PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2015, **17**, 21622

The water-catalyzed mechanism of the ring-opening reaction of glucose

Wojciech Plazinski,*[a] Anita Plazinska[b] and Mateusz Drach[c]

Received 10th June 2015,
Accepted 17th July 2015

DOI: 10.1039/c5cp03357h

www.rsc.org/pccp

The hexopyranose mutarotation is an important focus for carbohydrate chemistry for more than 150 years. The paper describes the results of advanced computational studies aimed at elucidating the ring-opening reaction of glucose. Molecular simulations based on the combination of the DFT method with the molecular dynamics formalism allowed for a detailed insight into the mechanism of the process accompanied by the information of the kinetic and dynamic nature. The results indicate that the process is initiated by deprotonation of the anomeric hydroxyl group by water molecules and the subsequent proton transfer to the ring oxygen atom. The latter event has been identified as a 'bottleneck' of the process triggering the ring cleavage. The most time-consuming steps of the ring-opening reaction are the orientational rearrangements of water molecule(s) participating in the proton transfer(s) and the final extension of the newly-formed aldehyde chain. The orientational preferences of the aldehyde group present in the acyclic form of D-hexopyranoses are responsible for the anomeric equilibrium characteristics.

1. Introduction

Glucose (D-glucopyranose) is one of the most significant compounds for both chemistry- and biology-related sciences. The 'textbook' reaction of glucose mutarotation has been known for over 150 years[1–3] but some of its details still remain unresolved. The interconversion from the α to β anomer and contrariwise occurring *via* the acyclic aldehyde form is of great importance not only for glucose but also for other carbohydrates. The glucose mutarotation can be catalyzed by either bases or acids but between pH 4 and 6 the water-catalyzed reaction becomes dominant and sometimes referred to as 'uncatalyzed' mutarotation due to the absence of a large amount of either hydronium or hydroxide ions.[1] According to the experimental data, the presence of a water environment is essential for the anomeric glucose interconversion.[4] The ring opening-reaction can also be catalyzed by natural enzymes and is an important component of glycans biochemistry.[5,6]

During the last few years various computational studies accompanied by the experimental measurements allowed for verifying a series of possible mechanisms of mutarotation.[7–11] Lewis *et al.* concluded[9] that the most possible mechanism of the ring-opening reaction occurs with contribution of only one water molecule. The earlier studies[7,8] concern the contribution of the larger number of water molecules. Recent Car–Parrinello molecular dynamics simulations devoted to the Brønsted acid-catalyzed reaction of the glucose mutarotation[10] allowed for estimating the free energy surface for opening the ring structure. In spite of these achievements, there are still a number of uncertainties related to the dynamic features of the ring opening reaction. The partial reason for that are the shortcomings of the theoretical methods commonly used to model the course of the chemical reaction: they are usually based on the geometry optimization routines or require defining the arbitrary reaction coordinates which are believed to capture the most essential features of the reaction.

In the preset work we focus on the ring-opening reaction of the molecules of α- and β-glucose, catalyzed by water. We apply the computational method based on the combination of the transition path sampling (TPS) approach[12,13] with the molecular dynamics protocols employing the *ab initio* potentials. Such a combination allows us to study the unbiased, molecular dynamics trajectories representing the favorable reaction paths of the process of interest. Additionally, it does not require defining any reaction coordinate which may be useful from the perspective of discovering the alternative paths which are not reflected by the small set of predefined coordinates. During the simulations described here we applied the 'hybrid' potentials of interactions according to which the particularly crucial parts of the systems (*i.e.* glucose molecules and some selected water molecules)

[a] *Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, ul. Niezapominajek 8, 30-239 Cracow, Poland. E-mail: wojtek_plazinski@o2.pl, wojtek@vega.umcs.lublin.pl; Fax: +48 815375685; Tel: +48 815375685*
[b] *Laboratory of Medicinal Chemistry and Neuroengineering, Department of Chemistry, Faculty of Pharmacy, Medical University of Lublin, W. Chodzki Str., 4a, 20-093 Lublin, Poland*
[c] *Department of Theoretical Chemistry, Faculty of Chemistry, UMCS, pl. M. Curie-Sklodowskiej 3, 20-031 Lublin, Poland*

are treated at the *ab initio* level of accuracy whereas the rest of the system (*i.e.* the remaining water molecules) are modeled by using the classical potentials. Although we apply this methodology only for the ring-opening reaction, the remaining aspects of the mutarotation process are studied in the context of the conformational features of the acyclic carbohydrates and their influence on the final product of the ring-closing reaction.

The results allowed for the detailed description of the full process of water-catalyzed ring-opening; this includes determining the number of water molecules participating in the reaction, types of proton transfers occurring and identifying the structural and conformational features which influence the process.

2. Methods

All molecular dynamics (MD) simulations were performed by using the GROMACS 4.55 package.[14] During simulations of the ring-opening reaction, the density functional theory (DFT)[15,16] potentials were calculated every step by applying the 6-31+G* basis set[17] and BP86 functional.[18,19] The ORCA 2.9 software[20] was used for that purpose. The DFT-based potentials were used to describe the interactions within the selected parts of the system (*i.e.* the whole glucose molecule and the water molecules located close to the C1–O5 bond, surrounding the 'reactive site' and able to participate in the reaction), while the 'DFT' part of the system interacts with the rest of the system (*i.e.* with the remaining water molecules) *via* 'classical' potentials. The atom labels are presented in Fig. 1. The number of water molecules treated at the DFT level of theory was equal to ~8 in the case of the TPS study; their locations were monitored during simulations. The Lennard-Jones parameters and partial charges of glucose molecules (used only to describe the interactions with classically modelled water molecules) were adopted from the CHARMM36 force field.[21] The CHARMM-compatible TIP3P model[22] was accepted to simulate water molecules. The all-atom CHARMM force field was chosen instead of united-atom one (*e.g.* GROMOS, used in the second part of the study) to avoid the problematic treatment of aliphatic hydrogen atoms which are not explicitly accounted for in GROMOS but must be present in the system at the DFT level. Note also that the choice of the classical force field for this stage of the study is of secondary importance with respect to the calculations performed at the DFT level of theory; the role of the CHARMM-derived Lennard-Jones parameters and partial charges of glucose molecule is to account for the interaction of the 'reactive' part of the system with the solvent environment.

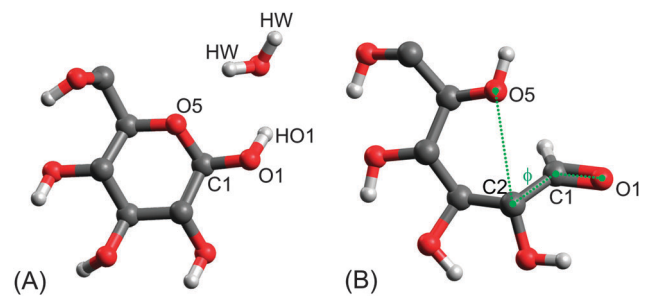

Fig. 1 (A) The glucose molecule (cyclic form), water molecules participating in the ring-opening reaction and accepted atom labels used in the description of the TPS protocol. (B) The acyclic form of glucose molecules with corresponding atom labels and the definition of the ϕ dihedral angle which are used to describe the metadynamics simulations.

The simulations were carried out under periodic boundary conditions and under the NPT conditions. The temperature was maintained close to its reference value (298 K) by applying the V-rescale thermostat,[23] whereas the Parrinello–Rahman barostat[24] was used to control the pressure (1 bar). The centre of mass motion was removed every step. The equations of motion were integrated using the leapfrog scheme[25] with a time step of 0.75 fs. Nonbonded interactions were computed using plain cutoff at a distance of 1.4 nm.

Initially, one molecule of glucose (either α of β anomer) in the $^4C_1$ ring conformation was placed in the centre of simulation cell of dimensions 3.0 × 3.0 × 3.0 nm^3. The system was subjected to the energy minimization (steep descent algorithm) and short MD (~5 ps) simulation based on the 'hybrid' protocol. The initial paths required by the TPS procedure were generated by applying various bias potentials on the distance between selected pairs of atoms (purpose of which was to *e.g.* increase or decrease the distance within the C1–O5 and HW–O5 pairs, respectively). The most effective procedure appeared to be the forced attachment of the HW hydrogen to O5 which resulted in spontaneous completion of the whole process, however, other methods were tested as well, *e.g.* forced stretching of the O5–C1 bond in the glucose molecule surrounded by the 'DFT' waters, able to participate in the proton exchange process. The procedure was repeated for various arrangements of water molecules surrounding glucose and led to discovering the two alternative paths common for both anomers. These paths were used in the subsequent TPS analysis.

Additionally, we performed several short MD runs initiated from the six structures (including glucose and catalytic water molecule) identified as potential transition states for the ring-opening reaction in ref. 9. The coordinates of glucose and catalytic water were initially frozen and the arrangement of remaining water molecules was randomized by running 50 ps long MD run. Then, some randomly chosen frames were used to initiate the 'hybrid' type of MD simulations, during which only glucose and one water molecule (that present in the original structure) were treated at the DFT level of theory.

The TPS method was applied to sample the ''reactive' trajectories, *i.e.* those connecting predefined initial and final states ('basins of attraction' determined by the free energy minima). The random walk through the trajectory space is performed by generating new trajectories from old ones by a shooting move algorithm combined with the Metropolis rule.[12,13] The randomly chosen timeframe (the shooting point) of the old trajectory is the starting point for creating a new trajectory by integrating the equations of motion forward and/or backward in time by using the conventional MD. We have used the TPS algorithm as described in ref. 26, *i.e.* the one-way, flexible path length algorithm, which was also applied in the

case of our previous studies.[27,28] The V-rescale (Bussi) thermostat[23] was used as a generator of stochastic noise. The other details of generating and accepting the new trajectories were based on the procedures described in ref. 26–28 and on the MD parameters listed above. In the case of all sampled trajectories, the data were collected every timestep, as the longer time space does not allow us to observe the system evolution in the vicinity of the free energy barriers. We defined the initial and final states (*i.e.* cyclic and acyclic forms of glucose molecules) based on the interatomic distances calculated for the pairs: O5–C1 (δ) and O5–HW (ω) where HW is any hydrogen atom belonging to water molecules at the beginning of the process. Initially (*i.e.* for simulations performed for one possible reaction path and for α-glucose), the glucose molecule was treated as cyclic when $\delta \leq 0.15$ nm and $\omega > 0.3$ nm. At the same time, the acyclic form was defined by $\delta > 0.3$ nm and $\omega \leq 0.1$ nm. After recognizing that a large amount of trajectory length is devoted to the conformational reorganization occurring after ring-opening reaction (*e.g.* conformational relaxation of the aldehyde chain) we decided to decrease the maximum values of both δ and ω to 0.25 nm in the three remaining cases.

The free energy profile aimed at explaining the possibility of the ring-opening reaction for the anionic form of glucose (lacking HO1 atom) was calculated for the α anomer by using the umbrella sampling approach.[29] The C1–O5 distance was used as the coordinate (δ). The pull code in GROMACS was used to generate snapshots for the US simulations from an initial pulling trajectory. Along the reaction coordinate, 7 windows were selected with a distance of ~0.02 nm between the adjacent positions for each of the systems. At this stage all water molecules were treated classically. The data within each window were collected for every 4 fs (12 ps for each window). After removing the first 2 ps for equilibration, the free energy profile was constructed with the weighted histogram analysis method (WHAM)[30] as implemented in GROMACS (*g_wham*).[31] Statistical uncertainties were estimated using the Bayesian bootstrapping of complete histograms.

The rotational preferences of the aldehyde moiety in the acyclic form of hexopyranoses were studied without using the DFT-based potentials; instead the GROMOS 53a6 force field[32,33] was used to model the glucose molecule in combination with the SPC water model.[34] The parameters for the aldehyde form of carbohydrates were generated by Automated Topology Builder.[35] The considered hexopyranoses have the acyclic form, therefore the usual, carbohydrate-dedicated force fields are not applicable. The simulation systems contained one hexopyranose molecule placed in the cubic ($3.0 \times 3.0 \times 3.0$ nm^3) simulation box containing about ~900 water molecules. Before the enhanced sampling protocol, the standard, unbiased MD simulations were performed for ~5 ns to equilibrate the systems. The equations of motion were integrated with a timestep of 2 fs. The LINCS algorithm[36,37] was applied to constrain all bond lengths. Nonbonded interactions were computed using a twin-range scheme,[38] with the short- and long-range cutoff distances of 0.8 and 1.4 nm, respectively, and a frequency of 5 timesteps for the update of the short-range pairlist and intermediate-range interactions. To account for electrostatic interactions beyond the long-range cutoff radius, a reaction field correction was applied using a relative dielectric permittivity of 61.[39] If not indicated otherwise, the other simulation parameters were identical to those used during the TPS protocol. The calculated free energy landscape was associated with the two coordinates: the distance between C1 and O5 atoms (δ) and the dihedral angle based on the O5–C2–C1–O1 quadruplet of atoms (ϕ). We applied parallel-tempering[40] combined with well-tempered metadynamics[41,42] which allows for sampling of all degrees of freedom (improving metadynamics accuracy) and improves the exploration of low probability regions. The calculations were performed by GROMACS combined with PLUMED 1.3.[43] Sixteen metadynamics simulations were calculated in parallel at different temperatures (298, 302.01, 306.07, 310.18, 314.33, 318.53, 322.77, 327.06, 331.41, 335.80, 340.24, 344.73, 349.23, 353.85, 358.49, 363.19 K) with configuration exchange according to the replica exchange scheme. The starting deposition rate was set to be 0.01 kJ mol^{-1} ps^{-1} with the Gaussian width equal to 0.05 nm (δ) or 0.2 rad (ϕ) and a temperature parameter ΔT (see eqn (2) in ref. 41) equal to 1788 K.

Note that the issues considered in the first (the ring-opening reaction studied by using the TPS approach) and the second (rotational feature of the aldehyde moiety studied by using the metadynamics method) stages of the study: (i) are of different nature; (ii) are studied by using diverse computational methodologies ('hybrid' potentials with the DFT component *vs.* classical force field); (iii) concern results which are not directly compared with each other. Therefore, we used two different classical force fields (CHARMM and GROMACS) with water models compatible with them to address each of the problems separately.

3. Results and discussion

3.1 'Static' *vs.* 'dynamic' transition states

Before initiating the TPS-based harvesting of reactive trajectories we checked the relation between the transition state (TS) structures reported by other authors[9] and the dynamic trajectories which can be developed from them by using the computational methodology used in the present study. By initiating multiple, unbiased MD trajectories differing in initial velocities and arrangements of the non-catalytic water molecules, we concluded that only one basin of attraction was possible to reach in each of the considered cases. However, the type of this basin is dependent on the initial structure. More specifically, only the structures which exhibit the water molecule contact with the O5 atom initiate MD trajectories leading to the open-ring form of glucose. In the remaining cases (water molecules in the contact with either one or two hydroxyl groups) only the cyclic form of glucose was possible to reach. This observation suggests that the water contact with O5 may be an important structural pattern which triggers the ring-opening process. Note also that each of the initial structures was very close to the cyclic glucose form as no significant cleavage was observed for the O5–C1 bond. Independently of the type of free energy minima associated with the given type of TS structure, one has to appreciate that the 'static' TS structures determined based on the saddle

points on the free energy surface do not necessarily correspond to the 'transition states' defined based on the dynamic properties of the system (*e.g.* the commitment probabilities[44]). Note that based on such a simple comparative procedure it is not possible to propose the alternative definitions of TS based on committor analysis.[26,28,44]

3.2 Initial trajectories

The multiple attempts of generating the initial trajectories needed for the subsequent TPS procedure lead to producing the two distinct types of reaction paths, common both for α and β anomers. The first of them involves the participation of one water molecule which plays a role of both the proton acceptor (accepts the HO1 proton) and the donor (donates one of its protons to the O5 atom). This type of mechanism will be referred to as 1-WM. In the case of the mechanism of the second type the two different water molecules are necessary to complete the ring-opening reaction; one of them plays a role of a proton acceptor (from O1) whereas the other acts as a proton donor (to O5). This kind of mechanism will be called 2-WM. The molecular details of both these reaction paths will be given in the further paragraphs whereas their symbolic depictions (corresponding to the already relaxed paths) can be found in Fig. 2. The selected 'reactive' trajectories corresponding both to 1-WM and 2-WM paths were taken as the initial ones in the subsequent TPS procedure. Although the selection of the initial paths may seem to be somehow arbitrary, the further trajectories are assumed to quickly diverge from the initial one and to relax to those representing favorable reaction paths. Note also that we were not able to obtain the MD trajectories representing the alternative mechanisms, such as: (i) the direct, intramolecular proton transfer from O1 to O5, not involving the water molecule(s); (ii) proton transfers involving more than two water molecules. Despite such reaction paths have been considered by other authors,[7,8] they most likely do not correspond to the unbiased, dynamic reactive trajectories. The mechanism excluding the water contribution is not likely to play a role as, according to the experimental data, in dry benzene, anomeric interconversion is virtually nonexistent.[4] As noted in Methods section, the most effective procedure for generating the reactive path was the forced attachment of the HW hydrogen to O5. Other tested procedures (*e.g.* forced stretching of the O5–C1 bond) also led to attachment of HW to O5 before the completion of the ring-opening process which speaks against alternative mechanisms (*e.g.* the concerted attachment of HW to O5 and breaking of the O5–C1 bond). It is possible that such behavior was just an artifact of the applied external force but such a force was absent during the subsequent, unbiased TPS simulations and it was possible for the reactive paths to relax into a more favorable state. It should be stressed that any type of mechanism was not assumed *a priori* and the only limitation were the accepted definitions of initial and final states. Finally, while analyzing the TPS results we observed the significant divergences from the initial paths (described in the further parts of the paper), which may be treated as indicators of good sampling of reactive trajectories.

3.3 Reaction mechanism

The relaxed trajectories harvested during the TPS procedure differ significantly from the initial ones; apart from the changed sequence of the molecular events (discussed later), the second difference is the spontaneous change of the mechanism from 2-WM to 1-WM observed in the case of β-glucose (also discussed further). As the definition of the basins of attraction varied slightly from one system to another we do not perform a detailed analysis of the length of the reactive trajectories. However, independently of the mechanism, the ring opening reaction is a very rapid event: the time needed to complete the transition from one to another basin varies from 0.035 to 0.3 ps. Apart from the definitions of the stable states, these differences result mainly from the initial arrangements of the water molecule(s) which catalyze the ring-opening reaction (this is stated based on the insight into the visualized MD trajectories). Furthermore, in some cases the majority of the trajectory length are devoted to conformational rearrangements occurring at the very end of the process (*e.g.* slow extension of the acyclic glucose chain). Thus, the overall lengths of the trajectories are sensitive to the boundary conditions (which are, at least to some extent, arbitrary). The most favorable initial structures, leading to the fastest occurrence of ring-opening, are those exhibiting the specific pattern of hydrogen

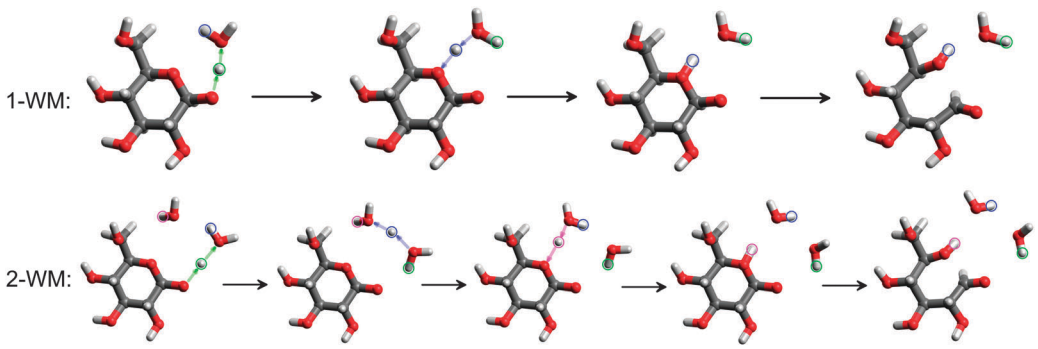

Fig. 2 Schematic depiction of two possible mechanisms of the ring-opening reaction distinguished based on the initial paths used to initiate the TPS procedure. Note that: (i) the schemes represent the already relaxed paths; (ii) in the case of β-glucose the 2-WM mechanism converges to the 1-WM one after further path relaxation. Further details in the text.

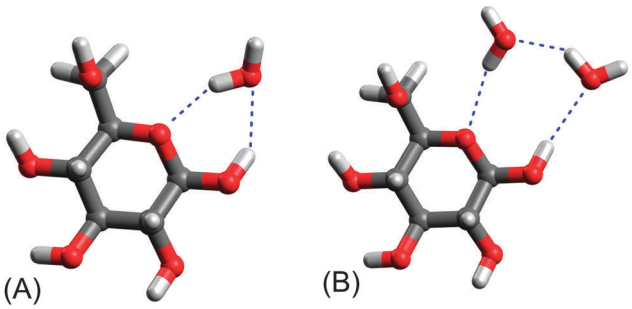

Fig. 3 The specific networks of hydrogen bonds facilitating the occurrence of the ring-opening reaction for: (A) 1-WM mechanism and (B) 2-WM mechanism. Note that the existence of such a network is not a necessary condition for the reaction to occur. See Section 3.3 for more details.

bonds (HB) between catalytic water molecule(s) and the reactive 'site' of glucose (including the hydroxyl group being the potential proton donor and the O5 atom, *i.e.* the proton acceptor). The two arrangements typical for 1-WM and 2-WM are shown in Fig. 3. If such an arrangement occurs in the initial frame of the reactive MD trajectory, the subsequent stages of the process occur within less than 0.05 ps. Otherwise, some more time is required for orientational rearrangement of at least one water molecule.

Independently either of the system or of the starting path there are number of common, subsequent molecular steps of the ring opening reaction briefly described below:

(i) Formation of a hydrogen bond between the anomeric hydroxyl group (–O1–HO1) and the water molecule; the hydroxyl group play the role of a HB donor. At the same time it is necessary for the water molecule to be located close (<0.45 nm) to the O5 ring atom.

(ii) The transfer of the HO1 proton to that water molecule combined with the subsequent transfer of other protons belonging to the same water molecule. In the latter case the type of acceptor depends on the reaction mechanism; it can be either O5 (1-WM) or other water molecule, located closer to O5 (2-WM).

(iii) In many cases both proton transfers occur nearly simultaneously, providing that there exists a related HB network (see Fig. 3). If not, the additional time is spent on the (usually very slightly) conformational changes of water molecule(s) allowing for the second proton transfer event. These changes include mainly the reorientation of the water molecule and formation of the appropriate HB with the potential proton acceptor (*i.e.* with another water molecule or O5).

(iv) In the case of the 1-WM mechanism, the proton is attached to the O5 ring atom which enables the (subsequent) breakage of the O5–C1 bond.

(v) In the case of the 2-WM mechanism, the proton is transferred to the second water molecule and yet another proton is attached to the O5 atom. Also in this case the rearrangement of the second water molecule may be required to form HB with O5. The next step (breakage of the C1–O5 bond) is analogous to that existing in the 1-WM mechanism.

(vi) The slow conformational rearrangement of the aldehyde molecule associated with the extension of its chain. Note that this step was not effectively sampled during TPS simulations but it is a logical consequence of the previous stages and is confirmed by the results of the MD simulations devoted to the acyclic form of glucose (described in the next section).

It should be stressed that the above sequence of the molecular events was not in any way imposed on the final results by the initial reactive trajectories. Rather, the proton attachment to the O5 atom was required to initiate the other steps of the process; however, such a mechanism quickly relaxed into those described earlier and presented in Fig. 2 as 1-WM and 2-WM. No other mechanisms than the stepwise 1-WM and 2-WM described above were observed (*e.g.* the concerted one, associated with simultaneous breaking of the O5–C1 bond and attachment of HW to O5) despite that it was fully allowed by both the TPS procedure and the accepted limiting values of interatomic distances δ and ω, defining the cyclic and acyclic forms of glucose.

Independently of the mechanism, the role of water is crucial for the process. During the first stages of the reaction water molecule acts as a Brønsted base attaching the OH1 proton whereas the resulting hydronium ion becomes a Brønsted acid donating (other) protons to the O5 atom.

In each of the studied cases the formation of HB between water molecules and anomeric hydroxyl groups was a necessary step to initiate the ring-opening reaction. Interestingly, we did not observe any other, specific patterns of hydrogen bonding as being essential for the further steps which would be formed before initiating the proton transfer processes. However, such a pattern can be an opportunistic cause of shortening the time needed for occurrence of one or two proton transfer events. Conversely, in the majority of the cases the hydrogen bond bridges have to be formed after rearrangements of water molecule(s). Thus, there is no contradiction between our results and observation made in ref. 9 stating that 'the water molecule participating in ring-opening must not bridge OH1 and O5'.

The main factor triggering the bond cleavage at O5–C1 is the attachment of the proton to the O5 atom. The path density plots (Fig. 4(A and B)) clearly indicate that the breaking of the O5–C1 bonds is not possible in the absence of proton binding O5. On the other hand, the deprotonation of the O1 oxygen does not (directly) influence the O5–C1 distance. The weak correlation between O1–HO1 and C1–O5 distances (see Fig. 4(B)) is a consequence of the molecular events following the deprotonation of O1. This result is in agreement with the previous stage of the study, in which the ring-opening process was possible to complete only for the structure exhibiting the O5–H_2O contact. The significance of the O5 protonation suggests that the main energetic barriers for the ring-opening reaction may be induced by the high affinity of the water molecules for protons (compare ref. 10).

The initial steps of the process involve the deprotonation of glucose molecules which donates the OH1 proton to water molecules. The stability of such anionic form of glucose was additionally confirmed by the umbrella sampling (US) simulations performed for the cyclic form of glucose, lacking HO1

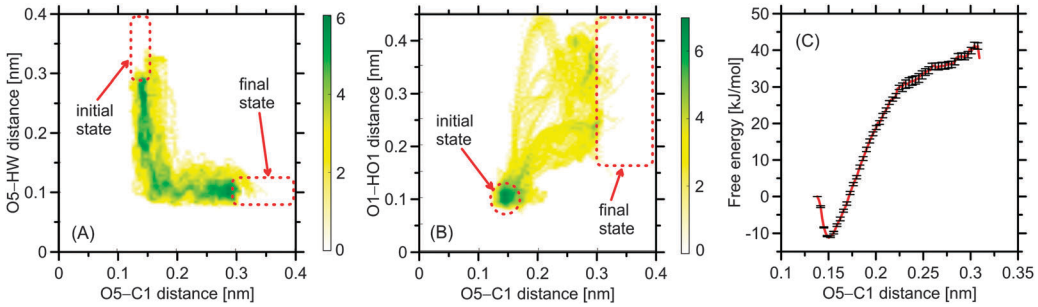

Fig. 4 (A and B) Path density plots for the two selected descriptors of the ring opening reaction. The paths shown here and representing the α-glucose case were selected due to the widest range of coordinate values but the analogous plots for β-glucose are essentially identical. The scale refers to the natural logarithm of the number of trajectories per one bin (C) the one-dimensional free energy profile corresponding to the non-catalyzed ring-opening reaction in the α-glucose anion lacking the HO1 proton. Further details in the text.

atoms and with the C1–O5 distance (δ) accepted as the single coordinate of the ring-opening reaction. The calculated free energy profile is given in Fig. 4(C). The free energy accompanied by the decyclization process monotonously increases reaching $\sim$50 kJ mol^{-1} at δ = 0.3 nm and shows no additional minima for $\delta > 0.15$ nm; it confirms that such a process is highly unfavorable without protonated O5. Note that the notable increase of the C1–O5 distance is possible to exist only if water molecules are not involved in the proton-transfer interactions with O5. Otherwise, stretching of the O5–C1 bond ends up with completing the conversion to aldehyde form associated with deprotonation of one of the water molecules.

The significance of the O5 protonation is in accordance with the recent Car–Parrinello MD simulations of the Brønsted acid-catalyzed ring-opening reaction.[10] According to ref. 10 such protonation was found to be the rate-limiting step for the overall process. This indicates that the bottleneck processes for both acid- and water-catalyzed mechanisms may be the same and the presence of a potential proton donor in the vicinity of O5 facilitates the O5–C1 bond cleavage. Furthermore, our results explain why it is impossible to observe the breaking of the O5–C1 bond without full protonation of O5.[9]

In spite of some features (discussed above) which are common for both the 1-WM and 2-WM mechanisms, they still represent quite different reaction paths. However, based on the results obtained for the set of trajectories representing the 2-WM mechanism and β-glucose one can state that the 1-WM mechanism is the more probable one. Namely, we have observed the spontaneous change of the favorable path from (initial) 2-WM to 1-WM. The trajectories collected after path 'switching' represent an essentially identical mechanism to that predicted by the 1-WM set of paths generated from the initial 1-WM-like trajectories. Therefore, we concluded that the mechanism involving only one water molecule is more favorable in comparison to those in which more water molecules are required to catalyze the process of interest. Interestingly, no analogous path-relaxing event was observed in the case of α-glucose. The possible interpretations include the existence of the two different mechanisms of ring-opening for α- and β-glucose but this would be contradictory to the results presented in ref. 9 and indicating that both the free energies and their enthalpic and entropic contributions are weakly dependent on the anomer type. Thus, we assume that the reason for this difference may be the sampling-related issues connected with the existence of free energy barriers impossible to reach by the TPS procedure. In spite of that this particular aspect of the results is somehow ambiguous, the remaining results (*e.g.* the role of water in the initial and final steps of the reaction, importance of the O5 protonation, *etc.*) are independent of the number of water molecules participating in the proton transfer events.

3.4 The orientation of the aldehyde group

The subsequent stage of the study was focused on the rotational features of the aldehyde moiety in the open-chain (aldehyde) form of hexopyranose molecules. Fig. 5 shows the 2D free energy surface (FES) related to the two following structural variables: (i) orientation of the formyl group with respect to the O5 oxygen atom (ϕ); (ii) distance between O5 and C1 atoms (δ) (see the Methods section and Fig. 1(B) for the definition of ϕ). The role of ϕ is to describe the conformational preferences of the aldehyde group when the O5–C1 distance is small, *i.e.* when the ring-closing reaction is possible to occur. From the geometric reasons, the positive values of ϕ represent the structure closer to the β anomer (the O1 oxygen oriented upwards with respect to the ring plane), whereas negative ϕ confirms that the structure is

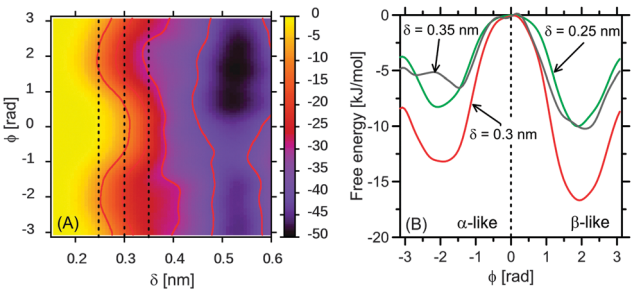

Fig. 5 (A) Two dimensional free energy profile related to the orientation of the aldehyde group in the acyclic form of glucose (ϕ) and the O5–C1 distance (δ). The three vertical dotted lines represent the three fixed δ values used to generate plots shown in panel (B). The scale is in [kJ mol^{-1}]. (B) The one-dimensional projections of the profile shown in panel (A) along the selected δ values. The vertical dotted line separates the two basins corresponding to the distinct orientational preferences of the aldehyde group.

Table 1 The comparison of the theoretical and experimental values of the relative free energy differences (in [kJ mol^{-1}]) between β and α anomers of D-hexopyranoses. The simulations were performed either for the acyclic (aldehyde) forms of hexopyranoses within the GROMOS 53a6 force field[32,33] of for their cyclic forms within the GROMOS 56a6$_{CARBO}$ force field[46]

	Allose	Altrose[a]	Glucose	Mannose	Gulose	Idose[a]	Galactose	Talose[a]
Experiment[b]	4.1	1.1	1.2	−1.7	4.8	0.25	1.7	−1.0
Simulations, acyclic form[c]	4.3	−0.9	2.2	−4.5	2.9	−2.9	1.1	−5.0
Simulations, cyclic form, 56a6$_{CARBO}$[d]	7.0	7.7	−1.2	−0.1	5.2	4.4	−0.4	−1.9

[a] Hexopyranoses with exceptionally high populations (∼30%) of furanose isomers. [b] Data taken from ref. 45. [c] Calculated from the projection of the 2D δ *vs.* ϕ free energy landscape for the arbitrary accepted δ value of 0.25 nm; the values correspond to the difference between the two free energy minima (see Fig. 5(B)). [d] Data taken from ref. 46.

closer to the α anomer (downward orientation of O1). The favorable orientation of the aldehyde group may indicate what will be the type of the main product of the ring-closing reaction (*i.e.* α or β anomers), as the rotation along the C2–C1 bond is hindered when the C1–O5 distance is small enough to initiate the ring-closing processes (which are assumed to be reversed with respect to those described here earlier). Note that FES was calculated by applying the classical force field, thus, no chemical reaction is considered at this point. The most preferred conformation of the aldehyde molecule is the extended chain which corresponds to the global minimum of the free energy located at δ = 0.50–0.58 nm. However, the ϕ variable is able to distinguish the orientation of the aldehyde moiety only for smaller values of δ. In the case of glucose, both the 2D FES and its 1D projections calculated for several values of δ confirm that the preferred orientation of aldehyde is that which can eventually lead to the β anomer. The experimentally measured $\alpha:\beta$ ratio for glucose is estimated to be equal to 38 : 62;[45] when expressing this ratio as the free energy value, it is equal to ∼1.2 kJ mol^{-1}. The related free energy differences vary from ∼0.7 kJ mol^{-1} (for δ = 0.2 nm) to ∼4 kJ mol^{-1} (for δ = 0.3 nm), depending on the fixed value of δ. The range of these values is in agreement not only with the experimental values[45] but also with the anomeric equilibrium predicted based on MD simulations of cyclic hexopyranoses.[46]

This remarkably small deviation is somehow surprising, considering the approximate nature of the atomistic description (classical force field) and neglecting the subsequent molecular steps of the process. To check if such agreement is possible to exist also for the case of other hexopyranoses we performed an analogous investigation for the remaining seven hexopyranoses of the D series. The summarized results are given in Table 1. They indicate that such an approximate approach can fairly well predict the preferred anomeric equilibrium in most of the cases (experimental data based on ref. 45); the deviations inverting the preferred ratio (*i.e.* characterized by the opposite signs of the free energy differences) are characteristic of altrose and idose, *i.e.* for the two compounds exhibiting the exceptionally large (>25%) population of furanose isomers. The furanose interconversion cannot be described in terms of the two accepted δ and ϕ coordinates and thus was not considered here. Even for those two cases the divergence between simulation results and experiment does not exceed 3 kJ mol^{-1}. Interestingly, the accuracy of the presented results is approximately the same compared to the force fields designed for correctly predicting the anomeric equilibrium[46] (see Table 1). This indicated the significance of the conformational preferences of the aldehyde residue as a crucial factor which determines the final product of the ring-closing reaction (α or β hexopyranose anomer). The good agreement of the classical simulations with the experimental data speaks either for the limited significance of the subsequent molecular events (*e.g.* proton transfers and the formation of the C1–O5 bond) on the final product of the ring-closing reaction or for the close similarity of these events for both types of the aldehyde group orientations.

4. Summary

The transition path sampling simulations of the ring-opening, water-catalyzed reaction of α and β glucose anomers allowed us to elucidate the molecular details of the process. This includes: (i) the description of the catalytic role of water; (ii) identifying the particularly crucial steps of the process; (iii) analysis of the conformational preferences of the aldehyde residue in the aldehyde form of glucose and its influence on the product of the ring-closing reaction. These results indicate that the most probable mechanism of the ring-opening reaction involves the contribution of only one water molecule. In the initial steps of the reaction the water molecule deprotonates the O1 atom acting as a Brønsted base. Then, after series of proton transfers, the O5 ring oxygen atom becomes protonated. This protonation is identified as the 'bottleneck' of the whole process which triggers the cleavage of the C5–O1 bond in analogy to the Brønsted acid-catalyzed reaction.[10] Such a reaction pattern is common for both α and β glucose anomers. According to the additional series of simulations, the orientational preferences of the aldehyde moiety determine the product of the ring-closing reaction. The calculated free energies remain in agreement with the experimentally measured $\alpha:\beta$ ratios for glucose and other D-hexopyranoses.

Acknowledgements

The authors acknowledge the financial support of the Polish Ministry of Science and Higher Education (contract financed in 2013–2015 under Project No. IP2012 006372) and the Polish National Science Centre (contract financed in 2012–2015 under Project No. 2011/03/D/ST4/01230 SONATA).

References

1 M. L. Sinnott, *Carbohydrate Chemistry and Biochemistry: Structure and Mechanism*, RSC Publishing, Cambridge, UK, 2007.

2 D. Horton, in *Carbohydrate Chemistry, Biology and Medical Applications*, ed. H. G. Garg, M. K. Cowman and C. A. Hales, Elsevier, 2008.
3 G. P. Moss, *Pure Appl. Chem.*, 1996, **68**, 2193.
4 C. G. Swain and J. F. Brown Jr, *J. Am. Chem. Soc.*, 1952, **74**, 2534.
5 J. A. Beebe, A. Arabshahi, J. G. Clifton, D. Ringe, G. A. Petsko and P. A. Frey, *Biochemistry*, 2003, **42**, 4414.
6 J. F. Bailey, P. H. Fishman, J. W. Kusiak, S. Mulhern and P. G. Pentchev, *Methods Enzymol.*, 1975, **41**, 471.
7 A. M. Silva, E. C. da Silva and C. O. da Silva, *Carbohydr. Res.*, 2006, **341**, 1029.
8 S. Yamabe and T. Ishikawa, *J. Org. Chem.*, 1999, **64**, 4519.
9 B. E. Lewis, N. Choytun, V. L. Schramm and A. J. Bennet, *J. Am. Chem. Soc.*, 2006, **128**, 5049.
10 X. H. Qian, *J. Phys. Chem. B*, 2013, **117**, 11460.
11 X. H. Qian, *Top. Catal.*, 2012, **55**, 218.
12 C. Dellago, P. G. Bolhuis and P. L. Geissler, *Adv. Chem. Phys.*, 2002, **123**, 1.
13 P. G. Bolhuis, D. Chandler, C. Dellago and P. L. Geissler, *Annu. Rev. Phys. Chem.*, 2002, **53**, 291.
14 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 and E. Lindahl, *Bioinformatics*, 2013, **29**, 845.
15 P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.
16 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.
17 R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724.
18 A. D. Becke, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1988, **38**, 3098.
19 J. P. Perdew, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1986, **33**, 8822.
20 F. Nesse, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2012, **2**, 73.
21 O. Guvench, S. S. Mallajosyula, E. P. Raman, E. Hatcher, K. Vanommeslaeghe, T. J. Foster, F. W. Jamison and A. D. MacKerell, *J. Chem. Theory Comput.*, 2011, **7**, 3162.
22 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
23 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
24 M. Parinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182.
25 R. W. Hockney, *Methods Comput. Phys.*, 1970, **9**, 136.
26 J. Vreede, J. Juraszek and P. G. Bolhuis, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 2397.
27 W. Plazinski and M. Drach, *J. Comput. Chem.*, 2012, **33**, 1709.
28 W. Plazinski and M. Drach, *RSC Adv.*, 2014, **4**, 25028.
29 G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187.
30 S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman and J. M. Rosenberg, *J. Comput. Chem.*, 1992, **13**, 1011.
31 J. S. Hub, B. L. de Groot and D. van der Spoel, *J. Chem. Theory Comput.*, 2010, **6**, 3713.
32 C. Oostenbrink, A. Villa, A. E. Mark and W. van Gunsteren, *J. Comput. Chem.*, 2004, **25**, 1656.
33 C. Oostenbrink, T. A. Soares, N. F. A. van der Vegt and W. van Gunsteren, *Eur. Biophys. J.*, 2005, **34**, 273.
34 H. J. C. Berendsen, J. P. M. Postma, W. van Gunsteren and J. Hermans, in *Handbook of Intermolecular Forces*, ed. B. Pullman, Springer-Science + Business Media, B.V., Israel, 1981.
35 A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink and A. E. Mark, *J. Chem. Theory Comput.*, 2011, **7**, 4026.
36 B. Hess, *J. Chem. Theory Comput.*, 2008, **4**, 116.
37 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463.
38 W. van Gunsteren and H. J. C. Berendsen, *Angew. Chem., Int. Ed.*, 1990, **29**, 992.
39 T. N. Heinz, W. van Gunsteren and P. H. Hűnenberger, *J. Chem. Phys.*, 2001, **115**, 1125.
40 Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.*, 1999, **314**, 141.
41 A. Barducci, G. Bussi and M. Parrinello, *Phys. Rev. Lett.*, 2008, **100**, 020603.
42 G. Bussi, F. L. Gervasio, A. Laio and M. Parrinello, *J. Am. Chem. Soc.*, 2006, **128**, 13435.
43 M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. Broglia and M. Parrinello, *Comput. Phys. Commun.*, 2009, **180**, 1961.
44 B. Peters and B. L. Trout, *J. Chem. Phys.*, 2006, **125**, 054108.
45 Y. Zhu, J. Zajicek and A. S. Serianni, *J. Org. Chem.*, 2001, **66**, 6244.
46 H. S. Hansen and P. H. Hünenberger, *J. Comput. Chem.*, 2011, **32**, 998.