

Cite this: *Chem. Sci.*, 2021, 12, 2060

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 20th October 2020  
Accepted 17th December 2020

DOI: 10.1039/d0sc05785a

rsc.li/chemical-science

# How many water molecules are needed to solvate one?<sup>†</sup>

Alessandro Rognoni,<sup>ID</sup> Riccardo Conte<sup>ID</sup> and Michele Ceotto<sup>ID</sup>\*

Many efforts undertaken to study the solvation process have led to general theories that may describe mean properties, but are unable to provide a detailed understanding at the molecular level. Remarkably, the basic question of how many solvent molecules are necessary to solvate one solute molecule is still open. By exploring several water aggregates of increasing complexity, in this contribution we employ semiclassical spectroscopy to determine on quantum dynamical grounds the minimal network of surrounding water molecules to make the central one display the same vibrational features of liquid water. We find out that double-acceptor double-donor tetrahedral coordination constituting the standard picture is necessary but not sufficient, and that particular care must be reserved for the quantum description of the combination band due to the coupling of the central monomer bending mode with network librations. It is actually our ability to investigate the combination band with a quantum-derived approach that allows us to answer the titular question. The minimal structure eventually responsible for proper solvation is made of a total of 21 water molecules and includes two complete solvation shells, of which the whole first one is tetrahedrally coordinated to the central molecule.

## 1. Introduction

The famous sorites paradox was first stated by Eubulides of Miletus in the 4th century B.C. In its original version it amounts to the question of down to which size a heap of sand can be still considered as such, when grains are removed one by one. A chemical-physical variant of the paradox can be formulated by counting how many water molecules make up the smallest water droplet, *i.e.* a supramolecular aggregate featuring the same properties as liquid water. Strictly related to this issue is the ubiquitous concept of solvation. Solvation is characterized by the nature of solute–solvent interactions, and often implies a structural reorganization of both. In the case of hydration, *i.e.* solvation performed by water, a specialized network of hydrogen bonds is built around the solute with solvent molecules arranged in shells. In practice, crucial phenomena like protein folding or reactivity in the condensed phase are strictly dependent on the mechanism of solvation. The structural organization of the solvent is also fundamental in explaining the hydrophobic force, sometimes described as the aggregation or desolvation of water repelling solutes that is thought to favor water-displacement equilibria when a ligand, such as a drug, approaches a receptor site in a hydrated protein. The principles of hydration are still debated with a constantly evolving

consensus, which has moved over time from the original iceberg model to more recent and refined descriptions.<sup>1</sup> Therefore, it is not an overstatement to affirm that the investigation of the spatial organization and number of water molecules needed to hydrate a central target species is of utmost importance and still an open issue.<sup>2–6</sup>

The conclusions that can be drawn from such a study certainly depend on the polarity and structural characteristics of the solute, but perhaps the first “solute” worth studying is water itself. In spite of the many attempts, the issue has not been settled yet. The reason for this lack of success is essentially two-fold. From the experimental point of view, it is difficult to determine with sufficient precision the number of molecules constituting a water cluster of medium-large size or to identify the specific cluster arrangement under investigation. Theoretically, it is hard to include key quantum effects in a simulation involving several dozens of degrees of freedom. To investigate the solvent at the molecular level one should neither employ rigid-molecule models within coarse-grained approaches nor invoke uniform medium approximations returning average properties. Conversely, each and every solvent molecule must be described explicitly and accurately according to the laws of quantum mechanics.

Vibrational spectroscopy provides a powerful experimental and theoretical tool to get insights into the molecular structure.<sup>7</sup> In liquid water, the infrared (IR) spectrum is characterized by a broad band for the stretches, located between 3200 and 3550 cm<sup>-1</sup>, which is strongly red shifted with respect to the corresponding lines of the isolated water molecule at 3657 and

Dipartimento di Chimica, Università degli Studi di Milano, Via Golgi 19, 20133 Milano, Italy. E-mail: michele.ceotto@unimi.it

<sup>†</sup> Electronic supplementary information (ESI) available: Theory, methodological approach, and benchmark calculations. See DOI: 10.1039/d0sc05785a



3756 wavenumbers.<sup>8</sup> This is determined by the unique and complex network of intermolecular hydrogen bonds which weakens the intramolecular stretches. Other relevant spectral features of liquid water include the band of hindered translations and rotations (librations) below  $1000\text{ cm}^{-1}$ , the bending signal (at around  $1650\text{ cm}^{-1}$ ), and the combination band of librations and bendings spanning the region around  $2100\text{ cm}^{-1}$ . The features of the liquid water IR spectrum are shown in the bottom half of Fig. 1.

We stipulate that for a water molecule to be labeled ‘solvated’ it must possess spectral features that resemble those of the bulk. In particular, besides the intramolecular bending and stretching signals, it is also crucial that the combination band of bending with librations matches the corresponding band of the bulk spectrum. A simple description of bulk liquid water depicts the  $\text{H}_2\text{O}$  molecules as tetrahedrally coordinated, which means that each individual acts as a double donor and double acceptor of hydrogen bonds. Reality is more complex, because liquid water is characterized by a variety of hydrogen-bond lengths and coordination numbers.<sup>12,13</sup> For instance, it is possible to find locally overcoordinated subunits, featuring three-acceptor two-donor configurations, which are responsible for the spectral tails at high frequency (on the acceptor side) and low frequency (on the donor side) of the stretching band.<sup>14</sup> Tri-, tetra-, and penta-coordinations are presumably the most common instances, and by means of Raman multivariate curve

resolution spectroscopy (MCR-Raman), it has been possible to identify the contribution of tetra-coordinated molecules within the wide stretching band of liquid water.<sup>12</sup> Another powerful investigation tool, THz spectroscopy, covers instead the typical energy range of intermolecular interaction in condensed phases. It is able to characterize the collective motion in liquids and has permitted spectroscopic signals at  $200\text{ cm}^{-1}$  and  $80\text{ cm}^{-1}$  to be assigned to first and second solvation shell dynamics, respectively.<sup>15</sup>

The study of water clusters of increasing size provides a way to get molecular level insights into condensed phase systems like liquid water and ice.<sup>16</sup> Many experimental and theoretical investigations have been undertaken to describe the intermolecular interactions of these systems, going from small quasi-planar (with reference to the oxygen atoms) clusters to larger ones.<sup>17–20</sup> On the theoretical side, this has led to the construction of accurate potential energy surfaces (PESs),<sup>21,22</sup> search for geometric and energy minima in a very complicated energy landscape, and spectroscopic calculations.<sup>23,24</sup> For instance, the structural characteristics of water clusters beyond the transition between “all-surface” and “interior” ones have been studied. The transition takes place at  $(\text{H}_2\text{O})_{17}$  when one monomer is included in the framework built by the ensemble of the other water molecules.<sup>25</sup> Another remarkable and very recent study, aimed at assessing the minimum system size showing bulk-like properties, has been performed on ice I.<sup>26</sup> The main target of that investigation was to find the minimum dimensionality needed by a water cluster to show the hydrogen-bond pattern of ice I. Experimental studies and classical molecular dynamics simulations concerning infrared spectra and free energy estimates have been weighed in to conclude that in the range between 90 and 150 water molecules ice I is present in a mixture with amorphous clusters. In spite of such a variety of research efforts, some important aspects have not yet been sorted out and are still controversial. For instance, is tetrahedral coordination strictly necessary for a target water molecule to display the spectral properties of bulk liquid water? Is it sufficient? And, finally, what is the minimum number of water molecules required to fully solvate a single one?

## 2. Results and discussion

To give a precise answer to these questions we performed a quantum investigation by means of semiclassical spectroscopy, a theoretical and computational technique suitable for dealing with high dimensional or complex systems and reproducing quantum effects starting from classical trajectories.<sup>27–31</sup> The upper panel of Fig. 1 demonstrates the high accuracy of SC spectroscopy in an application to the water dimer when compared with experimental results. However, it also shows that, as expected, the spectroscopic features of the water dimer (especially at high frequency) are not compatible with those of bulk water, confirming that more elaborate structures are necessary to mimic the bulk. Semiclassical methods have been applied successfully in a range going from small molecules,<sup>32–34</sup> including fluxional ones,<sup>35</sup> to medium-large systems.<sup>36</sup> For this work we employed the multiple coherent states divide-and-

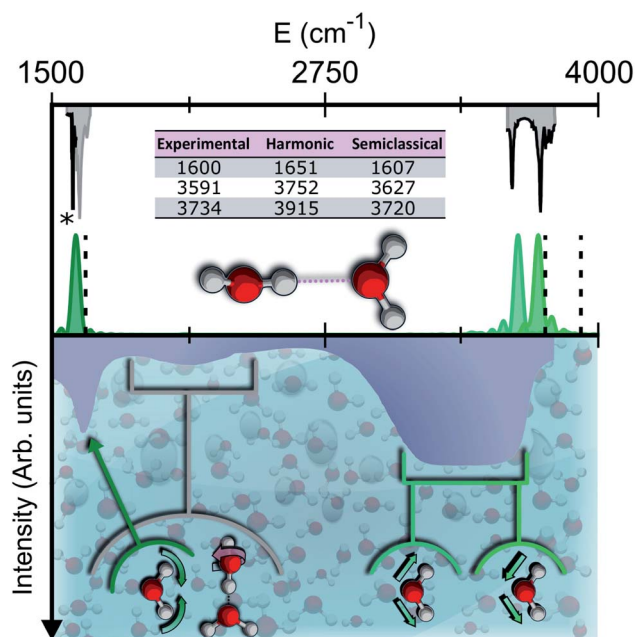


Fig. 1 Semiclassical and experimental spectra of the water dimer. Top half: comparison between the experimentally detected fundamental frequencies of vibration for the water dimer (grey)<sup>9–11</sup> and their theoretical semiclassical counterpart (green). The dashed vertical lines locate the harmonic estimates. The asterisk helps the eye detect the exact position of the reference experimental bending signal. Bottom half: the experimental IR spectrum of liquid water (solid contour). The normal modes involved in several signals are indicated with arrows and associated with the corresponding portions of the spectrum.



conquer semiclassical initial value representation technique (MC-DC SCIVR),<sup>37</sup> which has permitted open issues concerning protonated glycine supramolecular systems to be solved<sup>38</sup> and has already been applied to the vibrational investigation of small water clusters.<sup>39</sup> The interested reader may find more theoretical information on semiclassical dynamics and spectroscopy in Section 1 of the ESI.† For our semiclassical calculations we employed the water many-body MB-pol potential energy surface.<sup>40</sup> We ran single-trajectory simulations about 0.7 ps long with a time step of about 0.12 fs. The initial energy of vibration was set equal to zero for the low-frequency modes (those with harmonic frequencies below the bending frequency), while the rotational and the translational motion of each monomer were eliminated at each step and the velocities were rescaled accordingly. More details to permit the reproducibility of results are provided in Section 2 of the ESI.†

Our analysis starts by studying three small water clusters: the dimer, which we have already anticipated; the trimer; the hexamer, this one in its prism configuration. The remarkable information collected from these simulations is that the frequencies of asymmetric stretches are blue shifted with respect to the experimental band of liquid water, in spite of the wide range it spans, and the bending signal is shifted from its experimental counterpart too. These spectral features demonstrate that solvation is not achieved for this set of small clusters, as confirmed by Fig. 2, which reports for the trimer and the hexamer prism the power spectra obtained for the bending and stretches of each water monomer superimposed on the IR spectrum of liquid water. Detailed frequency values can be found in Section 3 of the ESI.† We remark that the height of our peaks is not related to the experimental IR intensity (see Section 1 of the ESI).† Therefore, we have chosen to scale the intensities to ease the comparison between theory and experiments. From a structural point of view, the monomers in these small clusters are not tetrahedrally coordinated, hinting at the necessity to have an appropriate shell of coordination in order to display solvation properties, as suggested by Raman and THz

spectroscopy studies. Fig. 2 also points out a key characteristic of calculated power spectra. They are able to reproduce all vibrational energy levels including those not involved in active IR or Raman transitions. For this reason our semiclassical simulations provide plenty of spectroscopic information and incorporate not only the target signals for bending and stretch fundamentals, but also combined excitations, Fermi resonances, overtones, and secondary signals related to other vibrational modes strongly coupled to the target ones. It is indeed the high density of vibrational states in association with inter-monomer interactions that shapes the wide IR bands typical of liquid water, an aspect also confirmed experimentally.<sup>41,42</sup>

Before moving to more complex systems, we note that there are two water oligomers that do have (in one of their several isomers) tetrahedral coordination around one of the monomers. They are the water pentamer and heptamer, for which the target stretches are well inside the corresponding IR band of liquid water (see Fig. 3). However, this does not constitute sufficient evidence of solvation. In fact, the IR spectrum of liquid water is also characterized by the presence of a band in the 1900–2300  $\text{cm}^{-1}$  range due to combined excitations of intra-monomer bendings and inter-monomer librations, as mentioned above. This feature can be explored by exploiting the capability of semiclassical spectroscopy to recover quantum effects. Specifically, the signal of the combined excitation of the bending of the target monomer with the libration most coupled to it (according to the analysis of the Hessian matrix at the equilibrium configuration) has been simulated showing that it is located outside the corresponding band in liquid water. This proves that even the presence of a 4-monomer coordination shell is not sufficient to achieve solvation, and an investigation at the quantum level able to reproduce combination bands is fundamental to point out this aspect. Furthermore, for tetrahedrally coordinated water molecules we can employ the MCR-Raman spectrum presented in ref. 12 to narrow the stretching band of liquid water. In Fig. 3, the MCR-Raman spectrum has been highlighted.

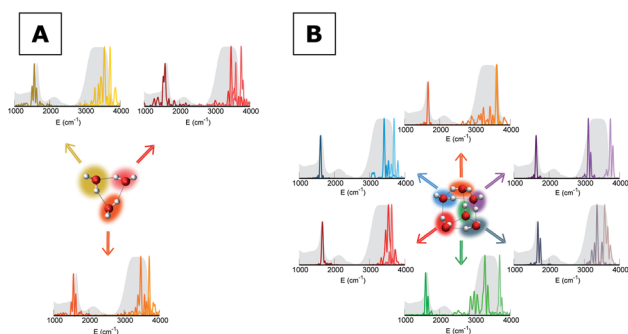


Fig. 2 Power spectra for bending and stretches of individual monomers of small water clusters. (A) Trimer and (B) hexamer prism. The colored haloes link each monomer to the appropriate power spectrum. Color nuances are darker for bendings and lighter for symmetric and asymmetric stretches. The IR spectrum of liquid water is reported as shaded gray areas. The intensities of power spectra have been scaled to match the maximum of the corresponding experimental IR band.

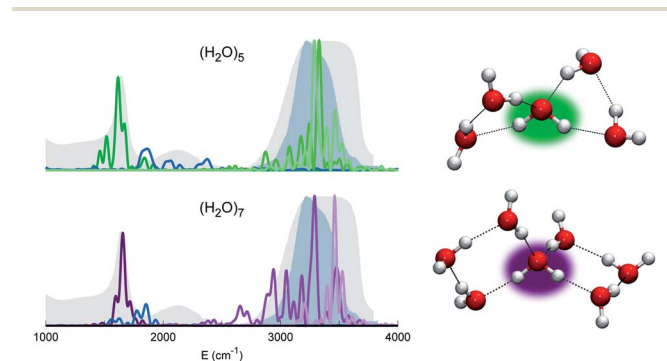


Fig. 3 Power spectra of the tetrahedrally coordinated monomer in different water clusters. Top left column: pentamer; bottom left column: heptamer. The combined excitation of the bending with the most coupled libration mode is also reported in dark blue. The liquid water IR spectrum is shown as shaded gray areas, and the MCR-Raman band (from ref. 12) for the stretches of tetrahedrally coordinated monomers in liquid water (light blue) is also presented.



We note that, while the pentamer with its unperturbed tetrahedral coordination has both stretches well within the band, the asymmetric stretch for the heptamer is displaced toward the border of it. This is interpreted as the effect due to two additional monomers not directly coordinated to the central molecule, a first clear indication that tetrahedral coordination is not sufficient, and that a more extended network is required to reproduce the solvation features.

Therefore, our quest for the minimal solvation structure proceeds by investigating more sizable systems for which a double shell of coordination around a central, target monomer is present. This is the case, for instance, of  $(\text{H}_2\text{O})_{19}$ . The spectral features of the central monomer of this aggregate are shown in Fig. 4. Even if the results are closing in toward the expected spectroscopic characteristics of a solvated molecule, some inconsistencies are still present. The bending signal is shifted from the corresponding band in liquid water, and the stretches, even if located at still reasonable frequencies, are quite blue shifted with respect to the tetrahedral coordination band. We reckon that another, more complex structure needs to be studied and move to  $(\text{H}_2\text{O})_{21}$ . In this case, the spectroscopic features of the bending and stretches of the target monomer are exactly where they are expected to be for a solvated water molecule. The final check is given by the combination peak, which is located at  $1950\text{ cm}^{-1}$ , *i.e.* still within the range of the IR experimental combination band. For these reasons our guess is that  $(\text{H}_2\text{O})_{21}$  represents the smallest water aggregate in which one monomer is solvated in the same way as molecules in the liquid.

To confirm this hypothesis we also analyzed a bigger structure, *i.e.*  $(\text{H}_2\text{O})_{23}$ . In this case, even the combination signal moves to the center of the liquid water band, confirming that a solvation structure has indeed been reached. This shift is due to the higher frequency of the libration mode most coupled to

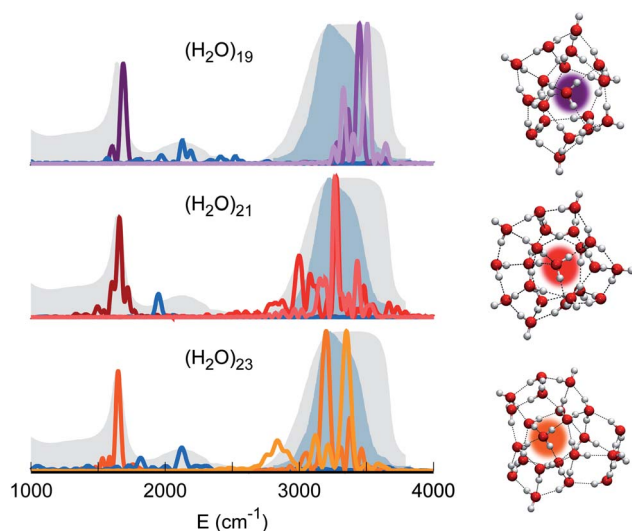


Fig. 4 Power spectra of the target monomer for larger clusters. The peaks corresponding to bendings, stretches, and combined bending-libration bands of  $(\text{H}_2\text{O})_{19}$  (top),  $(\text{H}_2\text{O})_{21}$  (middle), and  $(\text{H}_2\text{O})_{23}$  (bottom) are reported.

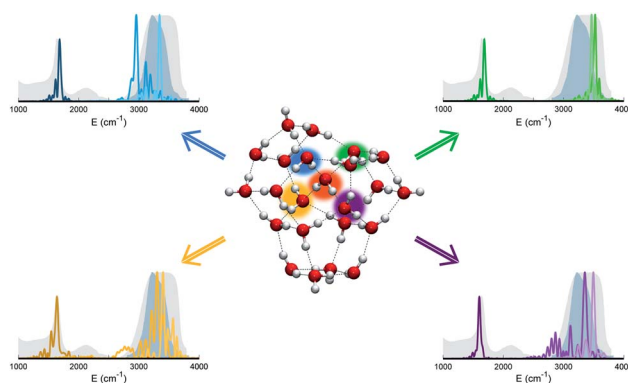


Fig. 5 Power spectra of bending and stretch signals for the four monomers that are hydrogen bonded to the central monomer in  $(\text{H}_2\text{O})_{23}$ . Different monomers and the corresponding power spectra are indicated with different colors (blue, green, violet, and yellow). The central monomer is highlighted in orange. Despite the tetrahedral coordination of these monomers, none of them matches the liquid water experimental spectra.

the target bending (about  $240\text{ cm}^{-1}$  at the harmonic level) in the case of  $(\text{H}_2\text{O})_{23}$  with respect to  $(\text{H}_2\text{O})_{21}$ . In the  $(\text{H}_2\text{O})_{23}$  cluster, similarly to the heptamer but in a less prominent way, the asymmetric signal moves toward the edge of the tetrahedral coordination band revealing that, even if solvation is maintained, the two additional monomers have slightly influenced the coordination and solvation network. Detailed results can be found in Section 3 of the ESI.†

The persistent influence of outer additional monomers triggered our interest in having a closer look at the structural properties of  $(\text{H}_2\text{O})_{23}$ . In particular, Fig. 5 focuses on the four monomers constituting the first coordination shell for the central one. Each of these four water molecules is itself tetrahedrally coordinated, but several spectral features do not match the liquid water ones. This is further confirmation that tetrahedral coordination alone does not warrant solvation, and that it is necessary to account for the overall network of hydrogen interactions which provides a complex and cooperative way for solvation in water.

### 3. Conclusions

Even though the bands characterizing the IR spectrum of liquid water are necessarily broad, a comparison with the results of accurate quantum spectroscopic simulations is able to discriminate the effective solvation properties of a central molecule embedded in clusters of increasing size. Small water clusters like the dimer, trimer, and hexamer are spectroscopically incompatible with solvation and their structure obviously lacks an appropriate coordination shell, which is found instead in the pentamer and heptamer that display tetracoordination at the central molecule. However, even in these latter oligomers the target spectral features do not match those of liquid water with reference to the combination band of bending and librations. Eventually, the size of the water aggregate had to be raised to 21 monomers in order to find the first example of a structure



providing a proper solvation of its central molecule. This finding was confirmed by the study of a 23-monomer cluster. The bottom line is that tetrahedral coordination is a necessary but not sufficient condition; the minimal solvation structure, the primary aim of our quest, is found to consist of a central water molecule surrounded by at least 20 water monomers organized in a double coordination shell. Substantial cooperative effects through the extended bonding manifold are at work in determining the spectral properties of water solvation.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

The authors thank Professor A. Gavezzotti for his careful reading of the manuscript. The authors acknowledge financial support from the European Research Council [Grant Agreement No. (647107) – SEMICOMPLEX – ERC-2014-CoG] under the European Union's Horizon 2020 research and innovation programme and from the Italian Ministry of Education, University, and Research (MIUR) (FARE programme Grant No. R16KN7XBRB-project QURE).

## Notes and references

- H. S. Frank and M. W. Evans, *J. Chem. Phys.*, 1945, **13**, 507–532.
- O. Mishima and H. Stanley, *Nature*, 1998, **396**, 329–335.
- J. Errington and P. Debenedetti, *Nature*, 2001, **409**, 318–321.
- D. Marx, *Science*, 2004, **303**, 634–636.
- C. A. Angell, *Science*, 2008, **319**, 582–587.
- J. A. Fournier, C. J. Johnson, C. T. Wolke, G. H. Weddle, A. B. Wolk and M. A. Johnson, *Science*, 2014, **344**, 1009–1012.
- C. Aieta, M. Micciarelli, G. Bertaina and M. Ceotto, *Nat. Commun.*, 2020, **11**, 1–9.
- A. G. Csaszar, E. Matyus, T. Szidarovszky, L. Lodi, N. F. Zobov, S. V. Shirin, O. L. Polyansky and J. Tennyson, *J. Quant. Spectrosc. Radiat. Transfer*, 2010, **111**, 1043–1064.
- R. H. Page, J. G. Frey, Y. Lee and Y.-R. Shen, *Chem. Phys. Lett.*, 1984, **106**, 373–376.
- G. Ayers and A. Pullin, *Spectrochim. Acta, Part A*, 1976, **32**, 1629–1639.
- Y. Bouteiller and J. Perchard, *Chem. Phys.*, 2004, **305**, 1–12.
- D. Ben-Amotz, *J. Am. Chem. Soc.*, 2019, **141**, 10569–10580.
- J. A. Fournier, C. T. Wolke, C. J. Johnson, M. A. Johnson, N. Heine, S. Gewinner, W. Schöllkopf, T. K. Esser, M. R. Fagiani, H. Knorke and K. R. Asmis, *Proc. Natl. Acad. Sci.*, 2014, **111**, 18132–18137.
- T. Morawietz, A. S. Urbina, P. K. Wise, X. Wu, W. Lu, D. Ben-Amotz and T. E. Markland, *J. Phys. Chem. Lett.*, 2019, **10**, 6067–6073.
- M. Heyden, J. Sun, S. Funkner, G. Mathias, H. Forbert, M. Havenith and D. Marx, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 12068–12073.
- N. Yang, C. H. Duong, P. J. Kelleher, A. B. McCoy and M. A. Johnson, *Science*, 2019, **364**, 275–278.
- Y. Wang and J. M. Bowman, *J. Chem. Phys.*, 2011, **134**, 154510.
- R. Schwan, M. Kaufmann, D. Leicht, G. Schwaab and M. Havenith, *Phys. Chem. Chem. Phys.*, 2016, **18**, 24063–24069.
- J. O. Richardson, C. Pérez, S. Lobsiger, A. A. Reid, B. Temelso, G. C. Shields, Z. Kisiel, D. J. Wales, B. H. Pate and S. C. Althorpe, *Science*, 2016, **351**, 1310–1313.
- A. Rakshit, P. Bandyopadhyay, J. P. Heindel and S. S. Xantheas, *J. Chem. Phys.*, 2019, **151**, 214307.
- Y. Wang, B. C. Shepler, B. J. Braams and J. M. Bowman, *J. Chem. Phys.*, 2009, **131**, 054511.
- Y. Wang and J. M. Bowman, *Phys. Chem. Chem. Phys.*, 2016, **18**, 24057–24062.
- D. J. Wales, *J. Am. Chem. Soc.*, 1993, **115**, 11180–11190.
- R. J. Saykally and D. J. Wales, *Science*, 2012, **336**, 814–815.
- C. C. Pradzynski, C. W. Dierking, F. Zurheide, R. M. Forck, U. Buck, T. Zeuch and S. S. Xantheas, *Phys. Chem. Chem. Phys.*, 2014, **16**, 26691–26696.
- D. R. Moberg, D. Becker, C. W. Dierking, F. Zurheide, B. Bandow, U. Buck, A. Hudait, V. Molinero, F. Paesani and T. Zeuch, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 24413–24419.
- W. H. Miller, *Adv. Chem. Phys.*, 1974, **25**, 69–177.
- E. J. Heller, *Acc. Chem. Res.*, 1981, **14**, 368–375.
- M. F. Herman and E. Kluk, *Chem. Phys.*, 1984, **91**, 27–34.
- A. L. Kaledin and W. H. Miller, *J. Chem. Phys.*, 2003, **118**, 7174–7182.
- W. H. Miller, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6660–6664.
- A. L. Kaledin and W. H. Miller, *J. Chem. Phys.*, 2003, **119**, 3078–3084.
- R. Conte, A. Aspuru-Guzik and M. Ceotto, *J. Phys. Chem. Lett.*, 2013, **4**, 3407–3412.
- F. Gabas, R. Conte and M. Ceotto, *J. Chem. Theory Comput.*, 2017, **13**, 2378.
- G. Bertaina, G. Di Liberto and M. Ceotto, *J. Chem. Phys.*, 2019, **151**, 114307.
- F. Gabas, R. Conte and M. Ceotto, *J. Chem. Theory Comput.*, 2020, **16**, 3476–3485.
- M. Ceotto, G. Di Liberto and R. Conte, *Phys. Rev. Lett.*, 2017, **119**, 010401.
- F. Gabas, G. Di Liberto, R. Conte and M. Ceotto, *Chem. Sci.*, 2018, **9**, 7894–7901.
- G. Di Liberto, R. Conte and M. Ceotto, *J. Chem. Phys.*, 2018, **148**, 104302.
- G. R. Medders, V. Babin and F. Paesani, *J. Chem. Theory Comput.*, 2014, **10**, 2906–2910.
- B. Wang, W. Jiang, Y. Gao, Z. Zhang, C. Sun, F. Liu and Z. Wang, *RSC Adv.*, 2017, **7**, 11680–11683.
- N. Yang, C. H. Duong, P. J. Kelleher and M. A. Johnson, *Nat. Chem.*, 2020, **12**, 159–164.

