

# Structure-activity relationships in carbohydrates revealed by their hydration

Laura Maugeri,<sup>†</sup> Sebastian Busch,<sup>‡</sup> Sylvia E. McLain,<sup>¶</sup> Luis Carlos Pardo,<sup>§</sup> Fabio  
Bruni,<sup>†</sup> and Maria Antonietta Ricci<sup>\*,†</sup>

*<sup>†</sup>Dipartimento di Scienze, Università degli Studi Roma Tre,  
Via della Vasca Navale 84, 00146 Roma, Italy.*

*<sup>‡</sup>German Engineering Materials Science Centre (GEMS)  
at Heinz Maier-Leibnitz Zentrum (MLZ),  
Helmholtz-Zentrum Geesthacht GmbH, Lichtenbergstr. 1  
85747 Garching bei München, Germany.*

*<sup>¶</sup>Department of Biochemistry, University of Oxford,  
South Park Road, Oxford, Oxfordshire OX1 3QU, UK*

*<sup>§</sup>Departament de Física i Enginyeria Nuclear,  
Universitat Politècnica de Catalunya  
08028 Barcelona, Catalonia, Spain.*

E-mail: mariaantonieta.ricci@uniroma3.it

Phone: +39(06) 5733-7226

## Abstract

One of the more intriguing aspects of carbohydrate chemistry is that despite having very similar molecular structures, sugars have very different properties. For instance, there is a sensible difference in sweet taste between glucose and trehalose, even though trehalose is a disaccharide comprised of two glucose units, suggesting a different ability of these two carbohydrates to bind to sweet receptors. Here we have looked at the hydration of specific sites and at the three-dimensional configuration of water molecules around three carbohydrates (glucose, cellobiose, and trehalose), combining neutron diffraction data with computer modeling. Results indicate that identical chemical groups can have radically different hydration patterns depending on their location on a given molecule. These differences can be linked with the specific activity of glucose, cellobiose, and trehalose as a sweet substance, as building block of cellulose fiber, and as a bioprotective agent, respectively.

**Keywords:** water; water-carbohydrate interactions; structure-function relationship; neutron diffraction; computer modeling; sweet taste.

## Introduction

Carbohydrates are among the most important naturally occurring biomolecules, with the premier role in metabolism of all living species. These molecules also have a myriad of other functions, working as bioprotective agents against dehydration and other environmental pressures and are important to the food industry as well as being a potential source of fuels.

One of the more intriguing aspects of carbohydrate chemistry is that despite having very similar molecular structures – which often only vary with respect to stereochemistry – sugars have very different properties which affect their ability to bind to protein receptors. For instance, there is a sensible difference in sweet taste between glucose and trehalose,<sup>1</sup> even though trehalose is a disaccharide comprised of two glucose units (linked by an  $\alpha$  – 1,1-glycosidic bond). Further, the intermolecular hydrogen bonding between glycol OH

groups, obviously present both in glucose and in trehalose, and the taste bud receptor site is considered the primary mechanism for sweet taste response by humans<sup>2</sup> and as such should be sufficient to elicit sweet response,<sup>2</sup> even though in practice it does not.

In the present work, we have considered the possibility that the different properties of carbohydrates, namely their ability to interact with receptors and/or other proteins, might be related to their different hydration shells. To tackle this issue, the hydration shell of glucose in solution has been determined using neutron diffraction experiments, enhanced by isotopic substitution (NDIS). Results are compared with the hydration shell of two disaccharides, namely trehalose<sup>3</sup> and cellobiose,<sup>4</sup> which are both made up by two glucose units yet linked by a different glycosidic bond. The rationale behind this approach is to compare the hydration pattern of the same building block, namely a single glucose unit, in three different situations: for glucose as a monomer, and when it is in two glucose-containing disaccharides with a different glycosidic bond between the glucose monomers, namely an  $\alpha - 1, 1$ -glycosidic bond for trehalose and a  $\beta - 1, 4$ -glycosidic bond for cellobiose. Interestingly and relevant in the present context, this latter carbohydrate is reported as tasteless. NDIS can measure the hydrogen bonding of molecules in aqueous solutions on the atomic length scale and as such is one of the premier techniques for structural determinations of hydrogen bonding in hydrogen-containing solutions,<sup>5-8</sup> including a number of biological molecules<sup>9-17</sup> in aqueous solutions. This is due, in part, to the fact that hydrogen and deuterium give a strong yet different signal in a neutron diffraction pattern. H/D substitution techniques allow the average structure and number of hydrogen bonding interactions to be assessed, with no need for extrapolation of structure at the atomic scale from dynamical properties of the sample.

The NDIS experiments have been augmented by two different tools: Empirical Potential Structure Refinement (EPSR)<sup>18</sup> and ANGULA.<sup>19,20</sup> EPSR is a computational technique which has been explicitly designed to provide atomic scale detail for disordered systems, using a set of NDIS data to constrain the 3-dimensional model of the solution in question.<sup>21-23</sup> While EPSR does not necessarily provide the only possible interpretation of the structural

data, it does provide a model that is consistent with the measured diffraction data. The program ANGULA was used to characterize the position of neighboring water molecules around the solute molecules - glucose, trehalose or cellobiose. The rationale behind such a combined experimental and computer simulation approach is to reveal similarities and differences between the local hydration of molecular sites and the hydration shells of glucose, cellobiose, and trehalose, linking the hydrated structure of each solute, on the atomic length scale, with its specific activity as a sweet substance, as building block of cellulose fiber, and as a bioprotective agent, respectively.

Linking the hydration of a sugar to its subsequent function could not only unveil the link between structure and function of carbohydrates in solution, but it could also help to unravel the biochemistry of sweet taste, a process that is yet to be fully understood.<sup>24</sup> It is relevant here to underline that to provide a full picture of the structure-function relationship presented by carbohydrates in solution, dynamical aspects of carbohydrate-water interactions and their implications in the sweetness reception should also be taken into consideration. While a very large body of information is available in the literature on the dynamical aspects of carbohydrate hydration (see for instance<sup>25,26</sup> and references therein), nothing is available on the sugar-sweet receptor dynamics of docking, possibly due to the lack of crystallographic structure of the receptor-ligand complex.

## Experimental and Data Analysis

Neutron diffraction experiments have been performed using the SANDALS neutron diffractometer, installed at the ISIS Facility (U.K.),<sup>35</sup> on an isotopomeric series of glucose-water solutions each at two concentrations corresponding to 1 solute molecule per 12.5 water molecules ( $\sim 4.4$  M), and to 1 solute molecule per 50 water molecules ( $\sim 1.1$  M), at ambient temperature and pressure (298 K, 1 bar). NDIS data for trehalose<sup>3</sup> and cellobiose<sup>4</sup> have been already published. Glucose ( $\text{C}_6\text{H}_{12}\text{O}_6$ ) samples were purchased from Sigma-Aldrich

and used without further purification. The sample list included fully hydrogenated and fully deuterated glucose (labelled as H and D, respectively), along with partially deuterated glucose samples, labelled H7 (all hydroxyl groups deuterated), and D7 with all hydrogen atoms, but those in the hydroxyl groups, deuterated. To fully exploit the advantages of isotopic substitution, a set of isotopically labelled glucose solutions were prepared (see Table 1). Neutron diffraction data have been collected for the empty instrument, empty container and vanadium standard, in order to normalize the data for all investigated samples to an absolute scale. These data sets have been processed using the “Gudrun” suite of programs,<sup>27,28</sup> which performs corrections for multiple scattering, absorption, inelasticity effects, and scattering from the samples. “Gudrun” also verifies that the measured scattered intensity is consistent with sample density and composition.

**Table 1.** Isotopic composition of the seven glucose solutions samples measured by neutron diffraction with isotopic substitution. The samples are labelled according to the format “xx/solvent” or “xx/yy/solvent”; for instance, the label of the first sample, xx/solvent = H7/D<sub>2</sub>O, indicates a glucose solution made with C<sub>6</sub>H<sub>7</sub>D<sub>5</sub>O<sub>6</sub> in D<sub>2</sub>O, whose H/D composition is given by 42% D, 58% H, and 100% D<sub>2</sub>O. Similarly, sample label xx/yy/solvent = D/H7/D<sub>2</sub>O indicates a glucose solution obtained by mixing C<sub>6</sub>D<sub>12</sub>O<sub>6</sub> and C<sub>6</sub>H<sub>7</sub>D<sub>5</sub>O<sub>6</sub> in D<sub>2</sub>O, whose H/D composition is given by 71% D, 29% H, and 100% D<sub>2</sub>O, and so on. All samples have been obtained by mixing H<sub>2</sub>O, D<sub>2</sub>O, C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>, C<sub>6</sub>D<sub>12</sub>O<sub>6</sub>, C<sub>6</sub>H<sub>7</sub>D<sub>5</sub>O<sub>6</sub>, and C<sub>6</sub>H<sub>5</sub>D<sub>7</sub>O<sub>6</sub> to obtain the desired solution concentration.

Sample label	%D	%H	% D <sub>2</sub> O	% H <sub>2</sub> O
H7/D <sub>2</sub> O	42	58	100	0
D/H7/D <sub>2</sub> O	71	29	100	0
D/D <sub>2</sub> O	100	0	100	0
D7/D/HDO	79	21	50	50
HD/HDO	50	50	50	50
H/H <sub>2</sub> O	0	100	0	100
D7/H <sub>2</sub> O	58	42	0	100

The outputs of “Gudrun” are the total neutron-weighted interference differential cross sections (IDCS) defined as

$$F(Q) = \sum_{\alpha} \sum_{\beta \geq \alpha} w_{\alpha\beta} [S_{\alpha\beta}(Q) - 1] \quad (1)$$

where  $\alpha$  and  $\beta$  label the atomic sites,  $Q$  is the magnitude of the change in the momentum vector by the scattered neutrons, defined as  $Q = 4\pi \sin\theta/\lambda$  where  $2\theta$  represents the scattering angle and  $\lambda$  the wavelength of scattered radiation. The functions

$$S_{\alpha\beta}(Q) = 4\pi\rho \int_0^\infty r^2(g_{\alpha\beta}(r) - 1) \frac{\sin Qr}{Qr} dr, \quad (2)$$

called partial structure factor (PSF), are the Fourier transforms of individual site-site radial distribution function (RDF)  $g_{\alpha\beta}(r)$ , and  $\rho$  is the atomic number density of the solution. The individual PSFs are weighted in Eq. 1 by  $w_{\alpha\beta} = c_\alpha c_\beta b_\alpha b_\beta (2 - \delta_{\alpha\beta})$ , where  $c_\alpha$  and  $c_\beta$  are the concentrations of the  $\alpha$  and  $\beta$  nuclei, and  $b_\alpha$  and  $b_\beta$  are their scattering lengths,<sup>29</sup> respectively.

Thus, each experimental IDCS is a linear combination of many PSF of the individual site-site radial distribution functions. In liquids with a small number of distinct atoms, like H<sub>2</sub>O, by measuring an array of different isotopically labeled samples, it is possible to directly extract all of the pair correlation functions from the experiment, giving a direct assessment of the hydrogen bonding present in the measured liquid. However, in more complex samples, like those investigated in the present report, it is not possible to isotopically label every atomic site. For this reason, we employ a simulation-assisted procedure that has been developed to convert IDCS data to real space, and extract a whole set of radial distribution functions. This is called empirical potential structure refinement (EPSR)<sup>18,30</sup> and is similar in principle to the methods routinely used in crystallography, which attempt to systematically refine a structural model to give best overall agreement with the diffraction data. Moreover, it should be noted that the larger is the number of isotopic contrast samples measured, the larger the number of constraints for the EPSR procedure. EPSR is required to fit all of the data sets, ensuring a physically reasonable model which is consistent with a set of measured diffraction data at the appropriate density and composition of each sample. Glucose in solution (at ambient pressure and temperature) undergoes mutarotation to give a 36:64 mixture of the anomers  $\alpha$ -D-glucose and  $\beta$ -D-glucose, respectively. To reproduce

the experimental sample, our EPSR simulation box contains both glucose anomers at a ratio of  $1\alpha/1.8\beta$  that closely resembles a glucose solution at the anomeric equilibrium.<sup>9</sup> The number of water molecules in the box and the box size is determined by the experimental sample concentration and density. As shown in Fig. 1, the measured diffraction data and the EPSR fits are in very good agreement; this ensures that the simulation box is indeed a good model of the real sample.

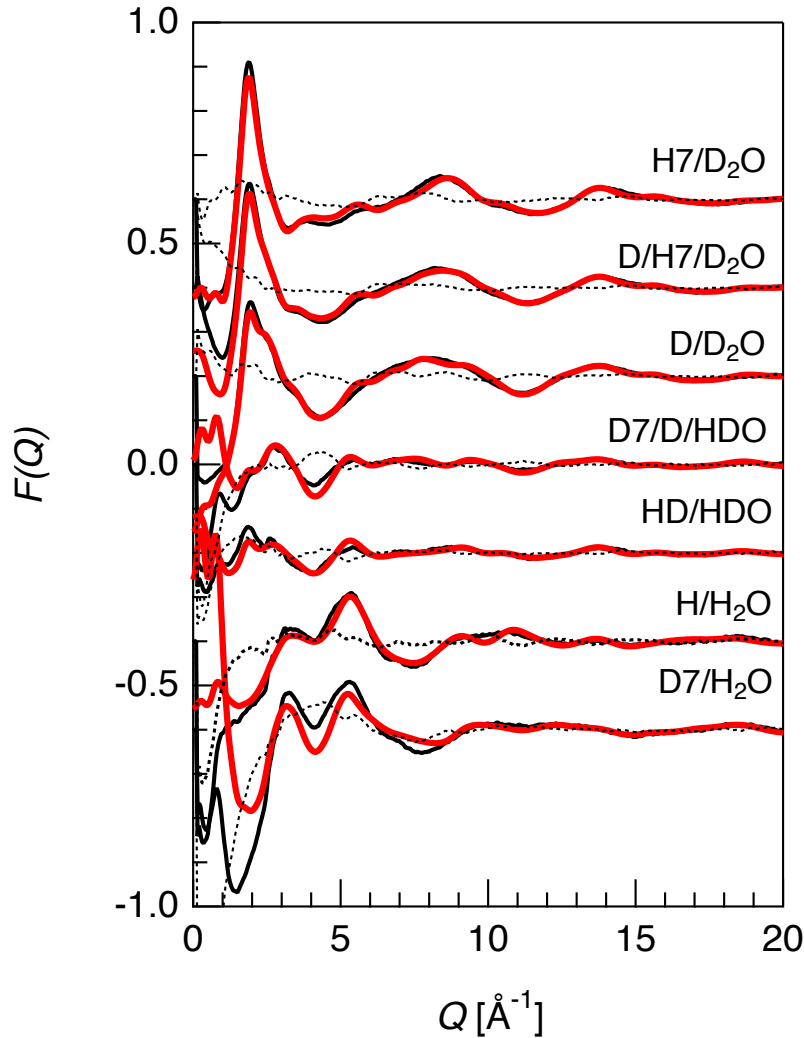


Figure 1: Color Online. The measured differential cross section of the 4.4 M glucose solution with different isotopic substitutions, as listed in Table 1. The solid black lines are the measured data, the red solid lines are the EPSR fits, and the dotted lines are the residuals. EPSR fits of similar quality have been obtained for all samples examined.

A set of EPSR configurations, all of which are consistent with the experimental data, has been analysed using the program ANGULA. This tool has been recently applied to understand the hydration of small biomolecules giving an unprecedented level of detail.<sup>31,33,34</sup> Two types of analysis were performed in the current work using ANGULA – namely nearest neighbor hydration of the -OH groups<sup>20</sup> and whole molecule analysis<sup>32</sup> In order to produce the nearest neighbor hydration, spatial density maps (SDM) around the hydroxyl and hydroxymethyl groups for the three sugars, orthonormal coordinate systems were assigned to one -OH and one -CH<sub>2</sub>OH group for glucose, cellobiose and trehalose. The distribution of the center of mass of the nearest neighbor water molecules, defined as the oxygen, was plotted relative to the center of mass of the orthonormal coordinate systems to each group. This procedure was performed for 5000 snapshots of the cellobiose and trehalose EPSR simulation boxes as well as 1500 snapshots of the glucose simulation, such that the aggregate nearest neighbor water molecule distribution was plotted *via* a spatial density map.<sup>31,33,34</sup> In addition, whole molecule analysis (WMA) to determine the most probable hydration around the entire molecule was also performed.<sup>32</sup> In this analysis the center of mass of any molecule within a chosen distance range from any atom of a central molecule can be extracted. WMA was performed for a distance range of 0 – 4 Å for water molecules around the sugar molecules, enabling the aggregate distribution of water around these sugars to be plotted. In order to generate the normalized distribution of molecules around the sugars ( $g^s(r)$  - the whole molecule surface RDF) - the probability distribution of the distances of an oxygen atom of water to the closest atom of the sugar molecules have been calculated. In a similar manner to the site-site radial distribution functions,  $g^s(r)$  approaches infinity as the radius increases. This function has been normalized using the asymptotic behavior of the calculated probability distribution function which is related with the surface of 'sheets' of hydration around the sugar molecules. It should be emphasized that the calculated  $g^s(r)$  does not have radial symmetry, and as a result it is not possible to calculate a radial distribution function for the surface by dividing to  $4\pi r^2$ .



## Results and discussion

Figure 2 shows the spatial density maps for water around a selected -OH group (representative of all other similar groups on the glucose ring) for the three carbohydrates (top row) compared with the SDMs for hydration about a selected hydroxymethyl group ( -CH<sub>2</sub>OH) for the three carbohydrates (bottom row). These SDMs represent the nearest neighbor hydration shells around specific sites on the solute molecules which have been extracted from the EPSR simulation box using the program ANGULA. Full details of the procedure required to calculate these SDMs are described elsewhere in the literature.<sup>31</sup> In both panels, the oxygen atom of either -CH<sub>2</sub>OH or -OH groups is at the origin of the reference frame, with its hydrogen along the  $z$ -axis, and the carbon atom just below the  $xy$ -plane in the negative  $x$ - direction.

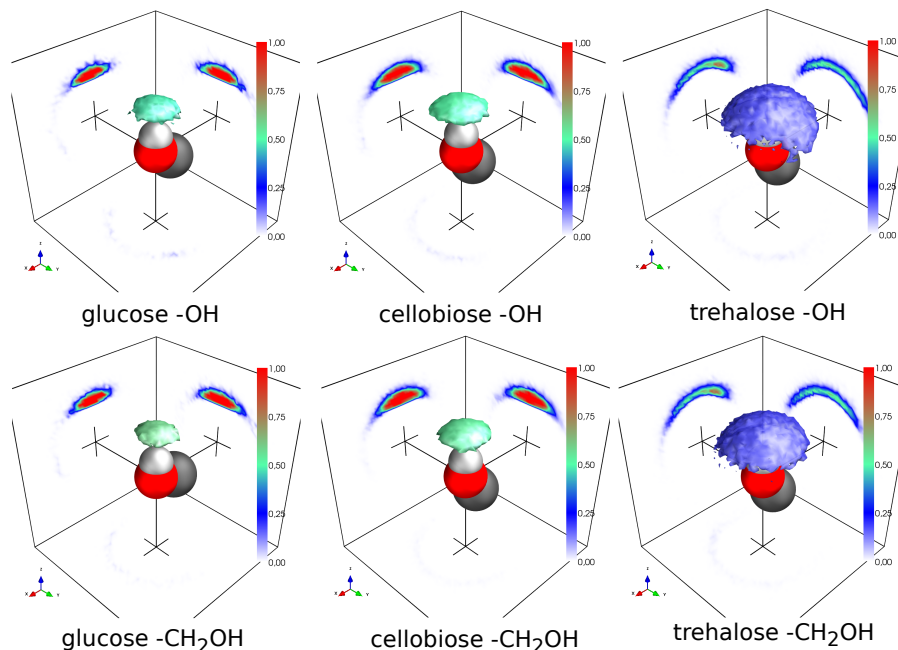


Figure 2: Color Online. TOP row. Spatial density function (SDMs) for nearest neighbor water molecules around the hydroxyl group (-OH) on glucose, cellobiose and trehalose. BOTTOM row: SDMs for water around the hydroxymethyl group (-CH<sub>2</sub>OH) on glucose, cellobiose and trehalose. In each case, the surrounding shells show the regions where the probability of finding a nearest neighbor water molecule is higher than 50%. In addition, for each SDM both cut throughs and heatmaps are shown to emphasize the highest probability location for water molecules in each closest hydration shell. While the shells show the top 50% of molecules, the cut throughs show the total number of molecules, averaged over the number of snapshots. The cut throughs are taken through the origin, i.e. the oxygen atom, and are subsequently displaced by 5 Å to the back for better visibility.

The surrounding density clouds in the SDMs indicate the regions where the probability of finding the nearest neighbor water molecule around the central oxygen atom, is higher than 50%. These shells predict the first nearest neighbor water (from the -OH hydrogens in each case) to these groups throughout the simulation and as such do not represent a particular distance range. A comparison between the SDMs, both for the hydroxymethyl and hydroxyl group, shows that the nearest neighbor water molecules are in a much more well defined location for both cellobiose and glucose compared with trehalose where the surrounding solvent shell is much broader. This is further supported by the cut-throughs in Fig. 2 which show a higher density of nearest neighbor waters directly on top of these group (in the  $+z$  direction) for both glucose and cellobiose compared with the same for trehalose which shows water distributed more broadly in a cap of density around these hydroxyl motifs. This suggests that both glucose and cellobiose form strong, directional hydrogen bond between the terminal hydrogen and the water oxygen, while an almost uniform hydration shell, at larger distances from the central oxygen, is found for trehalose, suggesting the presence of weakly hydrogen bonded water molecules. Distances of nearest neighbor water molecules from the the -OH and -CH<sub>2</sub>OH site can be obtained looking at Figure 3. This figure shows the nearest neighbor radial distribution functions for the O<sub>S</sub>O<sub>W</sub> correlations for the selected hydroxyl and the hydroxymethyl groups for the SDMs shown in Fig. 2 (red line) for each sugar, where O<sub>S</sub> is the oxygen atom on the solute molecule, either that of the hydroxyl (top row), or the hydroxymethyl group, (bottom row) and O<sub>W</sub> is the water oxygen. In this figure the RDFs for the total radial distribution functions, calculated from ANGULA, which encompasses the normalized distribution of all the water molecules around these particular groups are also shown (black line) as a comparison. Nearest neighbor water oxygen is located at about 2.8 Å from O<sub>S</sub> for both glucose and cellobiose -OH and -CH<sub>2</sub>OH groups, while the distance between O<sub>S</sub> and O<sub>W</sub> for trehalose is slightly larger (about 3 Å), for both its hydroxyl and hydroxymethyl sites. This confirms the presence of relatively weaker hydrogen bonded water molecules around these specific groups of trehalose, as noted above.

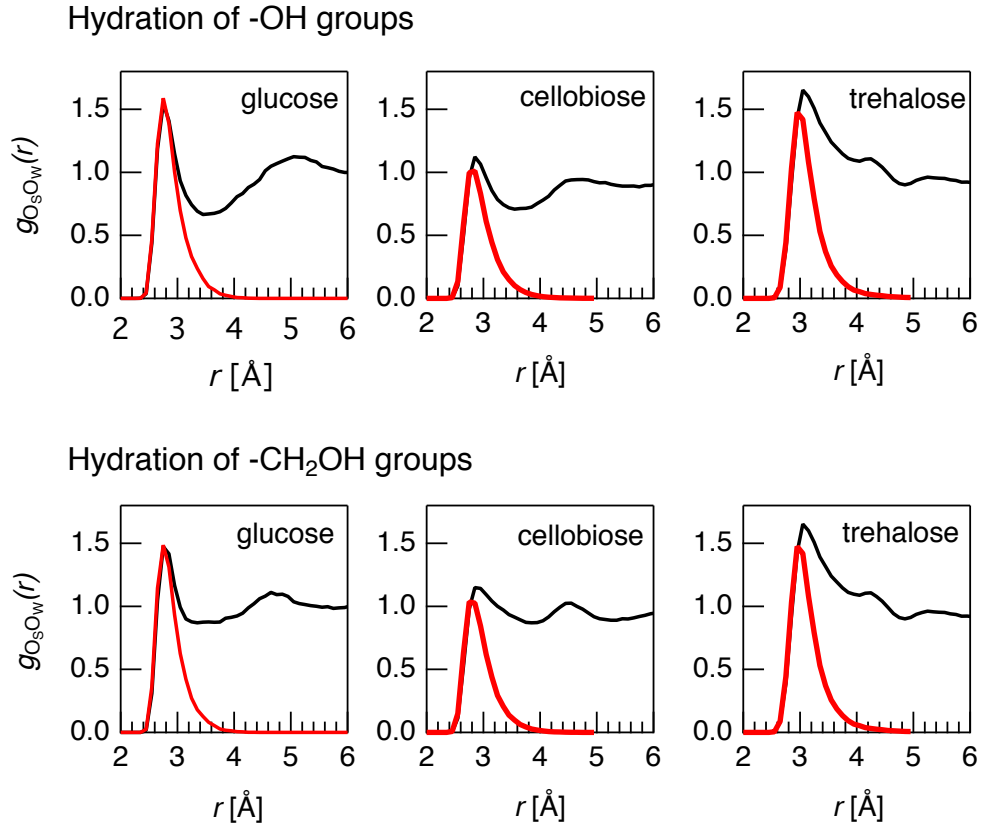


Figure 3: Color Online. Site-site correlation between the nearest neighbor water oxygen,  $O_w$ , and the oxygen atom of the solute,  $O_S$ , hydroxyl (top row, red line) and hydroxymethyl group (bottom row, red line). Radial distribution functions for the  $O_S$ - $O_W$  interactions for the SDMs shown in Fig. 2 are shown as a black line. The closest water oxygens for glucose and cellobiose are located about 2.8 Å away from the solute hydroxyl and hydroxymethyl sites, a distance compatible with strong hydrogen bonding. The closest water oxygens around trehalose are located instead at about 3.0 Å away from both -OH and -CH<sub>2</sub>OH groups, suggesting relatively weaker hydrogen bonding.

The distance between water molecules and the central solute molecule is shown in Figure 4. This plot reports the distribution function of water molecules around the "surface" of a solute. To calculate  $g_{\text{surface}}(r)$  we have taken the minimum distance between the oxygen atom of water and any atom of the solute molecule. This plot reveals that the closest water molecules for glucose and cellobiose are located about 1.8 Å away from the solute surface, a distance compatible with hydrogen bonding between hydroxymethyl and hydroxyl groups and the solvent. Conversely, the closest water molecules around trehalose are located at about 2.5 Å away from the solute surface, a distance too large to suggest strong hydrogen bonding between the solute and the solvent. All these observations are consistent with directional hydrogen bonding between glucose (or cellobiose) and the water molecules within its first hydration shell, while a much less directional hydrogen bonding pattern is found between trehalose and water molecules, suggesting a larger positional disorder and weaker bond strength.

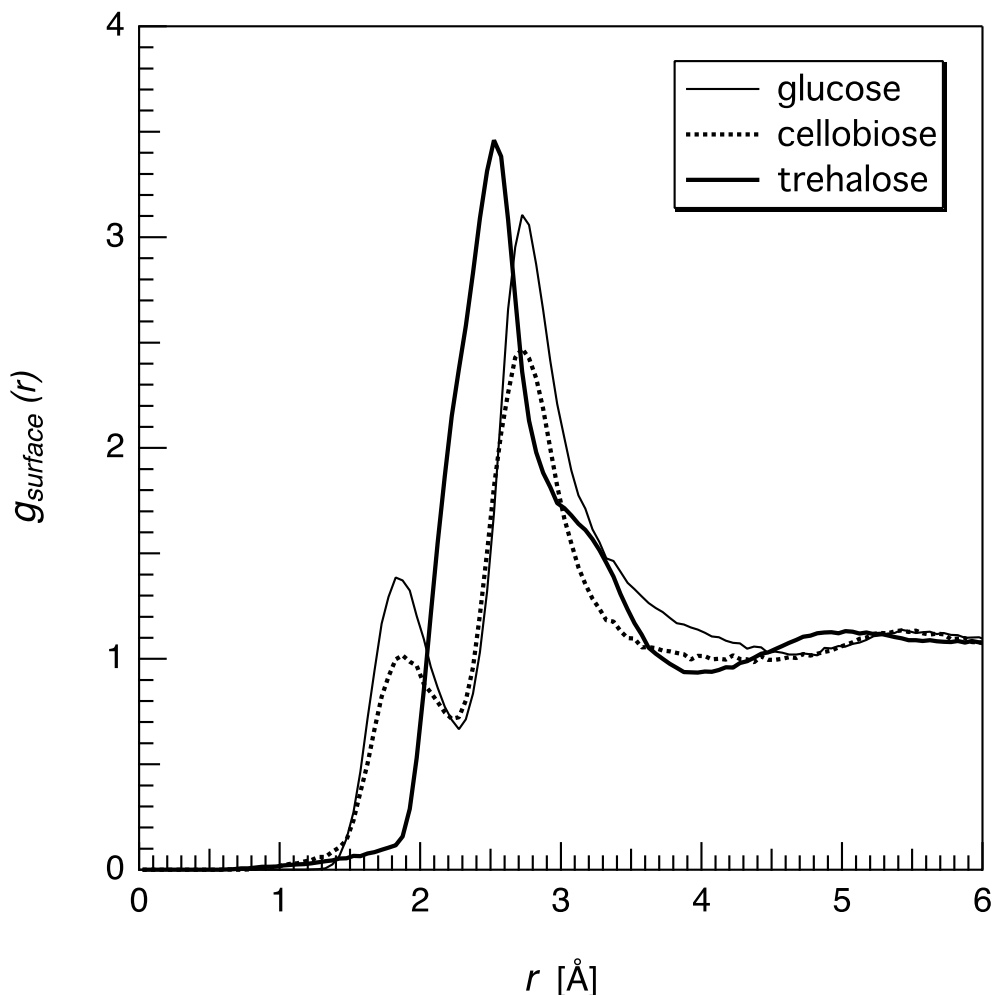


Figure 4: The distance between water molecules and the central solute molecule. This plot reports the distribution function of water molecules around the “surface” of a solute. It reveals that the closest water molecules for glucose and cellobiose are located about 1.8 Å away from the solute surface, a distance compatible with hydrogen bonding between hydroxymethyl and hydroxyl groups and the solvent. Conversely, the closest water molecules around trehalose are located at about 2.5 Å away from the solute surface, a distance too large to suggest strong hydrogen bonding between the solute and the solvent.

The three-dimensional arrangement of water molecules around the entire solute molecule has been obtained with ANGULA<sup>32</sup> and is shown in Figure 5. The SDMs have been calculated for glucose, cellobiose, and trehalose over the distance range 0 – 2.28 Å at 37% of the maximum probability. These highlight the most probable positions of all the water molecules around the three different solutes. As expected, the hydration patterns of glucose and cellobiose are not only quite similar, but are also consistent with the spatial density functions discussed above (Figure 2) showing the arrangement of water molecules around specific solute groups. In particular, the probability of finding a first neighbor water molecule around glucose and cellobiose is high close to the hydroxymethyl and hydroxyl groups, and low close to the oxygen atom and other sites on their rings. Notably, the hydration shell around one -OH site of glucose overlaps with that of a neighbor -OH site, resulting in a somewhat broader water density cloud compared to those of cellobiose.

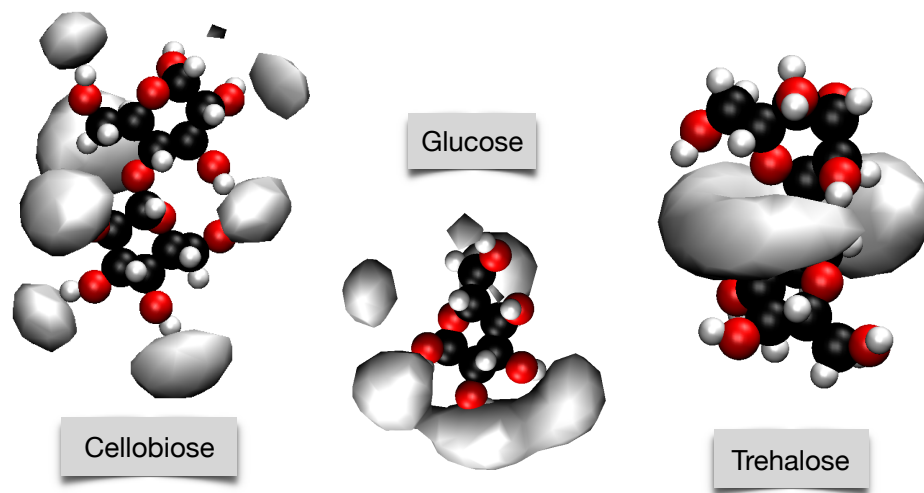


Figure 5: Color Online. The three-dimensional arrangement of water molecules around the entire solute molecule. These spatial density maps have been calculated for glucose, cellobiose, and trehalose over the distance range 0 – 3.5 Å at 37% probability.



In contrast, the hydration pattern of trehalose is significantly different. We recall that this is the only of the three molecules to rotate freely around the glycosidic linkage and the commercially available trehalose commonly crystallizes as a dihydrate, and in solution two “structural” water molecules bridging between its two rings were previously observed.<sup>3</sup> The SDM presented here for trehalose is consistent with this observation. These water molecules are likely those located between the two trehalose rings (Fig. 3 and Fig. 4). However, looking at this SDM it would not be correct to state that trehalose is hydrophobic but rather that trehalose is not involved in a large number of hydrogen bonding interactions with water compared with glucose or cellobiose and that its solubility is likely due to the presence of these “structural” waters around the glycosidic bond. It is interesting to notice that trehalose is about 6 times more soluble than cellobiose, suggesting that hydrogen bonding is not essential in describing solubility of a given carbohydrate.

We have used a combination of neutron diffraction experiments with isotopic substitution, computer simulation, and a new analysis tool to study the water-solute interactions at the atomic scale for three carbohydrates, namely glucose, cellobiose, and trehalose. Using this approach, possible connections between the hydration shells of each carbohydrate and their specific function could be obtained, as discussed in the following.

Starting with glucose, its ability to elicit sweet response has been attributed to the presence of a “functional structure” (AH-B-X, or “glucophore”) being compatible with a complementary tripartite AH-B-X site in the taste bud receptor.<sup>36</sup> According to this model, the AH site is a H-bond donor, the B site an H-bond acceptor, while X is an hydrophobic site. As a consequence, it has been suggested that sweetness originates from loose and temporal forces, such as hydrogen bonds, between sugars and taste receptors in a steric specific way.<sup>2,37</sup> At the present state of knowledge, this theory might appear too simplistic to explain all the observations; nevertheless, the tripartite AH-B-X concept, or its evolution called Multipoint Attachment theory,<sup>38</sup> has had its merits as a unifying criterion and proved useful in rationalizing structure-sweetness relationships in diverse classes of compounds After

40 years, the theory based on the presence of a glucophore structure still appears to be the best explanation for the ligand binding chemistry that induces sweet taste response, and it is also consistent with prevailing sweet taste transduction hypotheses.<sup>39</sup> Remarkably, the role of water in sweet taste chemoreception has been almost totally neglected, if exception is made for a macroscopic description of the sweetness of different sugar solutions in terms of viscosity of sugar solutions, apparent specific volume, and apparent molar volume.<sup>1,37,40,41</sup> While it is not clear how to relate these macroscopic quantities to the different sweetness of different sugars, it is likely that taste effects are mediated by water and the solution properties of carbohydrate molecules therefore give clues to the possible mechanisms of taste chemoreception.<sup>37</sup>

A sensible step forward in our comprehension of the sweet taste came when computer modeling<sup>42</sup> described the AH-B-X functional structure in terms of hydrophobicity potential profiles,<sup>43,44</sup> suggesting the role of water-mediated interactions between the sweet substance and the receptor.<sup>42</sup> In this scenario, water affinity of specific sugar sites and sugar molecule overall hydration should play an important role as they control the fitness between a given carbohydrate and the receptor, along with the activity in the receptor microenvironment,<sup>40</sup> even though the details of these interactions are still to be elucidated. At this stage, it would be fair to conclude that glucose is sweet because has the right structure to interact either directly, or indirectly (water-mediated) with the receptor. We have shown at the atomic length scale that glucose hydroxyl and hydroxymethyl sites (that can act as both hydrogen bond donor and acceptor, AH and B sites), do hydrogen bond with water; the hydrophobic site (X site, oxygen in the ring) completes the glucose glucophore. In this context, it is tempting to associate the possibility of the B site (-OH) to act both as a H-bond acceptor (through its O atom) with a receptor H atom (for instance that of a -NH group) and as a donor through its H atom with a nearest neighbor water molecule with the broad water density cloud shown in Fig. 5 The only issue still open is whether its hydrogen bonding with the receptor is direct or water mediated. Clearly, glucose can do both. In this context, it is

important to mention the findings of recent studies on the characterization of the binding site for the dipeptide sweetener aspartame,<sup>45</sup> and other low molecular weight sweet compounds.<sup>46</sup> It was found that all these ligands interact with the sweet taste receptor through salt bridges, hydrogen bonds, and hydrophobic interactions, and that two water molecules may play an important role in sweetener binding by forming hydrogen bond bridges between the sweetener and the receptor.<sup>45</sup> Unfortunately, the binding mode of sugars, such as sucrose, could not be determined.<sup>46</sup>

Cellobiose is a disaccharide made up by two glucose units (linked by a  $\beta - 1, 4$ -glycosidic bond).<sup>4</sup> At the atomic length scale, the spatial density maps of water around its hydroxyl and hydroxymethyl sites (Fig. 2), and the SDM of water around the entire cellobiose molecule (Fig. 5) are both quite similar to those of the single glucose unit with strong and directional hydrogen bonds. As found for glucose, both rings of cellobiose have a low water density region located near the ring oxygen (Fig. 5). The obvious question is why cellobiose is not sweet, given its ability to form a similar hydrogen bond pattern with water, and possibly with the sweet receptor? The analysis proposed by Lichtenthaler and Immel,<sup>44</sup> indicated a sensible folding of the two glucose units to assume the right glucophore structure required by the AH-B-X model. The presence of a strong intramolecular H-bond found in cellobiose<sup>4</sup> might reduce its flexibility and prevent docking of this carbohydrate in the sweet receptor bud. To prove that cellobiose is tasteless for this reason, it would be interesting to investigate an aqueous solution of lactose (sweet), whose structure is almost identical to that of cellobiose but with only one hydroxyl group flipped (with respect to the ring plane) compared to cellobiose, and with a relatively weaker intramolecular H-bond between the two glucose units. Clearly, the rigidity of cellobiose makes it good as a building block of cellulose fibers.

Trehalose is a disaccharide made up by two glucose units (linked by a  $\alpha - 1, 1$ -glycosidic bond), different from the glycosidic bond usually present in many sweet carbohydrates.<sup>47</sup> At the atomic level, the spatial density maps of water around trehalose hydroxyl and hydroxymethyl sites are quite different from that of the same sites of glucose and cellobiose

(Fig. 2), as discussed above. It is important to notice the striking difference between the SDM of water around glucose (or cellobiose) or around trehalose (Fig. 5). We have already commented on this difference, but let us discuss its implications for trehalose sweet response and for its specific role as a bioprotective agent.<sup>47</sup> The fact that trehalose is not able to form strong and directional hydrogen bonds with water, but rather coordinates water molecules at a relatively larger distance and with a relatively reduced degree of directionality compared to glucose (or cellobiose), suggests that water-mediated hydrogen bonding between a sugar and the receptor is a relevant requirement for sweet response. It is reasonable to assign the observed differences to the different type of bond between the two glucose units of trehalose. In comparison with the usual  $\beta - 1,4$ -glycosidic bond, the different linkage between the two glucose units of trehalose results in a configuration likely not compatible with the AH-B-X model, and, at the atomic level, in a totally different ability to interact with water molecules, possibly due to a rearrangement of the molecular orbitals on the rings. The low water affinity of trehalose might leave room to trehalose-trehalose interactions, resulting in a disordered matrix at low water content, that could explain the still unclear ability of this disaccharide to form glassy phases with an unusually high glass transition temperature.<sup>48-54</sup> As a consequence, trehalose can efficiently act as a bioprotective agent against dehydration by trapping residual water molecules in the glassy matrix and/or limiting water loss from biomaterials.

## Conclusions

In summary, we have shown that careful analysis of the hydration shell, at the atomic length scale, can unveil issues and provide information related to the link between structure and function of three carbohydrates, namely, glucose, cellobiose, and trehalose. We found that identical chemical groups can have radically different hydration patterns depending on their location on a given molecule. Moreover, the same glucose unit, present in all

carbohydrates investigated, has a different overall hydration if it is in solution as a single molecule or paired with an other glucose units but with different glycosidic bond. The observed difference between the overall hydration of the carbohydrates investigated might explain their different sweetness, as ligand size and its ability to interact trough hydrogen bonds with the microenvironment of the receptor could play an important role in sweet chemoreception. The possibility that these interactions are mediated by water molecules, bridging between the ligand and the receptor, could be an important feature in solvent recognition processes, and can have general relevance, as similar water mediated functions have been proposed from atomic scale investigations into the hydration shell of the peptide glutathione,<sup>17</sup> and the GPG peptide,<sup>16</sup> both model systems for protein folding and solvent accessibility.

## Acknowledgement

This work has been performed within the Agreement No.01/9001 between CCLRC and CNR, concerning collaboration in scientific research at the spallation neutron source ISIS and with partial financial support of CNR. SEM would like to thank the UK EPSRC for funding (EP/J002615/1).

## References

- (1) Portmann, M-O.; Birch, G. Sweet taste and solution properties of  $\alpha,\alpha$ -trehalose. *J. Sci. Food Agric.* **1995**, *69*, 275-281.
- (2) Shallenberger, R. S.; Acree, T. E.; Molecular theory of sweet taste. *Nature* **1967**, *216*, 480-482.
- (3) Pagnotta, S. E.; McLain, S. E.; Soper, A. K.; Bruni, F.; Ricci, M. A. Water and trehalose: How much do they interact with each other? *J. Phys. Chem. B* **2010**, *114*, 4904-4908.

- (4) O'Dell, W. B.; Baker, D. C.; McLain, S. E. Structural Evidence for Inter-Residue Hydrogen Bonding Observed for Cellobiose in Aqueous Solution. *PLoS ONE* **2012**, *7*, e45311.
- (5) Dixit, S.; Crain, J.; Poon, W. C. K.; Finney, J. L.; Soper, A. K. Molecular segregation observed in a concentrated alcohol-water solution. *Nature* **2002**, *416*, 829-832.
- (6) McLain, S. E.; Soper, A. K.; Luzar, A. Investigations on the structure of dimethyl sulfoxide and acetone in aqueous solution. *J. Chem. Phys.* **2007**, *127*, 174515/1-174515/12.
- (7) Mancinelli, R.; Botti, A.; Bruni, F.; Ricci, M. A.; Soper, A. K. Perturbation of water structure due to monovalent ions in solution. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2959-2967.
- (8) Mancinelli, R.; Botti, A.; Bruni, F.; Ricci, M. A.; Soper, A. K. Hydration of sodium, potassium, and chloride ions in solution and the concept of structure maker/breaker. *J. Phys. Chem. B* **2007**, *111*, 13570-13577.
- (9) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Sabounji, M. L.; Brady, J. W. Structure of aqueous glucose solutions as determined by neutron diffraction with isotopic substitution experiments and molecular dynamics calculations. *J. Phys. Chem. B* **2005**, *109*, 13104-13111.
- (10) Hulme, E. C.; Soper, A. K.; McLain, S. E.; Finney, J. L. The hydration of the neurotransmitter acetylcholine in aqueous solution. *Biophys. J.* **2006**, *91*, 2371-2380.
- (11) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Sabounji, M. L.; Cuello, G.; et al. Neutron diffraction and simulation studies on the exocyclic hydroxymethyl conformation of glucose. *J. Chem. Phys.* **2006**, *125*, 224505/1-224505/9.
- (12) McLain, S. E.; Soper, A. K.; Diadone, I.; Smith, J. C.; Watts, A. Charged-Based Interactions between Peptides Observed as the Dominant Force for Association in Aqueous Solution. *Angew. Chem. Int. Ed.* **2008**, *47*, 9059-9062.

- (13) Foglia, F.; Lawrence, M. J.; Lorenz, C. D.; McLain, S. E. On the hydration of the phosphocholine headgroup in aqueous solution. *J. Chem. Phys.* **2010**, *133*, 145101-145110.
- (14) Malardier-Jugroot, C.; Bowron, D. T.; Soper, A. K.; Johnson, M. E.; Head-Gordon, T. Structure and water dynamics of aqueous peptide solutions in the presence of co-solvents. *Phys. Chem. Chem. Phys.* **2010**, *12*, 382-392.
- (15) Rhys, N.H.; Soper, A.K.; Dougan, L. The Hydrogen-Bonding Ability of the Amino Acid Glutamine Revealed by Neutron Diffraction Experiments. *J. Phys. Chem. B* **2012**, *116*, 13308-13319.
- (16) Busch, S.; Bruce, C. D.; Redfield, C.; Lorenz, C. D.; McLain, S. E.. Water mediation essential to nucleation of  $\beta$ -turn formation in peptide folding motifs. *Angew. Chem.* **2013**, *52*, 13091-13095.
- (17) Scoppola, E.; Sodo, A.; McLain, S. E.; Ricci, M. A.; Bruni, F. Water-peptide site-specific interactions: A structural study on the hydration of glutathione. *Biophys. J.* **2014**, *106*, 1701-1709.
- (18) Soper, A. K. Empirical potential Monte Carlo simulation of fluid structure. *Chem. Phys.* **1996**, *202*, 295-306.
- (19) Pardo, L. C.; Rovira-Esteva, M.; Tamarit, J. L.; Veglio, N.; Bermejo, F. J.; Cuello, G. J. In *Metastable Systems Under Pressure*; Rzoska, S. Drozd-Rzoska, A., Mazur, V., Eds.; NATO Science for Peace and Security Series A: Chemistry and Biology; Springer: Dordrecht, Netherlands, 2010.
- (20) <https://gcm.upc.edu/en/members/luis-carlos/angula/ANGULA>
- (21) Soper, A. K. Partial structure factors from disordered materials diffraction data: An approach using empirical potential structure refinement. *Phys. Rev. B* **2005**, *72*, 104204.

- (22) Soper, A. K. On the uniqueness of structure extracted from diffraction experiments on liquids and glasses. *J. Phys.: Condens. Matter* **2007**, *19*, 415108.
- (23) Soper, A. K. Joint structure refinement of x-ray and neutron diffraction data on disordered materials: application to liquid water. *J. Phys.: Condens. Matter* **2007**, *19*, 335206.
- (24) DuBois, G. E. Unraveling the biochemistry of sweet and umami tastes. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13972-13973.
- (25) Lerbret, A.; Bordat, P.; Affouard, F.; Guinet, Y.; Hedoux, A.; Paccou, L.; Prevost, D.; Descamps, M. Influence of homologous disaccharides on the hydrogen-bond network of water: complementary Raman scattering experiments and molecular dynamics simulations. *Carbohydrate Res.* **2005**, *340*, 881-887.
- (26) Shiraga, K.; Ogawa, Y.; Kondo, N.; Irisawa, A.; Imamura, M.. Evaluation of the hydration state of saccharides using terahertz time-domain attenuated total reflection spectroscopy. *Food Chem.* **2013**, *140*, 315-320.
- (27) <https://www.facebook.com/disord.matt/>
- (28) Soper, A. K. GudrunN and GudrunX: programs for correcting raw neutron and X-ray diffraction data to differential scattering cross section. *Rutherford Appleton Laboratory Technical Report, RAL-TR-2011-013*, 2011.
- (29) Sears, V. F. *Neutron News* **1992**, *3*, 26-37.
- (30) Soper, A. K. Determination of the orientational pair correlation function of a molecular liquid from diffraction data. *J. Mol. Liq.* **1998**, *78*, 179-200.
- (31) Busch, S.; Lorenz, C. D.; Taylor, J.; Pardo, L. C.; McLain, S. E. Short-range interactions of concentrated proline in aqueous solution. *J. Phys. Chem. B* **2014**, *118*, 14267-14277.



- (32) Johnston, A.J.; Busch, S.; Pardo, L. C.; Callear, S. K.; Biggin, P. C.; McLain, S. E. On the Atomic Structure of Cocaine in Solution. *Phys. Chem. Chem. Phys.* **2016**, *18*, 991-999.
- (33) Johnston, A.J.; Zhang, Y.; Busch, S.; Pardo, L. C.; Imberti, S.; McLain, S. E. Amphipathic solvation of indole: Implications for the role of tryptophan in membrane proteins. *J. Phys. Chem. B* **2015**, *119*, 5979-5987.
- (34) Gillams, R.J.; Busto, J.V.; Busch, S.; Goñi F.M.; Lorenz, C.D.; McLain, S.E. Solvation and Hydration of the Ceramide Headgroup in a Non-Polar Solution. *J. Phys. Chem. B* **2015**, *119*, 128-139.
- (35) Soper, A. K. In *Proceedings of the Conference on Advanced Neutron Sources 1988*, IOP Conf. Proc. 97, Hyer, D. K., Institute of Physics and Physical Society, London, 1989.
- (36) Kier, L.B. A molecular theory of sweet taste. *J. Pharm. Sci.* **1972**, *61*, 1394-1397.
- (37) Birch, G. G.; Shamil, S. Structure, sweetness and solution properties of small carbohydrate molecules. *J. Chem. Soc., Faraday Trans.* **1988**, *1*, 84, 2635-2640.
- (38) Nofre, C.; Tinti, J-M. Sweetness reception in man: The multipoint attachment theory. *Food Chem.* **1996**, *56*, 263-274.
- (39) Eggers, S. C.; Acree, T. E.; Shallenberger, R. S. Sweetness chemoreception theory and sweetness transduction. *Food Chem.* **2000**, *68*, 45-49.
- (40) Birch, G. G.; Karin, R.; Lopez, A. Novel aspects of structure-activity relationships in sweet taste chemoreception. *Food Quality Preference* **1994**, *5*, 87-93.
- (41) Birch, G. G. Role of water in sweet taste chemoreception. *Pure Appl. Chem.* **2002**, *74*, 1103-1108.

- (42) Astley, T.; Birch, G. G.; Drew, M. G. B.; Rodger, P. M.; Wilden, G. R. H. Computer modelling studies of the water-structuring properties of carbohydrates and the sweetness response. *Food Chem.* **1996**, *56*, 231-240.
- (43) Lichtenthaler, F. W.; Immel, S.; Kreis, U. Evolution of the structural representation of sucrose. *Starch/Starke* **1991**, *43*, 121-132.
- (44) Lichtenthaler, F. W.; Immel, S. Sucrose, sucralose, and fructose: correlations between hydrophobicity potential profiles and AH-B-X assignments. In: Sweet Taste Chemoreception (Eds.: M. Mathlouthi, J. A. Kanter, and G. G. Birch), Elsevier Appl. Science, London/New York, 1993, pp. 21- 53.
- (45) Maillet, E. L.; Cui, M.; Jiang, P.; Mezei, M.; Hecht, E.; Quijada, J.; Margolskee, R. F.; Osman, R.; Max, M. Characterization of the Binding Site of Aspartame in the Human Sweet Taste Receptor. *Chem. Senses* **2015**, *40*, 577-586.
- (46) Masuda, K.; Koizumi, A.; Nakajima, K.; Tanaka, T.; Abe, K.; Misaka, T.; Ishiguro, M. Characterization of the Modes of Binding between Human Sweet Taste Receptor and Low-Molecular-Weight Sweet Compounds. *PLoS ONE* **2012**, *7*, e35380.
- (47) Elbein, A. D.; Pan, Y. T.; Pastuszak, I.; Carroll, D. New insights on trehalose: a multifunctional molecule. *Glycobiology* **2003**, *13*, 17R-27R.
- (48) Green, J.L.; Angell, C. A. Phase relations and vitrification in saccharide-water solutions and the trehalose anomaly. *J. Phys. Chem.* **1989**, *93*, 2880-2882.
- (49) Donnamaria, M. C.; Howard, E. I.; Grigera, J. R. Interaction of water with  $\alpha, \alpha$ -trehalose in solution: molecular dynamics simulation approach. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 2731-2735.
- (50) Conrad, P. B.; de Pablo, J. J. Computer Simulation of the Cryoprotectant Disaccharide  $\alpha, \alpha$ -Trehalose in Aqueous Solution. *J. Phys. Chem. A* **1999**, *103*, 4049-4055.

- (51) Molinero, V.; Cagin, T.; Goddard III, W. A. Sugar, water and free volume networks in concentrated sucrose solutions. *Chem. Phys. Lett.* **2003**, *377*, 469-474.
- (52) Bordat, P.; Lerbret, A.; Demaret, J.-P.; Affouard, F.; Descamps, M. Comparative study of trehalose, sucrose and maltose in water solutions by molecular modelling. *Europhys. Lett.* **2004** *65*, 41.
- (53) Lerbret, A.; Bordat, P.; Affouard, F.; Descamps, M.; Migliardo, F. How Homogeneous Are the Trehalose, Maltose, and Sucrose Water Solutions? An Insight from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2005**, *109*, 11046-11057.
- (54) Choi, Y.; Cho, K. W.; Jeong, K.; Jung, S. Molecular dynamics simulations of trehalose as a "dynamic reducer" for solvent water molecules in the hydration shell. *Carbohydrate Res.* **2006**, *341*, 1020 - 1028.