1358.

[23] Barrows, S.E.; Dulles, F.J.; Cramer, C.J.; French, A.D.; Truhlar, D.G. Relative stability of alternative chair forms and hydroxymethyl conformations of β-D-glucopyranose. *Carbohydr. Res.* **1995**, *276*, 219-251.

[24] Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. The conformational free energy landscape of β-D-glucopyranose. Implications for substrate preactivation in β-glucoside hydrolases. *J. Am. Chem. Soc.* **2007**, *129*(35), 10686-10693.

[25] Spiwok, V.; Králová, B.; Tvaroška, I. Modelling of β-D-glucopyranose ring distortion in different force fields: a metadynamics study. *Carbohydr. Res.* **2010**, *345*(4), 530-537.

[26] Hansen, H.S.; Hünenberger, P.H. Using the local elevation method to construct optimized umbrella sampling potentials: calculation of the relative free energies and interconversion barriers of glucopyranose ring conformers in water. *J. Comput. Chem.* **2010**, *31*(1), 1-23.

[27] Spiwok, V.; Tvaroška, I. Conformational free energy surface of α-*N*-acetylneuraminic acid: an interplay between hydrogen bonding and solvation. *J. Phys. Chem. B* **2009**, *113*(28), 9589-9594.

[28] Tvaroška, I.; Taravel, F.R.; Utille, J.P.; Carver, J.P. Quantum mechanical and NMR spectroscopy studies on the conformations of the hydroxymethyl and methoxymethyl groups in aldohexosides. *Carbohydr. Res.* **2002**, *337*, 353-367.

[29] Kirschner, K.N.; Woods, R.J. Solvent interactions determine carbohydrate conformation. *Proc. Natl. Acad. Sci. USA* **2001**, *98*(19), 10541-10545.

[30] Spiwok, V.; Tvaroška, I. Metadynamics modelling of the solvent effect on primary hydroxyl rotamer equilibria in hexopyranosides. *Carbohydr. Res.* **2009**, *344*(12), 1575-1581.

[31] IUPAC-IUB joint commission on biochemical nomenclature (JCBN) Symbols for specifying conformation of polysaccharide chain. *Eur. J. Biochem.* **1983**, *113*, 5-7.

[32] Csonka, G.I.; Kolossváry, I.; Császár, P.; Éliás, K.; Csizmsdia, I.G. The conformational space of selected aldo-pyrano-hexoses. *J. Mol. Struct. (THEOCHEM)* **1997**, *395-396*, 29-40.

[33] Klein, R.A. Electron density topological analysis of hydrogen bonding in glucopyranse and hydrated glucopyranose. *J. Am. Chem. Soc.* **2002**, *124*(46), 13931-13937.

[34] French, A.D.; Johnson, G.P. Quantum mechanics studies of cellobiose conformations. *Can. J. Chem.* **2006**, *84*(4), 603-612.

[35] Strati, G.L.; Willett, J.L.; Momany, F.A. *Ab initio* computational study of β-cellobiose conformers using B3LYP/6-311++G**. *Carbohydr. Res.* **2002**, *337*(20), 1833-1849.

[36] Cocinero, E.J.; Gamblin, D.P.; Davis, B.G.; Simons, J.P. The building blocs of cellulose: The intrinsic conformational structures of cellobiose, its epimer lactose, and their singly hydrated complexes. *J. Am. Chem. Soc.* **2009**, *131*(31), 11117-11123.

[37] Shen, T.; Langan, P.; French, A.D.; Johnson, G.P.; Gnanakaran, S. Conformational flexibility of soluble cellulose oligomers: Chain length and temperature dependencies. *J. Am. Chem. Soc.* **2009**, *131*(41), 14786-14794.

[38] Liu, Q.; Schmidt, R.K.; Teo, B.; Karplus, P.A.; Brady, J.W. Molecular dynamics studies of the hydration of α,α-trehalose. *J. Am. Chem. Soc.* **1997**, *119*(33), 7851-7862.

[39] Engelsen, S.B.; Pérez, S. Unique similarity of the asymmetric trehalose solid-state hydration and the diluted aqueous-solution hydration. *J. Phys. Chem. B* **2000**, *104*(39), 9301-9311.

[40] French, A.D.; Johnson, G.P.; Kelterer, A.-M.; Dowd, M.K.; Cramer, C.J. Quantum mechanics studies of the intrinsic conformation of trehalose. *J. Phys. Chem. A* **2002**, *106*(19), 4988-4997.

[41] Kuttel, M.M.; Naidoo, K.J. Ramachandran free-energy surfaces for disaccharides: trehalose, a case study. *Carbohydr. Res.* **2005**, *340*(5), 875-879.

[42] Kuttel, M.M.; Naidoo, K.J. Free energy surfaces for the α-(1→4)-glycosidic linkage: implications for polysaccharide solution structure and dynamics. *J. Phys. Chem. B* **2005**, *109*(15), 7468-7474.

[43] Kuttel, M.M. Conformational free energy maps for globobiose (α-D-Gal*p*-(1→4)-β-D-Gal*p*) in implicit and explicit aqueous solution. *Carbohydr. Res.* **2008**, *343*(6), 1091-1098.

[44] Johnson, G.P.; Stevens, E.D.; French, A.D. Octa-*O*-propanoyl-β-maltose: crystal structure, acyl stacking, related structures, and conformational analysis. *Carbohydr. Res.* **2007**, *342*(9), 1210-1222.

[45] French, A.D.; Kelterer, A.-M.; Johnson, G.P.; Dowd, M.K.; Cramer, C.J. Constructing and evaluating energy surfaces of crystalline disaccharides. *J. Mol. Graph. Mod.* **2000**, *18*(2), 95-107.

[46] Grootenhuis, P.D.J.; Haasnoot, C.A.G. A CHARMm based force field for carbohydrates using the CHEAT approach: carbohydrate hydroxyl groups represented by extended atoms. *J. Mol. Simul.* **1993**, *10*(2) 75-95.

[47] Minke, W.E.; Diller, D.J.; Hol, W.G.; Verlinde, C.L. The role of waters in docking strategies with incremental flexibility for carbohydrate derivatives: heat-labile enterotoxin, a multivalent test case. *J. Med. Chem.* **1999**, *42*(10), 1778-1788.

Received: May 24, 2010 Revised: August 06, 2010 Accepted: August 24, 2010