Original

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

In: Physical Chemistry Chemical Physics (2015) Royal Society of Chemistry

DOI: 10.1039/C5CP06090G

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 991

On the atomic structure of cocaine in solution†

Andrew J. Johnston,[a] Sebastian Busch,[b] Luis Carlos Pardo,[c] Samantha K. Callear,[d] Philip C. Biggin[a] and Sylvia E. McLain*[a]

Received 9th October 2015,
Accepted 27th November 2015

DOI: 10.1039/c5cp06090g

www.rsc.org/pccp

Cocaine is an amphiphilic drug which has the ability to cross the blood–brain barrier (BBB). Here, a combination of neutron diffraction and computation has been used to investigate the atomic scale structure of cocaine in aqueous solutions. Both the observed conformation and hydration of cocaine appear to contribute to its ability to cross hydrophobic layers afforded by the BBB, as the average conformation yields a structure which might allow cocaine to shield its hydrophilic regions from a lipophilic environment. Specifically, the carbonyl oxygens and amine group on cocaine, on average, form ~5 bonds with the water molecules in the surrounding solvent, and the top 30% of water molecules within 4 Å of cocaine are localized in the cavity formed by an internal hydrogen bond within the cocaine molecule. This water mediated internal hydrogen bonding suggests a mechanism of interaction between cocaine and the BBB that negates the need for deprotonation prior to interaction with the lipophilic portions of this barrier. This finding also has important implications for understanding how neurologically active molecules are able to interact with both the blood stream and BBB and emphasizes the use of structural measurements in solution in order to understand important biological function.

1 Introduction

Designing drugs which can cross the blood–brain barrier (BBB) presents a significant challenge for the pharmaceutical industry. That a large number of putative drugs fail to cross this barrier[1–3] has prompted a push towards development of both *in vitro* and computational models to better assess drug permeation into the central nervous system.[4,5] Despite a wide range of theories regarding the molecular motifs which must be incorporated into small molecules to enhance BBB permeability, it is still unknown how molecules which cross the BBB do so *in vivo*, especially on the atomic scale where these interactions necessarily occur.

Molecules which successfully cross the BBB are traditionally thought to passively diffuse across this largely hydrophobic barrier,[6] usually being of low molecular weight (<400 Da)[2,3] and high lipid solubility.[3,7] Other theories, however, have argued that drugs enter cells solely *via* carriers normally used for the transport of nutrients and intermediary metabolites.[8,9] Nevertheless, in order to exert a pharmacological function, after crossing the BBB the drug must then partition into the aqueous environment of the interstitial brain fluid.[7] Highly lipid soluble molecules can be sequestered by the capillary bed and not reach the cells beyond the BBB. As a result, effective molecules must somehow achieve the correct balance between lipophilicity or hydrophobicity and hydrophilicity in order to penetrate deeply enough into the BBB in order to perform their function.

The naturally occurring alkaloid, cocaine ($C_{17}H_{21}NO_4$), is a small molecule which easily crosses the BBB and is believed to function once it crosses this barrier by inhibiting dopamine reuptake.[10–12] Cocaine is highly lipophilic, so much so that treatment of cocaine overdose includes the administration of lipid emulsions to dampen its effects.[13] Cocaine is also hydrophilic; the hydrochloride salt is readily soluble in aqueous solutions. Administered both as the hydrochloride salt and in its freebase form, both forms of cocaine are very addictive and have similar pharmacological effects.[14,15] Although it has been stated that cocaine is active only in its protonated form and that cocaine crosses the BBB only as a freebase (or in its deprotonated form),[16] this is not necessarily the case,[17] especially given that crystallographic investigations of cocaine molecules bound into a protein receptor the cocaine molecule appear to be deprotonated[12] and measurement of cocaine as it crosses the BBB *in vivo* is not yet possible.

To date there is limited information concerning the structure of BBB penetrating molecules in solution, the physical milieu in

[a] *Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK. E-mail: sylvia.mclain@bioch.ox.ac.uk*
[b] *German Engineering Materials Science Centre (GEMS) at Heinz Maier-Leibnitz Zentrum (MLZ), Helmholtz-Zentrum Geesthacht GmbH, Lichtenbergstr. 1, 85747 Garching bei München, Germany*
[c] *Departament de Física i Enginyeria Nuclear, Escola Tècnica Superior d'Enginyeria Industrial de Barcelona (ETSEIB), Universitat Politècnica de Catalunya, 08028 Barcelona, Catalonia, Spain*
[d] *ISIS Facility, Rutherford Appleton Laboratory (STFC), Chilton, Didcot OX110QX, UK*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5cp06090g

which they operate *in vivo*. However, recently it has been shown that investigations into the structure and hydration of small biological molecules in solution can yield valuable insights into how these molecules may function in nature.[18–20] In the current work, the structure and conformation of cocaine hydrochloride has been investigated in aqueous solution on the atomic scale using a combination of neutron diffraction enhanced by isotopic substitution and computation. Understanding the details of drug interactions with aqueous environments will increase the comprehension of the mechanisms responsible for how small molecules permeate into the BBB. Importantly, both how cocaine interacts with water and its average conformation have been assessed in detail, as estimates of the ability of molecules to cross the BBB have been linked to molecular shape and hydrogen bonding descriptors, such as the number of hydrogen bond donors and acceptors.[21]

2 Methods

2.1 Sample preparation

Cocaine hydrochloride (CAS 53-21-4) was purchased from Sigma Aldrich and used without further purification, after verification by ^{1}H NMR. Cocaine·HCl was deuterated by dissolution in D_2O (99.9%; Sigma Aldrich), subsequently removing the solvent by vacuum ($\sim 1.0 \times 10^{-3}$ mbar). This process was repeated three times to ensure adequate deuteration. The success of this deuteration was verified by ^{1}H NMR (see ESI†). The cocaine molecule was roughly 93% deuterated as a result of this exchange process. All isotopomeric solutions of cocaine·HCl in water for the neutron and X-ray diffraction experiments were prepared by weight under an N_2 atmosphere to ensure sample purity.

2.2 Neutron diffraction with isotopic substitution

Neutron diffraction with isotopic substitution (NDIS) is a method by which the average local structure of hydrogen containing solutions can be determined. The use of neutrons as a probe is particularly useful when determining the structural interactions between biomolecules and water. This is due, in part, to the fact that H and D give rise to different scattering intensities when using neutrons as a probe.[22] Measurement of several isotopomeric samples of biomolecules at the same chemical concentration allows for multiple diffraction patterns of a specific system to be measured. NDIS is a well-established technique which has been used successfully to determine the atomistic interactions between a range of molecules and water in solution.[18–20,23–29]

Neutron diffraction measurements provide a direct measure of the structure in reciprocal space – the static structure factor, $F(Q)$. $F(Q)$ can be written as:

$$F(Q) = \sum_{\alpha,\beta \geq \alpha} (2 - \delta_{\alpha\beta}) \cdot c_\alpha c_\beta b_\alpha b_\beta \left[S_{\alpha\beta}(Q) - 1\right], \qquad (1)$$

where Q is the vector between the incident and scattered radiation ($Q = 4\pi\sin(\theta)/\lambda$; λ is the incident wavelength and 2θ is the scattering angle), c and b are the concentrations and the scattering lengths, respectively, of atom types, α and β and $\delta_{\alpha\beta}$ is the Kronecker delta. $F(Q)$ is the weighted sum of the partial structure factors, $S_{\alpha\beta}(Q)$. $S_{\alpha\beta}(Q)$ is related to the radial distribution functions (RDFs), $g_{\alpha\beta}(r)$, which give the distances between each atom pair in real space, *via* Fourier transformation:

$$S_{\alpha\beta}(Q) = 1 + 4\pi\rho \int r^2 \cdot \left[g_{\alpha\beta}(r) - 1\right] \cdot \frac{\sin(Qr)}{Qr} \mathrm{d}r, \qquad (2)$$

where ρ is the atomic number density of the sample (atoms Å^{-3}).

In the present work, six different isotopomers of cocaine hydrochloride in water, which varied only with respect to their H/D composition at a relative ratio of 1 cocaine hydrochloride: 65 waters ($\sim$0.8 M), were measured in SiO_2 cells with a sample thickness of 1 mm and sample cell wall thickness of 1 mm. Diffraction data were collected for between 8 and 9.5 hours per sample at a temperature of 298 K on the SANDALS diffractometer at located at the ISIS facility (STFC, UK). It should be noted that although this temperature is not the average body temperature (310 K), NMR measurements at both 298 K and 310 K (shown in the ESI†) indicate that the structure of cocaine is similar at both of these temperatures. Data were also collected for the empty cells, the empty instrument and a vanadium standard for background subtraction and normalization. All of the data for samples, cells, empty instrument and vanadium were corrected for absorption, multiple scattering and inelasticity effects and then subsequently converted to $F(Q)$ using the GUDRUN program,[30] which is freely available for download.[31] Details of all the isotopomeric samples measured are included in the ESI.†

2.3 X-ray diffraction

A sample was prepared for X-ray diffraction (XRD) at the same concentration as the samples measured by neutron diffraction and loaded into a borosilicate glass capillary with a diameter of 1 mm. The sample was measured at 298 K on a Panalytical X'pert Pro X-ray diffractometer equipped with a silver source at ISIS (UK), which has a Q range of 0.7–20 Å^{-1}, and the data were corrected using GudrunX. Similar to NDIS, the XRD measurements also provide a measure of $F(Q)$, where the individual $S(Q)$s are weighted by their atomic concentrations and form factors, $f(Q)$.[32]

2.4 Empirical potential structure refinement and ANGULA

Empirical Potential Structure Refinement (EPSR) is a reverse Monte Carlo technique which can be used to determine local interactions present in disordered materials, where the EPSR model is constrained by a set of diffraction data. EPSR uses a box of molecules, whose geometry (bond lengths which are defined by both the molecular angles and torsional angles) is defined at the outset, but imposes no further constraints on the movement of rotational groups, for instance, at the concentration, density and temperature of diffraction measurements. In addition, 'seed' or starting potentials are given to each unique atom, where these starting potentials consist of a Lennard-Jones potential, defined by σ (the distance at which the potential is zero) and ε (the well depth) as well as appropriate atomic charges (q_e). During the EPSR fitting process, these 'seed' potentials are iteratively

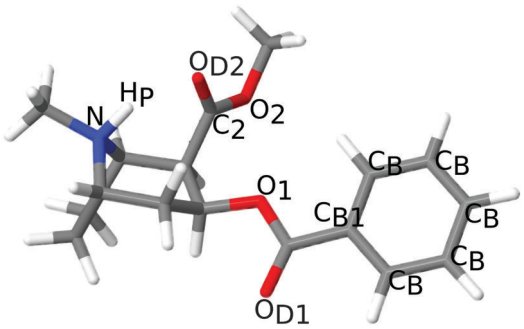

Fig. 1 The molecular structure of protonated cocaine, where the labels correspond the salient atomic sites discussed below.

refined until a good 'fit' to the diffraction data is obtained. Further details on the EPSR technique can be found elsewhere.[33–35]

In the current EPSR simulation, the modeling box contained 25 protonated cocaine molecules (Fig. 1), 25 Cl^- ions and 1625 water molecules at the measured density (ρ = 0.101 atoms Å^{-3}) at 298 K. Parameters for water molecules were taken from the SPC/E water model.[36] Parameters for the cocaine molecule and Cl^- ion were generated in AMBER format with ANTECHAMBER.[37,38] A full list of the parameters used for the starting potentials is included in the ESI.† Among other things, the individual site-site $g(r)$s (eqn (2)) can be extracted from the EPSR model. Coordination numbers ($n_\alpha^\beta(r)$), *i.e.* the average number of β atoms around a central α atom at a distance between r_{min} and r_{max}, can then be calculated by integration of these $g(r)$ functions *via*

$$n_\alpha^\beta(r) = 4\pi\rho c_\beta \int_{r_{min}}^{r_{max}} r^2 g_{\alpha\beta}(r)\mathrm{d}r. \quad (3)$$

In addition to the $g(r)$s, the three-dimensional arrangements of molecules relative to one another can also be extracted from the EPSR simulation box using the program ANGULA (described below).[39] Orthonormal coordinate systems were assigned to different fragments of the cocaine molecule (as defined in the ESI†) and to the water molecules. 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 assigned to the cocaine molecules.[40,41] This procedure was performed for ~4000 snapshots of the simulation box, such that the aggregate nearest neighbor water molecule distribution was plotted *via* a spatial density map (SDM).[24,41–43] The angles between the sets of axes assigned to the solute and the solvent molecules were also used to find the orientation of the water molecules relative to cocaine. A more detailed outline of how this procedure was used to find the position and orientation of the most probable nearest neighbor water molecules of cocaine is provided in the ESI.† Whole molecule analysis (WMA) was also performed using ANGULA. 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 cocaine molecules, enabling the aggregate distribution of water around cocaine to be plotted with reference to the whole molecule.

3 Results and discussion

The measured X-ray and neutron diffraction data ($F(Q)$) along with the EPSR fits to these data are shown in Fig. 2. In this figure, the EPSR fits to the neutron data are generally good with some small differences at low Q values due to the difficulty in correcting for inelastic scattering this region of the data.[44] The X-ray data fits are also good, with again some differences at lower Q values. This is likely because polarizability effects are not taken into account, as the weighted sum of atomic form factors are currently used to obtain the static structure factor, so some error in the data correction is probable.[45] The Fourier transformations of these data and EPSR fits in real space ($G(r)$) are included in the ESI.†

3.1 Cocaine conformation

Understanding the conformation and structure of molecules is important in the context of understanding how drugs interact with the hydrophilic and hydrophobic components of the BBB, as 3D structure has been linked to permeation and led to structure based drug design.[46,47] Fig. 3 shows some of the relevant intramolecular $g(r)$s between atoms in the cocaine molecule. The most prominent peak is found in $g_{O_{D2}H_P}(r)$, which shows a peak at ~1.83 Å. That this intramolecular peak is so well defined, suggests a highly directed intramolecular hydrogen bond (IHB) between these two groups in cocaine. The presence of this IHB will confer conformational stability to cocaine, as it will prevent rotation of the carbomethoxy group. The constraints afforded by IHBs are thought to influence the function of molecules.[48] Interestingly, the benzoate ester oxygen-amine hydrogen $g_{O_1H_P}(r)$ also shows a prominent peak at ~3.8 Å. While this distance

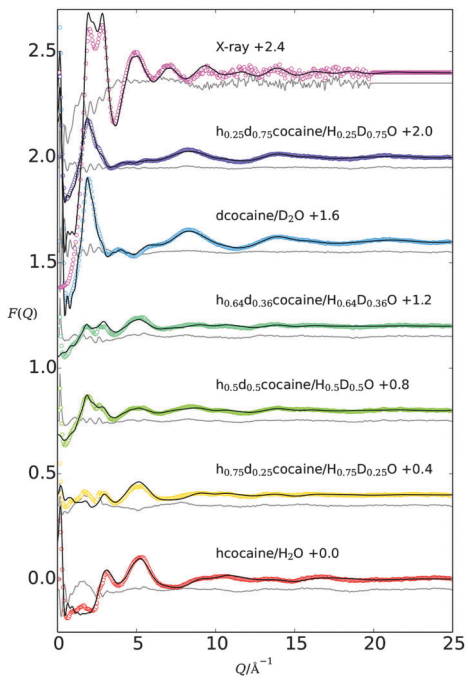

Fig. 2 Measured X-ray and neutron $F(Q)$ (colored lines) and EPSR fits to the data (black lines) in reciprocal space. The differences between data and fits (gray lines) are also shown, shifted by −0.05 for clarity.

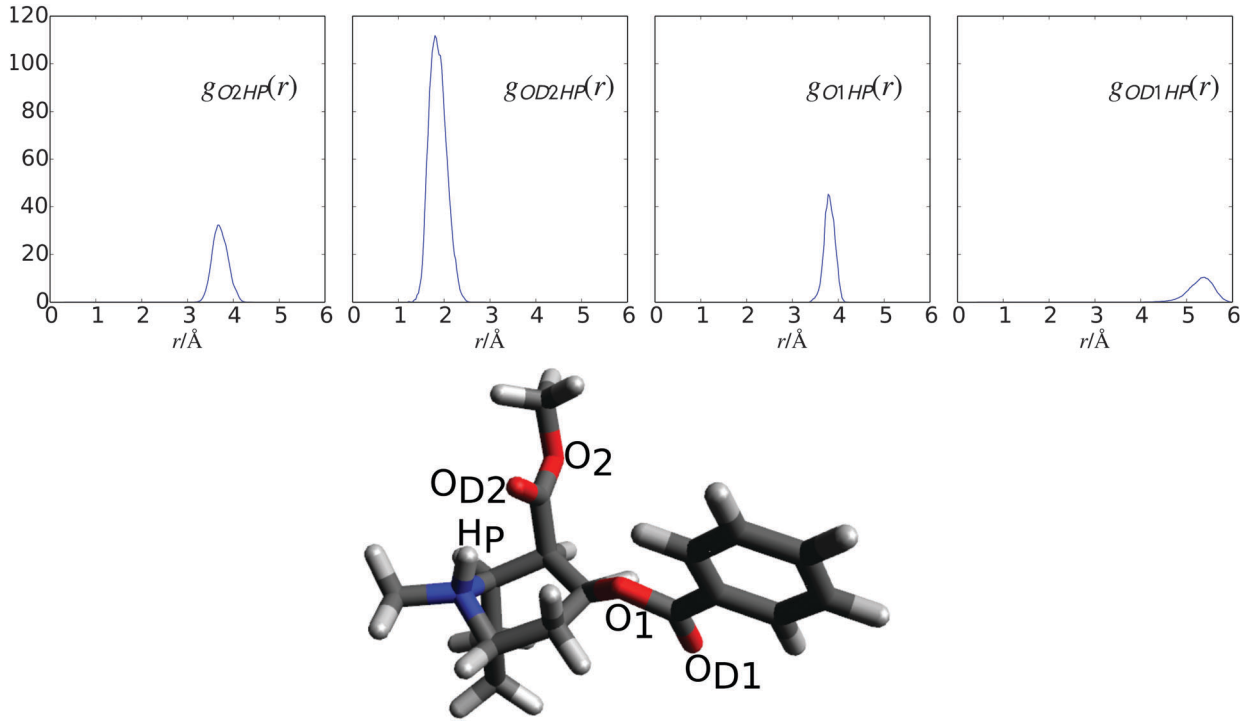

Fig. 3 Intramolecular RDFs for cocaine oxygen atoms with the amide hydrogen atom (top). The most probable conformation of cocaine based on intramolecular distances and the angular distributions extracted from EPSR (bottom).

is too far to be formally considered as a hydrogen bond, that this peak is relatively sharp indicates that this portion of the molecule is additionally constrained, as it might be expected that a broader range of distances would be observed, such that occur for the carbomethoxy ester oxygen (O_2).

A depiction of the conformation of cocaine deemed the most probable based on torsional angle distributions (see ESI†) and constrained to reflect the average distances in the top panel of Fig. 3 is shown in the lower panel of this figure. Relative to the other parts of the cocaine molecule, the benzene ring is the most mobile in the simulation. It has the broadest range of distances relative to the center of the molecule and shows a broad torsional angle distribution relative to other parts of the molecule (see ESI† and WMA below).

The molecular constraints of the tropane ring, coupled with the rigidity of the carbomethoxy group due to the formation of an IHB, ensure that cocaine rarely 'folds back' on itself, meaning the benzene ring does not orient itself towards the tropane ring but rather is tilted at an angle of $\sim 30^{\circ}$ relative to the plane of the piperidine ring in Fig. 3. In addition, the presence of a sharp peak in the $g_{O_1H_P}(r)$ and the average cocaine conformation shown in Fig. 3 also suggest that the amine group, the carbomethoxy carbonyl and the benzoate ester linkage form an almost ring-like structure towards interior of the cocaine molecule. These groups are oriented in such a way that leaves them less solvent accessible compared with, for instance, the carbonyl oxygen on the benzoate group.

3.2 Cocaine hydration

The carbonyl oxygen-water $g(r)$s shown in Fig. 4 indicate that water is highly oriented around these groups in cocaine. Specifically, $g_{O_{D1}O_w}(r)$ and $g_{O_{D2}O_w}(r)$ both show sharp peaks at ~ 2.8 Å and both $g_{O_{D1}H_w}(r)$ and $g_{O_{D2}H_w}(r)$ functions show first sharp peaks at 1.77 Å, well within the accepted length range of a short, strong hydrogen bond.[49] The carbonyl oxygens are slightly undersaturated (n_x^{β}, Table 1) compared with pure water[35] and for carbonyl groups which are largely solvent accessible.[20,50,51] A preference for water to interact with the carbonyl oxygens is expected given the high electronegativity of O_{D1} and O_{D2}, and the coordination numbers (Table 1) show the benzoate O_{D1} having slightly more waters present compared with the carbomethoxy O_{D2}. This greater number of waters is consistent with O_{D2} forming an IHB.

Fig. 4 also shows the $g(r)$s for the ether oxygens, where the broad peaks are indicative of water being comparatively much less oriented around these groups. This is particularly clear for the benzoate ester oxygen O_1, which has a broader function between ~ 2 Å and ~ 6 Å than O_2. The coordination numbers for these functions in Table 1 show that there are fewer waters around the benzoate ester oxygen (O_1) compared with the benzoate carbonyl oxygen (O_{D1}; Fig. 1). In contrast, there are slightly more waters around the carbomethoxy ester (O_2) compared with carbonyl oxygen of this same group, as evidenced by the O_x-O_w funtions. This hydration is consistent with the observation from the most probable cocaine conformation depicted in Fig. 3 that H_P, O_{D2} and O_1, given the average cocaine conformation, are less solvent accessible (Fig. 3).

SDMs (Fig. 4, lower panel) for the nearest neighbor waters around the carbonyl oxygens in cocaine show the areas of highest density are more broadly distributed around O_{D1} compared with O_{D2}. Specifically, the carbomethoxy carbonyl

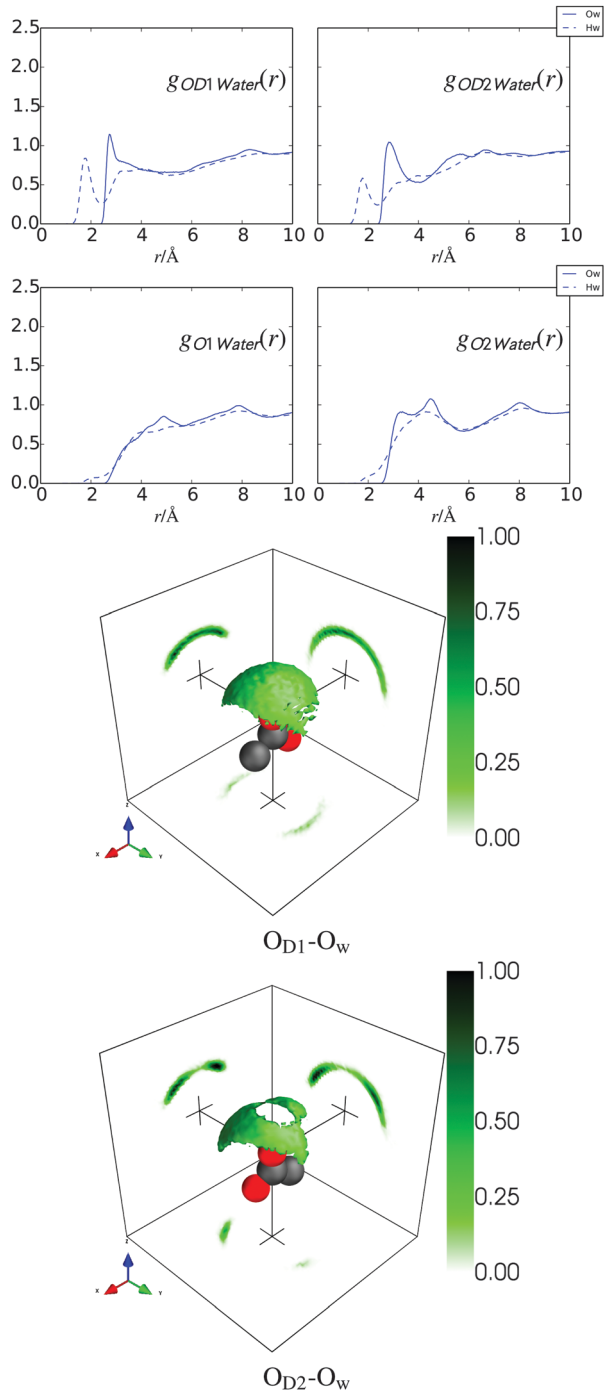

Fig. 4 $g(r)$s (top) for the interactions between the carbonyl oxygens of cocaine (O_{D1} and O_{D2}) and water molecules. $g(r)$s (middle) for the interactions between the ester oxygens of cocaine (O_1 and O_2) and water molecules. SDMs (bottom) showing the distribution of nearest neighbor water oxygens around O_{D1} and O_{D2}. The isocontour surface of the SDMs has been cut to include the top 50% of molecules and the cut-throughs from the origin, have been displaced to the background by 6 Å for clarity. Red, green and blue arrows in the bottom left of the SDMs indicate the *x*, *y* and *z* directions, respectively.

shows a much more well defined distribution, signifying that the surrounding waters are more localized in certain positions around this group, with the highest density of waters observed

Table 1 Coordination numbers of interactions between atoms of cocaine and water, corresponding to the RDFs shown in Fig. 4–6. Also included is the position at which the coordination number was taken, r_{max}, and the first peak maxima, r_{peak}. No discernible peak was observed for $g_{O_1H_w}(r)$

	n_z^β	r_{max}/Å	r_{peak}/Å
$n_{O_{D1}}^{O_w}$	4.15	3.93	2.76
$n_{O_{D1}}^{H_w}$	1.29	2.46	1.77
$n_{O_{D2}}^{O_w}$	3.63	3.93	2.85
$n_{O_{D2}}^{H_w}$	0.98	2.46	1.77
$n_{O_1}^{O_w}$	2.07	3.93	4.86
$n_{O_1}^{H_w}$	0.16	2.46	—
$n_{O_2}^{O_w}$	3.80	3.93	3.33
$n_{O_2}^{H_w}$	0.28	2.46	4.38
$n_{H_P}^{O_w}$	0.90	2.85	2.19
$n_N^{O_w}$	0.73	3.33	3.00
$n_B^{O_w}$	6.04	4.32	3.54
$n_B^{H_w}$	21.69	5.19	4.29

in the -*x* direction (as observed in the cut-throughs of the density in the side panels) towards the tropane ring on cocaine, where this O_{D2} atom is likely to be involved in an IHB with the amine hydrogen. This SDM also shows a relatively low probability of waters being directly above this oxygen in the *z*-direction. In contrast, O_{D1} shows a broader distribution of nearest neighbor waters, where the density is largely uniformly distributed around this atomic site.

The $g_{H_PO_w}(r)$ function in Fig. 5 shows a distinct peak at ~2.16 Å, where the coordination number for this peak shows that ~0.9 water molecules are bound to H_P. This bond length is indicative of a moderately strong hydrogen bond[49] from the amine hydrogen (H_P) to the surrounding water molecules. The pK_a for this proton was calculated from pH measurements of cocaine in water at different concentrations as 8.5. This value is in agreement with other values in the literature[10] and implies that all of the cocaine molecules are fully protonated, as solutions measured here have a pH of ~3.3. As a reference, the pH of neuronal cells of the vertebrate central nervous system ranges from around 6 to 8,[52] so that, similar to the current work, cocaine would exist in its protonated form over this range of different pHs within the central nervous system. Hydrogen bonding from the amine group to water is also visible in the $g_{NO_w}(r)$ function at ~2.97 Å. The peaks at larger distances (~4.4 Å and ~5.5 Å) are likely related to a second and third hydration shell around cocaine.

Fig. 5 also shows the SDM for the nearest neighbor waters around the amine group, where all of the water density is located above the H_P atom, slightly offset in the *x* direction. The corresponding side bar (which is a measure of relative probability) confirms this high density of water in this precise location. This relatively confined water density is suggestive of highly oriented water molecules which bind at highly localized positions above the amine group in cocaine. This localization of nearest neighbor water density suggests that the IHB in cocaine is either water mediated to some degree or strongly coordinates water at specific locations around this hydrogen-bonding moiety.

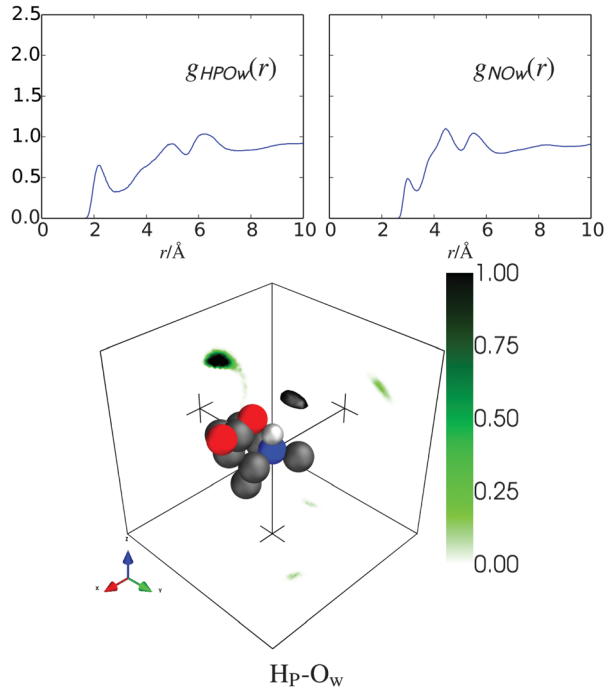

Fig. 5 RDFs (top) for the interaction between the amine group of cocaine and water molecules. SDM (bottom) showing the distribution of nearest neighbor water oxygens around the amine group of cocaine. The isocontour surface of the SDM has been cut to include the top 50% of molecules. Red, green and blue arrows in the bottom left of the SDM indicate the *x*, *y* and *z* directions, respectively.

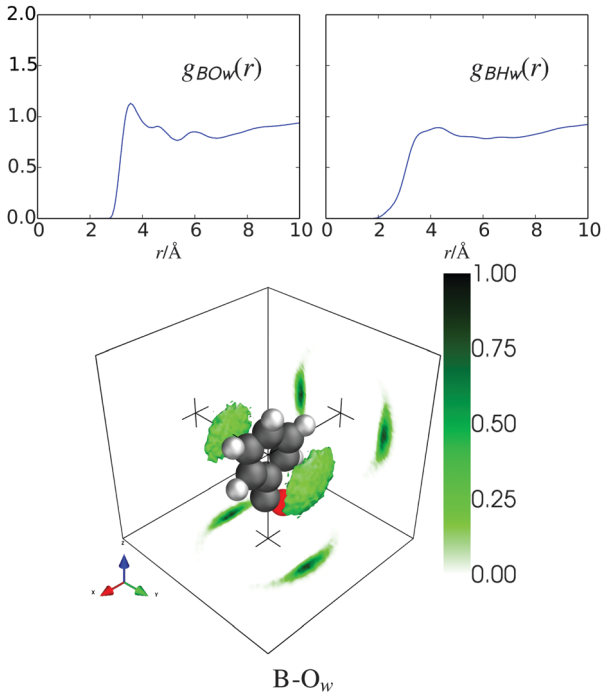

Fig. 6 RDFs (top) for the interactions between the benzene ring of cocaine and the water molecules (*i.e.* the weighted average of interactions between C_B and C_{B1} with O_w or H_w). SDM (bottom) showing the distribution of nearest neighbor water oxygens around the benzene ring. The isocontour surface of the SDM has been cut to include the top 50% of molecules. Red, green and blue arrows in the bottom left of the SDM indicate the *x*, *y* and *z* directions, respectively.

This view is also consistent with location of the most probable nearest neighbor waters in the O_{D2}–O_w SDM (Fig. 4), where the highest probability density indicates that water is more likely to be found between the amine and carbomethoxy carbonyl groups.

Fig. 6 shows the interactions between water and the benzene ring of cocaine. The broad peak at ~3.6 Å in the $g_{BO_w}(r)$ suggests that the waters surrounding benzene might have some preferred orientation. However, the much broader peak in $g_{BH_w}(r)$ indicates that this is not the case. Although water hydrogens are within a closer distance range, these hydrogens have a broad distribution of possible nearest neighbor distances around benzene, which is indicative of disordered interactions around this part of the cocaine molecule. This observation is at variance with previous investigations on indole in methanol/water solutions, where benzene-water interactions showed water directly hydrogen bonding to benzene,[43] and with other investigations on benzene-containing molecules which suggest more directed interactions between water and benzene.[53,54] From the SDM in Fig. 6, it is clear that the nearest neighbor waters are preferentially located above and below the plane of the benzene ring. However, it is clear that the density of water around benzene is more diffuse compared with the carbonyl and the amine groups. It may be that the waters around benzene are not as localized simply because cocaine affords more direct electrostatic hydrogen bonding opportunities compared with indole in solution.

3.3 Whole molecule analysis of cocaine hydration

The whole molecule analysis (WMA) for the most probable water location around cocaine in Fig. 7 (left panel) shows that water occupies the space between O_{D2} and H_P in the tropane ring over other parts of the molecule. These nearest neighbor water molecules appear to preferentially envelop the cavity between the IHB. In this figure, the limited movement of the carbomethoxy group is highlighted as the cocaine molecules at the center represent a distribution of the conformations from the EPSR modeling box. In contrast, these central molecules show that the benzene ring is able to move comparatively freely and samples a wider range of conformations throughout the simulation.

The most probable position and orientation of the nearest neighbor water around H_P, extracted from the EPSR simulation box using the ANGULA (see ESI,† for a full description of this procedure) is shown in Fig. 7 (right panel). The position and orientation of this water emphasizes that the IHB between the amine hydrogen and the carbomethoxy carbonyl group is likely to be mediated by or strongly coordinated with a water molecule, as suggested by the water density shown around this group in Fig. 5, where this water is oriented toward the IHB. On the surface this result is seemingly contradictory to the slightly higher solvent accessibility observed for O_{D1} relative to O_{D2} in Fig. 4. However, it should be noted that the WMA represents the most probable water location around all of the atoms in the

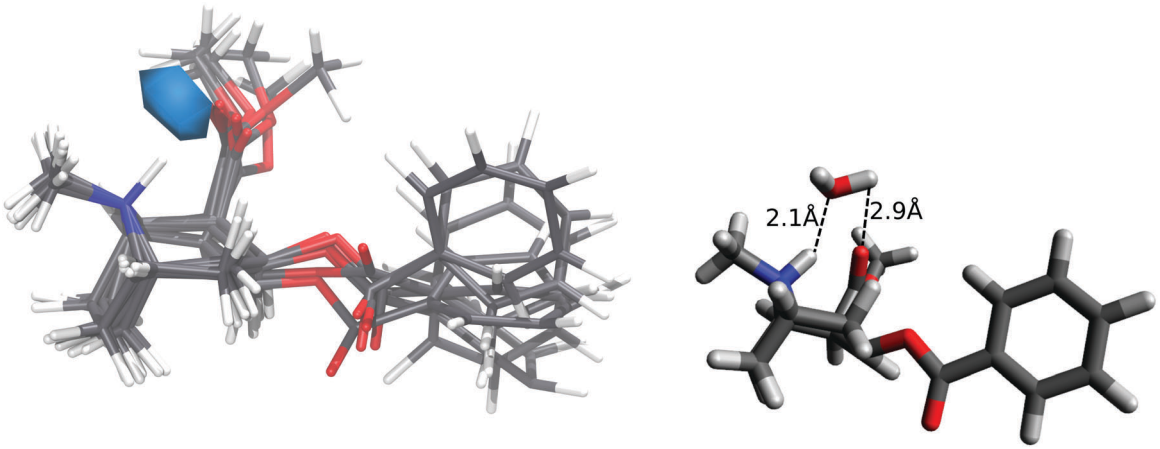

Fig. 7 Left panel: Whole molecule analysis for water (blue) around cocaine. The isocontour surface is cut to show the top 30% of molecules within a distance range between 0–4 Å, in order to illustrate the most probable areas of interaction between water and cocaine. Different cocaine molecules taken from the EPSR simulation box are shown at the origin to display the range of conformations adopted by the molecule. Right panel: The position and orientation of the most probable nearest neighbor water molecule of H_P. The cocaine molecule shown in this figure corresponds to the most probable conformation of cocaine, as shown in Fig. 3.

entire cocaine molecule within a distance range of 0–4 Å rather than being specific to a single site. As stated above, the most probable location of waters around the entire molecule is between the O_{D2} and H_P atoms, so this distribution is not solely representative of the hydration of water around one or the other of these atoms but rather shows that the most favored location of water is between these atoms. Further, a more highly solvent accessible atom would show a broader distribution of waters around a given group whereas a lower solvent accessibility indicates that there are fewer locations that the water molecules can bond to around a given atom but does not necessarily mean that there is a lower probability of finding water molecules in this location relative to the entire molecule.

4 Conclusions

Cocaine is very water soluble, yet water seems to be strongly associated with only a relatively small number of the atoms of the molecule. Highly directed hydrogen bonds only occur between water and the carbonyl oxygens or the amine group, with very diffuse interactions being observed between water and other portions of the cocaine molecule. Given the high solubility of cocaine, it might be expected that water would form highly directed hydrogen bonds with the ester oxygens or the benzene ring in solution as has been previously observed for some molecules,[43,55] yet these interactions are apparently absent for cocaine. This is not to say that there are no waters present but rather that the waters present around these groups are disordered, as is clear from the $g(r)$s in Fig. 4 and 6. It has previously been suggested that once the cumulative number of hydrogen bonds a drug forms with water exceeds 8–10, the BBB transport of the molecule will probably be pharmacologically insignificant.[3] In this system, significant hydrogen bonds are accepted from water from O_{D1}, O_{D1} and N atoms and summing the nearest-neighbor coordination numbers these atoms gives a value of 5.4, below the reported threshold above which molecules may not cross the BBB. The H_P atom on the average also donates 0.9 hydrogen bonds to the surrounding water solvent.

Torsional angle information and intramolecular $g(r)$s indicate that the carbomethoxy group on cocaine is rotationally restricted, held in place due to an IHB between the carbonyl oxygen and amine hydrogen. The results here also indicate that this IHB is strongly associated with water, where the waters around this hydrogen bonding motif are highly oriented. In addition, the carbomethoxy ether oxygen (O_2) and benzoate carbonyl oxygen (O_{D1}) show significantly higher hydration, compared with the benzoate ether oxygen (O_1) and the carbomethoxy carbonyl oxygen (O_{D2}), respectively. The reason for this is apparent from inspection of the most probable orientation of cocaine in solution (Fig. 3), where both of these hydrated oxygen atoms are oriented such that they are easily accessible to the surrounding water solvent. Conversely, the other hydrophilic atoms (H_P, O_{D2} and O_1) are oriented so that they are facing towards each other and, importantly, largely away from the surrounding water solvent. These groups are comparatively less hydrated, but the waters around these atoms are highly oriented, suggesting a very strong association between water and these more 'internal' hydrophilic atoms. Perhaps most remarkably, the IHB in cocaine appears to either be mediated by a water molecule whose hydrogen atoms are $\sim$2.8 Å from O_{D2} and H_P or to strongly bind to a water molecule rather than excluding water as might be expected from the formation of an IHB. It should be noted that previous crystallographic investigations of cocaine hydrochloride do not show the presence of an IHB, emphasizing the need for investigations in the liquid phase.[56]

The combination of intramolecular bonding and the apparent affinity for cocaine to tightly bind water to the atoms associated with this non-covalent bond provides a possible explanation for the mechanism by which cocaine is able to cross the BBB. In opposition to previous thought that cocaine must be deprotonated in order to effectively passively diffuse into the BBB,[16] the hydration

structure observed here suggests another possibility. In solution, the charged regions associated with the IHB present in cocaine appear to be effectively shielded by the conformation of cocaine. In addition, the formation of this internal hydrogen bond also strongly coordinates water molecules, which may also allow the benzoate ester oxygen to be pulled towards the inside of the cocaine molecules. This water-mediated conformation would effectively shield charges from an external hydrophobic environment such that the cocaine molecule becomes sufficiently lipophilic to cross the BBB. Water does not appear to be tightly bound to either the benzene ring or around the carbomethoxy ester, both of which are comparatively more solvent accessible. This might indicate that water would be easily shed from these regions of the molecule when cocaine is in a lipophilic environment.

It has been previously noted that IHBs can confer lipophilicity to drugs by shielding potential hydrogen bonding motifs from the solvent[57] and that IHBs were most probable when they resulted in the formation of a 5 or 6 membered ring.[58] The most probable conformation of cocaine observed here, coupled with the water mediation interactions, certainly suggest that cocaine could easily adopt a more complex ring structure, where these rings would include water molecules, and would thus help confer lipophilicity to the cocaine molecule. Water association with the IHB in cocaine may also have a more direct biological function than solely aiding diffusion across the BBB. Water-mediated interactions in cocaine have been previously suggested to help facilitate its function, where it has been proposed that water mediates cocaine interactions with antibodies *in vivo*.[59] These conserved waters are in a similar location around the tropane ring to those observed here.

The ability of cocaine to shield several of its potential hydrogen bonding atoms, perhaps even trapping a molecule within its structure, is suggestive that cocaine is highly lipophilic even when it is protonated. However, contrary to previous suggestions that more conformationally open forms drugs should be relatively more water-soluble,[57] the results here also show that cocaine is still readily water soluble in a more closed conformation. Furthermore, the results here suggest that cocaine can easily achieve both lipophilicity and hydrophilicity simultaneously without having to adopt a different conformation in either physical regime or indeed having to go through a protonation-deprotonation reaction to effectively permeate the BBB.

Acknowledgements

We thank the ISIS Facility (Rutherford Appleton Laboratories, STFC, UK) for the allocation of neutron beam time, the UK Engineering and Physical Sciences Research Council for (EP/J002615/1) for funding, Dr Richard Gillams (Oxford) for assistance with data visualization and Christopher Sowden for performing some of the NMR measurements. We also thank the Ministerio de Ciencia y Tecnología (FIS2014-54734-P) and the Catalonia Government (2014SGR-0581) for funding.

References

1 W. J. Geldenhuys, D. D. Allen and J. R. Bloomquist, *Expert Opin. Drug Metab. Toxicol.*, 2012, **8**, 647–653.
2 A. B. Haberman, *Genet. Eng. Biotechnol. News*, 2009, **29**, 48–49.
3 W. M. Pardridge, *J. Cereb. Blood Flow Metab.*, 2012, **32**, 1959–1972.
4 B. Wilson, *Nanomedicine*, 2009, **4**, 499–502.
5 S. D. Campbell, K. J. Regina and E. D. Kharasch, *J. Biomol. Screening*, 2013, **19**, 437–444.
6 U. Norinder and M. Haeberlein, *Adv. Drug Delivery Rev.*, 2002, **54**, 291–313.
7 W. A. Banks, *BMC Neurol.*, 2009, **9**, S1–S3.
8 D. B. Kell, P. D. Dobson and S. G. Oliver, *Drug Discovery Today*, 2011, **16**, 704–714.
9 D. B. Kell, P. D. Dobson, E. Bilsland and S. G. Oliver, *Drug Discovery Today*, 2013, **18**, 218–239.
10 C. G. Zhan, S. X. Deng, J. G. Skiba, B. A. Hayes, S. M. Tschampel, G. C. Shields and D. W. Landry, *J. Comput. Chem.*, 2005, **26**, 980–986.
11 D. A. Rincón, M. Cordeiro and R. A. Mosquera, *J. Phys. Chem. A*, 2009, **113**, 13937–13942.
12 S. B. Hansen and P. Taylor, *J. Mol. Biol.*, 2007, **369**, 895–901.
13 R. Jakkala-Saibaba, P. G. Morgan and G. L. Morton, *Anaesthesia*, 2011, **66**, 1168–1170.
14 M. Perez-Reyes, S. D. Guiseppi, G. Ondrusek, A. R. Jeffcoat and C. E. Cook, *Clin. Pharm. Ther.*, 1982, **32**, 459–465.
15 D. K. Hatksukami and M. W. Fischman, *JAMA, J. Am. Med. Assoc.*, 1996, **276**, 1580–1588.
16 F. Zheng and C. Zhan, *PLoS Comput. Biol.*, 2012, **8**, 1–10.
17 H. Pajouhesh and G. R. Lenz, *J. Am. Soc. Exp. Neurother.*, 2005, **2**, 541–553.
18 E. Scoppola, A. Sodo, S. E. McLain, M. A. Ricci and F. Bruni, *Biophys. J.*, 2014, **106**, 1701–1709.
19 N. H. Rhys, A. K. Soper and L. Dougan, *J. Phys. Chem. B*, 2012, **116**, 13308–13319.
20 S. Busch, C. D. Bruce, C. Redfield, C. D. Lorenz and S. E. McLain, *Angew. Chem., Int. Ed.*, 2013, **49**, 13091–13095.
21 H. WaterBeemd, G. Camenisch, G. Folkers, J. R. Chretien and O. A. Raevsky, *J. Drug Targeting*, 1998, **6**, 151–165.
22 V. F. Sears, *Neutron News*, 1992, **3**, 29.
23 S. Dixit, J. Crain, W. C. K. Poon, J. L. Finney and A. K. Soper, *Nature*, 2002, **416**, 829–832.
24 R. J. Gillams, J. V. Busto, S. Busch, F. M. Goñi, C. D. Lorenz and S. E. McLain, *J. Phys. Chem. B*, 2015, **119**, 128–139.
25 R. Hayes, S. Imberti, G. G. Warr and R. Atkin, *Angew. Chem., Int. Ed.*, 2013, **52**, 4623–4627.
26 R. Mancinelli, F. Bruni, M. A. Ricci and S. Imberti, *J. Chem. Phys.*, 2013, **138**, 204503.
27 J. Hladílková, H. E. Fischer, P. Jungwirth and P. Mason, *J. Phys. Chem. B*, 2015, **119**, 6357–6365.
28 T. F. Headen, C. A. Howard, N. T. Skipper, M. A. Wilkinson, D. T. Bowron and A. K. Soper, *J. Am. Chem. Soc.*, 2010, **132**, 5735–5742.
29 W. B. O'Dell, D. C. Baker and S. E. McLain, *PLoS One*, 2012, **7**, e45311.

30 A. K. Soper, *GudrunN and GudrunX: Programs for correcting raw neutron and X-ray diffraction data to differential scattering cross section*, Rutherford Appleton Laboratory, STFC, UK, 2011.
31 https://www.facebook.com/disord.matt.
32 S. E. McLain, C. J. Benmore, J. E. Siewenie, J. Urquidi and J. F. C. Turner, *Angew. Chem., Int. Ed.*, 2004, **43**, 1952–1955.
33 A. K. Soper, *Mol. Simul.*, 2012, **38**, 1171–1185.
34 A. K. Soper, *Chem. Phys.*, 1996, **202**, 295–306.
35 A. K. Soper, *ISRN Phys. Chem.*, 2013, **2013**, 1–67.
36 H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *J. Chem. Phys.*, 1987, **91**, 6269.
37 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
38 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, **25**, 247–260.
39 https://gcm.upc.edu/en/members/luis-carlos/angula/ANGULA.
40 L. C. Pardo, M. Rovira-Esteva, J. L. Tamarit, N. Veglio, J. F. Bermejo and G. J. Cuello, in *Metastable Systems under Pressure*, ed. R. Sylwester, D. Aleksandra and M. Victor, Springer, Netherlands, 2010, pp. 79–91.
41 S. Busch, C. D. Lorenz, J. Taylor, L. C. Pardo and S. E. McLain, *J. Phys. Chem. B*, 2014, **118**, 14267–14277.
42 S. Busch, L. C. Pardo, W. B. O'Dell, C. D. Bruce, C. D. Lorenz and S. E. McLain, *Phys. Chem. Chem. Phys.*, 2013, **15**, 21023–21033.
43 A. J. Johnston, Y. Zhang, S. Busch, L. C. Pardo, S. Imberti and S. E. McLain, *J. Phys. Chem. B*, 2015, **119**, 5979–5987.
44 A. K. Soper, *Mol. Phys.*, 2009, **107**, 1667–1684.
45 A. K. Soper, *J. Phys.: Condens. Matter*, 2007, **19**, 335206.
46 P. Crivori, G. Cruciani, P. Carrupt and B. Testa, *J. Med. Chem.*, 2000, **43**, 2204–2216.
47 I. D. Kuntz, *Science*, 1992, **257**, 1078–1082.
48 G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures*, Springer, Berlin Heidelberg, 1991, pp. 147–155.
49 T. Steiner, *Angew. Chem., Int. Ed.*, 2002, **41**, 48–76.
50 E. C. Hulme, A. K. Soper, S. E. McLain and J. L. Finney, *Biophys. J.*, 2006, **91**, 2371–2380.
51 F. Foglia, J. M. Lawrence, C. D. Lorenz and S. E. McLain, *J. Chem. Phys.*, 2010, **133**, 145103.
52 M. Chesler, *Physiol. Rev.*, 2003, **83**, 1183–1221.
53 S. Suzuki, P. G. Green, R. E. Bumgarner, S. Dasgupta, W. A. Goddard and G. A. Blake, *Science*, 1992, **257**, 942–945.
54 W. Zhang, B. N. Markiewicz, R. S. Doerksen, A. B. S. III and F. Gai, *Phys. Chem. Chem. Phys.*, 2015, DOI: 10.1039/c5cp04413h.
55 S. E. Pagnotta, S. E. McLain, A. K. Soper, F. Bruni and M. A. Ricci, *J. Phys. Chem. B*, 2010, **114**, 4904–4908.
56 E. J. Gabe and W. H. Barnes, *Acta Crystallogr.*, 1963, **16**, 796–801.
57 B. Kuhn, P. Mohr and M. Stahl, *J. Med. Chem.*, 2010, **53**, 2601–2611.
58 L. Infantes and W. D. S. Motherwell, *Z. Kristallogr.*, 2005, **220**, 333–339.
59 N. A. Larsen, B. Zhou, A. Heine, P. Wirsching, K. D. Janda and I. A. Wilson, *J. Mol. Biol.*, 2001, **311**, 9–15.