EBSA
Biophysics in Europe

ORIGINAL PAPER

Intramolecular hydrogen-bonding in aqueous carbohydrates as a cause or consequence of conformational preferences: a molecular dynamics study of cellobiose stereoisomers

Dongqi Wang · Maria Lovísa Ámundadóttir · Wilfred F. van Gunsteren · Philippe H. Hünenberger

Abstract It is often assumed that intramolecular hydrogen-bonding (H-bonding) exerts a significant influence on the conformational properties of aqueous (bio-)polymers. To discuss this statement, one should, however, distinguish between solvent-exposed and buried H-bonds, and between their respective roles in promoting stability (i.e., as a driving force) and specificity (for which the term steering force is introduced here). In this study, the role of solvent-exposed H-bonding in carbohydrates as a driving or steering force is probed using explicit-solvent molecular dynamics simulations with local elevation umbrella sampling in the simple context of cellobiose stereoisomers. More specifically, four $\beta(1\rightarrow4)$-linked D-aldohexopyranose disaccharides are considered, which present a different stereochemisty of the potentially H-bonding groups neighboring the glycosidic linkage. Although the epimerization may largely alter the intramolecular *trans*-glycosidic H-bonding pattern, it is found to have only very limited influence on the Ramachandran free-energy map of the disaccharide, a loss of intramolecular H-bonding being merely compensated for by an enhancement of the interaction with the solvent molecules. This finding suggests that solvent-exposed *trans*-glycosidic H-bonding (and in particular the $HO_3'\rightarrow O_5$ H-bond) is not the cause of the 2_1-helical secondary structure characteristic of cellooligosaccharides, but rather the opportunistic consequence of a sterically and stereoelectronically dictated conformational preference. In other words, for these compounds, solvent-exposed H-bonding appears to represent a minor (possibly adverse) conformational driving as well as steering force.

Keywords Computer simulation · Molecular dynamics · Conformational driving forces · Hydrogen bonds · Carbohydrates · Cellobiose

Introduction

Biopolymers such as proteins, nucleic acids and polysaccharides do not adopt random conformations in aqueous solution (Anfinsen 1973; Saenger 1984; Jaenicke 1991; Dobson et al. 1998; Rao et al. 1998; Pérez and Kouwijzer 1999). On a segmental basis, or considering short oligomers, the successive residues along a chain tend to adopt preferential relative orientations at the level of their linkages, resulting in secondary-structure patterns. On a nonlocal basis, and considering longer polymers or multiple chains, the secondary-structure elements may further pack against each other in a specific fashion, resulting in tertiary-structure (intra-chain) or quaternary-structure (inter-chain) patterns. For aqueous proteins and nucleic acids, these conformational preferences are often sufficiently strong to promote the appearance of a unique native state, i.e., of a conformational ensemble with very limited fluctuations around a sequence-defined native structure. For aqueous polysaccharides, the tertiary and quaternary arrangements are typically weaker and less specific, and the secondary-structure patterns more labile and short-ranged, although still far from random.

The quantitative description of these conformational equilibria relies on thermodynamics. It is typically formulated in terms of free-energy changes associated with specific conformational processes, e.g., unfolded to native

D. Wang · M. L. Ámundadóttir · W. F. van Gunsteren · P. H. Hünenberger (✉)
Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zurich, Switzerland
e-mail: phil@igc.phys.chem.ethz.ch

🖄 Springer

state considering tertiary-structure formation, and free to bound state considering quaternary-structure formation or, by extension, ligand binding. In practice, however, the determination of such a free-energy change requires two assumptions: (1) the identification of the considered conformational states with values of specific experimental observables (e.g., spectroscopic signal or biochemical activity); (2) the emulation of the state change under constant environmental conditions by means of a corresponding change induced by a perturbation of these conditions (e.g., temperature, pressure, pH or cosolute concentration), possibly followed by extrapolation to zero perturbation. Considering that the relationship between experimental observables (averages) and molecular conformations (ensembles) is not unique, and that the environmental perturbation may alter the conformational distribution (even after extrapolation), these two factors represent a first important source of ambiguity in the thermodynamic analysis of conformational changes concerning aqueous biopolymers.

The qualitative interpretation of this thermodynamic information typically relies on the concept of driving forces, i.e., classes of interactions or model effects assumed to contribute to a free-energy change in an additive fashion. Classes of interactions refer here to specific components of the potential energy function of the system (e.g., electrostatic, van der Waals or stereoelectronic interactions between specific atom groups), assumed to map directly to an enthalpic driving force. Model effects refer to prototypical behaviors concerning simple systems and extrapolated to the more complex biomolecular situation (e.g., backbone and sidechain entropy, hydrophobic effect), and generally include both an enthalpic and an entropic component. The driving forces commonly invoked to rationalize the conformational preferences of aqueous biopolymers are (Dill 1990; Honig and Yang 1995; Doig and Sternberg 1995; Yang and Honig 1995, 1996; Dill and Bromberg 2003; Baldwin 2007): (1) the hydrophobic effect, promoting the segregation of low-polarity solute residues from the aqueous environment; (2) packing constraints, penalizing any overlap between solute atoms; (3) chain entropy, favoring less structured and, thus, typically less compact conformations; (4) steric and stereoelectronic effects, favoring specific relative orientations of functional groups in close covalent proximity, e.g., across linkages; (5) electrostatic effects, namely charge-charge, charge-dipole and dipole-dipole interactions, including hydrogen-bonding (H-bonding). These factors are extremely difficult to cast into a quantitative form for all but the simplest model systems, especially for those involving an entropic component, which represents a second important source of ambiguity in the thermodynamic analysis of conformational changes. Nevertheless, they remain extremely useful

in a qualitative sense, and represent the basic vocabulary of conformational reasoning throughout polymer science and biochemistry. However, a clear distinction should still be made between the discussion of their role in terms of affinity (stabilization of the native relative to the unfolded conformations, or of the bound relative to the unbound configurations), for which the term driving force will be retained, or of specificity (stabilization of a given native conformation relative to alternative folds, or of a given binding pair relative to other possible pairs), for which the term steering force will be introduced here. The concepts of driving and steering forces are further discussed in an Appendix. The present article questions the relevance of one of the above factors as a driving or steering force, namely H-bonding (Jeffrey 1997).

The formation of secondary, tertiary or quaternary structure in biopolymers, as well as the binding of ligands, are very often accompanied by the formation of new solute-solute H-bonds. The early predictions concerning the secondary-structure elements of proteins (Pauling et al. 1951), nucleic acids (Watson and Crick 1953) and polysaccharides (Gardner and Blackwell 1974; Kolpak and Blackwell 1976; Sarko and Wu 1978) largely relied on the consideration of optimal H-bonding geometries, which proved a successful strategy. Nowadays, computational methods for structure prediction (Samudrala et al. 1999; Al-Lazikani et al. 2001; Cymerman et al. 2004), structure refinement (Brünger and Nilges 1993; Takashima 2006; Adams et al. 2010; Allison et al. 2012) or ligand docking (Sousa et al. 2006; Huang and Zou 2010) typically include a component favoring H-bond formation, which appears to be an essential ingredient also in this context. These observations, along with the fact that H-bonding typically represents a strong interaction in vacuum (Zheng and Merz 1992; Deshmukh et al. 2008; Paton and Goodman 2009; Korth 2010; Sun and Wang 2010; Li et al. 2011; Csonka and Kaminsky 2011), naturally lead to the view that H-bonding ought to be a major factor influencing the conformational and binding preferences of biopolymers, in terms of both affinity and specificity. However, a number of experimental and theoretical results (Dill 1985; Yang and Honig 1995, 1996; Jeffrey 1994; Kräutler et al. 2007; Hünenberger et al. 1999; de Bakker et al. 1999; Seebach et al. 2006; Wang et al. 2009; Spiwok and Tvaroska 2009; Pašalić et al. 2010; Perić-Hassler et al. 2010; Hansen and Hünenberger 2011) have suggested that this interpretation is probably too simplistic.

When the conformational process under study is accompanied by the association of two solvent-exposed partners to form a buried H-bond, the cost associated with the removal of the H-bonding partners from the solution environment (desolvation) offsets the electrostatic gain upon formation of the interaction itself. However, if the

desolvation of two potentially H-bonding groups is not accompanied by H-bond formation in a given conformation, this conformation will be penalized by the desolvation term. Considering the high polarity (dielectric permittivity) of water, a reasonable ansatz for first-order reasoning is that the desolvation and electrostatic terms are of comparable magnitudes. As a result, the formation of a buried H-bond can be thought of as representing a minor (possibly negligible or even, in some cases, opposing) conformational driving force, but an important conformational steering force.

When the conformational process under study is accompanied by the association of two solvent-exposed partners to form a solvent-exposed H-bond, the H-bonding partners remain in a highly polar environment, and their interaction is screened by the solvent dielectric response as well as subject to H-bonding competition by the solvent molecules. In the case of an aqueous environment, a reasonable ansatz for first-order reasoning appears to be that solvent-exposed H-bond formation is a neutral (no gain, no cost) process. As a result, the formation of a such an H-bond is expected to represent a minor conformational driving as well as steering force. According to this interpretation, solvent-exposed intramolecular H-bonding in an aqueous environment should be viewed as an opportunistic consequence of the close proximity of two H-bonding groups in a given molecular conformation, rather than as a force contributing to the stability of this specific conformation.

The latter issue is particularly relevant in the context of aqueous carbohydrates, for which intramolecular H-bonds are typically solvent-exposed. For example, the leading secondary-structure pattern for $\beta(1{\to}4)$-linked D-glucopyranose chains, as occurring, e.g., in cellobiose (Fig. 1) and longer cellooligosaccharides, is the 2_1-helix (Pérez and Vergelati 1985; Pérez and Samain 2010), compatible with a high-occurrence $HO_3'{\to}O_5$ trans-glycosidic H-bond, both also found in crystalline cellobiose (Jacobson et al. 1961; Chu and Jeffrey 1968), cellotetraose (Gessler et al. 1994) and cellulose (Gardner and Blackwell 1974; Kolpak and Blackwell 1976; Sarko and Wu 1978; Pérez and Samain 2010). Cellobiose and larger cellooligomers have been the target of numerous theoretical investigations nearly all of which have evidenced a dominant 2_1-helical conformation with a high-occurrence $HO_3'{\to}O_5$ H-bond in an aqueous environment. Note that the most recent quantum-mechanical calculations, along with some experimental evidence, suggest the dominance of another conformer in vacuum and low-polarity solvents (Strati et al. 2002; Jockusch et al. 2004; Bosma et al. 2006; French and Johnson 2006; Cocinero et al. 2009; Momany and Schnupf 2010). However, in spite of its high occurrence in aqueous solution, the

above discussion would suggest that the solvent-exposed $HO_3'{\to}O_5$ H-bond is not the cause for the 2_1-helical conformational preference, but rather a consequence of this preference, itself dictated by other effects, predominantly steric and stereoelectronic effects specific to the $\beta(1{\to}4)$ glycosidic linkage (Naidoo and Chen 2003; Kräutler et al. 2007; Perić-Hassler et al. 2010). This interpretation does not exclude the possibility of an important role for this H-bond as a determinant of the physico-chemical and mechanical properties (e.g., stability, rigidity, and insolubility) of crystalline cellulose (Kadla and Gilbert 2000; Umemura et al. 2005; Pérez and Samain 2010; Bergenstråhle et al. 2010), considering that it changes from a solvent-exposed to a buried status upon chain association (Peri et al. 2011). Note, that one should be cautious when referring to the factors determining the structure of crystalline cellulose, considering the existence of multiple allomorphs from different natural (biosynthesis) or industrial (treatment) origins (Pérez and Samain 2010), i.e., one cannot refer to a unique structure produced under strict thermodynamic control.

The aim of the present article is to test these ideas using explicit-solvent molecular dynamics (MD) simulation in the simple context of aqueous disaccharides. More specifically, four $\beta(1{\to}4)$-linked D-aldohexopyranose disaccharides are considered (Fig. 1), namely cellobiose and its C_2 or/and C_3' epimers, which present a different stereochemisty for the potentially H-bonding groups neighboring the glycosidic linkage. If solvent-exposed H-bonding represents a significant conformational driving force, the epimerization at C_2 or/and C_3' is expected to induce important changes in the distribution of the glycosidic dihedral angles ϕ and ψ, depending on the specific trans-glycosidic H-bonding patterns accessible to each of the four disaccharides. Conversely, the absence of significant differences would be compatible with the above suggestion of a minor conformational influence for solvent-exposed H-bonds.

The testing of this influence via isomerization presents two key advantages over a corresponding investigation via functional group alteration. First, the "removal" of a H-bond by deletion (e.g., hydroxyl group $\to$ hydrogen atom) or by modification (e.g., hydroxyl group $\to$ fluoride atom or alkyl group) of one or both of the H-bonding partners, as sometimes applied theoretically (Mendonca et al. 2002; French et al. 2005; Gattin et al. 2007; Eichenberger et al. 2010) or experimentally (Streefkerk and Stephen 1976; Withers et al. 1988; Bock et al. 1988; Abraham et al. 1994; López de la Paz et al. 2002; Seebach et al. 2002; Deechongkit et al. 2004), also unavoidably involves a major alteration of the interaction and solvation pattern, in the non-H-bonded as well as in the H-bonded
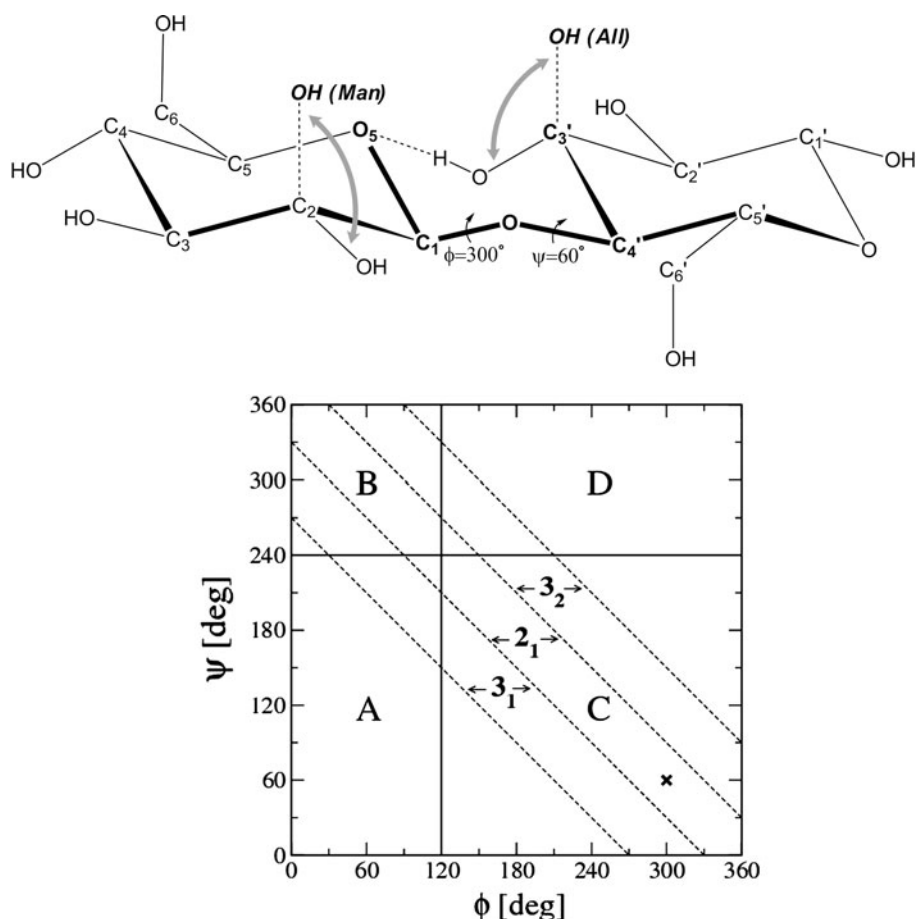
Fig. 1 Four $\beta(1\rightarrow4)$-linked D-aldohexopyranose disaccharides considered in the present study (*top*) and definition of the main conformational regions on the corresponding Ramachandran (ϕ, ψ) maps (*bottom*). The four disaccharides considered are Glc$p$-$\beta(1\rightarrow4)$-Glc$p$-β (GG, cellobiose), and its C_2 or/and C'_3 epimers, namely Glc$p$-$\beta(1\rightarrow4)$-All$p$-β (GA), Man$p$-$\beta(1\rightarrow4)$-Glc$p$-β (MG), and Man$p$-$\beta(1\rightarrow4)$-All$p$-β (MA), where Glc$p$, Man$p$ and All$p$ stand for D-glucopyranose, D-mannopyranose and D-allopyranose, respectively. The disaccharides are simulated with β anomery at the reducing residue (primed atom labels) and nearly exclusively found in the $^4C_1$ chair conformation of the two rings during the simulations (as displayed). The glycosidic dihedral angles ϕ and ψ are defined by the atom sequences O_5–C_1–O_1–C'_4 and C_1–O_1–C'_4–C'_4, respectively. The conformational regions A, B, C and D of the Ramachandran maps are defined by cutoff values of 120 and 240° for ϕ and ψ, respectively, according to Perić-Hassler et al. (2010) Approximate areas compatible with the formation of 3_1-, 2_1- or 3_2-helical secondary-structure motives are also indicated, defined by cutoff values of 270, 330, 390 and 450° for $\psi - \phi$. The limiting lines were determined empirically by analyzing the correlation between ϕ and ψ values and local helical conformation (as determined by the interresidue turn angle θ) in previous simulations of (formally-infinite) polyuronate chains (Perić et al. 2008; Perić-Hassler and Hünenberger 2010). The top drawing approximately corresponds to the conformation at the free-energy minimum for cellobiose, located in region C of the map (indicated by a cross). More details on the labelling can be found in Hansen and Hünenberger (2011) (see section "nomenclature and definitions" therein)

conformations. Second, from a computational point of view, it also allows for the use of closely similar force-field descriptions in terms of bonding, torsional and non-bonded interaction parameters (including atomic charges) for all compounds considered, resulting in cancellation of force-field errors. Note that another approach for this investigation is to alter the solvent polarity, as is sometimes applied computationally (Tvaroška and Kožar 1986; Spiwok and Tvaroska 2009; Pašalić et al. 2010) and experimentally (Lemieux and Brewer 1973; de Vries and Buck 1987; Rockwell and Grindley 1998; Roën et al. 2002; Mayato et al. 2004).

Computational details

All MD simulations were carried out using the GROMOS md^{++} program (Schmid et al. 2012) and the GROMOS 45A4 force field (Lins and Hünenberger 2005; see also Kräutler et al. 2007; Hansen and Hünenberger 2011), involving a united-atom description of aliphatic groups, along with the simple point charge (SPC) water model (Berendsen et al. 1981). The simulations involved the four alternative $\beta(1\rightarrow4)$-linked D-aldohexopyranose disaccharides (with β anomery at the reducing residue) illustrated in Fig. 1, namely Glc$p$-$\beta(1\rightarrow4)$-Glc$p$-β (GG, cellobiose),

Glc$p$-β(1→4)-All$p$-β (GA), Man$p$-β(1→4)-Glc$p$-β (MG) and Man$p$-β(1→4)-All$p$-β (MA), where Glc$p$, Man$p$ and All$p$ stand for D-glucopyranose, D-mannopyranose and D-allopyranose, respectively. Considering the most stable $^4C_1$ chair conformation of the two residues and labelling the atoms of the reducing residue with a prime, these disaccharides present equatorial orientations of the hydroxyl groups at C_3, C_4, C'_1 and C'_2 as well as of the hydroxymethyl groups at C_5 and C'_5, along with different orientations of the hydroxyl groups at C_2 and C'_3, namely equatorial-equatorial (GG), equatorial-axial (GA), axial-equatorial (MG) or axial-axial (MA). Note that the parameters of the four disaccharides in the 45A4 force field differ exclusively in terms of the sign of the reference improper-dihedral angle controlling the stereochemistry at carbon atoms C_2 and C'_3, i.e., the covalent, torsional and non-bonded interaction parameters of the four compounds are otherwise rigorously identical.

The simulations were performed under periodic boundary conditions based on cubic computational boxes of initial edge length 4 nm, containing one disaccharide molecule and 2081 (GG) or 2083 (GA, MG and MA) water molecules, and at 1 atm and 300 K. The leap-frog algorithm (Hockney 1970) was used to integrate Newton's equations of motion with a timestep of 2 fs. Solute bond-length constraints as well as the full rigidity of the water molecules were enforced by application of the SHAKE procedure (Ryckaert et al. 1977) with a relative geometric tolerance of 10^{-4}. The non-bonded interactions were calculated using a twin-range scheme (Berendsen et al. 1986), with short- and long-range cutoff distances set to 0.8 and 1.4 nm, respectively, and an update frequency of 5 timesteps for the short-range pairlist and intermediate-range interactions. A reaction-field correction (Barker and Watts 1973; Tironi et al. 1995) was applied to account for the mean effect of electrostatic interactions beyond the long-range cutoff distance, using a relative dielectric permittivity of 61 as appropriate (Heinz et al. 2001) for the SPC water model. The pressure was maintained close to its reference value of 1 atm by weakly coupling the particle coordinates and box dimensions (isotropic scaling) to an external bath (Berendsen et al. 1984) using a relaxation time of 0.5 ps and an isothermal compressibility of (van Gunsteren et al. 1996) 4.575×10^{-4} kJ^{-1} mol nm^3. The temperature was maintained close to its reference value of 300 K by weakly coupling solute and solvent degrees of freedom separately to external baths (Berendsen et al. 1984) using a relaxation time of 0.1 ps. The translation of the box center of mass was removed every timestep.

Two simulations were undertaken for each of the four disaccharides considered (Fig. 1). In a first set of simulations, plain MD was applied for a duration $t_{MD} = 50$ ns. In a second set of simulations, the local elevation umbrella sampling (LEUS) scheme (Hansen and Hünenberger 2010a; see also Perić-Hassler et al. 2010; Hansen et al. 2010; Hansen and Hünenberger 2010b; Hansen and Hünenberger 2011) was applied to enhance the conformational sampling in the space of the glycosidic dihedral angles ϕ and ψ (Fig. 1). This approach combines the advantages of the local elevation (Huber et al. 1994) (LE) conformational searching method and of the umbrella sampling (Torrie and Valleau 1977; Valleau and Torrie 1977) (US) conformational sampling method. More specifically, the LEUS scheme relies on two steps: (1) a LE build-up (searching) phase, that is used to construct an optimized biasing potential within a subspace of conformationally relevant degrees of freedom; (2) an US sampling phase, that is used to generate a biased ensemble with extensive coverage of the selected conformational subspace. The LE build-up phase consists of a MD simulation involving the progressive construction of a memory-based (i.e., time-dependent) biasing potential that penalizes the resampling of previously visited regions within the considered conformational subspace. The US sampling phase consists of a MD simulation involving a "frozen" (i.e., time-independent) biasing potential, which is set equal to the biasing potential reached at the end of the LE build-up phase. A successful build-up phase will produce a biasing potential that is approximately equal to the negative of the free-energy hypersurface within the considered conformational subspace, so that a sufficiently long sampling phase will result in a nearly homogeneous coverage of this subspace. In addition, because the biasing potential in this second phase is time-independent, thermodynamic information relevant for the physical (unbiased) ensemble can be recovered from the simulated data by means of a simple reweighting procedure (Torrie and Valleau 1977; Valleau and Torrie 1977; Hansen and Hünenberger 2010a, b). These LEUS simulations relied on truncated-polynomial basis functions to represent the biasing potential as described in Hansen et al. (2010) (appendix A, Eqs. (A.5), (A.6) and (A.9) therein). They involved a LE phase of duration $t_{LE} = 50$ ns, followed by an US phase of duration $t_{US} = 50$ ns. The LEUS parameters were similar to those employed in a previous work on disaccharides (Perić-Hassler et al. 2010) (except for the choice of different basis functions), namely $N_l = 2$ dimensions for the (ϕ, ψ) subspace where the LEUS enhancement was applied, $N_g = 32$ grid points for the discretization of this subspace along each of its N_l dimensions, and $k_{LE} = 10^{-4}$ kJ mol^{-1} for the force-constant increment per visit. Note that the latter value refers to Eq. (A.5) in Hansen et al. (2010) and that the polynomial widths d are set equal to the grid spacing as suggested therein.

Idealized conformations with $\phi = 300°$, $\psi = 60°$, and both rings in the $^4C_1$ chair conformation (Fig. 1) were used as initial structures for all disaccharides. For each disaccharide, immersion into the solvent, energy minimization and assignment of atomic velocities was followed by equilibration simulations of 0.1 ns at 1 atm and 300 K. The resulting configurations were used as starting points for the 12 simulations (plain MD as well as successive LE and US phases of the LEUS runs). Atomic coordinates were saved every 0.5 ps (plain MD simulations) or 0.2 ps (US phase of the LEUS simulations) for subsequent analysis.

Because ring conformational transitions (chair ↔ inverted-chair or chair ↔ boat) occur on the 0.05–1 µs timescale (Polacek et al. 2002; Hagen and Kaatze 2004; Behrends and Kaatze 2005), and although the corresponding alternative conformations contribute negligibly to the conformational ensembles of most aldohexopyranoses under ambient conditions (Angyal 1969), their occasional occurrence may compromise the statistical accuracy of the simulated data on the 50 ns timescale. For this reason, it was verified that the two pyranose rings of the different disaccharides remained in the $^4C_1$ conformation during all simulations. Two simulations initially evidenced the occurrence of a transition, and were, therefore, repeated with different initial velocities. According to the criterion provided in Hansen and Hünenberger (2010a) [AD assignment scheme with maximal deviations of 15° in terms of the Pickett and Strauss coordinates (Pickett and Strauss 1970)], the occurrence of the $^4C_1$ conformation is of at least 98.5 % for all the simulations (data not shown), the remainder representing slightly distorted $^4C_1$ chair conformations.

The LEUS simulations were analyzed in terms of: (1) free-energy maps $G(\phi, \psi)$ in the space of the glycosidic dihedral angles ϕ and ψ, along with corresponding minima G_m and their locations ϕ_m and ψ_m; (2) populations p of the different states (conformational basins), along with corresponding relative free energies G_p, average glycosidic dihedral angles $\overline{\phi}$ and $\overline{\psi}$, and root-mean-square fluctuations $\delta\phi$ and $\delta\psi$; and (3) occurrences f_i of intramolecular H-bonds in the different states. The corresponding analysis procedures, involving in particular an appropriate re-weighting of the configurations so as to remove the effect of the biasing potential energy term, are detailed in Perić-Hassler et al. (2010) and will not be repeated here. The free-energy maps were calculated using a grid spacing of 6° and anchored ($G = 0$ kJ mol^{-1}) at their global minimum. The value of G at grid points that were never visited during a simulation, which is formally infinite, was arbitrarily set to the maximal value G_{max} of G over all grid points that were visited at least once. For the ease of discussion, the free-energy maps are partitioned into four conformational

regions (A, B, C and D). The corresponding local minima were determined by quadratic interpolation based on the nearest-neighbor grid values. Approximate areas of the maps compatible with the formation of 3_1-, 2_1- or 3_2-helical secondary-structure motives are also indicated (see region definitions in Fig. 1 and its caption). The occurrence of intramolecular H-bonds was also analyzed, separately for each of the conformational regions A, B, C and D, considering all hydroxyl groups as potential H-bond donors and all hydroxyl or ring oxygen atoms as potential acceptors. The presence of an H-bond was defined by a maximal hydrogen–oxygen distance of 0.25 nm and a minimal oxygen–hydrogen–oxygen angle of 135°.

The plain MD simulations were analyzed in terms of the free-energy maps $G(\phi, \psi)$ in the space of the glycosidic dihedral angles ϕ and ψ, for comparison with the LEUS results. The occurrences f_s of solute-solvent H-bonds involving the atoms HO$'_3$, O$'_3$ and O$_5$ were also monitored based on the entire conformational ensemble (no resolution into regions A, B, C and D) and distinguishing between configurations presenting H-bonds involving one or two water molecules.

Results and discussion

The free-energy maps $G(\phi, \psi)$ obtained from the 50 ns plain MD simulations of the four disaccharides considered (Fig. 1) are displayed in Fig. 2. The corresponding maps calculated from the 50 ns sampling phases of the LEUS simulations are shown in Fig. 3.

As observed in previous work on the glucose-based disaccharides (Pereira et al. 2006; Pereira et al. 2007; Perić-Hassler et al. 2010), plain MD sampling on this timescale involves very scarce transitions around the two glycosidic dihedral angles. The simulations predominantly visit a single free-energy basin in the neighborhood of the global minimum, here in region C of the maps. Although alternative metastable conformational states are apparent in regions A and D, the latter for GG and GA only, the number of C↔A and C↔D transitions is insufficient to permit a reliable estimation of the corresponding relative populations and free energies.

As seen previously (Perić-Hassler et al. 2010), the comparison of Figs. 2 and 3 reveals the sampling enhancement afforded by the application of the LEUS protocol at identical sampling times. While the plain MD simulations explored conformations up to a free energy of about 30 kJ mol^{-1} above the global minimum, this threshold is extended to about 50 kJ mol^{-1} in the LEUS simulations. Although the maps issued from the two types
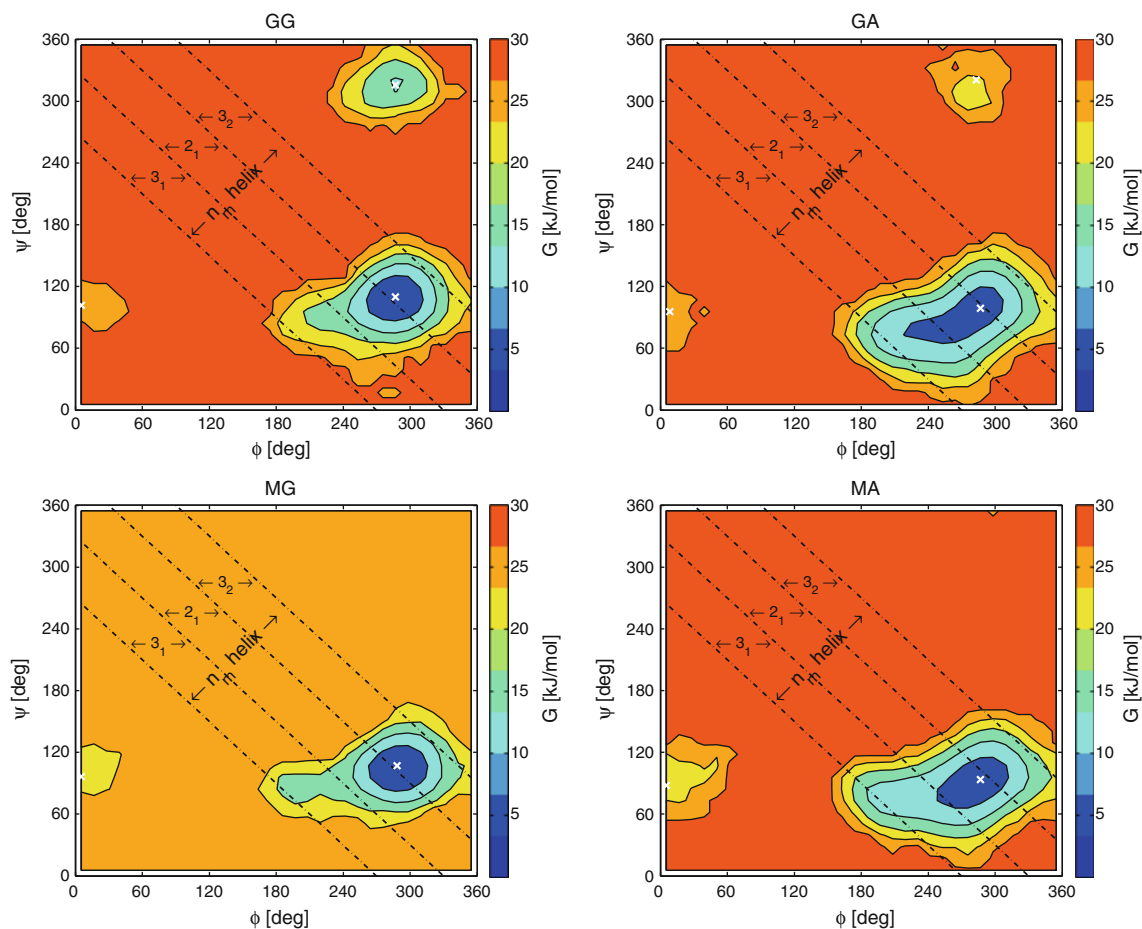
Fig. 2 Free energy maps $G(\phi, \psi)$ in the space of the glycosidic dihedral angles ϕ and ψ for the four $\beta(1 \rightarrow 4)$-linked D-aldohexo-pyranose disaccharides considered (Fig. 1), calculated based on the corresponding 50 ns plain MD simulations in water at 1 atm and 300 K. The conformational regions A, B, C and D of the maps, and the approximate areas compatible with the formation of 3_1-, 2_1- or 3_2-helical secondary-structure motives, are defined in Fig. 1. The locations of the global and local minima are marked with white crosses (no minimum found in region D for MG and MA). The maps are anchored to $G = 0 \, \text{kJ mol}^{-1}$ at the location of their global minimum, and the value of G at grid points that were never visited during the simulations is arbitrarily set to the maximal value G_{max} of G over all grid points that were visited at least once

of simulations closely resemble each other below about 10 kJ mol^{-1}, the LEUS enhancement is required for an appropriate sampling of the higher free energy area of region C and of the alternative metastable states A and D, and in particular of basin D for MG and MA, which was not visited in the plain MD simulations. The numbers of C↔A and C↔D transitions in the biased simulation, on the order of 50–100 over 50 ns, are now sufficient to permit a reliable estimation of the corresponding relative populations and free energies.

Considering the above observations, only the results of the LEUS simulations will be further discussed. For convenience, the fractional populations p and relative free energies G_p of the different states, the locations (ϕ_m, ψ_m) and relative heights G_m of the associated free-energy minima, and the corresponding average values $(\overline{\phi}, \overline{\psi})$ and

fluctuations $(\delta\phi, \delta\psi)$ of the glycosidic dihedral angles are reported in Table 1.

Visual inspection of the free-energy maps in Fig. 3 reveals the following main features. All maps are characterized by a lowest free-energy basin in region C, with a global minimum at about $(\phi_m, \psi_m) = (287°, 100 \pm 10°)$. The corresponding average values of the glycosidic dihedral angles over basin C are about $(\overline{\phi}, \overline{\psi}) = (280° \pm 10, 100 \pm 10°)$, with fluctuations $\delta\phi$ and $\delta\psi$ on the order of 10–30°. These values are close to the available experimental estimates for cellobiose (GG) in water as inferred from nuclear magnetic resonance data, namely (Cheetham et al. 2003) (272°, 99°), or in the crystal as inferred from X-ray crystallographic data, namely (Jacobson et al. 1961; Chu and Jeffrey 1968) (284°, 108°). They are also close to the values typically found using other recent empirical force
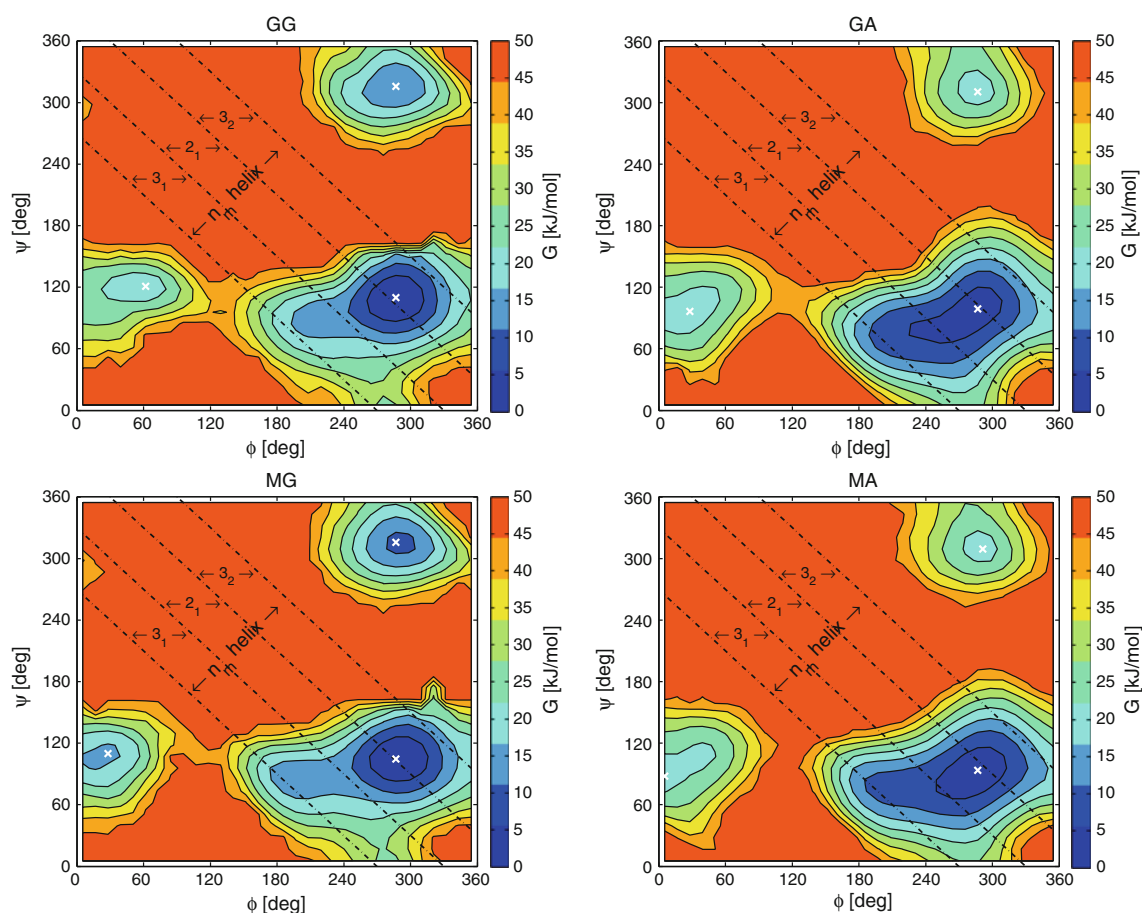
Fig. 3 Free energy maps $G(\phi, \psi)$ in the space of the glycosidic dihedral angles ϕ and ψ for the four $\beta(1 \rightarrow 4)$-linked D-aldohexopyranose disaccharides considered (Fig. 1), calculated based on the 50 ns sampling phases of the corresponding LEUS simulations in water at 1 atm and 300 K. The conformational regions A, B, C and D of the maps, and the approximate areas compatible with the formation of 3_1-, 2_1- or 3_2-helical secondary-structure motives, are defined in Fig. 1. The locations of the global and local minima are marked with white crosses. The maps are anchored to $G = 0 \, \text{kJ mol}^{-1}$ at the location of their global minimum, and the value of G at grid points that were never visited during the simulations is arbitrarily set to the maximal value G_{max} of G over all grid points that were visited at least once

fields (Stortz et al. 2009) (note that the ψ values therein should be decreased by 120° to match the definition adopted here). This conformation, sometimes referred to in the literature (Höög et al. 2001; Larsson et al. 2004; Perić-Hassler et al. 2010; Hansen and Hünenberger 2011) as the *syn-ϕ* and *syn-ψ* conformation, is overwhelmingly populated, with a fractional population p above 96 % in all cases.

The alternative metastable conformations A and D, associated with marginal populations of at most about 3 %, correspond to different values of ϕ (about 25–45° in terms of $\overline{\phi}$) or ψ (about 315° in terms of $\overline{\psi}$), respectively. These alternative metastable conformations A and D are referred to as *anti-ϕ* and *anti-ψ* conformations, respectively. The metastable conformation B (*anti-ϕ* and *anti-ψ* conformation) is apparently associated with a much higher relative free energy ($>50 \, \text{kJ mol}^{-1}$) and not visited at all in the present simulations.

The above observations can easily be rationalized on the basis of steric and stereoelectronic considerations. The *exo*-anomeric effect (Lemieux and Koto 1974; Pérez and Marchessault 1978; Thogersen et al. 1982; Graczyk and Mikolajczyk 1994; Tvaroška and Carver 1998; Rao et al. 1998) implies a stereoelectronic preference for $\phi \approx 60°$ (g^+) and $\phi \approx 300°$ (g^-). For a non-reducing residue involved in a β-linkage, the former value of the dihedral angle is disfavored by steric constraints, resulting in the dominance of conformation C over conformation A. Steric effects also explain the preference for $\psi \approx 60°$ (g^+) and $\psi \approx 300°$ (g^-) in disaccharides where the C_4' hydroxyl group is *anti* relative to the hydroxymethyl group in the reducing residue. For this configuration of the reducing residue, the preference for the latter value of the dihedral angle is probably also in part of stereoelectronic origin, considering that the same preference is already observed

Table 1 Conformational and hydrogen-bonding properties of the four $\beta(1 \rightarrow 4)$-linked D-aldohexopyranose disaccharides considered (Fig. 1), calculated based on the 50 ns sampling phases of the corresponding LEUS simulations in water at 1 atm and 300 K

	A	C	D
GG			
p (%)	0.16	98.06	1.78
G_p [kJ mol^{-1}]	16.0	0.0	10.0
G_m (kJ mol^{-1})	16.4	0.0	10.2
(ϕ_m, ψ_m) [°]	(62°, 121°)	(287°, 110°)	(287°, 316°)
$(\overline{\phi}, \overline{\psi})$ [°]	(47°, 118°)	(286°, 108°)	(286°, 316°)
$(\delta\phi, \delta\psi)$ [°]	(20°, 11°)	(16°, 12°)	(15°, 11°)
f_i [%]	HO2 → O$'_3$ (51.8)	HO$'_3$ → O$_5$ (80.1)	HO$'_3$ → O$_2$ (0.4)
	HO$'_3$ → O$_2$ (16.7)		
f_s [%]		HO$'_3$ → OW (12/0), HW → O$'_3$ (3/0), HW → O$_5$ (71/12)	
GA			
p [%]	0.16	99.74	0.10
G_p [kJ mol^{-1}]	16.0	0.0	17.3
G_m [kJ mol^{-1}]	15.5	0.0	15.1
(ϕ_m, ψ_m) [°]	(28°, 96°)	(287°, 96°)	(287°, 309°)
$(\overline{\phi}, \overline{\psi})$ [°]	(29°, 99°)	(267°, 89°)	(283°, 316°)
$(\delta\phi, \delta\psi)$ [°]	(17°, 16°)	(30°, 18°)	(13°, 15°)
f_i [%]	HO$'_6$ → O$_5$ (3.3)	HO2 → O$'_6$ (3.9)	HO$'_3$ → O$_5$ (0.2)
	HO6 → O$'_6$ (1.0)	HO$'_3$ → O$_5$ (0.3)	
		HO$'_6$ → O$_2$ (0.8)	
f_s [%]		HO$'_3$ → OW (88/0), HW → O$'_3$ (28/0), HW → O$_5$ (71/9)	
MG			
p [%]	0.48	96.25	3.27
G_p [kJ mol^{-1}]	13.2	0.0	8.4
G_m [kJ mol^{-1}]	13.9	0.0	8.4
(ϕ_m, ψ_m) [°]	(28°, 110°)	(287°, 104°)	(287°, 316°)
$(\overline{\phi}, \overline{\psi})$ [°]	(24°, 106°)	(289°, 104°)	(290°, 315°)
$(\delta\phi, \delta\psi)$ [°]	(14°, 12°)	(19°, 12°)	(13°, 10°)
f_i [%]	HO2 → O$'_3$ (64.4)	HO$'_3$ → O$_5$ (80.6)	
	HO$'_3$ → O$_2$ (34.0)		
	HO6 → O$'_6$ (0.3)		
f_s [%]		HO$'_3$ → OW (12/0), HW → O$'_3$ (6/0), HW → O$_5$ (70/11)	
MA			
p [%]	0.13	99.79	0.08
G_p [kJ mol^{-1}]	16.6	0.0	17.6
G_m [kJ mol^{-1}]	16.0	0.0	16.0
(ϕ_m, ψ_m) [°]	(6°, 88°)	(287°, 93°)	(291°, 309°)
$(\overline{\phi}, \overline{\psi})$ [°]	(26°, 99°)	(276°, 89°)	(287°, 313°)
$(\delta\phi, \delta\psi)$ [°]	(17°, 17°)	(29°, 16°)	(13°, 14°)
f_i [%]	HO$'_6$ → O$_5$ (3.4)	HO$'_4$ → O$_6$ (0.1)	HO$'_3$ → O$_5$ (0.2)
	HO2 → O$'_3$ (1.0)	HO$'_3$ → O$_5$ (0.2)	
	HO6 → O$'_6$ (1.1)		
f_s [%]		HO$'_3$ → OW (88/0), HW → O$'_3$ (23/0), HW → O$_5$ (75/9)	

The data is reported separately considering the conformational regions A, C and D of the corresponding free-energy maps in the space of the glycosidic dihedral angles ϕ and ψ (Fig. 3), as defined in Fig. 1. State B was never sampled and is therefore omitted. The reported quantities are the fractional populations p and relative free energies G_p of the different states, the locations (ϕ_m, ψ_m) and relative heights G_m of the associated free-energy minima, the corresponding average values $(\overline{\phi}, \overline{\psi})$ and fluctuations $(\delta\phi, \delta\psi)$ of the glycosidic dihedral angles, and the corresponding occurrences f_i and f_s of intramolecular and solute-solvent hydrogen bonds (H-bonds). The free energies G_m and G_p are reported relative to the corresponding lowest value (state C). The occurrences f_i of intramolecular H-bonds are reported on a per-state basis, i.e., relative to the overall population of the state (which is typically very small for states A and D). Only H-bonds with occurrences of at least 0.1 % are indicated. The occurrences f_s of solute-solvent H-bonds are reported based on the entire conformational ensemble, and were evaluated independently from 1 ns plain MD simulations of the four disaccharides. Only the atoms HO$'_3$, O$'_3$ and O$_5$ were considered for this analysis. The analysis distinguishes between configurations presenting H-bonds involving one (first value between parentheses) or two (second value between parentheses) water molecules. Configurations with H-bonds involving three or more water molecules had negligible occurrences in all cases. The calculation of all the above quantities (except f_s) involved appropriate reweighting of the configurations so as to remove the effect of the biasing potential energy term

for the χ_4 dihedral angle of the unfunctionalized mono-saccharide (Kräutler et al. 2007).

The free-energy maps of the four disaccharides considered are qualitatively very similar. However, closer inspection reveals that epimerization of cellobiose at C_2 (GG→MG or GA→MA) has almost no visible influence on the map, whereas epimerization at C_3' (GG→GA or MG→MA) slightly alters its detailed features. In the latter case, the change of the reducing residue from Glc to All induces a slight elongation the dominant basin of region C in the direction of lower ϕ and ψ values. For GG and MG, this basin is centered at the border between the regions compatible with 3_2- and 2_1-helical secondary-structure motives (see region definitions in Fig. 1 and its caption). As a result of the elongation, for GA and MA, the basin extends more significantly into the region compatible with a 2_1-helical motif, and even slightly into that compatible with a 3_1-helical motif. The global free-energy minimum is displaced towards lower ψ values (by about 10–15°) and the fluctuations $\delta\phi$ and $\delta\psi$ increase slightly (by about 10 and 5°, respectively). The epimerization also tendentially increases the relative free energies G_m of the metastable conformations. For GG and MG, the metastable conformations A and D are associated with free energies of about 14–16 and 8–10 kJ mol^{-1}, respectively, relative to C. These relative free energies are consistently lower by about 2 kJ mol^{-1} for MG compared to GG. For GA and MA, the metastable conformations A and D both correspond to free energies of about 15–16 kJ mol^{-1}, respectively, relative to C. These relative free energies are now higher by about 1 kJ mol^{-1} for MA compared to GA.

Of particular relevance in the present context are: (1) the absence of dramatic differences between the four maps; (2) the slight displacement of the lowest free energy basin towards the 2_1-helical region and the increase of about 5–8 kJ mol^{-1} in the relative free energy of conformation D resulting from the epimerization at C_3' (GG→GA or MG→MA).

The observed intramolecular H-bonds and the corresponding occurrences f_i are also reported in Table 1 on a per-conformation basis, i.e., relative to the overall population of the conformation (which is typically very small for conformations A and D). The occurrences f_s of solute-solvent H-bonds involving the atoms HO$_3'$, O$_3'$ and O$_5$ are also indicated, based on the entire conformational ensemble and distinguishing between configurations presenting H-bonds involving one or two water molecules.

In agreement with the results of previous simulation studies of cellobiose and longer cellooligosaccharides and with the structure of crystalline cellobiose and cellulose (Jacobson et al. 1961; Chu and Jeffrey 1968; Gardner and Blackwell 1974; Kolpak and Blackwell 1976; Sarko and

Wu 1978), GG presents a very high occurrence HO$_3'$→O$_5$ *trans*-glycosidic H-bond (80 % occurrence) in its dominant (98 % population) conformation C. However, conformation A, although marginally populated (0.2 % population), is also compatible with a high occurrence flip-flop O$_3'$↔O$_2$ *trans*-glycosidic H-bond (total 68 % occurrence). Similar considerations apply to MG, where conformation C (96 % population) presents the same high occurrence HO$_3'$→O$_5$ H-bond (81 % occurrence), although conformation A (<1 % population) is also compatible with the high occurrence flip-flop O$_3'$↔O$_2$ H-bond (total 98 % occurrence). The latter H-bond is actually enhanced in MG compared to GG, due to the axial orientation of the hydroxyl group at C_2. In both cases, the hydroxyl group at C_3' is not engaged in significant H-bonding with the solvent, while O$_5$ typically accepts H-bonds from one solvent molecule (occasionally two), in addition to the intramolecular HO$_3'$→O$_5$ H-bond.

Epimerization at C_3' (GG→GA or MG→MA) affects the H-bonding pattern very significantly. Neither the HO$_3'$→O$_5$ H-bond in the dominant conformation C, nor the flip-flop O$_3'$↔O$_2$ H-bond in the marginally populated conformation A can be formed in this case. As a result, the two disaccharides only present intramolecular H-bonds with very low occurrences (at most 4 %). The loss of intramolecular H-bonding is compensated for by an enhancement of solute-solvent H-bonding. The hydroxyl group at C_3' is now typically engaged in H-bonding with one solvent molecule, either as a donor or as an acceptor, while the solute-solvent H-bonding pattern of O$_5$ is not significantly altered.

A simple conformational inspection of the four disaccharides (supported by energy minimizations in vacuum; data not shown) suggests the existence of many other potentially H-bonding conformations. For example, the HO$_3'$→O$_5$ H-bond could in principle be recovered for GA and MA with (ϕ, ψ) values of about (275°, 35°) (region C, observed, marginal) or (295°, 330°) (region D, observed, marginal). Similarly, the flip-flop O$_3'$↔O$_2$ H-bond could also be achieved with (ϕ, ψ) values of about (255°, 315°) for GG (region D, observed, marginal), (160°, 285°) for GA (region D, well off minimum, not observed), (185°, 295°) for MG (region D, well off minimum, not observed), or (25°, 20°) for MA (region A, observed, marginal). Finally, many conformations could potentially present H-bonds involving the two hydroxymethyl groups at C_6 and C_6' along with other atoms, some of which are observed in the simulations with marginal populations.

If intramolecular H-bonding represented an important conformational driving force, one would expect the loss of the dominant HO$_3'$→O$_5$ H-bond in conformation C and of the flip-flop O$_3'$↔O$_2$ H-bond in conformation A resulting from the epimerization at C_3' (GG→GA or MG→MA) to

induce a relative destabilization of these two conformations. Conversely, since conformation D is in principle compatible with the recovery of the $HO'_3 \rightarrow O_5$ H-bond, it should be stabilized. However, as seen previously, the epimerization actually results in an increase of the relative free energy of conformation D by about 5–8 kJ mol^{-1}, instead of a decrease, while that of conformation A is essentially unaffected (GG→GA) or only slightly increased (MG→MA). In summary, it appears that the disruption of intramolecular H-bonding caused by epimerization at C'_3 does not induce any major conformational change driven by the possible recovery of these H-bonds (e.g., shift to conformation D) or the achievement of an alternative intramolecular H-bonding pattern (e.g., involving $O'_3 \leftrightarrow O_2$ or the hydroxymethyl groups at C_6 and C'_6).

As seen earlier, the epimerization also results in a slight displacement of the lowest free energy basin towards the 2_1-helical region, although it is associated with the disruption of the $HO'_3 \rightarrow O_5$ H-bond typical of cellobiose. This finding is clearly at odds with the commonly accepted view of a fundamental role for this hydrogen bond in stabilizing the 2_1-helical secondary structure motif typical of $\beta(1 \rightarrow 4)$-linked D-glucopyranose chains in cellooligosaccharides. It suggests instead that the formation of the $HO'_3 \rightarrow O_5$ H-bond is rather opportunistic, i.e., follows from the dominance of this conformation while representing a negligible (possibly even adverse) driving force towards its formation.

Finally, bundles of 20 structures extracted from the plain MD simulations of the four disaccharides are shown in Fig. 4. These bundles illustrate graphically the marginal effect of the strereochemistry differences on the sampled conformational ensembles. The only major difference resides in the orientational preferences of the hydroxyl group at C'_3, which is locked by the intramolecular $HO'_3 \rightarrow O_5$ H-bond in GG and MG, and essentially free to rotate in GA and MA. The two latter compounds also appear somewhat more flexible, a consequence of the slight broadening of the dominant free-energy basin, which results in enhanced fluctuations of the glycosidic dihedral angles.

Conclusion

The aim of the present article was to evaluate the relevance of solvent-exposed H-bonding as a driving or steering force in the conformational equilibria of aqueous biopolymers, using explicit-solvent MD simulation in the simple context of cellobiose stereoisomers. More specifically, four $\beta(1 \rightarrow 4)$-linked D-aldohexopyranose disaccharides were considered (Fig. 1), which present a different stereochemisty of the potentially H-bonding groups neighboring the

glycosidic linkage. The main findings of this work can be summarized as follows.

The epimerization of the potentially H-bonding groups neighboring the linkage may largely alter the intramolecular (trans-glycosidic) H-bonding pattern. Most prominently, the epimerization of cellobiose at C'_3 promotes the disappearance of the high-occurrence $HO'_3 \rightarrow O_5$ H-bond typical of $\beta(1 \rightarrow 4)$-linked D-glucopyranose chains. Yet, this epimerization has only very limited influence on the Ramachandran free-energy map of the disaccharide. Although an adjustment of the conformational distribution could in principle permit the recovery of the disrupted H-bonds or the achievement of alternative intramolecular H-bonding patterns, such an adjustment does not occur. Instead, the loss of intramolecular H-bonding is merely compensated for by an enhancement of the interaction with the solvent molecules. This finding suggests that the $HO'_3 \rightarrow O_5$ H-bond is not the cause of the 2_1-helical secondary structure motif typical of cellooligosaccharides, but rather a consequence of the dominance of this motif. In fact, the limited effect of disrupting the $HO'_3 \rightarrow O_5$ H-bond of cellobiose on the free-energy map consists of a slight displacement of the lowest free energy basin towards the 2_1-helical region. This suggests that this H-bond might even represent a slightly adverse driving force with respect to the formation of a 2_1-helical pattern.

Based on these and previous (Kräutler et al. 2007; Perić-Hassler et al. 2010) observations, it is suggested that the conformational preferences of oligosaccharides in a dilute aqueous environment and, by extension, of other short biopolymers with the potential of forming solvent-exposed intramolecular H-bonds (e.g., peptides) is primarily dictated by steric and stereoelectronic effects. The formation of solvent-exposed intramolecular H-bonds is then to be viewed as an opportunistic consequence of these preferences (close proximity of two potentially H-bonding groups in a given molecular conformation), representing in itself a minor (possibly even, in some cases, adverse) conformational driving as well as steering force (towards a molecular conformation presenting this proximity). The reason is that upon formation of a solvent-exposed H-bond in water, the H-bonding partners remain in a highly polar environment, and their interaction is screened by the solvent dielectric response as well as subject to H-bonding competition by the solvent molecules. In other words, the lack of a specific H-bond is easily compensated for by a stronger interaction of the potentially H-bonding partners with the solvent molecules.

The above statement concerning the presumably weak conformational driving force associated with H-bonding pertains to small oligomers in a dilute aqueous environment. It may not be applicable to longer polymers and to
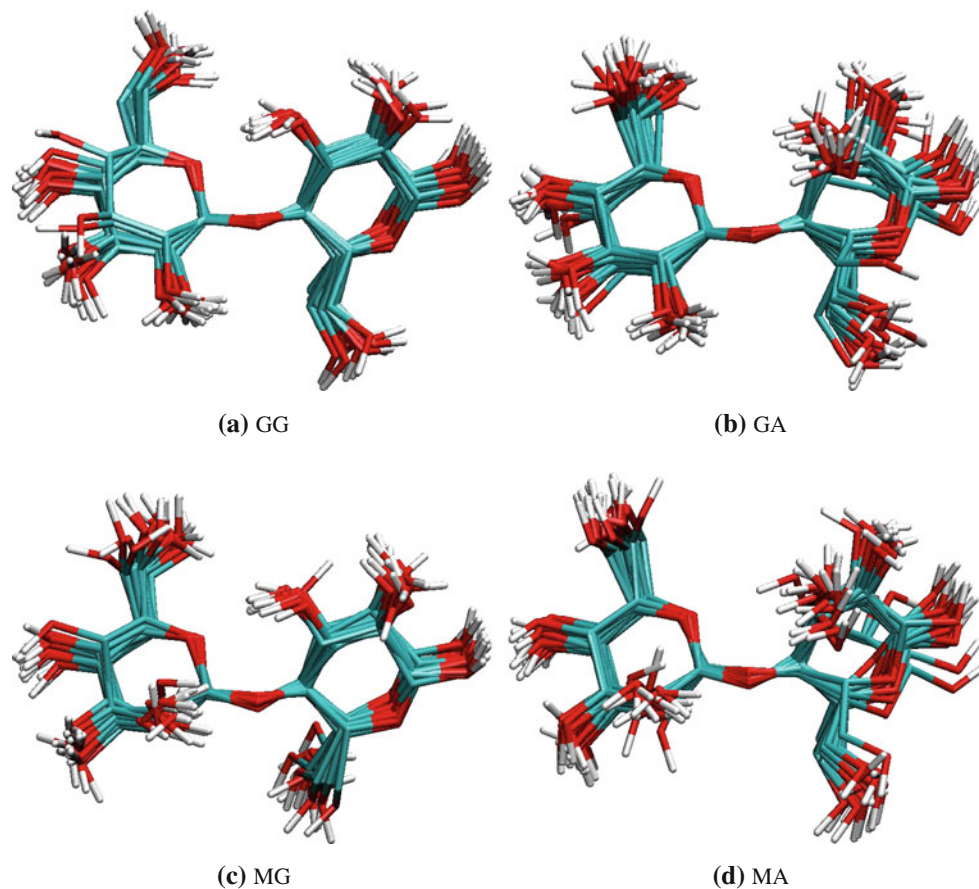
(a) GG

(b) GA

(c) MG

(d) MA

Fig. 4 Bundles of illustrative structures for the four $\beta(1 \rightarrow 4)$-linked D-aldohexopyranose disaccharides (Fig. 1), based on the corresponding plain MD simulations in water at 1 atm and 300 K. For each disaccharide, the 20 structures are sampled at an interval of 10 ps from the plain MD simulations, and correspond to the conformational region *C*. They are superimposed by minimization of their mutual atom-positional root-mean-square deviation considering all non-hydrogen atoms

other environments. For longer polymers, H-bonding cooperativity effect (Jeffrey 1990; Luque et al. 1998; Hawley et al. 2002; López de la Paz et al. 2002; Çarçabal et al. 2005; Simons et al. 2005; Dashnau et al. 2005; Deshmukh et al. 2008) may come into play. Furthermore, for extended chains or at finite concentrations, folding or intermolecular association (principally driven by the hydrophobic effect) lead to the occurrence of buried H-bonds. In contrast to the solvent-exposed ones and although they still probably represent a minor (possibly even, in some cases, adverse) conformational driving force, buried H-bonds may represent an important conformational steering force towards specific folding or association patterns. Finally, the proposed marginal conformational role of solvent-exposed intramolecular H-bonding does not imply that these have no effect on the physico-chemical properties of a specific polymer, because many of these properties are actually defined by a change of environment (e.g., to crystals, fibers, solutions with non-polar solvents, or vacuum) relative to the bulk aqueous environment at high dilution. The reason is that solvent-exposed intramolecular

H-bonding is expected to reduce the propensity of the molecule to interact with the solvent molecules (hydrophilicity), with itself (at finite concentration) or with other potentially H-bonding solute molecules (in the presence of other solutes). In the context of disaccharides (Perić-Hassler et al. 2010) [see also previous work on monosaccharides (Kräutler et al. 2007)], the absence of high-occurrence intramolecular H-bonds compensated for by a more extensive interaction with the solvent may lead to a higher hydrophilicity (Galema et al. 1994; Cheetham and Lam 1996) or, equivalently, a lower apparent hydrophobicity (Sivkama Sundari and Balasubramanian 1997; Koga et al. 2007; Simons et al. 2009) (relative affinity of the compound for less polar environments), a higher propensity of the compound to self-aggregate (Green and Angell 1989; Sun and Leopold 1994; Koster et al. 1994; Sun and Leopold 1997; Koster et al. 2000) (e.g., tendency to cluster, higher glass transition temperature, more limited solubility) or to interact with other polar solutes, and a slowing down of the dynamics in aqueous solution (Choi et al. 2006) (e.g., lower transport coefficients, higher viscosity).

In summary, evidence has been presented suggesting that solvent-exposed intramolecular H-bonding in aqueous carbohydrates is a consequence and not a cause of conformational preferences, i.e., it represents a negligible conformational driving as well as steering force. By extension, it is suggested that this interpretation is generally applicable to other short biopolymers in dilute aqueous solution (e.g., peptides). This suggestion immediately results from the high dielectric permittivity and strong hydrogen-bonding capacity of water. To further test this hypothesis, we are currently investigating the effect of these two factors by simulating the same (and other) molecular systems in different (physical and artificial) solvent with lower polarities.

Appendix: Driving versus steering force

In the introductory section of this article, reference was made to the concepts of affinity and specificity, an interaction or model effect contributing to affinity being referred to as a driving force, and an interaction or effect contributing to the specificity being referred to as a steering force. Although of fundamental importance, the discussion of the thermodynamic factors controlling (bio)chemical conformational processes in terms of these familiar concepts may be highly ambiguous, unless clear definitions are provided for: (1) the conformational states of the system that are considered and compared; (2) the reference
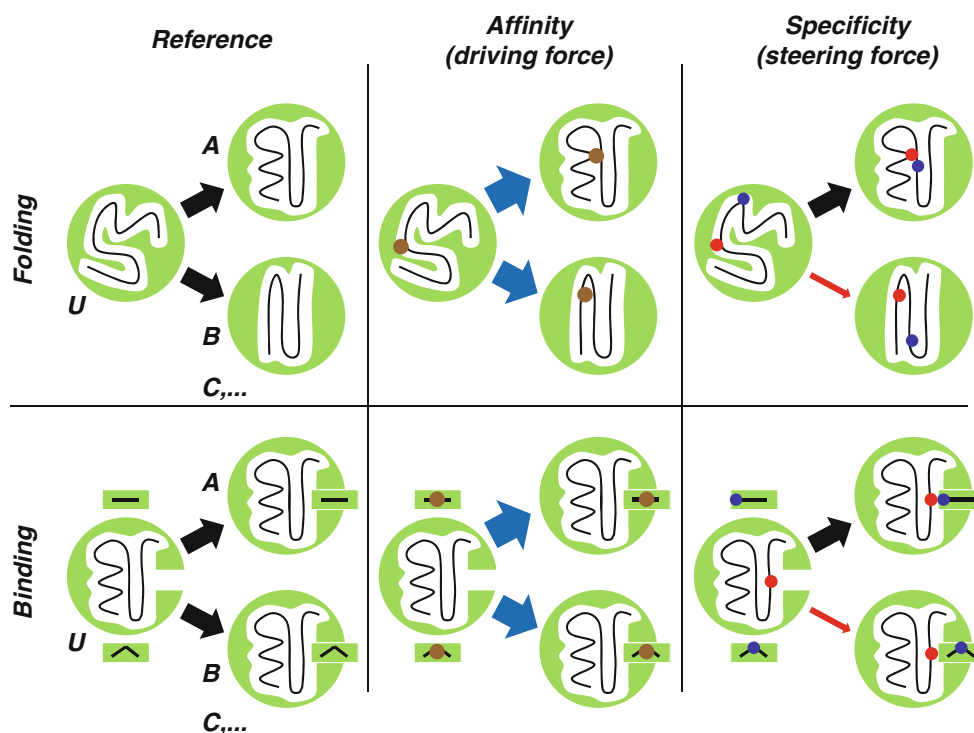

Fig. 5 Illustration of the concepts of driving and steering forces in the context of protein folding (*top*) and protein-ligand binding (*bottom*). The reference situation is representative of a system lacking the interaction or model effect of interest. Comparing the actual to the reference system, two extreme situations may occur, where the relative sizes and colors of the arrows indicate the magnitude of the negative free-energy change. If the additional interaction or effect promotes an increase in stability of all compact (folding) or bound (binding) states (*A*, *B*, *C*, …) relative to the ensemble of all non-compact (unfolded) free-protein free-ligand (unbound) conformations (*U*), it contributes to affinity and not to specificity, and is thus a driving force. This situation is illustrated in the figure by the introduction of an extra hydrophobic site (*brown ball*) which changes from solvent exposed to buried upon folding or binding. If the additional interaction or effect promotes a decrease in stability for all but one compact (folding) or bound (binding) states (*B*, *C*, …), without altering the stability of this specific state (*A*) relative to the entire set of non-compact (unfolded) or free-protein free-ligand (unbound) conformations (*U*), it contributes to specificity and not to affinity, and is thus a steering force. This situation is illustrated in the figure by the introduction of two oppositely charged functional groups (*red* and *blue* balls) which change from solvent exposed to buried upon folding or binding (in close contact in *A* but not in *B*). The reality will always be somewhere between these two extremes. Note that this description is only consistent provided that the states are defined in terms specified volumes within the 3*N*-dimensional space of atomic coordinates and that their definitions are the same for both systems (reference and actual)

situation considered as defining the absence of the interaction or effect of interest. These issues are briefly discussed here considering as examples the processes of protein folding and protein-ligand binding, as illustrated in Fig. 5.

The simplest statistical-mechanical definition of conformational states relies on considering specified volumes within the 3N-dimensional space of atomic coordinates, N being the number of atoms in the system. The relative free energies of these states are then directly related to the negative logarithm of the corresponding integrated populations within an equilibrium ensemble. This is the definition adopted in Fig. 5. In the context of folding, one considers states corresponding to distinct compact conformations (A, B, C, ...; only two of which are shown), along with the entire set of all non-compact (unfolded) conformations (U). In the context of (competitive) binding, one considers states corresponding to the specific protein-ligand complexes (A, B, C, ...; only two of which are shown), along with the entire set of free-protein free-ligand (unbound) conformations (U).

In both cases, the reference situation considered as defining the absence of the interaction or effect of interest must then be clearly defined. For example, in the context of folding, one may refer to a protein mutant lacking a side-chain presenting a specific characteristics (e.g., charged → polar or polar → hydrophobic mutation). In the context of binding, one may refer to a generic ligand presenting the common scaffold of all compounds considered, but lacking a specific functional group (e.g., functionalized → aliphatic group).

Comparing the actual to the reference system, two extreme situations may occur. If the additional interaction or effect promotes an increase in stability of all compact (folding) or bound (binding) states, it contributes to affinity and not to specificity, and is thus a driving force. If the additional interaction or effect promotes a decrease in stability for all but one compact (folding) or bound (binding) states, without altering the stability of this specific state A relative to U, it contributes to affinity and not to specificity, and is thus a steering force. The reality will always be somewhere between these two extremes but, for first-order reasoning, it is of great interest to assess which of the two extremes is closest to the reality. For example, in the present article, the claim is made that solvent-exposed intramolecular H-bonding is, as a first approximation, neither a significant driving force nor a significant steering force.

Besides the issues of defining conformational states and reference situations, there is a third issue that affects the above considerations, namely that of compatibility with experiment. Experimentally, a state is seldom defined as a specified volume within the 3N-dimensional space of atomic coordinates, but rather as a collection of conformations characterized by a non-negligible population, limited structural fluctuations, short interconversion times, a given spectroscopic signal or/and a specific functional activity. As a result, the state definition may become relatively fuzzy and, possibly, dependent of the system considered (e.g., reference vs. actual system in Fig. 5). For example, in the context of folding, alternative compact (misfolded) states (B, C, ...) will most often be experimentally undetectable. These conformations will then be lumped into the unfolded state, the distinction between driving and steering force becoming meaningless. Note that this problem is less serious in the context of binding, which is generally probed by means of separate experiments involving a single ligand (rather than in a competitive way as illustrated in Fig. 5).

References

Abraham RJ, Chambers EJ, Thomas WA (1994) Conformational analysis. Part 20. Conformational analysis of 4-deoxy-4-fluoro-D-glucose and 6-deoxy-6-fluoro-D-galactose in solution. Magn Reson Chem 32:248–254

Adams PD, Afonine PV, Grosse-Kunstleve RW, Read RJ, Richardson JS, Richardson DC, Terwilliger TC (2010) Recent developments in phasing and structure refinement for macromolecular crystallography. Curr Opin Struct Biol 19:566–572

Al-Lazikani B, Jung J, Xiang Z, Honig B (2001) Protein structure prediction. Curr Opin Struct Biol 5:51–56

Allison JR, Hertig S, Missimer JH, Smith LJ, Steinmetz O, Dolenc J (2012) Probing the structure and dynamics of proteins by combining molecular dynamics simulations and experimental NMR data. J Chem Theory Comput 8:3430–3444

Anfinsen CB (1973) Principles that govern the folding of protein chains. Science 181:223–230

Angyal SJ (1969) The composition and conformation of sugars in solution. Angew Chem Int Ed 8:157–166

Baldwin RL (2007) Energetics of protein folding. J Mol Biol 371:283–301

Barker JA, Watts RO (1973) Monte Carlo studies of the dielectric properties of water-like models. Mol Phys 26:789–792

Behrends R, Kaatze U (2005) Molecular dynamics and conformational kinetics of mono- and disaccharides in aquous solution. Chem Phys Chem 6:1133–1145

Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) Interaction models for water in relation to protein hydration. In: Pullman B (eds) Intermolecular forces. Reidel, Dordrecht, pp 331–342

Berendsen HJC, Postma JPM, van Gunsteren WF, di Nola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81:3684–3690

Berendsen HJC, van Gunsteren WF, Zwinderman HRJ, Geurtsen RG (1986) Simulations of proteins in water. Ann NY Acad Sci 482:269–285

Bergenståhle M, Wohlert J, Himmel ME, Brady JW (2010) Simulation studies of the insolubility of cellulose. Carbohydr Res 345:2060–2066

Bock K, Fernandez J, Guzman B (1988) A ^{1}H- and ^{13}C-NMR spectroscopic analysis of six pseudohexoses. Carbohydr Res 174:354–359

Bosma WB, Appell M, Willett JL, Momany FA (2006) Stepwise hydration of cellobiose by DFT methods. 2. Energy contributions

to the relative stabilities of cellobiose·(H$_2$O)$_{1-4}$ complexes. J Mol Struct Theochem 776:21–31

Brünger AT, Nilges M (1993) Computational challenges for macromolecular structure determination by X-ray crystallography and solution NMR-spectroscopy. Q Rev Biophys 26:49–135

Çarçabal P, Jockusch RA, Hünig I, Snoek LC, Kroemer RT, Davis BG, Gamblin DP, Compagnin I, Oomens J, Simons JP (2005) Hydrogen bonding and cooperativity in isolated and hydrated sugars: mannose, galactose, glucose and lactose. J Am Chem Soc 127:11414–11425

Cheetham NWH, Lam K (1996) Molecular dynamics simulations of glycosides in aqueous solution. Carbohydr Res 282:13–23

Cheetham NWH, Dasgupta P, Ball GE (2003) NMR and modelling studies of disaccharide conformation. Carbohydr Res 338:955–962

Choi Y, Cho KW, Jeong K, Jung S (2006) Molecular dynamics simulations of trehalose as a dynamic reducer for solvent water molecules in the hydration shell. Carbohydr Res 341:1020–1028

Chu SSC, Jeffrey GA (1968) The refinement of the crystal structures of β-D-glucose and cellobiose. Acta Crystallogr B 24:830–838

Cocinero EJ, Gamblin DP, Davis BG, Simons JP (2009) The building blocks of cellulose: the intrinsic conformational structures of cellobiose, its epimer, lactose, and their singly hydrated complexes. J Am Chem Soc 131:11117–11123

Csonka GI, Kaminsky J (2011) Accurate conformational energy differences of carbohydrates: a complete basis set extrapolation. J Chem Theory Comput 7:988–998

Cymerman I, Feder M, Pawowski M, Kurowski MA, Bujnicki JM (2004) Computational methods for protein structure prediction and fold recognition. Nucleic Acids Mol Biol 15:1–21

Dashnau JL, Sharp KA, Vanderkooi JM (2005) Carbohydrate intramolecular hydrogen bonding cooperativity and its effect on water structure. J Phys Chem B 109:24152–24159

de Bakker PIW, Hünenberger PH, McCammon JA (1999) Molecular dynamics simulations of the hyperthermophilic protein Sac7d from Sulfolobus acidocaldarius: contribution of salt bridges to thermostability. J Mol Biol 285:1811–1830

de Vries NK, Buck HM (1987) Solvent dependence of the rotamer population around the interglycosidic C(5)–C(6) bond of (1→6)-β-linked digalactopyranosides. Rec Trav Chim Pays-Bas 106:453–460

Deechongkit S, Dawson PE, Kelly JW (2004) Toward assessing the position-dependent contributions of backbone hydrogen bonding to β-sheet folding thermodynamics employing amide-to-ester perturbations. J Am Chem Soc 126:16762–16771

Deshmukh MM, Bartolotti LJ, Gadre SR (2008) Intramolecular hydrogen bonding and cooperative interactions in carbohydrates via the molecular tailoring approach. J Phys Chem A 112:312–321

Dill KA (1985) Theory for the folding and stability of globular proteins. Biochemistry 24:1501–1509

Dill KA (1990) Determinant forces in protein folding. Biochemistry 29:7133–7155

Dill KA, Bromberg S (2003) Molecular driving forces, statistical thermodynamics in chemistry and biology. Garland Science, New York

Dobson CM, Šali A, Karplus M (1998) Protein folding: a perspective from theory and experiment. Angew Chem Int Ed 37:868–893

Doig AJ, Sternberg MJE (1995) Side-chain conformational entropy in protein folding. Protein Sci 4:2247–2251

Eichenberger AP, Gattin Z, Yalak G, van Gunsteren WF (2010) Molecular dynamics simulation of ester-linked hen egg white lysozyme reveals the effect of missing backbone hydrogen bond donors on the protein structure. Helv Chim Acta 93:1857–1869

French AD, Johnson GP (2006) Quantum mechanics studies of cellobiose conformations. Can J Chem 84:603–612

French AD, Johnson GP, Kelterer A-M, Czonka GI (2005) Fluorinated cellobiose and maltose as stand-ins for energy surface calculations. Asymmetry 16:577–586

Galema SA, Howard E, Engberts JBFN, Grigera JR (1994) The effect of stereochemistry upon carbohydrate hydration. A molecular dynamics simulation of β-D-galactopyranose and (α, β)-D-talopyranose. Carbohydr Res 265:215–225

Gardner KH, Blackwell J (1974) The structure of native cellulose. Biopolymers 13:1975–2001

Gattin Z, Glättli A, Jaun B, van Gunsteren WF (2007) Simulation of β-depsipeptides: the effect of missing hydrogen-bond donors on their folding equilibria. Biopolymers 85:318–332

Gessler K, Krauss N, Steiner T, Betzel C, Sandmann C, Saenger W (1994) Crystal structure of β-D-cellotetraose hemihydrate with implications for the structure of cellulose II. Science 266:1027–1029

Graczyk PP, Mikolajczyk M (1994) Anomeric effect. Origin and consequences. Top Stereochem 21:159–349

Green JL, Angell CA (1989) Phase relations and vitrification in saccharide-water solutions and the trehalose anomaly. J Phys Chem 93:2880–2882

Hagen R, Kaatze U (2004) Conformational kinetics of disaccharides in aqueous solutions. J Chem Phys 120:9656–9664

Hansen HS, Daura X, Hünenberger PH (2010) Enhanced conformational sampling in molecular dynamics simulations of solvated peptides: fragment-based local elevation umbrella sampling. J Chem Theory Comput 6:2598–2621

Hansen HS, Hünenberger PH (2010a) 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 31:1–23

Hansen HS, Hünenberger PH (2010b) Ball-and-stick local elevation umbrella sampling: molecular simulations involving enhanced sampling within conformational or alchemical subspaces of low internal dimensionalities, minimal irrelevant volume and problem-adapted geometries. J Chem Theory Comput 6:2622–2646

Hansen HS, Hünenberger PH (2011) A reoptimized GROMOS force field for hexopyranose-based carbohydrates accounting for the relative free energies of ring conformers, anomers, epimers, hydroxymethyl rotamers and glycosidic linkage conformers. J Comput Chem 32:998–1032

Hawley J, Bampos N, Aboitiz N, Jiménez-Barbero J, López de la Paz M, Sanders JKM, Carmona P, Vicent C (2002) Investigation of the hydrogen bonding properties of a series of monosaccharides in aqueous media by ^{1}H NMR and IR spectroscopy. Eur J Org Chem 12:1925–1936

Heinz TN, van Gunsteren WF, Hünenberger PH (2001) Comparison of four methods to compute the dielectric permittivity of liquids from molecular dynamics simulations. J Chem Phys 115:1125–1136

Hockney RW (1970) The potential calculation and some applications. Methods Comput Phys 9:136–211

Honig B, Yang A-S (1995) Free energy balance in protein folding. Adv Prot Chem 46:27–58

Höög C, Landersjö C, Widmalm G (2001) Oligosaccharides display both rigidity and high flexibility in water as determined by ^{13}C NMR relaxation and ^{1}H,^{1}H NOE spectroscopy: Evidence of anti-ϕ and anti-ψ torsions in the same glycosidic linkage. Chem Eur J 7:3069–3077

Huang S-Y, Zou X (2010) Advances and challenges in protein-ligand docking. Int J Mol Sci 11:3016–3034

Huber T, Torda AE, van Gunsteren WF (1994) Local elevation: a method for improving the searching properties of molecular dynamics simulation. J Comput-Aided Mol Des 8:695–708

Hünenberger PH, Helms V, Narayana N, Taylor SS, McCammon JA (1999) Determinants of ligand binding to cAMP-dependent protein kinase. Biochemistry 38:2358–2366

Jacobson RA, Wunderlich JA, Lipscomb WN (1961) The crystal and molecular structure of cellobiose. Acta Crystallogr 14:598–607

Jaenicke R (1991) Protein folding: local structures, domains, subunits and assemblies. Biochemistry 30:3147–3161

Jeffrey GA (1994) The role of the hydrogen bond and water in biological processes. J Mol Struct 322:21–25

Jeffrey GA (1997) An introduction to hydrogen bonding. Oxford University Press, New York

Jeffrey GA (1990) Crystallographic studies of carbohydrates. Acta Crystallogr B 46:89–103

Jockusch RA, Kroemer RT, Talbot FO, Snoek LC, Çarçabal P, Simons JP, Havenith M, Bakker JM, Compagnon I, Meijer G, von Helden G (2004) Probing the glycosidic linkage: UV and IR ion-dip spectroscopy of a lactoside. J Am Chem Soc 126:5709–5714

Kadla JF, Gilbert RD (2000) Cellulose structure: a review. Cell Chem Technol 34:197–216

Koga Y, Nishikawa K, Westh P (2007) Relative hydrophobicity/hydrophilicity of fructose, glucose, sucrose and trehalose as probed by 1-propanol: a differential approach in solution thermodynamics. J Phys Chem B 111:13943–13948

Kolpak FJ, Blackwell J (1976) Determination of structure of cellulose II. Macromolecules 9:273–278

Korth M (2010) Third-generation hydrogen-bonding corrections for semiempirical QM methods and force fields. J Chem Theory Comput 6:3808–3816

Koster KL, Webb MS, Bryant G, Lynch DV (1994) Interactions between soluble sugars and POPC (1-palmitoyl-2-oleoylphosphatidylcholine) during dehydration: vitrification of sugars alters the phase behavior of the phospholipid. Biochim Biophys Acta 1193:143–150

Koster KL, Lei YP, Anderson M, Martin S, Bryant G (2000) Effects of vitrified and nonvitrified sugars on phosphatidylcholine fluid-to-gel phase transitions. Biophys J 78:1932–1946

Kräutler V, Müller M, Hünenberger PH (2007) Conformation, dynamics, solvation and relative stabilities of selected β-hexopyranoses in water: a molecular dynamics study with the GROMOS 45A4 force field. Carbohydr Res 342:2097–2124

Larsson EA, Staaf M, Söderman P, Höög C, Widmalm G (2004) Determination of the conformational flexibility of methyl α-cellobioside in solution by NMR spectroscopy and molecular simulations. J Phys Chem A 108:3932–3937

Lemieux RU, Brewer JT (1973) Conformational preferences for solvated hydroxymethyl groups in hexopyranose structures. Adv Chem 117:121–146

Lemieux RU, Koto S (1974) The conformational properties of the glycosidic linkages. Tetrahedron 30:1933–1944

Li Y, Jiang X-N, Wang C-S (2011) Rapid evaluation of the binding energies in hydrogen-bonded amide-thymine and amide-uracil dimers in gas phase. J Comput Chem 32:953–966

Lins RD, Hünenberger PH (2005) A new GROMOS force field for hexopyranose-based carbohydrates. J Comput Chem 26:1400–1412

López de la Paz M, Ellis G, Pérez M, Perkins J, Jiménez-Barbero J, Vicent C (2002) Carbohydrate hydrogen-bonding cooperativity—Intramolecular hydrogen bonds and their cooperative effect on intermolecular processes—binding to a hydrogen-bond acceptor molecule. Eur J Org Chem 5:840–855

Luque FJ, López JM, López de la Paz M, Vicent C, Orozco M (1998) Role of intramolecular hydrogen bonds in the intermolecular hydrogen bonding of carbohydrates. J Phys Chem A 102:6690–6696

Mayato C, Dorta RL, Vázquez JT (2004) Experimental evidence on the hydroxymethyl group conformation in alkyl β-ᴅ-mannopyranosides. Tetrahedron Asymmetry 15:2385–2397

Mendonca S, Johnson GP, French AD, Laine RA (2002) Conformational analyses of native and permethylated disaccharides. J Phys Chem A 106:4115–4124

Momany FA, Schnupf U (2010) DFTMD studies of β-cellobiose: conformational preference using implicit solvent. Carbohydr Res 346:619–630

Naidoo KJ, Chen JY (2003) The role of water in the design of glycosidic linkage flexibility. Mol Phys 101:2687–2694

Pašalić H, Aquino AJA, Tunega D, Haberhauer G, Gerzabek MH, Georg HC, Moraes TF, Coutinho K, Canuto S, Lischka H (2010) Thermodynamic stability of hydrogen-bonded systems in polar and nonpolar environments. J Comput Chem 31:2046–2055

Paton RS, Goodman JM (2009) Hydrogen bonding and π-stacking: How reliable are force fields? A critical evaluation of force field descriptions of nonbonded interactions. J Chem Inf Model 49:944–955

Pauling L, Corey RB, Branson HR (1951) The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. Proc Natl Acad Sci USA 37:205–211

Pereira CS, Kony D, Baron R, Müller M, van Gunsteren WF, Hünenberger PH (2006) Conformational and dynamical properties of disaccharides in water: a molecular dynamics study. Biophys J 90:4337–4344

Pereira CS, Kony D, Baron R, Müller M, van Gunsteren WF, Hünenberger PH (2007) Erratum to "Conformational and dynamical properties of disaccharides in water: a molecular dynamics study". Biophys J 93:706–707

Pérez S, Kouwijzer M (1999) Shapes and interactions of polysaccharide chains. In: Finch P (ed) Carbohydrates: structures, dynamics and syntheses. Kluwer, Dordrecht, pp 258–293

Pérez S, Marchessault RH (1978) The exo-anomeric effect: experimental evidence from crystal structures. Carbohydr Res 65:114–120

Pérez S, Samain D (2010) Structure and engineering of celluloses. Adv Carbohydr Chem 64:25–116

Pérez S, Vergelati C (1985) Unified representation of helical parameters: application to polysaccharides. Biopolymers 24:1809–1822

Peri S, Karim MN, Khare R (2011) Potential of mean force for separation of the repeating units in cellulose and hemicellulose. Carbohydr Res 346:867–871

Perić L, Pereira CS, Perez S, Hünenberger PH (2008) Conformation, dynamics and ion-binding properties of single-chain polyuronates: a molecular dynamics study. Mol Simul 34:421–446

Perić-Hassler L, Hünenberger PH (2010) Interaction of alginate single-chain polyguluronate segments with mono- and divalent metal cations: a comparative molecular dynamics study. Mol Simul 36:778–795

Perić-Hassler L, Hansen HS, Baron R, Hünenberger PH (2010) Conformational properties of glucose-based disaccharides using molecular dynamics simulations with local elevation umbrella sampling. Carbohydr Res 345:1781–1801

Pickett HM, Strauss HL (1970) Conformational structure, energy and inversion rates of cyclohexane and some related oxanes. J Am Chem Soc 92:7281–7290

Polacek R, Stenger J, Kaatze U (2002) Chair-chair conformational flexibility, pseudorotation, and exocyclic group isomerization of monosaccharides in water. J Chem Phys 116:2973–2982

Rao VSR, Qasba PK, Balaji PV, Chandrasekaran R (1998) Conformation of carbohydrates. Harwood Academic Publishers, Amsterdam

Rockwell GD, Grindley TB (1998) Effect of solvation on the rotation of hydroxymethyl groups in carbohydrates. J Am Chem Soc 120:100953–100963

Roën A, Padrón JI, Vásquez JT (2002) Hydroxymethyl rotamer populations in disaccharides. J Org Chem 68:4615–4630

Ryckaert J-P, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 23:327–341

Saenger W (1984) Principles of nucleic acid structure. Springer, New York

Samudrala R, Xia Y, Huang E, Levitt M (1999) Ab initio protein structure prediction using a combined hierarchical approach. Proteins Struct Funct Genet Suppl 3:194–198

Sarko A, Wu H-CH (1978) The crystal structures of A-, B- and C-polymorphs of amylose and starch. Starch 30:73–78

Schmid N, Christ CD, Christen M, Eichenberger AP, van Gunsteren WF (2012) Architecture, implementation and parallelisation of the GROMOS software for biomolecular simulation. Comput Phys Commun 183:890–903

Seebach D, Mahajan YR, Senthilkumar R, Rueping M, Jaun B (2002) β-depsipeptides—the effect of a missing and a weakened hydrogen bond on the stability of the β-peptidic 3_{14}-helix. Chem Commun 15:1598–1599

Seebach D, Jaun D, Sebesta R, Mathad RI, Plögel O, Limbach M, Sneller H, Cottens S (2006) Synthesis, and helix or hairpin-turn secondary structures of 'mixed' α/β-peptides consisting of residues with proteinogenic side chains and of 2-amino-2-methylpropanoic acid (Aib). Helv Chim Acta 89:1801–1825

Simons JP, Jockusch RA, Çarçabal P, Hünig I, Kroemer RT, McLeod NA, Snoek LC (2005) Sugars in gas phase. Spectroscopy, conformation, hydration, co-operativity and selectivity. Int Rev Phys Chem 24:489–531

Simons JP, Davis BG, Cocinero EJ, Gamblin DP, Stanca-Kaposta EC (2009) Conformational change and selectivity in explicitly hydrated carbohydrates. Tetrahedron Asymmetry 20:718–722

Sivkama Sundari C, Balasubramanian D (1997) Hydrophobic surfaces in saccharide chains Prog Biophys Mol Biol 67:183–216

Sousa SF, Fernandes PA, Ramos MJ (2006) Protein-ligand docking: current status and future challenges. Proteins Struct Funct Bioinf 65:15–26

Spiwok V, Tvaroska I (2009) Metadynamics modelling of the solvent effect on primary hydroxyl rotamer equilibria in hexopyranosides. Carbohyd Res 344:1575–1581

Stortz CA, Johnson GP, French AD, Csonka GI (2009) Comparison of different force fields for the study of disaccharides. Carbohyd Res 344:2217–2228

Strati GL, Willett JL, Momany FA (2002) Ab initio computational study of β-cellobiose conformers using B3LYP/6-311++G**. Carbohydr Res 337:1833–1849

Streefkerk DG, Stephen AM (1976) PMR studies on fully methylated aldohexopyranosides and their 6-deoxy analogues using lanthanide-shift reagents. Carbohydr Res 49:13–25

Sun WQ, Leopold AC (1994) Glassy state and seed storage stability: a viability equation analysis. Ann Bot 74:601–604

Sun WQ, Leopold AC (1997) Cytoplasmic vitrification and survival of anhydrobiotic organisms—a viability equation analysis. Comp Biochem Physiol A 117:327–333

Sun C-L, Wang C-S (2010) Estimation on the intramolecular hydrogen-bonding energies in proteins by the analytic potential energy function. J Mol Struct 956:38–43

Takashima H (2006) High-resolution protein structure determination by NMR. Annu Rep NMR Spectros 59:235–273

Thogersen H, Lemieux RU, Bock K, Meyer B (1982) Further justification for the exo-anomeric effect. Conformational analysis based on nuclear magnetic resonance spectroscopy of oligosaccharides. Can J Chem 60:44–57

Tironi IG, Sperb R, Smith PE, van Gunsteren WF (1995) A generalized reaction field method for molecular dynamics simulations. J Chem Phys 102:5451–5459

Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: umbrella sampling. J Comput Phys 23:187–199

Tvaroška I, Carver JP (1998) The anomeric and exo-anomeric effects of a hydroxyl group and the stereochemistry of the hemiacetal linkage. Carbohydr Res 309:1–9

Tvaroška I, Kožar T (1986) Theoretical studies on the conformation of saccharides. VII. Structure and stereochemistry of α- and β-D-glucopyranose in solution. Theor Chim Acta 70:99–114

Umemura M, Yuguchi Y, Hirotsu T (2005) Hydration at glycosidic linkages of malto- and cello-oligosaccharides in aqueous solution from molecular dynamics simulation: effect of conformational flexibility. J Molec Struct 730:1–8

Valleau JP, Torrie GM (1977) A guide to Monte Carlo for statistical mechanics: 1. Highways. In: Berne BJ (ed) Modern theoretical chemistrym, vol 5, Plenum Press, New York, pp 169–194

van Gunsteren WF, Billeter SR, Eising AA, Hünenberger PH, Krüger P, Mark AE, Scott WRP, Tironi IG (1996) Biomolecular simulation: the GROMOS96 manual and user guide. Verlag der Fachvereine, Zürich

Wang DQ, Jaun B, van Gunsteren WF (2009) Folding and unfolding of two mixed α/β peptides. ChemBioChem 10:2032–2041

Watson JD, Crick FHC (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. Nature 171:737–738

Withers SG, Street IP, Percival MD (1988) Fluorinated carbohydrates as probes of enzyme specificity and mechanism. ACS Symp Ser 374:59–77

Yang A-S, Honig B (1995) Free energy determinants of secondary structure formation: I. alpha-helices. J Mol Biol 252:351–365

Yang A-S, Honig B (1995) Free energy determinants of secondary structure formation: II. antiparallel beta-sheets. J Mol Biol 252:366–376

Yang A-S, Honig B (1996) Free energy determinants of secondary structure formation: III. beta-turns and their role in protein folding. J Mol Biol 259:873–882

Zheng Y-J, Merz KM Jr (1992) Study of hydrogen bonding interactions relevant to biomolecular structure and function. J Comput Chem 13:1151–1169