Manifolds of low energy structures for a magic number of hydrated sulfate : SO 4 2– (H 2 O) 24

Carine Clavaguéra, Florian Thaunay, Gilles Ohanessian

Journal Name

Manifolds of low energy structures for a magic number of hydrated sulfate : $SO_4^{2-}(H_2O)_{24}$ [†]

Carine Clavaguéra,[*a] Florian Thaunay,[b] and Gilles Ohanessian[*b]

Low energy structures of $SO_4^{2-}(H_2O)_{24}$ have been obtained using a combination of classical molecular dynamics simulations and refinement of structures and energies by quantum chemical calculations. Extensive exploration of the potential energy surface led to a number of low-energy structures, confirmed by accurate calibration calculations. An overall analysis of this large set was made after devising appropriate structural descriptors such as the numbers of cycles and their combinations. Low energy structures bear common motifs, the most prominent being fused cycles involving alternatively four and six water molecules. The latter adopt specific conformations which ensure the appropriate surface curvature to form a closed cage without dangling O-H bonds and at the same time provide 12-coordination of the sulfate ion. A prominent feature to take into account is isomerism via inversion of hydrogen bond orientations along cycles. This generates large families of ca. 100 isomers for this cluster size, spanning energy windows of 10-30 kJ.mol^{-1}. This relatively ignored isomerism must be taken into account to identify reliably the lowest energy minima. The overall picture is that the magic number cluster $SO_4^{2-}(H_2O)_{24}$ does not correspond to formation of a single, remarkable structure, but rather to a manifold of structural families with similar stabilities. Extensive calculations on isomerization mechanisms within a family indicate that large barriers are associated to direct inversion of hydrogen bond networks. Possible implications of these results for magic number clusters of other anions are discussed.

1 Introduction

Sulfate is known to be important in many contexts, for instance because of the numerous inorganic salts it makes and also as an agent of post-translational modifications of proteins. Even the simple hydrated sulfate dianion SO_4^{2-} is relevant in several contexts, e. g. it is the fourth most abundant anion in mammalian plasma and has a regulation role in many metabolic and cellular processes in vivo.[1,2] It also plays a key role in nucleation of cloud drops and ice particles in the upper troposphere.[3] Revealing some of the structural and energetic details of sulfate hydration is made possible in nanodroplets by a combination of experimental and computational studies. Experiments using mass spectrometers enable selecting ionic clusters of specific sizes, e.g. $SO_4^{2-}(H_2O)_n$ with well controlled n, and then obtaining signatures by e. g. photoelectron[4–6] and infrared photodissociation

(IRPD)[7] spectroscopies. These signatures must then be translated into structural features, generally requiring computational studies. Within this flourishing literature, the sulfate dianion in water nanodroplets $SO_4^{2-}(H_2O)_n$ has been one of the thoroughly studied hydrated gaseous ions. The cluster size range studied is impressive : n=3-550 for IRPD studies[8–10] and n=4-40 for photoelectron spectroscopy.[4,5,11,12] Water dissociation energetics have also been investigated.[13,14] For small sizes, typically n≤12, the experimental signatures are used to discriminate between structures whose properties can be calculated accurately. For larger sizes, signatures are used to detect structural transitions, such as the onset for the formation of the second solvation shell,[15] the appearance of dangling O-H bonds at the cluster surface[16] or structural population changes as a function of temperature (e. g., for n=6[15] and for n=550[9]).

Exploiting experimental results requires identification of the lowest energy structures for each size. For such species involving numerous non-covalent interactions, computations have shown that for virtually every cluster size investigated, there are numerous structures inside a few kJ.mol^{-1} window even for the smallest clusters, e. g. n=3-7,[17] which cannot be identified straightforwardly. As shown for e. g. small hydrated sulfate clusters[18]

[a] Institut de Chimie Physique, Université Paris-Saclay - CNRS, UMR 8000, 91405 Orsay, France; E-mail: carine.clavaguera@universite-paris-saclay.fr
[b] Laboratoire de Chimie Moléculaire (LCM), CNRS, École Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France; E-mail: gilles.ohanessian@polytechnique.edu

and pure water clusters (n=16, 17),[19] DFT methods can be used to obtain reliable relative energies as compared to accurate wavefunction methods such as CCSD(T), as long as great care is exercised in choosing the appropriate density functional.

The situation is less clear concerning the appropriate way to explore efficiently potential energy surfaces of such non-covalent systems. Molecular dynamics at the DFT level can only be carried out on time frames of tens of picoseconds. While this permits accurate local exploration and proper account of anharmonic and temperature effects on populations and vibrational spectra, it is not yet amenable to extensive exploration of structural space.[15,20–22] Global methods such as basin-hopping global optimization,[23–25] simulated annealing[26] and various genetic algorithms[27,28] have been developed to improve the energy-landscape exploration. Approximate quantum methods have been used to extend the range of system sizes in comparison with the DFT framework. Indeed, the semi-empirical PM7 method coupled to a simulated annealing algorithm has been successfully applied to identify the numerous isomers of $[Cu(H_2O)_n]^{2+}$ clusters, with n = 12, 16, and 18.[26] Moreover, the Self-Consistent-Charge Density-Functional Based Tight-Binding (SCC-DFTB) approach combined to a global optimization search turned out to be efficient in describing the potential energy surfaces of sulfate-water clusters up to 20 water molecules and protonated water-clusters (20 to 23 water molecules).[25,29] However, an extensive, even more an exhaustive, PES exploration for these medium-size clusters is likely to be achieved only using a classical representation of the energy with a force field.[24,30–34] The quality of the force field has an influence on the relative energies of the various isomers in such non-covalent systems and polarization effects have been included to improve energetics.[34–37] The TTM2.1-F and TTM3-F flexible polarizable potentials have been developed to reproduce accurately infrared spectra of liquid water and ice by improving the description of the hydrogen bond network.[34] This force field family has been used to reproduce accurately the structures and spectroscopic features of large water clusters (from 20 to 400) and protonated water clusters with 20 molecules.[27,38,39] The AMOEBA polarizable force field has the specific feature for the electrostatic treatment of combining multipolar expansion up to quadrupole moments and a self-consistent atomic dipole polarization procedure.[35] In recent years, AMOEBA has shown its capability to reproduce accurately the competition between different intermolecular interactions in solution and the solvent organisation around a solute.[35,40,41] The PES of $SO_4^{2-}(H_2O)_8$ was explored by molecular dynamics associated to the AMOEBA force field to extract low-energy isomers for DFT calculations in order to analyze the vibrational spectrum.[42]

In the last years, we have extended the AMOEBA force field to the finite temperature calculation of infrared spectra of gas phase molecules.[43–45] This force field has been used to study the zinc coordination number in water as a function of cluster size and temperature[46] and to identify the low-energy structures of $SO_4^{2-}(H_2O)_n$, n=9-13.[8] The latter work included discussion of the especially stable structure at n=12.[14] We showed that interpretation in terms of a single structure corresponding to closure of the first hydration shell, is questionable. We pursue this work

here, using a similar approach outlined in Section 2, by exploring the potential energy surface of $SO_4^{2-}(H_2O)_{24}$ with a combination of methods which, while without pretention of being exhaustive, can shed light on its low energy structures. It is shown in Section 3 that permutational isomerism of hydrogen bond cycles is a crucial addition to any exploration strategy. Describing the results then requires, for systems this large and fluxional, the use of appropriate structural descriptors, as described in Section 4. The results draw a different picture of structures underlying a magic number cluster.

2 Computational methodology

2.1 Force field

The AMOEBA polarizable force field has been shown to be particularly relevant to the computation of chemical properties of a number of hydrated mono- and divalent atomic cations and anions,[47–50] the Mg^{2+}:phosphate ion pair[51] or lanthanide and actinide cations.[52] It has the main advantage of combining a complex electrostatic model including permanent charge, dipole, and quadrupole moments on each atom, derived from quantum mechanical calculations, and many-body polarization effects, which are explicitly treated using a self-consistent atomic dipole polarization procedure.[53,54] Repulsion-dispersion interactions between pairs of non-bonded atoms are represented by a buffered 14-7 potential.[55] The intramolecular valence terms consist of anharmonic bond stretchings and angle bendings and torsions with the MM3 force field energy form, which also includes out-of-plane bending and bond angle cross-terms.[56] The AMOEBA14 parameter set for water[57] was used, however vibrational spectra were found to be greatly improved with the use of fluctuating atomic charges.[44] The charges on H are allowed to change in a $-0.05 : +0.06]$ range in response to distance and angle fluctuations, while the charge on O adapts to maintain overall neutrality. The intramolecular OHO bending force constant was also modified from 600 to 400 kcal.mol^{-1}.Å^{-2}. Modifications to sulfate parameters available in the literature[17] were used as described in our previous work[44] : the S-O bond stretching constant was modified to fit the experimental stretching frequencies of micro-hydrated sulfate.[10] Atomic multipoles for S and O were determined using the DMA procedure[58,59] using the MP2/aug-cc-pVTZ electron density of SO_4^{2-} and $SO_4^{2-}(H_2O)_6$.

2.2 Exploration of the potential energy surfaces

MD simulations were run with a modified version of TINKER-6[44,60] using a 1 fs time step in the canonical ensemble and the Nosé-Hoover thermostat. In order to model low temperature ions as generated experimentally, structural evolution was followed at 200 K. Initial MD tests were carried out from a manually constructed structure, after equilibration for 50 ps, in order to confirm that substantial modifications of the hydrogen-bond network were perceptible at 200 K. Trajectories were then propagated for 4 ns, starting from two very different structures, one with a free O-H bond and one without and more compact, to ensure diversity. Replica Exchange MD could have been used instead of several independent, longer trajectories as used here. It is difficult to make

a general comparison of the efficiencies of the two approaches. Structures were extracted every 200 ps to be subjected to DFT optimization. Whenever very low DFT energies were found for two successive points, the time step for extraction was reduced from 200 to 100 ps. Attempts at further reduction to 50 ps led to redundant results, thus the 200/100 ps time step for extraction was deemed satisfactory and applied to the entire 4 ns trajectories. The same procedure was further carried out starting from two additional starting points: the lowest energy found so far, and a manually constructed structure with the sulfate ion lying on the water cluster surface, to ensure diversity. This generated a total of 16 ns of trajectories, out of which over 100 structures were recomputed at the DFT level (see below). In addition, the structure previously published by Johnston et al. was used.[30] Finally, the AMOEBA search cannot sample salt bridge structures. A number of attempts at introducing hydrated $OH^- \cdots H_3O^+ \cdots SO_4^{2-}$ motifs, adjacent or water-separated, were unsuccessful as they collapsed to "canonical" water clusters around the sulfate ion when optimized at the DFT level. It may be noted that initial exploration by classical MD simulations using the AMOEBA force field, prior to using more accurate energy calculations, has been previously carried for related systems including $SO_4^{2-}(H_2O)_8$,[42] $I^-(THF)_n$ n=1-30[61] and $(H_2O)_n$ n=20-30.[62]

2.3 Quantum chemical calculations

Based on previous work, two density functionals were used for geometry optimizations and vibrational frequency calculations. The benchmarking study of Head-Gordon et al. on $SO_4^{2-}(H_2O)_n$ n=3-6 concluded that M11 is reliable (we avoid here double hybrid functionals for computational efficiency).[63] Since this calibration was done on small clusters for which sulfate-water interactions are energetically dominant, we also used the ωB97X-D functional which was found by Truhlar et al. to perform well for the relative energies of pure water clusters containing 16 or 17 water molecules.[19] The cc-pVTZ basis set was used for geometry optimizations and vibrational frequency calculations, while improved energetics used aug-cc-pVTZ. Reference energetics were obtained at the DLPNO-CCSD(T) level using both cc-pVTZ and aug-cc-pVTZ bases. All DFT calculations were carried out with Gaussian09[64] while for DLPNO- CCSD(T) calculations, ORCA 4.1[65] was used.

As seen in Table S1 for 9 low-energy structures of $SO_4^{2-}(H_2O)_{24}$, relative energies differ at most by 1.8 kJ.mol^{-1} between ωB97X-D/aug-cc-pVTZ+ZPE and DLPNO/aug-cc-pVTZ+ZPE except for one at 3.2 kJ.mol^{-1}. It is also clear that ωB97X-D performs much better than M11 for $SO_4^{2-}(H_2O)_{24}$, thus only ωB97X-D results are described below. A more extensive calibration of ωB97X-D/cc-pVTZ (i. e. without extra diffuse functions) against DLPNO-CCSD(T)/cc-pVTZ was carried out for 57 low-energy structures (see Figures S1 and S2). Again this functional was found to be a good compromise between tractability and accuracy.
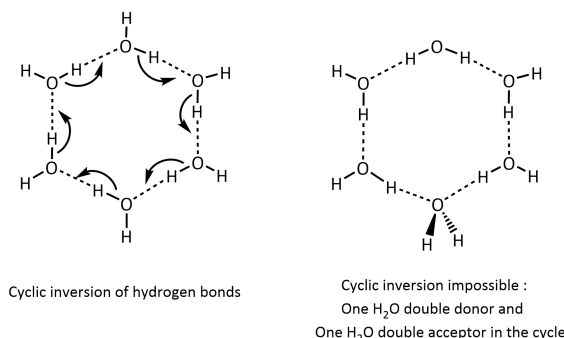
When relatively large numbers of structures of a fluxional system are generated, redundancies must be sorted out and it becomes an issue to decide when two structures are identical. On one hand, identical structures generated by distinct computational paths, e. g. geometry optimizations with different starting points, may have non negligible residual differences in their energies and structures.[66] Thus a similarity criterion should not be too tight. On the other hand, two truly different structures may turn out to have energies within less than 10^{-4} a.u., much below "chemical accuracy". We find that structures having energies within less than 10^{-4} a.u. at three different levels: DFT/cc-pVTZ, DFT/cc-pVTZ including ZPVE and DFT/aug-cc-pVTZ can be safely considered to be identical.

3 Results of PES exploration: energy spectrum for low energy structural families of $SO_4^{2-}(H_2O)_{24}$

The above exploration yielded 12 structures within an energy window of 10 kJ.mol^{-1}. Out of these twelve, three pairs of identical structures were identified. For the resulting 9 structures, relative energies were refined at the CCSD(T) level. The structure found to be the most stable by Johnston et al. was added and confirmed to be the most stable.[30] In agreement with the work of Xantheas et al. on water clusters,[39] a glimpse on the structural diversity may be obtained by calculating the numbers of 4-, 5- and 6-membered cycles made by water molecules, excluding sulfate (see section 4). This showed several identical sets of numbers ; looking more closely into these cases, it appears that some of these structures correspond to different water network skeletons, while others differ only by simultaneous inversion of hydrogen bonds along a cycle such as the one shown in Scheme 1 for a 6-membered cycle (Note that the inversion shown in the left part of Scheme 1 yields an identical structure for the cycle itself, but a different structure for the full cluster). The latter type of isomerism has been discussed before for water clusters,[27,67] small hydrated sulfate ions[8] and for several ions embedded in water clusters with n=3 - 50, including sulfate.[30–33] It was therefore deemed important to take this type of isomerism into account in a systematic way. As in reference[30], the NetworkX package[68] was used to generate directed graphs where "directed" refers to the orientation of hydrogen bonds. From all hydrogen-bonded cyclic fragments of any size having all hydrogen bonds in the same direction, i. e. such that their directions can all be inverted simultaneously, the isomers resulting from inversions are generated. In addition, all combinations of such inversions are generated ; all combinations of 2, 3 and 4 inversions were also considered in this work. This was carried out for the seven lowest energy structures previously found, after elimination of identical and inversion-related structures. The values in Table 1 indicate that for $SO_4^{2-}(H_2O)_{24}$, the numbers of combinations generating new structures drop sharply with the combination order, to the extent that there are often no combinations of 4 inversions generating new structures. Given the large number of structures thus generated (more than 770, see Table 1) this was done only for these seven structures. The ensembles of structures generated by ring inversions of cyclic hydrogen-bonding motifs and combinations thereof will be called "structural families" below. With all structures thus generated, geometry optimization

at the ωB97X-D/cc-pVTZ level was carried out, and low energy structures (typically 7 - 10 depending on the family, spanning a range of 3 to 5 kJ.mol^{-1}) were then optimized with tighter convergence criteria and subjected to frequency and improved energy calculations as described above.

Cyclic inversion of hydrogen bonds

Cyclic inversion impossible :
One H$_2$O double donor and
One H$_2$O double acceptor in the cycle

Scheme 1

As can be seen in Table 1 and Figure 1, each family contains 90 to 167 structures, spanning an energy range from 9 to 29 kJ.mol^{-1}. Clearly this facet of isomerism cannot be neglected when trying to identify the low energy structures in this size range (the larger the cluster, the more important this issue is likely to be). It is also possible that interconversion of such structures leads to significant entropy, provided that interconversion barriers are low, favoring families containing larger numbers of structures. This exhaustive study for the 7 lowest energy families of $SO_4^{2-}(H_2O)_{24}$ generated 772 structures to characterize at the DFT level.

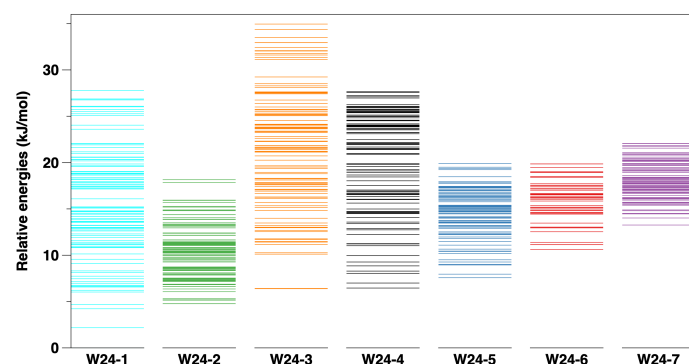**Fig. 1** Energy spectrum of $SO_4^{2-}(H_2O)_{24}$ structures for the seven most stable structural families. Energies at the ωB97X-D/cc-pVTZ level correspond to geometries optimized at the same level, not including ZPE corrections.

The results described thus far are for electronic energies at 0 K. In order to check the impact of ZPE and thermal corrections on relative energetics, energies at 0 K and free energies at 298 K are compared in Figure S3 for a set of 57 low-energy structures. It can be seen that thermal corrections do not introduce large changes, thus 0 K results will be used hereafter.

4 Low energy structures

In this section we describe the lowest energy structures found for $SO_4^{2-}(H_2O)_{24}$. MD simulations generated a large number of structures, out of which more than 100 were fully optimized at the DFT level using the protocol described in section 2.3. An energy window of 10 kJ.mol^{-1} was chosen to analyze structures in more details. While this is somewhat arbitrary, it allows getting a picture of the structural diversity which exists even within a narrow energy segment. This was followed by the hydrogen-bond permutation study detailed in section 3. Structures in this size range become rather complex and except for **W24-2** and to some extent **W24-4**, no symmetry property is available to reduce complexity. Thus an understanding of structural features must be based on structural descriptors, which are described in the next subsection.

4.1 Structural descriptors

In order to describe structures at this level of complexity, it is necessary to use criteria which can help distinguishing between structural families. A simple and common descriptor is the gyration radius. We found however that it was not discriminant enough (between 3.96 and 4.02 Å for the seven families, see Table 2) to be useful for such clusters. Rotational constants or moments of inertia were also considered. Although the structural differences do lead to significant variation (up to ca. 10%) of rotational constants, a limited sampling showed cases where structures of different families had sets of rotational constants which were not more different from one another than compared to other structures of the same family. More significant are differences of structural anisotropy between families. Thus the largest difference between the three principal moments of inertia was used as a simple index of structural compactness (see Table 2).

The numbers of reversible cycles and their combinations described in section 3.2 and Table 1 make, in principle, a set of quite specific descriptors. They vary significantly within a structural family however, while we need to distinguish between families. Thus we used instead the numbers of cycles irrespective of their reversible or non-reversible character. They were defined using the numbers of 4-, 5-, 6- and 7-H$_2$O rings they contain. Here an n–ring is a ring containing n hydrogen-bonded water oxygens, regardless of the orientation of the hydrogen bonds connecting them. Such ring counting has previously been used to characterize water clusters.[39] The values computed here (see Table 2) are much smaller than those found for $(H_2O)_{24}$ which are in the 5-17 range,[39] due to the large restructuration of the water cluster through introduction of sulfate inside. The hydrogen-bond network may be characterized by the AADD (doubly hydrogen-bond accepting and doubly hydrogen-bond donating), ADD, AD... types of water molecules. The total numbers of each type of molecule have also been used as descriptors below. In a conservative way, a hydrogen bond was considered to exist for any non-covalently bonded O···H pair with a distance of less than 2.4 Å, regardless of O-H···O angles. The numbers of dangling O-H bonds (usually in AAD molecules) turned out not to be discriminant here since none of the low energy families has any

Table 1 Numbers of possible inversions across hydrogen-bonded cyclic water networks, and combinations thereof, for the seven most stable structural families of $SO_4^{2-}(H_2O)_{24}$. Energy ranges are in kJ.mol^{-1}. The starting structure must be added to obtain the total numbers of structures.

Family name	W24-1	W24-2	W24-3	W24-4	W24-5	W24-6	W24-7
Number of inversions	21	79	21	24	51	42	41
Number of 2-comb	46	73	46	46	26	37	43
Number of 3-comb	33	15	33	26	12	12	5
Number of 4-comb	10	0	6	5	13	0	0
Total number of structures	111	168	107	102	103	92	90
Energy range of family	27.8	13.4	27.1	21.2	12.3	9.2	8.9

such dangling bonds. Other local motifs will also be used to describe structures below : "cubicity" and "Nc4-Mc6" fragments. We find that approximate cubes, reminiscent of the true cubes found in pure water clusters, are often present in $SO_4^{2-}(H_2O)_{24}$. Cubicity has been introduced previously for water clusters by Hartke et al.[27] as the tendency to form intermolecular OHO angles near 90°, a common feature in water clusters. For hydrated sulfate ions, we find that such motifs are common (see below) however not forming real cubes since sulfate oxygens are often involved. Thus we use "cubicity" as a trend property of structures to exhibit cube-like motifs which can be high, low or absent (see Table 2). Similar trends have already been documented for hydrated, oxygen-containing polyatomic anions including sulfate,[30] perchlorate,[31] chlorate and sulfite,[33] iodate,[6,69] CO_2^-[70] and also hydrated thiocyanate SCN^-.[6,32]

Another commonly found motif for $SO_4^{2-}(H_2O)_{24}$ is a series of fused rings including 4 and 6 water oxygens in alternation. We call these 4-rings and 6-rings, respectively, to distinguish them from usual 4- and 6-membered rings in which hydrogen atoms are counted. In these motifs, 4-rings have a butterfly conformation (thus rather different from those in cubes) and 6-rings are chair-like. We call "Nc4-Mc6" fragments including N 4-rings and M 6-rings in alternation, all grouped by sharing edges and with the specific conformational properties indicated above, providing local surface curvature around the sulfate ion. As described below (see Figure 2), one of the most stable families (**W24-2**) which has high sphericity, is made exclusively of such motifs. Most other structures involve such fragments. A series of python scripts was written to compute and combine the structural descriptors described above.

4.2 Description of low-energy structural families

The lowest energy structure is that published by Johnston et al.[30] It is remarkable that the use of a rather crude force field, when combined with a sophisticated search on the PES, was able to identify the lowest energy structure found so far. In addition, generating the 111 cyclic permutations of hydrogen bonds did lead to this structure being the most stable within its family. It should be noted however that ongoing work on neighboring sizes $(SO_4^{2-}(H_2O)_{23}$ and $SO_4^{2-}(H_2O)_{25}$ using the same methods as described above for $SO_4^{2-}(H_2O)_{24})$, yielded 3 and 2 structures more stable than those in reference[30], respectively. Clearly the identification of global minima in this size range remains an open challenge. The most stable families of structures will now be described in more details.

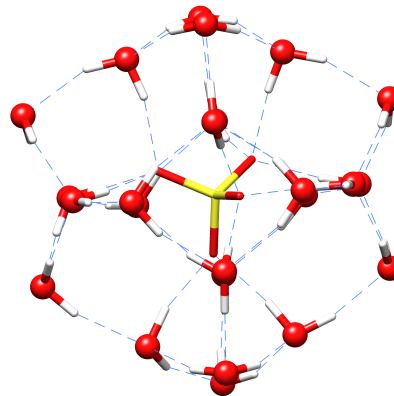**W24-2**. We start with this family since its water network is

Fig. 2 Structure of the **W24-2** family displaying its symmetry. Butterfly-shaped 4-rings are at the top, bottom, left, right and center of this representation.

highly symmetrical, exclusively made of alternating and fused six 4- and eight 6-rings (counting O's only). The most stable structure of this family is shown in Figure 2. All 4-rings have a butterfly shape while 6-rings are chair-like. Such ring conformations generate a network with the appropriate curvature to make a "spherical" cage around sulfate with 12 coordination sites of its tetrahedron in the first sphere. All rings have alternating 1st and 2nd sphere molecules. All 1st sphere molecules are AADD, D toward a sulfate oxygen and a 2nd sphere one, while all 2nd sphere molecules are ADD, donating to a 1st sphere oxygen and to a 2nd sphere one. The resulting network is highly symmetrical as shown by the small differences between the three moments of inertia (see mean Δmax in Table 2). Within this family, these differences nearly drop to zero for structures in which the orientation of hydrogen bonds is symmetrical.

W24-1 is the most stable family and indeed its 8 most stable structures are the most stable found for $SO_4^{2-}(H_2O)_{24}$ overall, in a 5.2 kJ.mol^{-1} energy window. As shown in Table 2, its distribution of water cycles is rather different from that of **W24-2**, including one 7-ring and five 5-rings. Although it has four 4-cycles and four 6-cycles, it has only a 3c4-3c6 motif. This still amounts to 16 out of the 24 water molecules, i. e. two thirds of **W24-1** are common with **W24-2**, as highlighted by Figure 3 where the 3c4-3c6 motif and the sulfate are shown explicitly, leaving the remaining eight water molecules in dashed lines where a pseudo-cube can be seen. This water skeleton is less symmetrical overall as shown by the higher difference between its moments of inertia. The 7-ring provides a flexible ligand to the 3 coordination

Table 2 Structural descriptors for the seven lowest-energy families of $SO_4^{2-}(H_2O)_{24}$. Mean values of the gyration radius R_{gyr} (Å) and of the maximum difference between moments of inertia (amu.Å^2) are computed for the most stable structures (8 on average) of each family.

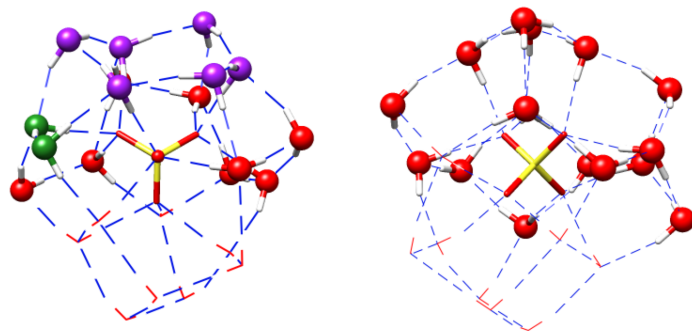| Family name	Mean R_{gyr}	Mean Δmax	4-rings	5-rings	6-rings	7-rings	Cubicity	Nc4-Mc6
W24-1	3.963	742.3	4	5	4	1	Low	3 - 3
W24-2	3.980	120.5	6	0	8	0	None	6 - 8
W24-3	3.981	870.8	4	4	6	0	None	3 - 4
W24-4	3.979	762.2	4	6	3	1	Low	2 - 2
W24-5	3.968	623.8	4	5	4	1	Low	3 - 3
W24-6	4.017	1125.0	5	6	2	2	High	None
W24-7	3.964	517.4	4	5	4	1	Low	3 - 3

Fig. 3 Structure of the **W24**-1 family. The sixteen H$_2$O engaged in a 3c4–3c6 motif are shown explicitly, while the remaining eight water molecules are shown with dashed lines. In the left view, a c4–c6 motif is highlighted with the chair c6 cycle oxygens in purple and the fused butterfly c4 cycle with two oxygens in purple and two in green. The right view highlights the similarity with structure **W24**-2 in Figure 2.

sites of one sulfate oxygen. This flexibility may be the reason for the greater stability of **W24-1** relative to **W24-2**.
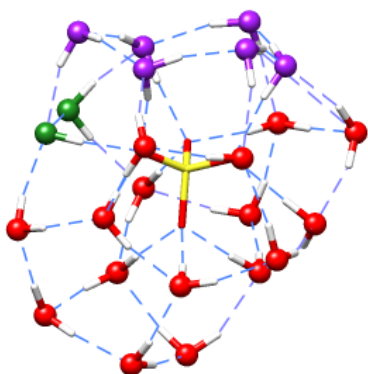

Fig. 4 Structure of the **W24**-3 family. A c4–c6 motif is highlighted with the same color code as for **W24**-1 in Figure 3.

W24-3 is the 3rd most stable family, although it extends across a large energy window and some of its structures are the least stable among the seven families. Figure 1 displays a large gap between the most stable structure and the second, however this disappears when ZPE is taken into account. This family has a 3c4-4c6 network involving 18 out of its 24 water molecules, underlining again the special stability conferred by the c4-c6 motif. Figure 4 shows the specific chair-like conformation of these 6-rings (in purple) and the butterfly conformation of the 4-ring fused to the
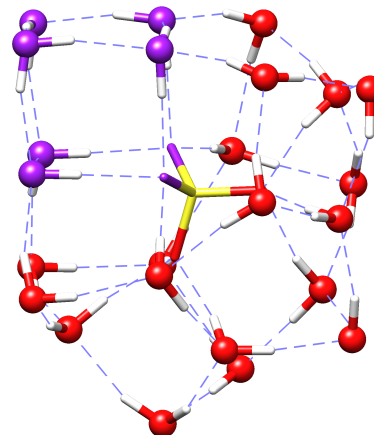
6-ring (in green and purple).

Fig. 5 Structure of the **W24**-4 family. A pseudo cube is illustrated with six water oxygens in purple. The pseudo cube is completed by two sulfate oxygens.

W24-4 is the 4th most stable family. It has a 2c4-2c6 motif which describes half of its structure. In addition it displays another structural motif, a cube involving 6 water and 2 sulfate oxygens as highlighted purple in Figure 5. The 4-rings are planar in this cube, rather different from the butterfly shape described above in **W24-2**. This tendency to form square planar arrays of water molecules within partial cubes is similar to "cubicity"[27] in water clusters. There is no symmetry in this structural family. Again all 1st sphere oxygens are AADD and all 2nd sphere oxygens are ADD.

W24-5 combines a 3c4-3c6 motif (involving 16 out of the 24 H$_2$O) and a cube involving another 4 water molecules (see Figure S4). It is rather similar to **W24-1**, with the same numbers of 4-, 5- and 6-cycles and a 7-ring providing full coordination to a sulfate oxygen. The 3c4-3c6 fragment is partially fused with the 7-ring, however in a way that is different than in **W24-1**. In the latter, the 3c4-3c6 fragment provides full coordination to 2 other sulfate oxygens, while it is only partial in **W24-5**.

W24-7 has strong similarity with **W24-5** since all descriptors are equal or very similar (see Table 2). Its structure involves a cube and a 3c4-3c6 motif partially fused to a 7-ring. The oxygen skeletons of **W24-5** and **W24-7** are nearly superimposable except for one water molecule, due to changes in hydrogen bonding in a local network of 4 molecules. With such small differences, **W24-7** and **W24-5** may be considered as members of a super-family.

W24-6 is the second least stable family and the only one lacking any c4-c6 fragment. It is the least compact of all, with the largest mean gyration radius and the largest difference between moments of inertia. The oxygen network has a symmetry plane (see Figure S5). There are 5 cubes or distorted cubes in this structure including pairs of fused cubes with nearly planar 4-rings. Two 7- and one 8- ring "hold" the structure. Although this family is quite different from the previous (as indicated by the numbers of cycles in Table 2), all 1st sphere oxygens are again AADD and all 2nd sphere are ADD.

The first conclusion of this panorama is that the c4-c6 motif is particularly efficient in shaping the water network tightly around sulfate. Indeed, except for **W24-6**, all families have at least half of their water molecules in such motifs. In addition to making a tightly packed drop, this motif provides appropriate binding sites to the four sulfate oxygens in their tetrahedral positions. The second conclusion is that there remains room for structural diversity, as a variety of 5- and 7-rings can also be accommodated in stable structures, and even some 4- and 6-rings with shapes that are different than in the c4-c6 motif. In all 7 families the twelve 1st sphere water molecules are of the AADD type while the twelve 2nd sphere H_2O are ADD. Thus at this cluster size the first sphere is similar to what is known in aqueous solution, and several possibilities exist with optimum numbers of hydrogen bonds in the partially filled second sphere. However, some structural motifs lead to AADD molecules with oxygen environments strongly away from an ideal tetrahedron, generating some energy penalty.

One may wonder if structures belonging to the same family bear identical or nearly identical properties. One such property is their IR spectrum, which can be recorded experimentally and is, according to the present results, likely to correspond to the sum of spectra of a number of structures. The spectra of the five most stable structures of **W24-1** and **W24-2** are shown in Figure S6 in the 3000-3800 cm^{-1} range. These structures are in an energy range of less than 5 kJ.mol^{-1}. The spectra are roughly similar, both when comparing the two families and within each family. In the latter cases, some noticeable differences exist due to structural changes associated with inversion of hydrogen bond cycles. For both families, there are two major active ranges at 3250-3400 and 3430-3580 cm^{-1}. For **W24-1**, the spectra of structures A and B are very similar while the other three spectra form a second group of similar spectra (C, D and E). The main differences are found near 3150 and 3300 cm^{-1}. In **W24-2**, the spectral differences are slightly easier to appreciate. In the 3200-3400 cm^{-1} range of A, there are three distinct bands that are not present in C for which the spectrum is broader, for example. Similarly in the 3430-3580 cm^{-1} range, there are two groups of bands for which one can see an inversion of intensity for the two most intense peaks at 3450 and 3550 cm^{-1} (A-B vs. C, D, E). However, the delocalisation of the modes makes the attribution of the differences difficult. The main conclusion is that cycle inversions can lead to non-negligible differences in vibrational signatures, both in intensities and in frequency shifts. These differences are of a magnitude similar to those between different families. Thus disregarding this type of isomerism may introduce significant errors in computed IR spectra.

4.3 Isomerization mechanisms

Transition states (TSs) for the interconversion of structures of hydrated sulfate have been considered previously by Johnston et al. based on disconnectivity graphs.[30] For n=9 and 12, it was found that isomers sharing oxygen skeletons and thus differing only by hydrogen bond directionality within rings are separated by relatively large barriers, resulting in a frustrated energy landscape. There was no indication provided on the isomerization mechanisms, which ought to be different from those considered herein since 3-rings are dominant in small clusters. Treatment by a non-reactive force field however implies that only relative translations and rotations of water molecules are involved. The interconversion via cyclic permutations of hydrogen bonds involves making and breaking of covalent O-H bonds. Since this type of isomerism remains relatively unexplored, we have looked into the transition states for cyclic inversion, at the same DFT level used to optimize energy minima.

The large number of low-energy structures and the large number of cycles within each structure made it necessary to choose a sampling set. The sampling can however not be restricted to the lowest energy structure since part of its cycles are not reversible and all types of cycles are not present. The most stable structure of **W24-1** bears a total of 14 cycles (see Table 2), out of which 8 are reversible and transition states were obtained for all eight. This set was complemented by a few cases taken from structures of other families, as described below.

Out of the four 4-cycles of **W24-1**, the three butterflies are reversible in the most stable structure and transition states were obtained for all three, while the square planar one is not. The three energy barriers were computed to be in the 100-106 kJ.mol^{-1} range. All mechanisms are in a single step and all transition states involve OH^- bound to a sulfate oxygen and proton transfer between two water molecules, out of which one is sulfate-bound. A qualitative description of the mechanism is depicted in the scheme in Figure 6. Another butterfly 4-cycle, taken in **W24-3**, led to a nearly identical result. This is also consistent with our previous result on the inversion of a butterfly 4-cycle in the most stable structure of W12 for which an energy barrier of 111 kJ.mol^{-1} was found, albeit at a slightly different DFT level.[8] Two square planar 4-cycles were taken from cubic portions of **W24-6**. Both have two water molecules bound to the sulfate, however they are contiguous while they were not in butterfly 4-rings. Interestingly the two mechanisms are different. One is similar to the butterfly cases above while the other does not involve charge separation, but is still asynchronous and concerted. The computed barriers are smaller, both of 85 kJ.mol^{-1}.

Out of the five 5-cycles of **W24-1**, four are reversible and were considered. They differ by their puckering angle and by the number of sulfate-bound oxygens, one, three or four. In all cases, the single-stepped mechanism involved a transition state which has charge separation with OH^- bound to the sulfate and ei-
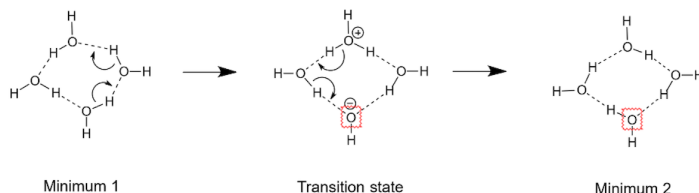
Fig. 6 Mechanism of concerted asynchronous cyclic inversion of hydrogen bonds along a butterfly 4-cycle.

---

ther H_3O^+ not bound, or two water molecules sharing the proton with only one bound. The energy barriers range from 97 to 116 kJ.mol^{-1}. Two other 5-cycles were taken from **W24-6**, in which two non contiguous H_2O are sulfate bound. With TS similar to those above, their computed energy barriers are 103 and 111 kJ.mol^{-1}.

Finally, inversion of a chair-like 6-cycle in the 6c4-8c6 motif of **W24-2** was studied. The transition state is very similar to that shown in Figure 6 for butterfly 4-cycles, with two water molecules separating the ions on each side instead of one. An energy barrier of 134 kJ.mol^{-1} was obtained.

In summary, all mechanisms are concerted, albeit asynchronous, in a single step. The associated transition states often involve OH^- bound to the sulfate and proton transfer between two water molecules that are distal from the OH^-. Sometimes the proton is localized in the TS, in which case the transient H_3O^+ is not bound to the sulfate. Square planar cycles involve slightly different, less asynchronous mechanisms. Activation barriers lie in the 97-116 kJ.mol^{-1} range, except for the square planar 4-rings for which they are only slightly smaller (85 kJ.mol^{-1}). The conclusion is that concerted inversion of rings, irrespective of their size, is a highly demanding process.

There are other possible mechanisms for inversion of hydrogen bonds along a cyclic fragment in a cluster. A track was opened by Hey et al. in the case of $ClO_3^-(H_2O)_8$.[33] The two most stable isomers found differ by inversion along a 4-cycle and their interconversion was suggested to follow a three-step mechanism which only involves relative motions of intact water molecules, rather than exchange of hydrogens between water molecules as considered herein. The three steps involve breaking hydrogen bonds and forming others. The second step was found to be rate-limiting with a barrier of 44 kJ.mol^{-1}. The entire path was recomputed at the ωB97X-D/cc-pVTZ level with full geometry reoptimisation and vibrational frequency calculations. The energy profile was found to be remarkably similar to that obtained at the force field level, with a rate-limiting barrier of 55 kJ.mol^{-1}. The one-step inversion of hydrogen bonds within the 4-cycle was also computed and the barrier was found to be 85 kJ.mol^{-1}. Thus at least in this case, the latter type of mechanism is not the most favorable. Quantum tunnelling effects described in small water clusters do not appear to be significant for such cyclic inversions.[71] More work is needed to ascertain the feasibility of isomerization between structures within a family, especially in closed structures such as the ones described herein for $SO_4^{2-}(H_2O)_{24}$.

5 Magic number clusters for hydrated anions

There are many cases of hydrated ions for which mass spectra display significant variations of intensity for clusters of neighboring sizes. The sizes with higher intensities are termed "magic numbers" although intensity differences are very large in only a few cases. The best-established magic number clusters (MNCs) are those made by singly charged cations $X(H_2O)_{20}$, X = H_3O^+, NH_4^+, Cs^+, Rb^+ and K^+.[9,37,38,72,73] They are due to the special stability of pentagonal dodecahedral cage structures of $(H_2O)_{20}$ bearing a central cavity which fits large alkali ions and small polyatomic cations. Protonated water and bulkier ions, e. g. protonated alkylamines,[73] can sit on the water cluster surface while there is no MNC for the smallest atomic ions (Li^+ and Na^+). Such structures are not expected to fit anions which would lie next to water oxygens and in fact, hydrated anions do not appear to display MNCs in this size range. In contrast they do for larger clusters, including $F^-(H_2O)_n$ for n=54 and 62, $Cl^-(H_2O)_n$, $Br^-(H_2O)_n$ and $I^-(H_2O)_n$ for n=49.[69] In the last 3 cases, intensity alternation is noticeable in the size range above n=49. For polyatomic ions, MNCs have been reported in some cases, however at different sizes : $Fe(CN)_6^{4-}(H_2O)_n$ for n=46 and 64,[9] $CO_2^{\circ-}(H_2O)_n$ for n=49,[70] $O^{\circ-}$ n=55.[74] In other cases such as $IO_3^-(H_2O)_n$, no such special cases appear.[69] $SO_4^{2-}(H_2O)_n$ also displays MNCs in this size range, for n=56 and 58. However in sharp contrast to all previous cases, it also shows MNCs for n=24 and 36, and to some extent for n=39.[75,76]

Sulfate appears to be distinct with its MNCs for n=24 and 36 and to some extent for n=12.[14] While it is tempting to consider that $SO_4^{2-}(H_2O)_{12}$ is somewhat more stable than its neighbours because it has a filled first hydration shell in analogy with what is known in the bulk, there has been some debate over the size at which the first shell is actually closed in the gas phase : 12, 13 or even more water molecules.[8,10,20,77] We have shown that experimental data may actually be interpreted with the most stable structure in which the sulfate ion is not fully encaged. The case is even more striking for n=24 which has a more pronounced MNC, yet we find a series of structural families whose most stable structures are within a few kJ.mol^{-1} of one another. Only one of these, **W24-2**, is structurally special enough to suggest an interpretation of the MNC of a single family, however it is only the second most stable. We conclude that n=24 is special because it allows a variety of structures to be "closed" in the sense that they have no dangling bond, twelve water molecules in the first hydration shell and twelve in the second ; twelve water molecules with AADD coordination and twelve with ADD. This does not hold for the neighboring sizes n=23 and 25 (work in progress).

We may further speculate on the existence of MNCs at n=24 with similar cage structures for other anions, if structures analogous to those found for $SO_4^{2-}(H_2O)_{24}$ also have special stability. This may be expected to occur for other tetragonal oxo anions XO_4^{n-} since coordination of each anion oxygen by three O-H bonds would be secured. This may hold true if ion-water interactions are not too different from those with sulfate, because of either the nature of the central atom or the total charge or both. It is difficult to predict how flexible the cage structure

is when it comes to hosting ions of different sizes and/or charges. In order to test this idea, a rapid survey was made by changing $SO_4^{2-}(H_2O)_{24}$ into $FO_4^-(H_2O)_{24}$, $PO_4^{3-}(H_2O)_{24}$, $ClO_4^-(H_2O)_{24}$, $MnO_4^-(H_2O)_{24}$, $SeO_4^-(H_2O)_{24}$, and $BrO_4^-(H_2O)_{24}$ in **W24-1** (the most stable for sulfate), **W24-2** (the full 6c4-8c6) and **W24-6** (the most cubic with no c4-c6) and optimizing their geometries at the ωB97X-D/cc-pVTZ level. In all cases except for $FO_4^-(H_2O)_{24}$, the overall structures were maintained during optimization. However in a number of cases, one or two hydrogen bonds between water molecules and the ion were weakened or broken with local restructuration of the water network and even sometimes a dangling O-H bond appearing (e.g. Figure S7). The only cases for which structures were maintained for the three isomers and geometry optimizations were straightforward are $PO_4^{3-}(H_2O)_{24}$ and $SeO_4^-(H_2O)_{24}$. These results indicate that these ions would be interesting candidates for mass spectrometric study of ion abundances as a function of cluster size, to search for possible MNCs in this size range.

6   Conclusions

Low energy structures of $SO_4^{2-}(H_2O)_{24}$ have been searched for in two steps, by first exploring the potential energy surface with classical molecular dynamics simulations using the polarizable AMOEBA force field. Ca. one hundred low energy structures were then refined with DFT calculations. The accuracy of the ωB97X-D functional associated with the cc-pVTZ and aug-cc-pVTZ basis sets has been demonstrated by extensive calibration on relative energetics at the DLPNO-CCSD(T) level. It turned out that the structures found took only little account of isomerism due to inversion of hydrogen bond networks along cycling arrays of water molecules. All such isomers were then obtained in a further step for the seven most stable structures previously found, using a systematic generation software. For these large families, each comprising between 90 and 167 isomers, geometry reoptimization at the DFT level showed that energy differences within a family can be as large as 30 kJ.mol^{-1}. These results demonstrate that this relatively ignored isomerism must be taken into account to identify reliably the lowest energy minima.

In order to describe the manifolds of structures thus generated, a number of possible descriptors were considered. Those related to the numbers of 4-, 5- and 6-cycles of water molecules turned out to be the most useful, as low energy structures were shown to bear common motifs. The most prominent motif is made of fused cycles involving alternatively four and six water molecules. The latter adopt specific conformations which ensure the appropriate surface curvature to form a closed cage and provide 12-coordination of the sulfate ion as in the bulk. An unexpected result is that there exists a whole variety of structure types which are able to provide this stable environment to the sulfate with nearly equal efficiency. Thus the magic number $SO_4^{2-}(H_2O)_{24}$ is not associated with a single, remarkable structure but rather to a whole series of structural families. Extensive calculations on isomerization mechanisms in 4-, 5- and 6-cycles indicate that large barriers of ca. 100 kJ.mol^{-1} are associated with direct inversion

of hydrogen bond networks in a single step. This would suggest that these families may not generate large entropies due to easy interconversion. It remains however to explore other paths which would only involve relative motions of intact water molecules. Mass spectra of other small anions do not display magic number clusters for n=24 or even in this size range. Exploratory calculations suggest that MNCs may exist at n=24 for other tetraoxo anions.

Supporting Information

Calibration of quantum chemical relative energies, structures of the **W24-5** and **W24-6** families, IR spectra of the most stable structures of **W24-1** and **W24-2**, a structure of $MnO_4^-(H_2O)_{24}$ derived from **W24-6** of sulfate, and Cartesian coordinates of most stable structure of each $W24-n$ family (n=1 to 7) are provided as supplementary files associated with this article.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 L. G. Tallgren, *Acta Med Scand Suppl*, 1980, **640**, 1–100.

2 A. Lee, P. A. Dawson and D. Markovich, *Int. J. Biochem. Cell Biol.*, 2005, **37**, 1350–1356.

3 V. Ramanathan, P. J. Crutzen, J. T. Kiehl and D. Rosenfeld, *Science*, 2001, **294**, 2119.

4 X. Wang and L. Wang, *Annu. Rev. Phys. Chem.*, 2009, **60**, 105–26.

5 L.-S. Wang, *J. Chem. Phys.*, 2015, **143**, 040901.

6 X.-B. Wang, *J. Phys. Chem. A*, 2017, **121**, 1389–1401.

7 K. R. Asmis and D. M. Neumark, *Acc. Chem. Res.*, 2012, **45**, 43–52.

8 F. Thaunay, A. A. Hassan, R. J. Cooper, E. R. Williams, C. Clavaguéra and G. Ohanessian, *Int. J. Mass spectrom.*, 2017, **418**, 15–23.

9 S. Heiles, R. J. Cooper, M. J. DiTucci and E. R. Williams, *Chem. Sci.*, 2017, **8**, 2973–2982.

10 J. Zhou, G. Santambrogio, M. Brümmer, D. T. Moore, L. Wöste, G. Meijer, D. M. Neumark and K. R. Asmis, *J. Chem. Phys.*, 2006, **125**, 111102.

11 X. Yang, X.-B. Wang and L.-S. Wang, *J. Phys. Chem. A*, 2002, **106**, 7607–7616.

12 X.-B. Wang, A. P. Sergeeva, J. Yang, X.-P. Xing, A. I. Boldyrev and L.-S. Wang, *J. Phys. Chem. A*, 2009, **113**, 5567–5576.

13 A. T. Blades and P. Kebarle, *J. Phys. Chem. A*, 2005, **109**, 8293–8298.

14 R. L. Wong and E. R. Williams, *J. Phys. Chem. A*, 2003, **107**, 10976–10983.

15 H. Knorke, H. Li, Z.-F. Liu and K. R. Asmis, *Phys. Chem. Chem. Phys.*, 2019, **21**, 11651–11659.

16 J. T. O'Brien, J. S. Prell, M. F. Bush and E. R. Williams, *J. Am. Chem. Soc.*, 2010, **132**, 8248–8249.

17 D. S. Lambrecht, G. N. I. Clark, T. Head-Gordon and M. Head-Gordon, *J. Phys. Chem. A*, 2011, **115**, 11438–11454.

18 D. S. Lambrecht, L. McCaslin, S. S. Xantheas, E. Epifanovsky and M. Head-Gordon, *Mol. Phys.*, 2012, **110**, 2513–2521.

19 H. R. Leverentz, H. W. Qi and D. G. Truhlar, *J. Chem. Theory Comput.*, 2013, **9**, 995–1006.

20 Q. Wan, L. Spanu and G. Galli, *J. Phys. Chem. B*, 2012, **116**, 9460–9466.

21 C. I. León-Pimentel, J. I. Amaro-Estrada, J. Hernández-Cobos, H. Saint-Martin and A. Ramírez-Solís, *J. Chem. Phys.*, 2018, **148**, 144307.

22 V. Brites, A. Cimas, R. Spezia, N. Sieffert, J. M. Lisy and M.-P. Gaigeot, *J. Chem. Theory Comput.*, 2015, **11**, 871–883.

23 D. J. Wales, *Energy landscapes*, Cambridge University Press, Cambridge, UK, 2003.

24 T. James, D. J. Wales and J. Hernández-Rojas, *Chem. Phys. Lett.*, 2005, **415**, 302–307.

25 K. Korchagina, A. Simon, M. Rapacioli, F. Spiegelman, J.-M. L'Hermite, I. Braud, S. Zamith and J. Cuny, *Phys. Chem. Chem. Phys.*, 2017, **19**, 27288–27298.

26 A. Monjaraz-Rodríguez, M. Rodriguez-Bautista, J. Garza, R. A. Zubillaga and R. Vargas, *J. Mol. Model.*, 2018, **24**, 187.

27 U. Buck, C. C. Pradzynski, T. Zeuch, J. M. Dieterich and B. Hartke, *Phys. Chem. Chem. Phys.*, 2014, **16**, 6859–6871.

28 D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley Longman Publishing Co., Inc., 1989.

29 K. A. Korchagina, A. Simon, M. Rapacioli, F. Spiegelman and J. Cuny, *J. Phys. Chem. A*, 2016, **120**, 9089–9100.

30 L. C. Smeeton, J. D. Farrell, M. T. Oakley, D. J. Wales and R. L. Johnston, *J. Chem. Theory Comput.*, 2015, **11**, 2377–2384.

31 J. C. Hey, L. C. Smeeton, M. T. Oakley and R. L. Johnston, *J. Phys. Chem. A*, 2016, **120**, 4008–4015.

32 L. Smeeton, J. Hey and R. Johnston, *Inorganics*, 2017, **5**, 20.

33 J. C. Hey, E. J. Doyle, Y. Chen and R. L. Johnston, *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.*, 2018, **376**, 20170154.

34 S. Imoto, S. S. Xantheas and S. Saito, *J. Chem. Phys.*, 2013, **138**, 054506.

35 P. Ren, C. Wu and J. W. Ponder, *J. Chem. Theory Comput.*, 2011, **7**, 3143–3161.

36 C. K. Egan and F. Paesani, *J. Chem. Theory Comput.*, 2019, **15**, 4816–4833.

37 Q. Yu and J. M. Bowman, *J. Phys. Chem. A*, 2020, **124**, 1167–1175.

38 S. S. Xantheas, *Can. J. Chem. Eng.*, 2012, **90**, 843–851.

39 J. A. Bilbrey, J. P. Heindel, M. Schram, P. Bandyopadhyay, S. S. Xantheas and S. Choudhury, *J. Chem. Phys.*, 2020, **153**, 024302.

40 Y. Shi, C. Wu, J. W. Ponder and P. Ren, *J. Comput. Chem.*, 2011, **32**, 967–977.

41 F. Thaunay, G. Ohanessian and C. Clavaguéra, *Chem. Phys. Lett.*, 2017, **671**, 131–137.

42 H. Knorke, H. Li, J. Warneke, Z.-F. Liu and K. R. Asmis, *Phys. Chem. Chem. Phys.*, 2020, **22**, 27732–27745.

43 D. Semrouni, A. Sharma, J.-P. Dognon, G. Ohanessian and C. Clavaguéra, *J. Chem. Theory Comput.*, 2014, **10**, 3190–3199.

44 F. Thaunay, C. Jana, C. Clavaguéra and G. Ohanessian, *J. Phys. Chem. A*, 2018, **122**, 832–842.

45 F. Thaunay, F. Calvo, E. Nicol, G. Ohanessian and C. Clavaguéra, *ChemPhysChem*, 2019, **20**, 803–814.

46 C. Jana, G. Ohanessian and C. Clavaguéra, *Theor Chem Acc*, 2016, **135**, 141.

47 A. Grossfield, P. Ren and J. W. Ponder, *J. Am. Chem. Soc.*, 2003, **125**, 15671–15682.

48 J.-P. Piquemal, L. Perera, G. A. Cisneros, P. Ren, L. G. Pedersen and T. A. Darden, *J. Chem. Phys.*, 2006, **125**, 054511.

49 J. C. Wu, J.-P. Piquemal, R. Chaudret, P. Reinhardt and P. Ren, *J. Chem. Theo. Comput.*, 2010, **6**, 2059–2070.

50 D. Semrouni, W. C. Isley, C. Clavaguéra, J.-P. Dognon, C. J. Cramer and L. Gagliardi, *J. Chem. Theo. Comput.*, 2013, **9**, 3062–3071.

51 M. Kumar, T. Simonson, G. Ohanessian and C. Clavaguéra, *ChemPhysChem*, 2015, **16**, 658–665.

52 A. Marjolin, C. Gourlaouen, C. Clavaguéra, P. Y. Ren, J. C. Wu, N. Gresh, J.-P. Dognon and J.-P. Piquemal, *Theor. Chem. Acc.*, 2012, **131**, 1198.

53 P. Ren and J. W. Ponder, *J. Comput. Chem.*, 2002, **23**, 1497–1506.

54 J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, M. Head-Gordon, G. N. I. Clark, M. E. Johnson and T. Head-Gordon, *J. Phys. Chem. B*, 2010, **114**, 2549–2564.

55 T. A. Halgren, *J. Am. Chem. Soc.*, 1992, **114**, 7827–7843.

56 N. L. Allinger, Y. H. Yuh and J. H. Lii, *J. Am. Chem. Soc.*, 1989, **111**, 8551–8566.

57 M. L. Laury, L.-P. Wang, V. S. Pande, T. Head-Gordon and J. W. Ponder, *J. Phys. Chem. B*, 2015, **119**, 9423–9437.

58 A. Stone, *Chem. Phys. Lett.*, 1981, **83**, 233–239.

59 A. Stone and M. Alderton, *Mol. Phys.*, 1985, **56**, 1047–1064.

60 J. W. Ponder, *TINKER - Software Tools for Molecular Design (version 6)*, 2014, http://dasher.wustl.edu/tinker.

61 R. M. Young, R. J. Azar, M. A. Yandell, S. B. King, M. Head-Gordon and D. M. Neumark, *Mol. Phys.*, 2012, **110**, 1787–1799.

62 Z. Yang, S. Hua, W. Hua and S. Li, *J. Phys. Chem. A*, 2010, **114**, 9253–9261.

63 N. Mardirossian, D. S. Lambrecht, L. McCaslin, S. S. Xantheas and M. Head-Gordon, *J. Chem. Theory Comput.*, 2013, **9**, 1368–1380.

64 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao,

H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision D.01*, Gaussian Inc. Wallingford CT 2009.

65 F. Neese, *WIREs Comput Mol Sci*, 2011, **2**, 73–78.

66 K. Joshi, D. Semrouni, G. Ohanessian and C. Clavaguéra, *J. Phys. Chem. B*, 2011, **116**, 483–490.

67 H. Takeuchi, *J. Chem. Inf. Model.*, 2008, **48**, 2226–2233.

68 A. A. Hagberg, D. A. Schult and P. J. Swart, Proceedings of the 7th Python in Science Conference, Pasadena, CA USA, 2008, pp. 11 – 15.

69 S. Chakrabarty and E. R. Williams, *Phys. Chem. Chem. Phys.*, 2016, **18**, 25483–25490.

70 A. Herburger, M. Ončák, C.-K. Siu, E. G. Demissie, J. Heller, W. K. Tang and M. K. Beyer, *Chem. Eur. J.*, 2019, **25**, 10165–10171.

71 M. T. Cvitaš and J. O. Richardson, *Phys. Chem. Chem. Phys.*, 2020, **22**, 1035–1044.

72 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.

73 T. M. Chang, R. J. Cooper and E. R. Williams, *J. Am. Chem. Soc.*, 2013, **135**, 14821–14830.

74 J. Lengyel, M. Ončák, A. Herburger, C. van der Linde and M. K. Beyer, *Phys. Chem. Chem. Phys.*, 2017, **19**, 25346–25351.

75 M. J. DiTucci, C. N. Stachl and E. R. Williams, *Chem. Sci.*, 2018, **9**, 3970–3977.

76 C. N. Stachl and E. R. Williams, *J. Phys. Chem. Lett.*, 2020, **11**, 6127–6132.

77 F. Thaunay, C. Clavaguéra and G. Ohanessian, *Phys. Chem. Chem. Phys.*, 2015, **17**, 25935–25945.