

Conference paper

Wojciech Plazinski* and Anita Plazinska

Molecular dynamics simulations of hexopyranose ring distortion in different force fields

DOI 10.1515/pac-2016-0922

Abstract: The four classical, biomolecular force fields designed to study hexopyranose-based carbohydrates (GROMOS 56a6_{CARBO}/56a6_{CARBO-R}, GROMOS 53a6_{GLYC}, CHARMM and GLYCAM06) have been tested in the context of ring-inversion properties. These properties were evaluated for both unfunctionalized monomers of all hexopyranoses of the D series and for residues in a chain composed of uniform units connected by $\alpha(1\rightarrow4)$ and $\beta(1\rightarrow4)$ glycosidic linkages. The results indicate that the tested force fields differ in their predictions of the ring-inversion properties of both monomers and residues in a chain. The comparison with the available experimental data and with the semi-empirical Angyal scheme reveals that, at the level of monomers, GROMOS 56a6_{CARBO}, GROMOS 53a6_{GLYC} and CHARMM correctly reproduce the ring-inversion free energies. However, due to the lack of analogous reference data we cannot state which force field is more or less accurate in the context of ring distortion of residues in a chain. Therefore, the use of ab initio potentials is recommended in the prospective, quantitative studies on the related subject.

Keywords: carbohydrates; conformational analysis; force fields; ICS-28; molecular dynamics; pyranoses; ring inversion.

Introduction

The flexibility of pyranose ring is a degree of freedom that often is neglected when studying the three-dimensional structure of carbohydrates. This is because in many cases the ring-inversion free energies are high enough to assume that the ring shape is 'conformationally locked' in the energetically favorable state (which, for hexopyranoses of the D series, is the 4C_1 conformation, as a rule). However, there is a growing number of reports pointing out that the ring shape and the related conformational rearrangements may play an important role in various, biologically-relevant processes [1–4]. Furthermore, there exists a significant collection of various hexopyranoses for which a large degree of ring flexibility has been proven in terms of the NMR-based studies, performed for molecules in solution [5–9]. In particular, such studies usually concern idose [5], altrose [6] and their derivatives [8, 9] as well as some functionalized pyranoses such as mannuronic [7] or iduronic [8] acids. Even if the population of non-standard ring geometries is not high, it may significantly increase under influence of external factors. The example are atomic force microscopy (AFM) experiments

Article note: A collection of invited papers based on presentations at the XXVIII International Carbohydrate Symposium (ICS-28), New Orleans, July 17–21 2016

***Corresponding author: Wojciech Plazinski**, Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, ul. Niezapominajek 8, 30-239 Cracow, Poland, Tel.: +48815375685, Fax: +48815375685, E-mail: wojtek_plazinski@o2.pl; wojtek@vega.umcs.lublin.pl

Anita Plazinska: Department of Biopharmacy, Faculty of Pharmacy, Medical University of Lublin, W. Chodzki Str., 4a, 20-093 Lublin, Poland

[10, 11] during which the carbohydrate chain is pulled by the mechanical force. Contrary to the case of monomers, there are no experimental studies that would quantitatively describe the ring flexibility for residues in carbohydrate chain and the available data are either of qualitative nature [12–14] or concern the static crystal structures (see discussion in Ref. [15]). Currently, there is no quantitative interpretation of the AFM ‘stretching curves’ in terms of population of particular ring conformers in the pulled chain. Thus, the results of such experiments (although giving an important assessment of the favorable ring conformation), provide only indirect evidences which can hardly be translated into the quantitative ring-inversion properties [10, 11].

Because of scarce number of experimental reports that provide the quantitative description of hexopyranose ring flexibility, the theoretical approaches based on the molecular dynamics (MD) technique seem to be an attractive alternative to investigate this subject. The results of recent theoretical studies suggest that the ring flexibility may be enhanced when inserting a given residue in a chain of specific linkage type or upon functionalization (β -O₁-methylation; see Fig. 1a for atom numbering) [17]. Furthermore, the ring shape may influence the favorable orientation of glycosidic linkages while the significance of this effect depends not only on the ring-inversion properties but also on the linkage type [19, 20].

The accuracy of the MD-based methods relies primarily on the quality of potentials that approximate various interactions present in the molecular system of interest. In spite of several cases of successful application of the ab initio potentials into MD calculations aimed at studying the carbohydrate conformation [21, 22], the high computational cost makes such attempts limited to molecular systems of small size. Larger number of studies focused on puckering of the hexopyranose ring were conducted by using classical, biomolecular force fields dedicated to study the conformational properties of carbohydrates [15, 17–20, 23–26]. So far, the accuracy of such force fields has been comparatively studied only in the context of small sets of molecular systems (e.g. β -D-glucopyranose) [26] or only for unfunctionalized hexopyranose monomers [24]. The collection of all hexopyranoses of the D series either in monomeric forms or connected by α (1→4) and β (1→4) linkages has never been considered.

The main aims of our investigation are:

1. Describing and comparing the ring inversion properties of unfunctionalized monomers of all hexopyranoses of the D series calculated within four different force fields (GROMOS 56a6_{CARBO}/56a6_{CARBO_R} [17, 18], GROMOS 53a6_{GLYC} [24], CHARMM [27, 28] and GLYCAM06 [29]). Contrary to already existing studies (see e.g. table 4 in Ref. [24]) these properties will be expressed here as the two separate quantities: (i) the free energy change associated with the full chair-chair ring inversion from 4C_1 to 1C_4 (ΔF_i); (ii) the free energy change associated with the ‘half’ ring inversion from 4C_1 to boat/skew-boat conformers (B/S) (ΔF_h).
2. Describing and comparing the ring inversion properties of residues in a chain composed of uniform monomers connected by glycosidic linkages of the α (1→4) and β (1→4) types calculated within three different force fields (GROMOS 56a6_{CARBO}/56a6_{CARBO_R}, GROMOS 53a6_{GLYC} and CHARMM). In this stage of the study the ring inversion properties are also investigated in the context of both the ΔF_i and ΔF_h values.

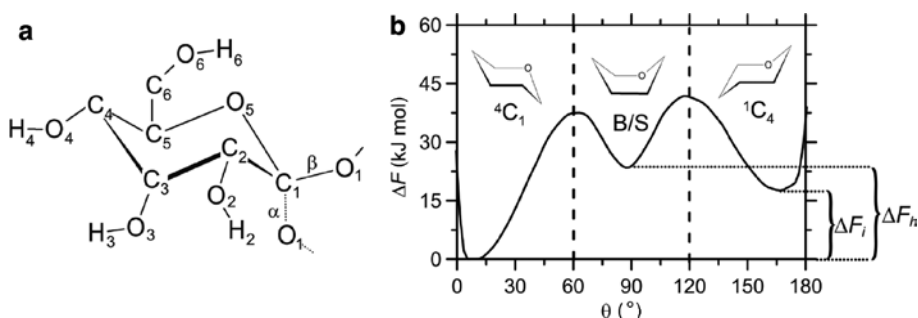


Fig. 1: (a) Molecular structure of the D-glucopyranose (Glc) molecule (α - or β -anomer) illustrating the atom labeling used in the article. (b) Illustrative free energy (ΔF) profile (FEP) along the Cremer-Pople θ parameter [16] corresponding to the unfunctionalized α -Glc monomer within the 56A6_{CARBO} force field [17, 18]; the dashed vertical lines at $\theta = 60^\circ$ and 120° separate the three free-energy basins (4C_1 , B/S, 1C_4).

Methods

All MD simulations were carried out with the GROMACS 4.5.5 program [30], using the following four force fields: GROMOS 56a6_{CARBO}/56a6_{CARBO_R} [17, 18], GROMOS 53a6_{GLYC} [24], CHARMM [27, 28] and GLYCAM06 [29]. Note that one of these force fields exists in two versions, namely: GROMOS 56a6_{CARBO} and GROMOS 56a6_{CARBO_R}; the latter version contains parameters applicable for simulating the alkylated and glycosylated carbohydrates whereas parameters for unfunctionalized monomers are identical in both variants. For the sake of simplicity, in the present article we will be using only the 56a6_{CARBO} term. The GROMACS-compatible CHARMM parameters were downloaded from the official CHARMM website (http://mackerell.umaryland.edu/charmm_ff.shtml); the GROMOS 56a6_{CARBO_R} parameters were prepared during the procedure of the force field parameterization (see Ref. [17]); the GROMOS 53a6_{GLYC} parameters were downloaded from the website of Biomaterial Modeling Group of Federal University of Pernambuco (<http://dqfnet.ufpe.br/biomat>); the GLYCAM parameters for monomers were provided by Dr V.H. Rusu. In the case of trisaccharides, the *pdb2gmx* program was used to produce the molecular topologies from respective monomeric building blocks.

For each of these force field an appropriate water model was applied to describe the explicit solvent present in the system, i.e. SPC [31] (GROMOS force fields) or TIP3P [32] (GLYCAM06 and CHARMM). The systems considered consisted of one solute molecule solvated by water molecules within a cubic computational box simulated under periodic boundary conditions at constant volume and at a temperature of 298 K. Given the selected numbers of water molecules (~900 to ~1500 for the different systems), the box edges (3.1–3.6 nm for the different systems) were all preoptimized by a 1 ns constant-pressure MD equilibration at 1 bar and 298 K, ensuring an effective solvent density appropriate for these conditions in the subsequent isochoric production simulations.

The solutes considered were mono- or trisaccharides. The monosaccharides included the eight hexopyranoses of the D-series, namely allose (All), altrose (Alt), glucose (Glc), mannose (Man), gulose (Gul), idose (Ido), galactose (Gal) and talose (Tal), considering both the α - and β -anomers. The trisaccharides considered systematically were $\alpha(1\rightarrow4)$ - and $\beta(1\rightarrow4)$ -linked homooligomers of the eight monosaccharides without any functionalization, the anomeric configuration at the reducing end being the same as that of the linkage.

The calculations involved the determination of one-dimensional free-energy profiles (FEPs) along the Cremer-Pople θ parameter [16] describing the conformation of specific rings. For the FEP analyses along θ , the intervals $[0-60^\circ]$, $[60-120^\circ]$ and $[120-180^\circ]$ were ascribed to regular chair (4C_1), boat/skew-boat (B/S) and inverted chair (1C_4) conformations, respectively, as illustrated in Fig. 1. The free-energy changes (ΔF_i and ΔF_h) associated with the ring inversions (${}^4C_1 \rightarrow {}^1C_4$ and ${}^4C_1 \rightarrow \text{B/S}$, respectively) were defined as the difference between the values at the FEP minima in the two corresponding regions. These analyses were carried out systematically for the monomers and for the central residue in the trimers.

The calculation of FEPs relied on an enhanced-sampling scheme combining [33] parallel tempering [34] and well-tempered metadynamics [35] as implemented in the PLUMED 1.3 plug-in [36] of the GROMACS program [30]. The well-tempered metadynamics [35] relied on Gaussian local functions of widths 2.86° , an initial deposition rate of $0.01 \text{ kJ mol}^{-1} \text{ ps}^{-1}$ and a temperature parameter ΔT (see eq. (2) in Ref. [35]) of 7152 K (simulations within the GROMOS 53a6_{GLYC} force field) or 1788 K (remaining force fields). The parallel-tempering [34] relied on 16 metadynamics simulations carried out in parallel at different temperatures ranging from 298.0 to 363.2 K in steps of about 4.3 K, along with replica-exchange attempts performed at 2 ps intervals.

After equilibration, all simulations were carried out until the convergence of ΔF_i and ΔF_h values were detected which resulted in total simulation time varying from 10 ns (GROMOS 56a6_{CARBO}, GLYCAM06) to 110 ns (GROMOS 53a6_{GLYC}). The convergence of the ΔF_i and ΔF_h values was checked during the course of simulations by using the PLUMED built-in *sum_hills* tool and hand-written scripts.

In all simulations, the equations of motion were integrated using the leap-frog scheme [37] with a timestep of 2 fs. The full rigidity of the water molecules was enforced by application of the *SETTLE* procedure [38]. The temperature was maintained close to its reference value by application of the V-rescale thermostat [39] with a relaxation time of 0.1 ps. The translational center-of-mass motion was removed every timestep. The solute bond lengths were constrained by application of the LINCS procedure [40, 41] with a relative

geometric tolerance of 10^{-4} to all covalent bonds (the GROMOS force fields) or to hydrogen-containing bonds (GLYCAM06 and CHARMM). In the case of GROMOS 56a6_{CARBO}, the non-bonded interactions were calculated using a twin-range scheme [42], with short- and long-range cutoff distances set to 0.8 and 1.4 nm, respectively, and an update frequency of five timesteps for the short-range pairlist and intermediate-range interactions. A reaction-field correction [43, 44] was applied to account for the mean effect of the electrostatic interactions beyond the long-range cutoff distance, using a relative dielectric permittivity of 61 [45]. In the case of GROMOS 53a6_{GLYC}, long-range electrostatic interactions were treated by the reaction-field method with a relative dielectric permittivity of 66. A 1.2 nm cutoff was used for the short range electrostatics and Lennard-Jones interactions. Within GLYCAM06-based simulations long-range electrostatic interactions were treated by using the particle mesh Ewald summation [46] (a 0.9 nm distance cutoff was used for direct space nonbonded calculation). In the case of the CHARMM force field, a force-switched smoothing function [47] was applied to Lennard-Jones interactions in the range of 1–1.2 nm. Long-range Coulomb interactions were handled using particle mesh Ewald with a real-space cutoff of 1.2 nm.

Results

The results of simulations are gathered in Tables 1 and 2 as the free energy values ΔF_i and ΔF_h . The correlation between selected sets of data is presented graphically in Figs. 2–5. All possible monomers of D-hexopyranoses as well as homotrimers composed of corresponding residues were considered, except of the case of GLYCAM06 where, due to the lack of available parameters, investigations were limited to non-Ido monomers.

Let us mention that both force fields belonging to the GROMOS family (56a6_{CARBO} and 53a6_{GLYC}) have been parametrized with attention paid to possible deformations of pyranose rings. The respective papers (Refs. [17, 18, 24]) contain a detailed description of the hexopyranose ring-inversion properties which were studied

Table 1: Chair-chair-inversion (${}^4C_1 \rightarrow {}^1C_4$) free energies ΔF_i in kJ mol⁻¹ for the unfunctionalized α - and β -D-hexopyranoses in water, calculated for monomers or central residue in homotrimer using the CHARMM [27, 28], GLYCAM06 [29], GROMOS 53a6_{GLYC} [24] and GROMOS 56A6_{CARBO} force-fields [17, 18] or estimated using the empirical, experimentally-inferred scheme by Angyal [48] (only for monomers, values rounded to one-digit precision).

	Monomers					Trimers		
	56A6 _{CARBO}	CHARMM	53a6 _{GLYC}	GLYCAM06	Angyal	56A6 _{CARBO}	CHARMM	53a6 _{GLYC}
α -All	13.4	4.1	11.1	-10.4	6.1	17.5	3.6	6.7
β -All	23.3	8.7	35.5	14.6	13.0	0.2	14	34.8
α -Alt	4	1.7	2.7	-25.4	0.8 (1.6)	11.7	-2.5	-6.8
β -Alt	12.1	11.1	21.6	1.3	8.4 (5.4)	3	9.1	3.2
α -Glc	18.1	24.6	44.3	21.5	17.4	4.8	14.2	20.1
β -Glc	28.3	32.9	71.4	48.7	24.9	6.8	37.4	59.6
α -Man	10.1	33.3	21.3	6.4	12.8	1.2	11.1	2.3
β -Man	20	22.7	41.8	31.5	19.7	2.5	27.8	39.7
α -Gul	7.7	1.5	4.1	-14.2	3.1	8.2	4.5	-12.2
β -Gul	19.1	7.3	18.9	10.7	10.0	21.3	18	17.6
α -Ido	-1.5	-4.2	-28.9	-	-2.1 (-3.4)	7.3	-1.6	-26.5
β -Ido	9.2	6.4	6	-	5.4 (2.7)	16	11.8	7.9
α -Gal	12.9	19	22.7	15	14.4	5.5	12.1	14.1
β -Gal	26.6	34.2	60.5	41.8	22.0	15.7	37.2	51.5
α -Tal	6.3	19.6	2.2	0.2	9.8	0.5	3.2	-5.2
β -Tal	17.5	26.5	29.8	19.4	16.7	6.8	31.5	8.9

Direct estimates based on NMR experiments are also available for D-Ido [5] and D-Alt [6] (values reported between parentheses). The calculated values rely on Cremer-Pople coordinates [16] for the state assignment. The correlation between the selected values is illustrated graphically in Figs. 2–5.

Table 2: Regular chair-boat/skew-boat-inversion (${}^4C_1\rightarrow B/S$) free energies ΔF_i in kJ mol^{-1} for the unfunctionalized α - and β -D-hexopyranoses in water, calculated for monomers or central residue in homotrimer using the CHARMM [27, 28], GLYCAM06 [29], GROMOS 53a6_{GLYC} [24] and GROMOS 56a6_{CARBO} force-fields [17, 18].

	Monomers				Trimers		
	56A6 _{CARBO}	CHARMM	53a6 _{GLYC}	GLYCAM06	56A6 _{CARBO}	CHARMM	53a6 _{GLYC}
α -All	17.1	13	51.3	8.4	22.5	13.6	44.8
β -All	20.9	21.4	59.6	13.5	13.1	18.9	54
α -Alt	5.9	5.5	26.3	-13.5	15.2	4.8	27.8
β -Alt	18.6	17.3	47.7	4.1	23.7	12.2	44.6
α -Glc	23.5	21.9	50.6	15.8	9.8	14.9	34.9
β -Glc	23.7	19.9	59.9	20.2	12.8	21.5	54.8
α -Man	14.3	25.2	46.9	9	3.8	18.4	33.1
β -Man	19.7	28	63	26.2	12.4	23.6	56.9
α -Gul	22.8	7.2	49.5	-2.8	26.8	8.7	31.1
β -Gul	27.6	17.6	59.9	8.6	29	15.3	62.3
α -Ido	9.3	-1.9	11.7	-	17.8	-3.5	6.2
β -Ido	19.4	13.8	34.1	-	16.4	7.2	39.9
α -Gal	30.9	23.3	56.6	15.7	24.8	19.1	42.3
β -Gal	28.1	22.6	65.8	28.8	19.5	27.2	58.4
α -Tal	20.9	25.8	53.6	10.8	16.9	12.8	39.3
β -Tal	25.3	26.7	72.5	28.7	17.1	27.2	65.4

The calculated values rely on Cremer-Pople coordinates [16] for the state assignment. The correlation between the selected values is illustrated graphically in Figs. 2–5.

in terms of all possible ring conformers [18] or just the chair-chair inversion [17, 24]. Initially, these analyses have been performed only at the level at which the ring flexibility-related features were adjusted, i.e. for unfunctionalized monomers; therefore, in the case of GROMOS 56a6_{CARBO}, the additional patch was needed to account for the more accurate description of the ring distortion in O₁-functionalized pyranoses [17]. Conversely, GROMOS 53a6_{GLYC}, GLYCAM06 and CHARMM have never been systematically tested with respect to the ring-puckering properties in the context of varying local environment of the distorted ring (i.e. unfunctionalized monomer vs. functionalized monomer vs. residue in a chain).

As the experimental data providing the desired quantities (i.e. the relative populations of at least the two most abundant ring conformers) are limited to idose [5] and altrose [6], for the rest of hexopyranoses the semi-empirical Angyal scheme [48] was used for comparative purposes (Table 1). Note that this scheme concerns only unfunctionalized hexopyranose monomers and the chair-chair inversion energy. Thus, for the second type of considered data, i.e. free energies of chair-boat/skew-boat transition, there is neither experimental data nor semi-empirical schemes that could be used to verify the results obtained by MD simulations within different force fields. The same is true for any type of ring inversions occurring for residues in a chain.

Both the GROMOS 56a6_{CARBO} and CHARMM force fields show an excellent agreement with both the Angyal scheme prediction and (where available) experimental data. The largest divergences are observed for compounds that are assumed to exhibit very rigid rings, independently of the method of calculations, e.g. β -Gal, β -Tal or α -Man. In the case of experimental data the largest divergence corresponds to the 56a6_{CARBO} force field and β -anomers of Ido and Alt and equals to ~ 6 kJ/mol. Note that for each compound the sign of the ring-inversion free energy remains uniform with that of the Angyal scheme and/or experimental data.

A similar level of qualitative agreement is offered by the GROMOS 53a6_{GLYC} force field. Also in this case the sign of ΔF_i values remains in agreement with the Angyal scheme prediction and the NMR data. However, now a much broader dispersion of the ring-inversion free energies is suggested. This results in both more extreme values of ΔF_i observed for most of compounds ($|\Delta F_i| > 20$ kJ/mol for 10 out of 16 compounds) and larger divergences between 53a6_{GLYC} and Angyal scheme which, in the extreme case of β -Glc, can reach ~ 50 kJ/mol. Moreover, in spite of the correct sign of ΔF_i for α -Ido, the absolute difference between the calculated and NMR-inferred values equals to ~ 25 kJ/mol. The related difference for β -Alt is also large (~ 12 kJ/mol).

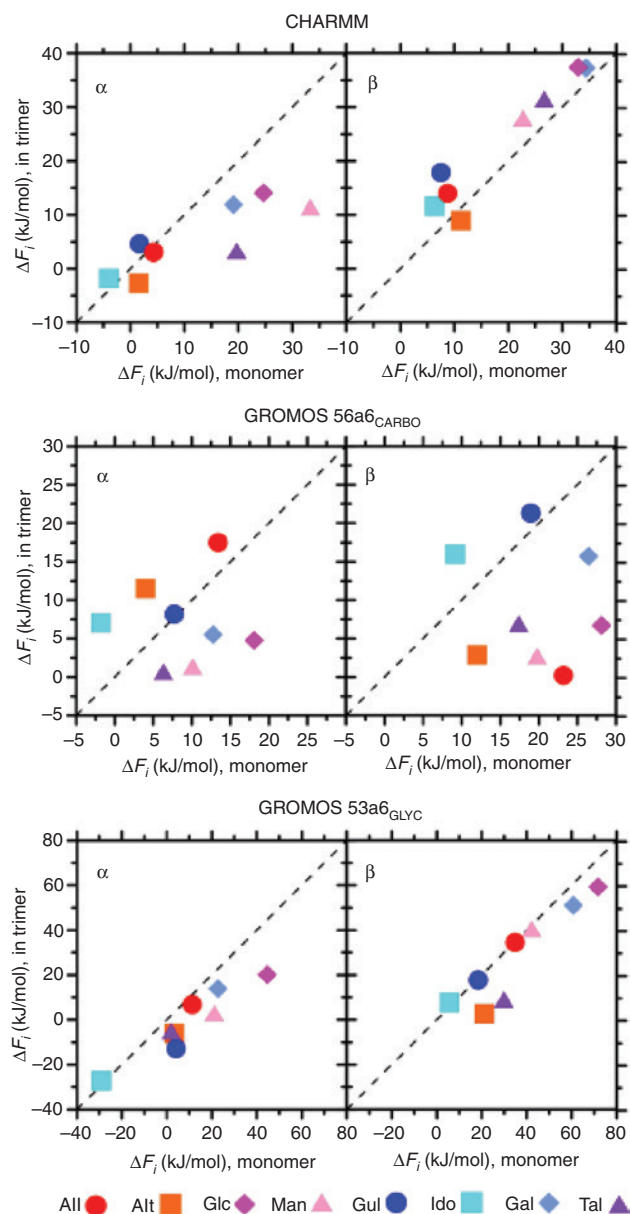


Fig. 2: Chair-chair-inversion (${}^4C_1 \rightarrow {}^1C_4$) free energies ΔF_i for the central residue of $\alpha(1 \rightarrow 4)$ - and $\beta(1 \rightarrow 4)$ -linked homotrimers in water calculated using three different force fields, compared to corresponding values for monomers.

The least degree of agreement between MD simulations and the available experimental data and/or the Angyal scheme was exhibited by the GLYCAM06 force field. GLYCAM06 is clearly at odd with the experimental data concerning α -Alt and suggests strongly dominating population of 1C_4 conformers ($\Delta F_i < -25$ kJ/mol). Furthermore, GLYCAM06 predicts the negative values of ΔF_i for α -All and α -Gul which is rather surprising in the context of both the semi-empirical scheme-inferred data [48] and the ab initio calculations [49]. The lack of parameters for Ido additionally hinders the comparison with the second set of existing experimental data.

The reasons for relatively worse GLYCAM06 performance in comparison with other force fields are most likely associated with the method of parametrization accepted during its development. Contrary to both GROMOS force fields and CHARMM, each molecule of unfunctionalized hexopyranose modeled by GLYCAM is associated with its own unique set of parameters (e.g. partial atomic charges) developed on the basis of ensemble-averaged conformation. As the dominating conformer is usually 4C_1 , the derived charges are suited

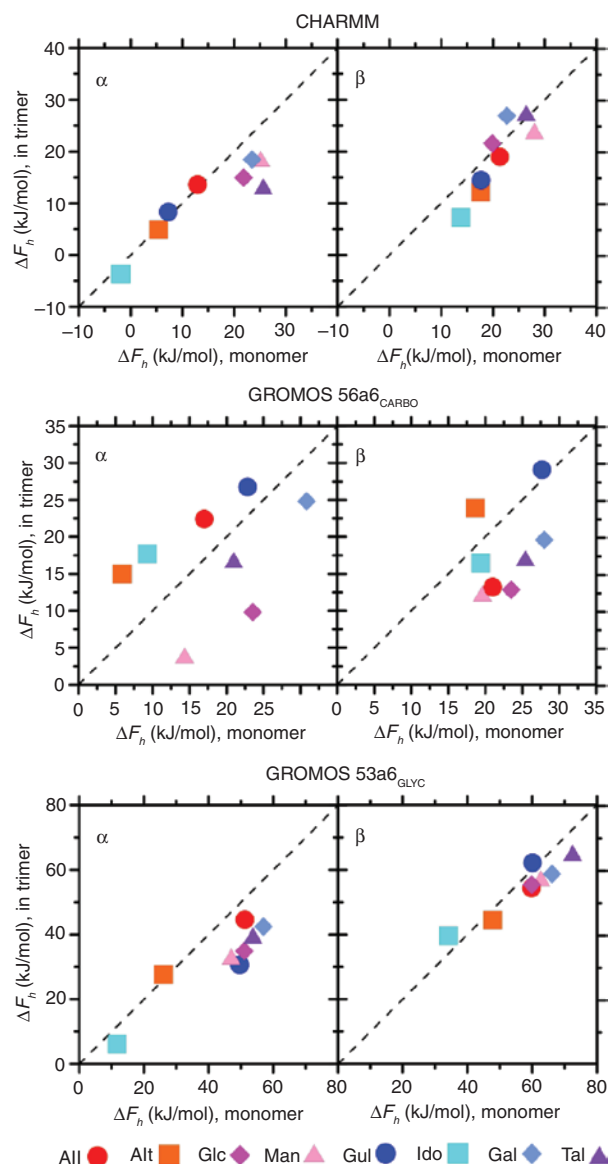


Fig. 3: Regular chair-boat/skew-boat-inversion ($^4C_1 \rightarrow B/S$) free energies ΔF_h for the central residue of $\alpha(1 \rightarrow 4)$ - and $\beta(1 \rightarrow 4)$ -linked homotrimers in water calculated using three different force fields, compared to corresponding values for monomers.

very well to describe the ensemble of states associated with the regular chair conformation but may fail to represent alternative ring shapes as well as their relative populations. GROMOS and CHARMM are designed to keep the set of parameters identical or nearly identical with respect to all hexopyranoses, thus, they suffer less from this effect.

Let us emphasize that the divergence between predictions of the given force field and the other types of data may influence the force field applicability only in some very special cases. This is due to the following reasons: (i) the behavior of the most frequently studied hexopyranose monomers (Glc, Man, Gal) is roughly independent of the applied force field and corresponds to strongly dominating population of the 4C_1 conformers; (ii) the free energy barriers separating particular ring conformers are high enough to prevent spontaneous ring ‘flips’ at the timescale of tens or even hundreds of ns of unbiased MD simulation.

Finally, let us refer to the ring-inversion free energies reported elsewhere [24]. The differences between our calculations and the data shown in table 4 of Ref. [24] are of quantitative nature and can be ascribed to different schemes of metadynamics-based calculations. Namely, calculations reported in Ref. [24] employed

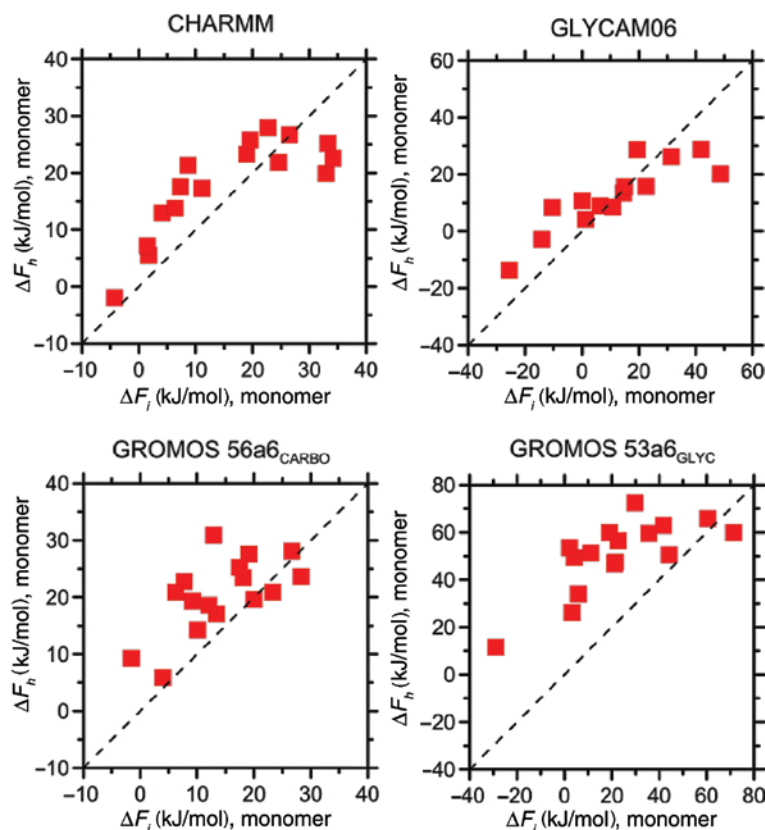


Fig. 4: Chair-chair-inversion (${}^4C_1 \rightarrow {}^1C_4$) free energies ΔF_i for the hexopyranose monomers in water calculated using three different force fields, compared to regular chair-boat/skew-boat-inversion (${}^4C_1 \rightarrow B/S$) free energies ΔF_h for the same compounds.

‘plain’ metadynamics, without bias scaling, whereas in the present case, the well-tempered metadynamics scheme was used in combination with the replica-exchange method. Appreciating the existence of these minor differences, let us note that also in this case, the signs of ΔF_i values remain uniform.

The tested force fields, compared in the context of correlation between ring-inversion properties of monomers and those of residues in a chain, reveal a series of differences in their predictions. As discussed in details in Ref. [17], in the case of GROMOS 56A6_{CARBO}, the values of ΔF_i are tendentially reduced in magnitude upon insertion of the free residue into a chain (for 12 out of 18 trisachcarides). Furthermore, the correlation between the ΔF_i values for the monomers and for the corresponding residues in a chain is very poor along the D-hexopyranose series. In contrast, there exists a much higher correlation between the same quantities calculated within both CHARMM and GROMOS 53a6_{GLYC} force fields. While the monomer- and trimer-related ΔF_i values for β -anomers are very close to each other (except of decrease of ΔF_i by ~ 20 kJ/mol for β -Alt and β -Tal simulated within GROMOS 53a6_{GLYC}), the tendential decrease of the trimer-related ΔF_i values is observed for 4–6 out of 8 α -anomers (i.e. for Glc, Gal, Man and Tal in the case of CHARMM and Glc, Gal, Man, Tal, Gul and Alt in the case GROMOS 53a6_{GLYC}). Interestingly, Glc, Gal, Man and Tal are the same compounds for which also the GROMOS 56A6_{CARBO} force field predicts the largest decrease of ΔF_i upon inserting residue into a $\alpha(1 \rightarrow 4)$ -linked chain. All of these compounds contain equatorial hydroxyl group at C_3 in the 4C_1 chair (see Fig. 1 for notation). Note that both the CHARMM and GROMOS 56A6_{CARBO} force fields agree in their predictions that the dominating ring conformation at the level of (non-Ido) monomer is the same as that of residue in a chain. On the contrary, in the case of GROMOS 53a6_{GLYC}, there are three compounds for which the monomer- and trimer-related ΔF_i values differ in their signs, i.e. α -Alt, α -Gul and α -Tal. The reason for that is a tendency to systematic decrease of ΔF_i exhibited by GROMOS 53a6_{GLYC} in the case of $\alpha(1 \rightarrow 4)$ -linked residues. Such decrease in combination with the residues of relatively small, monomer-related, ΔF_i values results in the negative values of ring-inversion free energies for residues in a chain.

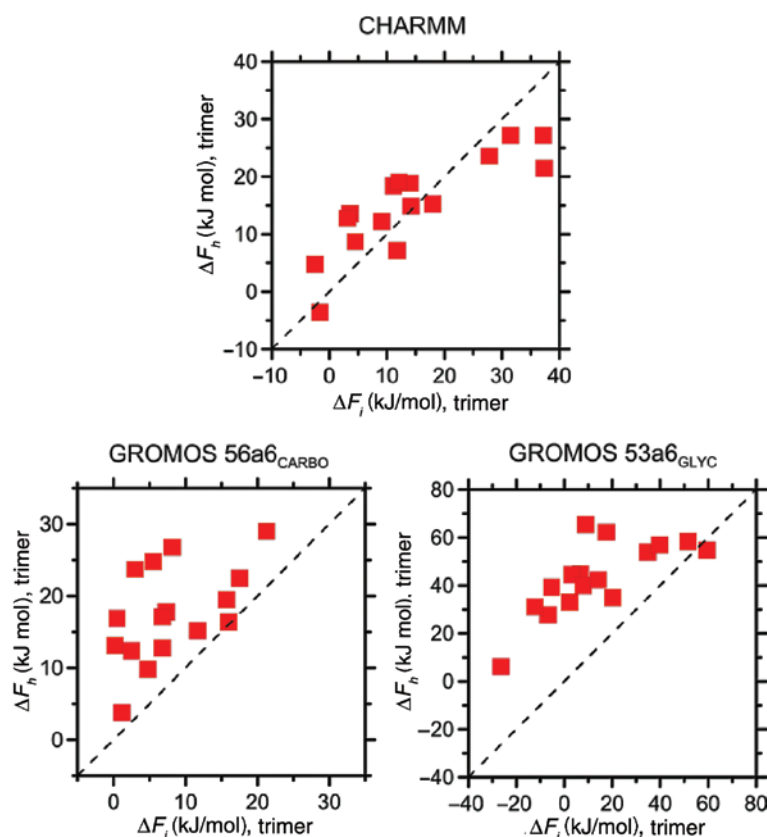


Fig. 5: Chair-chair-inversion (${}^4C_1 \rightarrow {}^1C_4$) free energies ΔF_i for the central residue of $\alpha(1 \rightarrow 4)$ - and $\beta(1 \rightarrow 4)$ -linked homotrimers in water calculated using three different force fields, compared to regular chair-boat/skew-boat-inversion (${}^4C_1 \rightarrow B/S$) free energies ΔF_h for the same compounds.

Analogous investigations focused on the possible correlation between the ΔF_h values calculated for the monomer and for the central residue in trimer give similar results, i.e. a poor correlation for GROMOS 56A6_{CARBO} and much better ones for CHARMM and GROMOS 53a6_{GLYC}. Interestingly, ΔF_h values characteristic of residues in a chain are again decreased in comparison to those of monomers for the same group of four α anomers containing equatorial hydroxyl group at C₃ in the 4C_1 chair conformer (the related decrease is additionally exhibited by α -Gul in the case of GROMOS 53a6_{GLYC}).

There is no quantitative experimental data that would allow to assign the ring-inversion free energy value to the residue being a part of a larger chain. The existing qualitative data determining the preferred conformation of such residues concern only selected systems and agree with the predictions of all tested force fields (e.g. Refs. [12–14] suggesting the 4C_1 chair conformation as dominating in residues of the amylose chain). Similarly, there is no semi-empirical schemes generalized for the case of glycosylated hexopyranoses and no ab initio methods have been used so far to systematically study this issue. For these reasons, currently there is no possibility to compare the quality of the force fields on the basis of the data collected in Table 2. However, the most striking difference between force fields is the presence or the absence of the correlation between ring-inversion properties of monomers and those of residues in a chain. This is especially important in the context of the lack of quantitative data concerning the ring conformation of residues in a chain and the possibility of predicting of such properties on the basis of relatively better recognized properties of respective monomers.

The mentioned difference can be explained when considering the monomer-inherent flexibility of the rings predicted by each of the tested force fields. The chair-chair ring-inversion free energies suggested by GROMOS 56A6_{CARBO} for monomers vary from ~ -2 to ~ 28 kJ/mol (i.e. in the range of ~ 30 kJ/mol). The corresponding ranges for the CHARMM and GROMOS 53a6_{GLYC} force fields are larger and equal to ~ 40 kJ/mol

and ~ 100 kJ/mol, respectively. Let us assume that the non-systematic alterations of the ring flexibility upon inserting residue into a chain, as suggested by GROMOS 56A6_{CARBO}, originate from the residue-residue interactions and that the strength of such interactions is of comparable magnitude for each of the force fields. Then, the ‘decorrelation effect’ will be more significant in the case of force field that predicts larger flexibility of rings, i.e. GROMOS 56A6_{CARBO}. The residue-residue interactions of the same strength are less important for rings of greater degree of rigidity, i.e. those described in terms of CHARMM and, especially, GROMOS 53a6_{GLYC}. Such hypothesis would link the possible decorrelation of ring-inversion free energies with the capability of the tested force fields to reflect very subtle differences between ring puckering at the level of monomers. As most of the currently available data allow only to assign a particular ring conformation as the dominating one and are not capable to characterize this domination by providing the ring-inversion free energy (or equivalent quantity), the possible alternative methods of choice would include MD-related techniques employing ab initio potentials of increased accuracy (see e.g. Ref. [21, 22]). However, due to the high significance of intramolecular hydrogen bonding [50] that strongly influences the conformation of hexopyranoses in the absence of solvent, such simulations would require the presence of explicit solvent, possibly modeled at the molecular-mechanics level of theory.

Interestingly, there exists a relatively good correlation of the chair-chair inversion free energies (ΔF_i) values with those of chair-boat/skew-boat ones (ΔF_h). The correlation is present for all tested force fields, however, it becomes very poor in the case of residue in a chain simulated within GROMOS 56A6_{CARBO} and is of (better) comparable quality for the remaining cases (see Figs. 4 and 5). In the case of monomers, with the exception of very rigid rings ($\Delta F_i > 20$ kJ/mol), the relation $\Delta F_i \leq \Delta F_h$ is always satisfied. For residues in a chain the same inequality holds except of α - and β -Ido simulated within the CHARMM force field. This means that the moderately flexible rings should be described in terms of the ${}^4C_1 \rightarrow {}^1C_4$ transition rather than the ${}^4C_1 \rightarrow B/S$ one. For more rigid rings the population of B/S conformers may be larger in comparison to 1C_4 , however both of them will be very scarce in relation to 4C_1 . Finally, the most flexible rings (All and Ido, according to predictions of all force fields) may require a more complex description involving both types of transitions or, possibly, introducing additional coordinate(s) accounting for the ring shape within the B/S set of conformers (e.g. Cremer-Pople ϕ).

Conclusions

Four different classical force fields have been tested in the context of their capabilities to simulate the ring distortion in hexopyranose molecules (in monomers and in residues in a chain exploiting either $\alpha(1 \rightarrow 4)$ or $\beta(1 \rightarrow 4)$ glycosidic linkages). Both the GROMOS 56a6_{CARBO} and CHARMM force fields offer very similar predictions for unfunctionalized monomers which, at the same time, remain in agreement with the available experimental data and with the semi-empirical Angyal scheme. GROMOS 56a3_{GLYC} also reflects this agreement at the qualitative level but differs quantitatively from GROMOS 56a6_{CARBO} and CHARMM by predicting a much broader dispersion of ring-inversion free energies varying significantly from one compound to another. GLYCAM06 exhibits the worst accuracy at the level of unfunctionalized monomers, being at odd with the NMR data for α -Alt and predicting the dominating population of inverted chairs for α -Gul and α -All. The most frequently studied hexopyranoses (Glc, Man, Gal) exhibit similar ring-inversion properties, independently of the force field used. All force fields predict alterations of the chair-chair inversion free energies upon inserting residue into a chain. However, the scale and direction of such alterations are dependent on the force field: GROMOS 56a3_{GLYC} and CHARMM suggest a strong correlation between monomer- and trimer-related properties and increased flexibility of some of the $\alpha(1 \rightarrow 4)$ -linked residues, relative to the properties of corresponding monomers. According to GROMOS 56a6_{CARBO}, the ring-inversion free energies for residues in a chain are typically smaller in magnitude compared to those of the monomers (for both the $\alpha(1 \rightarrow 4)$ - and $\beta(1 \rightarrow 4)$ -linkages) and correlate poorly with the latter. Except of very rigid rings, chair-boat/skew-boat inversion free energies are higher than those corresponding to the chair-chair inversion. For all tested force fields, there exists a correlation between chair-chair and chair-boat/skew-boat inversion free energy for both monomers and residues in a chain. However, it becomes rather poor for trimers simulated within GROMOS 56a6_{CARBO}.

Finally, as the availability of the reliable quantitative data related to the ring distortion of residues in a chain is very limited, the use of the force field-independent, ab initio potentials in combination with the presence of explicit solvent may be suggested as an alternative method to study this issue.

Acknowledgments: The authors acknowledge the financial support of the Polish National Science Centre (contract financed in 2012–2015 under Project No. 2011/03/D/ST4/01230 SONATA). We thank Dr V.H. Rusu for providing the topology files of some of the molecular structures used in this work.

References

- [1] B. M. Sattelle, J. Shakeri, A. Almond. *Biomacromolecules* **14**, 1149 (2013).
- [2] H. B. Mayes, L. J. Broadbelt, G. T. Beckham. *J. Am. Chem. Soc.* **136**, 1008 (2014).
- [3] J. Iglesias-Fernández, L. Raich, A. Ardèvolb, C. Rovira. *Chem. Sci.* **6**, 1167 (2015).
- [4] J. Topin, M. Lelimosin, J. Arnaud, A. Audfray, S. Pérez, A. Varrot, A. Imbert. *ACS Chem. Biol.* **11**, 2011 (2016).
- [5] J. R. Snyder, A. S. Seriani. *J. Org. Chem.* **51**, 2694 (1986).
- [6] E. Autieri, M. Sega, F. Pederiva, G. Guella. *J. Chem. Phys.* **133**, 095104 (2010).
- [7] J. Rönnols, M. T. C. Walvoort, G. A. van der Marel, J. D. C. Codée, G. Widmalm. *Org. Biomol. Chem.* **11**, 8127 (2013).
- [8] D. R. Ferro, A. Provalosi, M. Ragazzi, G. Torri, B. Casu, G. Gatti, J. C. Jacquinet, P. Sinay, M. Petitou, J. Choay. *J. Am. Chem. Soc.* **108**, 6773 (1986).
- [9] D. A. Horita, P. J. Hajduk, L. E. Lerner. *Glycoconjugate J.* **14**, 691 (1997).
- [10] P. E. Marszalek, A. F. Oberhauser, Y.-P. Pang, J. M. Fernandez. *Nature* **396**, 661 (1998).
- [11] P. E. Marszalek, Y.-P. Pang, H. Li, J. E. Yazal, A. F. Oberhauser, J. M. Fernandez. *Proc. Natl. Acad. Sci. USA* **96**, 7894 (1999).
- [12] H. Sugiyama, T. Nitta, M. Horii, K. Motohashi, J. Sakai, T. Usui, K. Hisamichi, J. I. Ishiyama. *Carbohydr. Res.* **325**, 177 (2000).
- [13] B. Casu, M. Reggiani, G. G. Gallo, A. Vigevani. *Tetrahedron* **22**, 3061 (1966).
- [14] V. S. R. Rao, J. F. Foster. *J. Phys. Chem.* **67**, 951 (1963).
- [15] B. M. Sattelle, A. Almond. *Phys. Chem. Chem. Phys.* **16**, 8119 (2014).
- [16] D. Cremer, J. A. Pople. *J. Am. Chem. Soc.* **97**, 1354 (1975).
- [17] W. Plazinski, A. Lonardi, P. H. Hünenberger. *J. Comp. Chem.* **37**, 354 (2016).
- [18] H. S. Hansen, P. H. Hünenberger. *J. Comput. Chem.* **32**, 998 (2011).
- [19] W. Plazinski, M. Drach. *Carbohydr. Res.* **415**, 17 (2015).
- [20] W. Plazinski, M. Drach, A. Plazinska. *Carbohydr. Res.* **423**, 43 (2016).
- [21] X. Biarnés, A. Ardèvol, A. Planas, C. Rovira, A. Laio, M. Parrinello. *J. Am. Chem. Soc.* **129**, 10686 (2007).
- [22] K. Govender, J. Gao, K. J. Naidoo. *J. Chem. Theory Comput.* **10**, 4694 (2014).
- [23] W. Plazinski, M. Drach. *Carbohydr. Res.* **416**, 41 (2015).
- [24] L. Pol-Fachin, V. H. Rusu, H. Verli, R. D. Lins. *J. Chem. Theory Comput.* **8**, 4681 (2012).
- [25] W. Plazinski, M. Drach. *RSC Adv.* **4**, 25028 (2014).
- [26] V. Spiwok, B. Králová, I. Tvaroška. *Carbohydr. Res.* **345**, 530 (2010).
- [27] O. Guvench, S. N. Greene, G. Kamath, J. W. Brady, R. M. Venable, R. W. Pastor, A. D. MacKerell, Jr. *J. Comput. Chem.* **29**, 2543 (2008).
- [28] O. Guvench, E. R. Hatcher, R. M. Venable, R. W. Pastor, A. D. MacKerell, Jr. *J. Chem. Theory Comput.* **5**, 2353 (2009).
- [29] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. Gonzalez-Outeirino, C. R. Daniels, B. L. Foley, R. J. Woods. *J. Comput. Chem.* **29**, 622 (2008).
- [30] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. Van Der Spoel, B. Hess, E. Lindahl. *Bioinformatics* **29**, 845 (2013).
- [31] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans. “Interaction models for water in relation to protein hydration”, in *Handbook of Intermolecular Forces*, B. Pullman (Ed.), Springer-Science + Business Media, B. V. Israel (1981).
- [32] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein. *J. Chem. Phys.* **79**, 926 (1983).
- [33] G. Bussi, F. L. Gervasio, A. Laio, M. Parrinello. *J. Am. Chem. Soc.* **128**, 13435 (2006).
- [34] Y. Sugita, Y. Okamoto. *Chem. Phys. Lett.* **314**, 141 (1999).
- [35] A. Barducci, G. Bussi, M. Parrinello. *Phys. Rev. Lett.* **100**, 020603 (2008).
- [36] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. Broglia, M. Parrinello. *Comput. Phys. Comm.* **180**, 1961 (2009).
- [37] R. W. Hockney. *Methods Comput. Phys.* **9**, 136 (1970).
- [38] S. Miyamoto, P. A. Kollman. *J. Comput. Chem.* **13**, 952 (1992).
- [39] G. Bussi, D. Donadio, M. Parrinello. *J. Chem. Phys.* **126**, 014101 (2007).
- [40] B. Hess. *J. Chem. Theory Comput.* **4**, 116 (2008).

- [41] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije. *J. Comput. Chem.* **18**, 1463 (1997).
- [42] H. J. C. Berendsen, W. F. van Gunsteren, H. R. Zwinderman, R. G. Geurtsen. *Ann. N. Y. Acad. Sci.* **482**, 269 (1986).
- [43] J. A. Barker, R. O. Watts. *Mol. Phys.* **26**, 789 (1973).
- [44] I. G. Tironi, R. Sperb, P. E. Smith, W. F. van Gunsteren. *J. Chem. Phys.* **102**, 5451 (1995).
- [45] T. N. Heinz, W. F. van Gunsteren, P. H. Hünenberger. *J. Chem. Phys.* **115**, 1125 (2001).
- [46] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen. *J. Chem. Phys.* **103**, 8577 (1995).
- [47] P. J. Steinbach, B. R. Brooks. *J. Comput. Chem.* **15**, 667 (1994).
- [48] S. J. Angyal. *Angew. Chem. Int. Ed.* **8**, 157 (1969).
- [49] U. Schnupf, J. L. Willett, F. Momany. *Carbohydr. Res.* **345**, 503 (2010).
- [50] D. Wang, M. L. Ámundadóttir, W. F. van Gunsteren, P. H. Hünenberger. *Eur. Biophys. J.* **42**, 521 (2013).