Special topic

Misako Aida* and Dai Akase

Hydrogen-bond pattern to characterize water network

https://doi.org/10.1515/pac-2018-0721

Abstract: Hydrogen-bond (HB) patterns correspond to topologically distinct isomers of water clusters, and can be expressed by digraphs. The HB pattern is used to divide the configuration space of water cluster at a finite temperature. The populations of the HB patterns are transformed into the relative Helmholtz energies. The method is based on the combination of molecular simulation with graph theory. At a finite temperature it can be observed that other isomers than local minimum structures on the potential energy surface are highly populated. The dipole moment of a constituent molecule in a water cluster is enhanced depending on the local HB network around the water molecule. Rooted digraph is used to represent topologically distinct isomers of protonated water (PW) clusters. O–H bonds of PW clusters are classified into 10 topological types based on the combination of the local HB types of the contributing water molecules to the O–H bond. If the topological type is the same, vibrational frequencies of those O–H bonds of PW clusters are similar even in different isomers; i.e. vibrational frequency of O–H bond is transferable, and can be used as a vibrational spectral signature of PW clusters.

Keywords: digraph; dipole moment of water; Distinguished Women in Chemistry and Chemical Engineering; HB pattern; O–H vibrational frequency; topology of network.

Introduction

The connection between chemistry and graph theory has begun in the 19th century: Cayley used graph theory to solve a problem concerning the number of isomers of alkanes [1]. He introduced trees to represent carbon frameworks of hydrocarbons. This work initiated the development of isomer enumeration methods [2]. And then in 1947, Wiener investigated the relationship of boiling points of substituted alkanes with structural arrangements of the constituent atoms [3]. In 1971, Hosoya showed the correlation between the boiling point of saturated hydrocarbons and the topological index (Hosoya index) [4, 5]. Their investigations indicated that we can predict a molecular property based on the topological structure, i.e. the connectivity of the components, of a chemical compound. It is now the age of big data: the concept of topological property will play more important role in chemistry [6–9], to grasp the gist from the large amount of data.

Here, we concentrate on the hydrogen-bond (HB) network in water cluster and in protonated water (PW) cluster. An HB network is also a target of graph theoretical approaches. A digraph rather than a graph (undirected graph) is appropriate to represent an HB network, because a digraph expresses not only the connectivity but also the direction of an edge: the direction proceeds from the proton donor to the proton acceptor between the hydrogen-bonded water molecules. The stability of a water cluster is dominated by not only the number of hydrogen bonds but also the HB network. The strength of a hydrogen bond in a water cluster

***Corresponding author: Misako Aida,** Department of Chemistry, Graduate School of Science, Hiroshima University, Higashi-Hiroshima, Japan; and Center for Quantum Life Sciences, Hiroshima University, Higashi-Hiroshima, Japan, e-mail: maida@hiroshima-u.ac.jp. https://orcid.org/0000-0001-8788-1071
Dai Akase: Department of Chemistry, Graduate School of Science, Hiroshima University, Higashi-Hiroshima, Japan; and Center for Quantum Life Sciences, Hiroshima University, Higashi-Hiroshima, Japan. https://orcid.org/0000-0002-4464-8808

© 2019 IUPAC & De Gruyter. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. For more information, please visit: http://creativecommons.org/licenses/by-nc-nd/4.0/

depends on the local HB network [9]. In other words, the cooperative effect of hydrogen bonds is included inherently in the HB network.

It is nowadays a routine work to obtain local minimum structures of a molecule of rather large size or a molecular assembly because of the development of computers and computer codes of quantum chemistry. And also, it is now a routine work to run a very long simulation [such as Monte Carlo (MC) sampling or molecular dynamics (MD)] for a large system such as aqueous solution or biological system at a finite temperature, because of the development of computers and computer codes of molecular simulation. In the former case (quantum chemistry), once a local minimum structure on a potential energy surface is obtained, various properties, such as IR spectra, NMR chemical shift, and so on, and also the thermodynamic quantities, including Gibbs energy, for the local minimum structure can be readily calculated on the basis of theoretical equations with several approximations. In the latter case (molecular simulation), a large amount of calculated data are used to obtain average values or transformed to thermodynamic quantities.

For a rigid molecule, a local minimum structure can be regarded as an averaged structure which is observed at a finite temperature. It may not be the case for a weakly-bound cluster, in which the interaction energy connecting monomers is of the order of kinetic energy. A molecular simulation technique (such as MC or MD) is used to generate structures of molecular clusters which are observed at a finite temperature. In the case of weakly-bound cluster, every structure which is generated in a molecular simulation is different, and it is not straightforward to categorize the generated structures to subgroups (i.e. isomers). A concept is needed to categorize structures rationally, which can grasp not only local minima but also many other structures. To this end, we propose to define a set of topological patterns to distinguish topologically distinct isomers. The condition for a set of topological patterns to be used to divide the configuration space into subspaces is that the set of topological patterns satisfies all the following four properties: (i) every pattern is different, (ii) a structure is categorized to only one pattern, (iii) any one of all possible structures can be categorized to one of all the patterns, and (iv) the total number of the topological patterns (the number of subspaces) is known.

Water cluster is a representative of weakly-bound clusters; the local minimum structures of water clusters have been the targets of various high-level quantum chemical calculations [10]. HB interaction energy is rather weak and of the order of kinetic energy; it may happen that other isomers than local minimum structures are observed at a finite temperature. Here in this article, we treat water clusters to demonstrate the importance of the new concept which categorize structures from the topological point of view. In the case of water cluster, a set of topological patterns which satisfies the above condition is the HB pattern. The HB patterns can be used to divide the configuration space of water clusters with the same number of water molecules into subspaces. Furthermore, we show that each of the constituent water molecules can be categorized to one of the HB types according to the local HB network around the water molecule, which is related to the enhancement of dipole moment of each water molecule.

In the case of PW cluster, we show that O–H bonds can be classified into 10 topological types based on the HB types of the water molecules; the O–H vibrational frequency can be rationalized according to the classification.

Theoretical background to graphs

First, we give a brief introduction of graph, digraph and rooted digraph, together with adjacency matrix and HB matrix. We refer only to some basic terms of graphs [11].

Graph and digraph

A graph is a set of vertices and edges. Vertices are often called nodes or points, and edges are also called lines. The edge donates the connectivity between two vertices. Vertices and edges may be distinguishable by labels. Graphs are widely used for schematic representations (Fig. 1). In chemistry, a chemical formula is described

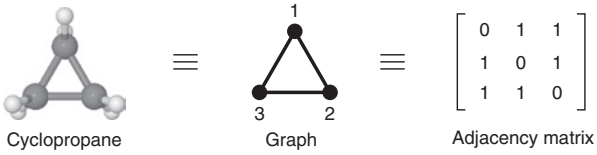


Fig. 1: A graph representation of cyclopropane and the corresponding adjacency matrix.

as a graph whose vertices represent atoms and edges represent chemical bonds: labels of vertices donate elements symbols and labels of edges mean kinds of bonds such as single bonds, double bonds, aromatic bonds, and so on.

Graphs can be represented by various ways. One convenient way is to use an adjacency matrix. A graph in which edges are not directed is called an undirected graph. The adjacency matrix of an undirected graph with n vertices is the n^{th} order square matrix, whose element a_{ij} is equal to 1, if there is an edge between vertex i and vertex j, and 0 otherwise. Hence the matrix is symmetric.

A graph is a mathematical structure which represents the topology of a molecule. Molecular structures can be represented with graphs, and each of different graphs corresponds to a different adjacency matrix. By imposing the conditions which need to be satisfied, all possible adjacency matrices can be created, and then all possible graphs can be obtained, thus all possible topologically distinct structures can be obtained. In the case of hydrocarbon, the graph must satisfy the following two conditions: (i) the maximum number of the edges connected to a vertex is 4, (ii) the graph is connected. All possible isomers with three carbon atoms can be obtained by counting up all possible graphs of $n=3$ with the imposed conditions satisfied.

A graph invariant is a property of graph. The number of vertices and the number of edges are two simple graph invariants. The Hosoya index [4] is also a kind of graph invariant. If two graphs are equivalent, the graphs are called isomorphic. Isomorphic graphs have the same graph invariants.

The most important feature of hydrogen bond is that it possesses the direction. An HB network can be represented by a digraph. A digraph (or directed graph) is a graph whose edges are directed (Fig. 2). In this case, vertices represent components forming hydrogen bonds, and arcs stand for hydrogen bonds from hydrogen donors to accepters.

A digraph contains more information than a graph, concerning the relation between two vertices, i.e. the direction from a vertex toward another vertex. A directed edge is called an arc. An HB network indicates how water molecules are connected. Since a water molecule possesses two hydrogen donors and two hydrogen acceptors, it can form up to four hydrogen bonds.

Radhakrishnan and Herndon [12] in 1991 used digraphs to represent hydrogen-bonded water clusters, and the interaction energy of water clusters were fitted as a function of the graph invariants. The correlation was very good if proper graph invariants were selected for the analyses. The selected graph invariants include not only the number of arcs, which represents the number of the hydrogen bonds formed in the cluster, but also information about the network, e.g. the numbers of paths of different lengths, the number of three- and four-membered rings, and the total number of double donation or acceptance of arcs at each vertex. This result indicates that the cooperativity of hydrogen bonds is inherently associated with the HB network.

Graph theoretical techniques were used to generate stable arrangements of water clusters by McDonald et al. [13] in 1998. Graph invariants were used to present the features of water clusters by Kuo et al. [14] in

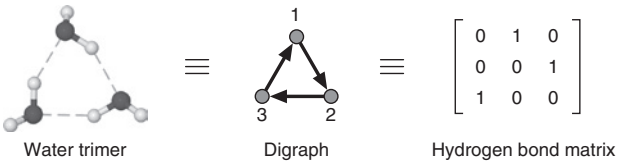


Fig. 2: A graph representation of water trimer and the corresponding hydrogen bond matrix.

2001, based on the bond variables b_{ij}, in which b_{ij} is equal to 1, if water i donates to water j, –1 if water j donates to water i, and 0 if there is no hydrogen bond between i and j. Physical properties of water clusters are fit to some graph invariants [14].

Digraph and HB matrix, and enumeration of HB patterns

An important aspect of the matrix representation of digraph is that it allows efficient enumeration of water clusters. In order to enumerate all possible digraphs corresponding to water clusters and also to distinguish topologically distinct isomers, a special adjacency matrix was defined by Miyake and Aida [15], where the element h_{ij} of the adjacency matrix is 1 if, and only if, the arc from vertex i toward vertex j exists. This matrix for digraph was named "hydrogen bond (HB) matrix" (Fig. 2).

A digraph satisfying the following three conditions is considered as representing a water cluster: (i) maximum number of arcs directed toward a vertex is 2 and maximum number of arcs directed from a vertex is 2, (ii) up to one arc is allowed between any two vertices, (iii) the digraph must be connected. A digraph which satisfies these conditions corresponds to HB network. Miyake and Aida [15, 16] enumerated the number of the digraph satisfying the above conditions by means of adjacency and HB matrices up to eight vertices. Vukičević et al. [17] extended up to 12 vertices using a more sophisticated algorithm. Then, a different algorithm was created to generate up to 16 vertices by Brinkmann [18]. The numbers guarantee the upper limits of the HB patterns of water clusters; in other words, there are no other topologically distinguishable isomers.

All possible graphs and digraphs which correspond to water clusters, for $n = 3$–5, are given in [16]. Here, all possible graphs and all possible digraphs with three vertices are shown in Fig. 3. The number of the graphs is two (one linear and one cyclic). The graphs with $n = 3$ correspond to all possible isomers of saturated hydrocarbon with three carbon atoms. Actually, the linear isomer is propane and the cyclic isomer is cyclopropane. The digraphs with $n = 3$ correspond to water trimer, and the number of those digraphs is 5 (3 linear and 2 cyclic). The linear digraphs are created from the linear graph. The cyclic digraphs are created from the cyclic graph. The important point here is that we know all the possible topologically distinct HB patterns for water trimer: there are no HB patterns for water trimer other than the digraphs shown in Fig. 3.

Rooted digraph and HB matrix

A rooted graph is a graph in which one vertex is different from the other vertices. Molecular structures with a heteroatom can be represented with rooted graphs, where vertices correspond to a heteroatom or carbon atoms, and edges correspond to bonds. Rooted digraph is used to represent the feature of protonated water (PW) cluster, since PW is distinguished from the other water molecules in the PW cluster [19, 20]. A rooted digraph is a set of vertices and arrows (Fig. 4). The HB matrix can be applied to a rooted digraph: for a rooted digraph with n vertices, the HB matrix H is the n^{th} order square matrix, whose element h_{ij} is equal to 1 for an arrow directed from vertex i to vertex j, and 0 otherwise.

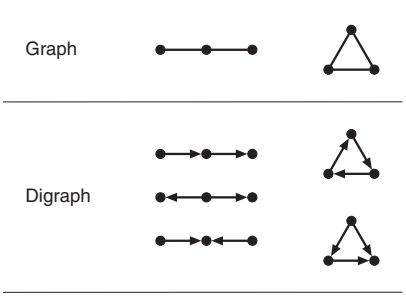


Fig. 3: All possible graphs and digraphs with three vertices.

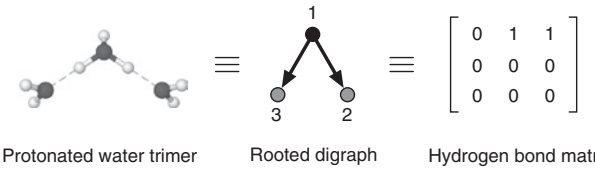


Fig. 4: A graph representation of protonated water trimer and the corresponding hydrogen bond matrix.

The conditions for a rooted digraph to be equivalent to a PW cluster are the following. In a PW cluster, one PW can accept up to one proton from other water molecules, and donate up to three protons to other water molecules. One water molecule can accept up to two protons from other water molecules or PW, and donate up to two protons to other water molecules or PW. In addition, in a PW cluster, all of the neutral water molecules and PW should be connected by hydrogen bonds.

HB pattern: subset of configuration space of water cluster

For a rigid molecule, a local minimum structure on the potential energy surface corresponds to a structure which is regarded to be observed at a finite temperature. Very often, partition functions are calculated using ideal gas approximation, rigid rotor approximation and harmonic oscillator approximation at the geometry of local minimum (left method in Fig. 5). Thermodynamic quantities, such as Gibbs energy, can be calculated along with these approximations. On the other hand, a molecular simulation technique such as MC or MD is used to generate coordinates of molecular clusters which are observed at a finite temperature, and the partition function can be calculated (right method in Fig. 5). Every structure which is generated in a molecular simulation is different, and we use the HB pattern to categorize the generated coordinates to subgroups (i.e. isomers).

The HB pattern satisfies all the following four properties: (i) every pattern is different, (ii) a structure is categorized to only one pattern, (iii) any one of possible structures can be categorized to one of all possible patterns, and (iv) the total number of the patterns is known. The HB pattern can be used for the categorization to distinguish topologically distinct water cluster isomers (Fig. 5) [21, 22].

Generation of water cluster structures

TTM2-R potential energy function [23] was used to calculate interaction energies of water clusters. The permanent dipole moment of this model originating from partial charges is 1.853 D, which reproduces the experi-

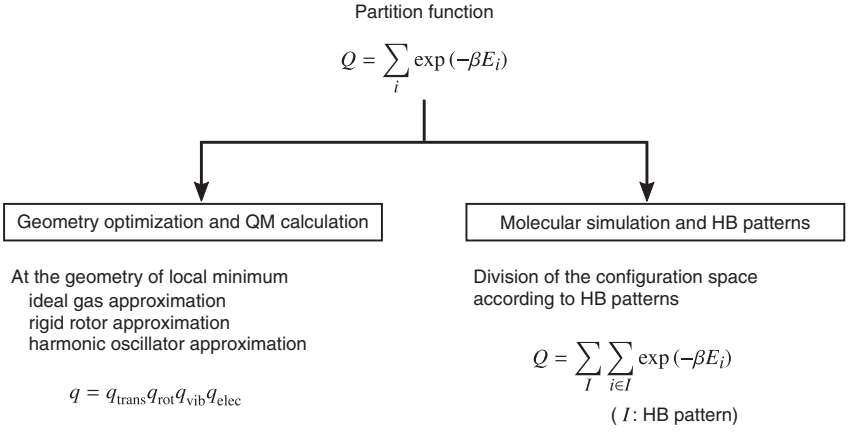


Fig. 5: Two different methods to calculate partition functions.

mental value of 1.8546 D [24] in the gas phase. In addition, atomic polarizabilities placed on O- and H-sites produce induced dipole moments influenced by surrounding molecules. The repulsive exponential term acting between the O-sites to the two-body pairwise terms was added [22] in order to correct the unphysical stabilization at a short O···O distance of the original TTM2-R function. The modified two-body pairwise terms are

$$V^{\text{pair}} = \sum_{i=1}^{N-1}\sum_{j>i}^{N}\left[\frac{A}{r^{12}} + \frac{B}{r^{10}} + \frac{C}{r^{6}} + De^{-Er}\right].$$

The values of parameters A, B, and C are taken from the original paper [23]. The parameters for the added terms are $D = 1.0 \times 10^{13}$ kcal/mol and $E = 13.2$ Å^{-1} [22].

The MC simulations of water clusters, $(H_2O)_n$, $n = 2$–10, were performed with the standard Metropolis algorithm [25]. The details concerning the simulation procedure are given in [22]. Configurations were sampled at every MC step, and 10^9 configurations were classified into HB patterns by solving the graph isomorphism problem for each system. We defined the formation of hydrogen bonding, if the hydrogen-bond distance r(OH···O) is less than 2.45 Å and at the same time the hydrogen-bond angle θ(O–H···O) is wider than 120°.

The relative Helmholtz energy of each HB pattern is evaluated from the population of HB patterns, because the MC simulations were performed in the canonical ensemble.

$$\Delta A_{IJ} = -RT \ln \frac{N_I}{N_J}.$$

Here, ΔA_{IJ} is the Helmholtz energy of HB pattern I relative to HB pattern J. R is the gas constant and T is temperature. N_I denotes the number of the configurations classified into the HB pattern I.

For each HB pattern, the energy average $\langle E \rangle_I$ is given by

$$\langle E \rangle_I = \frac{\sum_{i \in I} V_i}{N_I},$$

where the sum is over the configurations belonging to the HB pattern I, and V_i is the potential energy associated with a configuration i. The internal energy of HB pattern I relative to HB pattern J is obtained by

$$\Delta U_{IJ} = \langle E \rangle_I - \langle E \rangle_J.$$

The relative Helmholtz energy between HB patterns I and J is related to the relative internal energy and entropy,

$$\Delta A_{IJ} = \Delta U_{IJ} - T\Delta S_{IJ},$$

where ΔS_{IJ} is the relative entropy between the HB patterns I and J. The intramolecular vibrational entropic contributions of water molecules are not included, because the rigid water model is used. Note that the internal energy and the entropy of an HB pattern are implicitly temperature-dependent.

Distribution of HB patterns of water trimer

From the MC simulations of water clusters $(H_2O)n$, $n = 3$–8, the relative values of Helmholtz energy (ΔA), internal energy (ΔU), and entropic term ($-T\Delta S$) for several HB patterns of each n-water system were evaluated, and the results of some of low Helmholtz energy HB patterns for each system are given in Ref. [22].

Here, the results for water trimer are summarized. As shown in Fig. 3, the number of topologically distinct patterns of the system with three vertices is 5; i.e. it is certain that there are only five digraphs (HB patterns)

for water trimer. The relative and cumulative populations of HB patterns of water trimer are shown in Ref. [22]. At 200 K, a cyclic HB pattern **3-3A** is the most populated, namely, the lowest in Helmholtz energy. The Helmholtz energy (blue bar), the internal energy (red bar) and the entropic term (green bar) of other HB patterns relative to the corresponding values of **3-3A** at 200 K are shown in Fig. 6. The Helmholtz energy (in blue) of **3-2A** is 2.9 kJ/mol higher than that of **3-3A**, followed by **3-2B**, **3-2C** and **3-3B**.

At 300 K, on the other hand, the linear pattern **3-2A** is the most populated, lying 1.0 kJ/mol lower than the cyclic pattern **3-3A** in Helmholtz energy. The Helmholtz energy, the internal energy and the entropic term of each of other HB patterns relative to the corresponding values of **3-2A** at 300 K are shown also in Fig. 6. The internal energy of **3-2A** is 12.0 kJ/mol higher than that of **3-3A** which remains the lowest among the five patterns even at 300 K. The entropic term demonstrates that the cyclic patterns (**3-3A** and **3-3B**) are entropically less favorable than the linear patterns (**3-2A**, **3-2B**, and **3-2C**).

All of the global minimum structure and several stationary points of water trimer, which were obtained by means of high-level ab initio molecular orbital (MO) calculations [26, 27], are identified in the configuration space of **3-3A**. We have calculated the local minimum geometries of water trimer at the theoretical level of MP2/aug-cc-pVDZ using Gaussian 09 program package [28], and have found four local minima. They are shown in Fig. 7. The lowest two geometries correspond to **3-3A**. In the HB pattern, the orientation of dangling OH groups is not distinguished. The other two geometries correspond to **3-2C** and **3-2B**, respectively. Local minima do not exist corresponding to either **3-2A** or **3-3B**. It is worthy to note that there is no local minimum on the potential energy surface of water trimer which corresponds to the linear **3-2A**, which is the

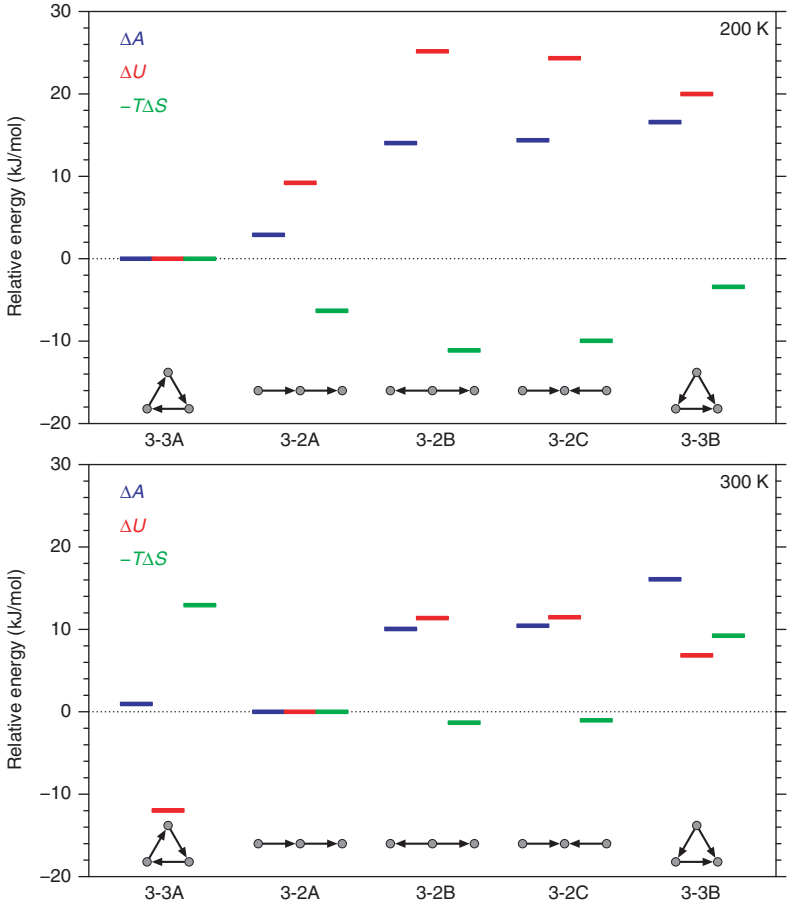


Fig. 6: Helmholtz energies, internal energies, entropic terms relative to the most populated HB pattern for water trimer at 200 and 300 K.

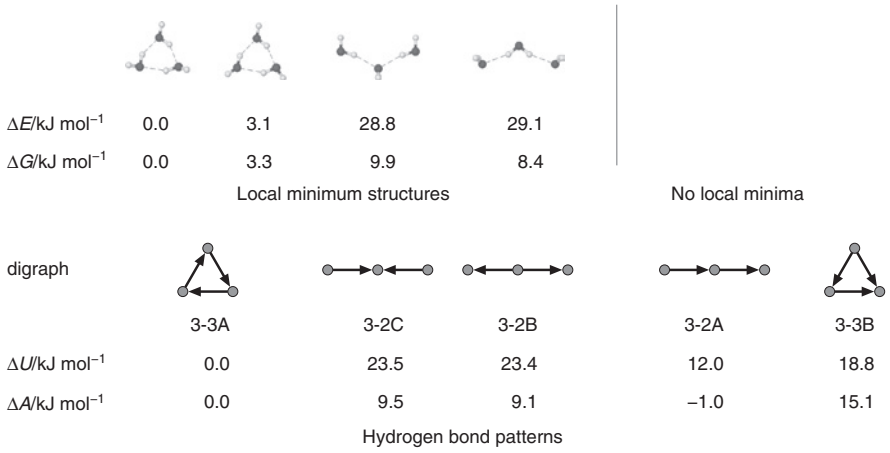


Fig. 7: Local minimum structures of water trimer with potential energy (ΔE) and Gibbs energy (ΔG) at 298.15 K relative to those of the global minimum structure, and the HB patterns of water trimer with internal energy (ΔU) and Helmholtz energy (ΔA) relative to those of **3-3A** at 300 K.

most popular HB pattern at 300 K. This is a very important finding; local minimum geometries cannot be always representative.

The internal energies of the HB patterns **3-2C** and **3-2B** relative to that of **3-3A** at 300 K are 23.5 and 23.4 kJ/mol, respectively, and they are similar to but lower than the relative energies of the corresponding local minima (28.8 and 29.1 kJ/mol, respectively, at MP2/aug-cc-pVDZ). This is fairly reasonable and it may indicate the reliability of the HB pattern analysis based on the MC simulation using TTM2-R potential energy function [23].

Relative Gibbs energy ΔG of each local minimum structure in Fig. 7 is a value relative to Gibbs energy of the global minimum geometry. Gibbs energy of each local minimum structure has been calculated at 298.15 K, at the theoretical level of MP2/aug-cc-pVDZ using Gaussian 09 program package [28]. On the other hand, relative Helmholtz energy ΔA between each HB pattern and the most populated HB pattern in Fig. 7 is a value evaluated from MC simulation at 300 K [22]. Since no local minima are found corresponding to either the linear HB pattern **3-2A** or the cyclic HB pattern **3-3B**, we cannot estimate ΔG of a structure corresponding to either HB pattern.

Quantum chemical evaluation of ΔG at 298.15 K of a molecular system is often performed based on the left method in Fig. 5. It must be noticed that other isomers which do not correspond to local minima on the potential energy surface of a molecular system can be observed at a finite temperature. On the other hand, the results of molecular simulation depend severely on the quality of the potential energy function used in the simulation. In the current case [22], the quality of the TTM2-R potential energy function [23] is reasonably high as described above and also later.

Distribution of HB patterns of water octamer

Here we show another example that an isomer which corresponds to the global minimum on the potential energy surface may not be observed at a finite temperature.

The number of topologically distinct patterns of the system with eight vertices is 164 461; i.e. it is certain that there are 164 461 digraphs (HB patterns) for water octamer at the most [15]. From the MC simulations of water clusters, the results for water octamer are shown here. The relative and cumulative populations of HB patterns of water octamer are given in ref [22].

At 200 K, 18 589 HB patterns were sampled in the MC simulation. At 200 K, one of the cubic HB pattern (S_4 symmetry) is the most populated, namely, the lowest in Helmholtz energy. The Helmholtz energy (blue

bar), the internal energy (red bar) and the entropic term (green bar) of other HB patterns relative to the corresponding values of the S_4 cubic HB pattern at 200 K are shown in Fig. 8, where two cubic HB patterns and three pseudo-cubic HB patterns in which one hydrogen bond is broken are included. In each of these pseudo-cubic HB patterns, the internal energy becomes higher, while the entropic term becomes lower than the corresponding cubic pattern. In Fig. 8, various HB patterns other than cubic pattern are included. In those patterns, the internal energy becomes higher, while the entropic term becomes lower than the cubic pattern. The spirocyclic HB pattern shows the lowest entropic term in the HB patterns in Fig. 8, while it is the seventh most stable in Helmholtz energy at 200 K.

At 300 K, 37 978 HB patterns were obtained in the MC simulation. At 300 K, the spirocyclic HB pattern, which is the lowest in entropy term at 200 K in Fig. 8, is the most populated, namely, the lowest in Helmholtz energy. The Helmholtz energy, the internal energy and the entropic term of each of other HB patterns relative to the corresponding values of the spirocyclic isomer at 300 K are shown in Fig. 9, where many other patterns are included. In those patterns, the entropic terms are further lower. It is worthy to note that the populations of cubic isomers are very small at 300 K.

For the cubic water octamers, there are 14 HB patterns, and 14 local minimum structures are found at the theoretical level of MP2/aug-cc-pVDZ. One local minimum is found for each HB pattern, and they are shown in Fig. 10. The symmetry of each structure (**1–14**) in Fig. 10 is as follows: **1** S_4, **2** D_{2d}, **3** C_i, **4** C_2, **8** C_s, **13** C_4 and **14** C_4 symmetry. Other structures in Fig. 10 are in C_1 symmetry. Relative Gibbs energy ΔG of each local minimum structure is plotted in Fig. 11, which is a value relative to Gibbs energy of the structure **1**. Gibbs energy of each local minimum structure has been calculated at 298.15 K, at the theoretical level of MP2/aug-cc-pVDZ using Gaussian 09 program package [28].

Relative Helmholtz energy ΔA between each HB pattern and the most populated HB pattern **1** is plotted in Fig. 11. We have evaluated the relative Helmholtz energies of the cubic HB patterns from the MC simulation at 300 K [22] using TTM2-R potential energy function [23]. Note that the structures corresponding to HB patterns **13** and **14** were not sampled in the MC simulation, which indicates that the energies of those HB patterns are too high to be observed at 300 K.

As shown in Fig. 11, the order of ΔG of the cubic structures and the order of ΔA of the cubic HB patterns are similar: it may indicate the reliability of the HB pattern analysis based on the MC simulation using TTM2-R potential energy function [23].

The cubic structure **2** (D_{2d} symmetry) is the global minimum on the potential energy surface of water octamer. The cubic structure **1** (S_4 symmetry) is the most populated in a low temperature. At higher temperature, entropically unfavorable cubic structures are less populated. The populations of cubic HB patterns are

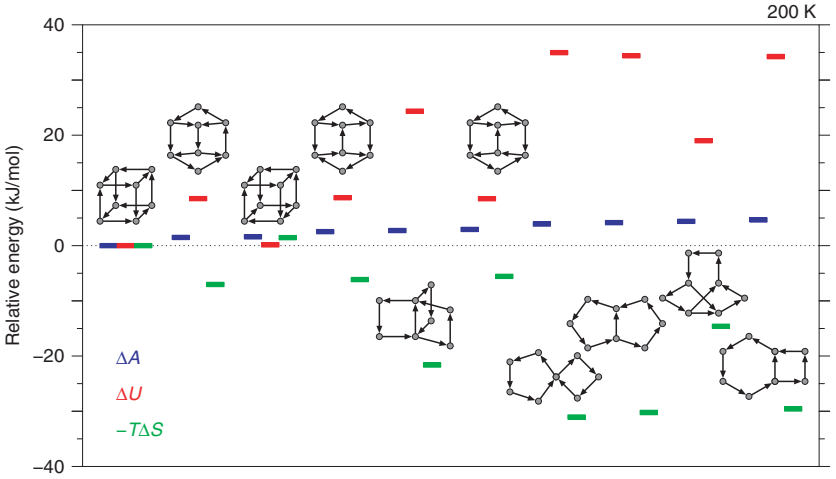


Fig. 8: Helmholtz energies, internal energies, entropic terms relative to the most populated HB pattern for the 10 most populated HB patterns of water octarmer at 200 K.

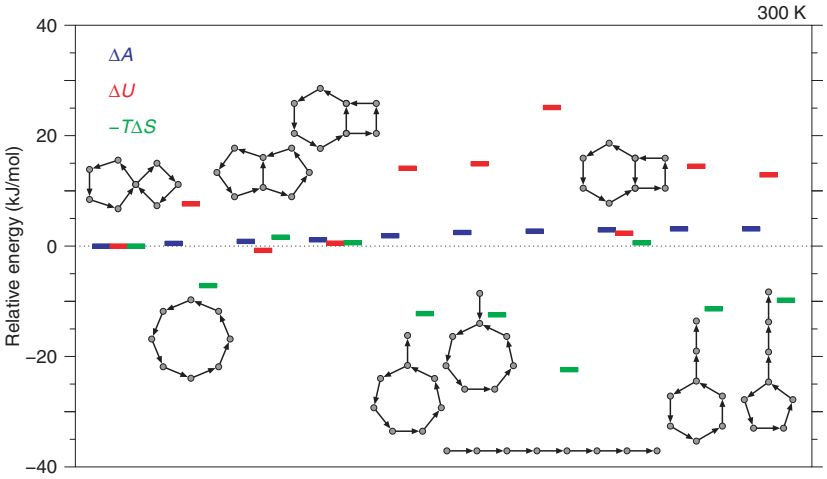


Fig. 9: Helmholtz energies, internal energies, entropic terms relative to the most populated HB pattern for the 10 most populated HB patterns of water octarmer at 300 K.

very small at 300 K: any cubic structures may not be observed at 300 K. Even though ΔG of the cubic structure **1** is the lowest among the 14 cubic structures at 298.15 K, it does not mean the cubic structure **1** is the most populated among water octamer at 298.15 K.

The relative stability in the 14 cubic water octamers is related to the HB patterns, and is attributed to the strong dependence of the HB strength on the local HB types (see Section "HB type of water molecule" and Fig. 12) of the contributing water molecules [9]. The hydrogen bond between the water molecules of the type *dda* and the type *daa* (i.e. *dda* water molecule is proton acceptor and *daa* water molecule is proton donor, *dda*←*daa*) is the strongest; the origin of the strong HB energy is the large charge transfer energy as well as the large dispersion energy [9]. The number of the *dda*←*daa* pair in each HB pattern (**1–14**) in Fig. 10 is as follows: 4 for HB patterns **1** and **2**, 2 for **3–8**, 1 for **9–12**, 0 for **13** and **14**.

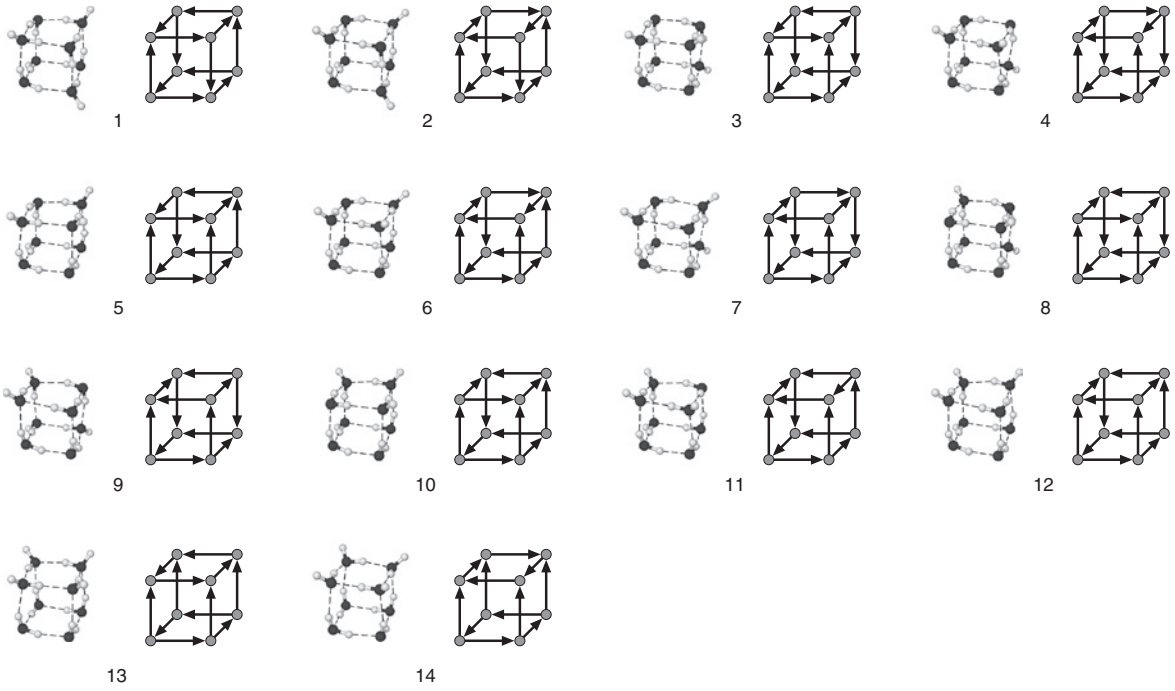


Fig. 10: Cubic water octamers.

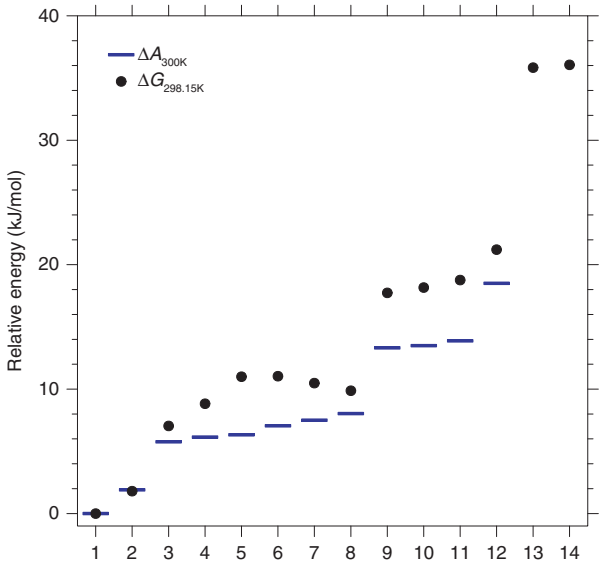


Fig. 11: Relative free energies of HB patterns and relative free energies of local minimum structures of cubic water octamers.

HB type of water molecule

It is well-known that the dipole moment of an isolated water molecule is different from that of bulk water. The dipole moment of a water molecule in the gas phase was obtained to be 1.8546 D from Stark effect measurements [24], whereas the molecular dipole moment of liquid water was estimated to be 2.9 ± 0.6 D from X-ray measurements [29]. The gas phase dipole moment has been reproduced in good agreement by ab initio MO calculations [30, 31]. Many theoretical studies estimated the enhancement of the dipole moment of a water molecule in clusters and in the condensed phases [32–39]. The estimated dipole moments depend on the water model for polarizable potentials and the way to assign the electron density into a molecule for ab initio MO calculations. The reason of the enhancement has been attributed to the polarization caused by the intermolecular interactions with surrounding water molecules via hydrogen bonding. The influence must be rationalized by HB type of each water molecule.

Local HB type of water molecule

The constituent water molecules in water cluster can be classified into the *local HB types* (*free*, *d*, *a*, *dd*, *aa*, *da*, *dda*, *daa*, and *ddaa*) (Fig. 12) according to the way how the water molecule participates in the HB network [22].

Although the TTM2-R model [23] has the fixed geometry of a water molecule, a dipole moment of a constituent water molecule of a water cluster changes because of the induced dipole moments which are generated by the interaction with surrounding molecules.

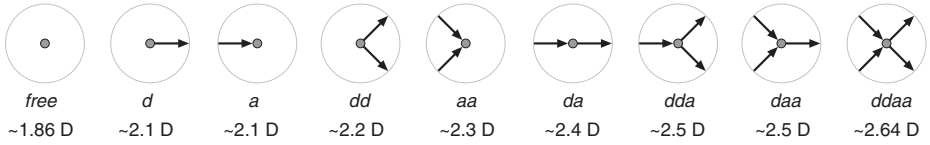


Fig. 12: Molecular local HB type and dipole moment of water molecule estimated from MC simulation of water decamer.

Molecular dipole moments of constituent water molecules

The ensemble averages of molecular dipole moments of water molecules in water clusters (in the MC simulation [22] at 300 K) for each system size (the number of water molecules in the system) are given in [22]. The dipole moments of hydrogen-bonded water molecules increase with the size of the water cluster. The molecular dipole moment depends on the local HB network around the water molecule [22]. The magnitude of the molecular dipole moment increases in the following order: *free* < *d* < *a* < *dd* < *aa* < *da* < *dda* < *daa* < *ddaa*. The average dipole moment for each local HB type of water decamer is shown in Fig. 12. The dipole moments of not-hydrogen-bonded (*free*) molecules are almost constant independent of the number of water molecules in the system, whereas the dipole moments of hydrogen-bonded molecules increase with the number of water molecules in the system.

The dipole moments of *free* molecules are close to that of an isolated water molecule (1.853 D). The dipole moment of a water molecule is enhanced by forming hydrogen bonds, and the enhancement depends on the type of the local HB environment. The dipole moments of *da* type molecules are larger than those of *aa* and *dd* types. It may indicate that a water molecule which accepts a hydrogen bond and at the same time donates a hydrogen bond is polarized by the continuous hydrogen bonds. The dipole moments of water molecules of *dda* and *daa* types are larger than those of *da* type molecules. The dipole moments of *ddaa* type water molecules are larger than those of *dda* or *daa* type molecules. The *ddaa* type water molecule which accepts two hydrogen bonds and at the same time donates two hydrogen bonds is highly polarized.

As was indicated in [22], it is noteworthy that the dipole moments of *ddaa* type water molecules in water decamer is close to that in bulk water calculated in the MD simulation with the TTM2-R water model (2.65 D) [23]. Thus, the large dipole moment of a water molecule in bulk water is attributed to the large population of the *ddaa* type water molecules, which is highly polarized inherently. A dipole moment of a water molecule depends on the local type of hydrogen bonding. The similar results have been obtained from DFT calculations of water clusters, in which the importance is stressed of the polarization and delocalization between water molecules [39]. In bulk water, *ddaa* type water molecules are dominant in average. The average number of hydrogen bonds per molecule in liquid water at 300 K was reported [40] to be 3.6 from the MD simulation with the SPC/E model.

Protonated water clusters

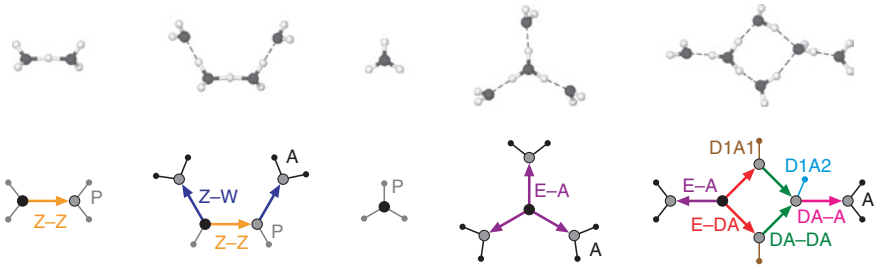
Protonated water (PW) clusters are common in nature and play important roles in many fields of science including atmospheric chemistry, solution chemistry, and biological chemistry. There are a lot of stable structures in PW clusters with various arrangements of the constituent water molecules. The number of local minima increases rapidly with the cluster size. Because of the proton included, the feature of O–H bond varies according to the arrangement of HB network.

Classification of O–H bond in PW cluster

All the O–H bonds in PW cluster are classified into 10 topological types (Fig. 13), utilizing the combination of the local HB types of the contributing water molecules in the O–H bond. Six of them are hydrogen-bonded O–H bonds. Four of them are free O–H bonds [43]. The free O–H bonds are explicitly represented by small vertices, each of which links to a larger vertex in the digraph representation in Fig. 13.

Distribution of PW octamer

Rooted digraph is used to represent topologically distinct isomers of PW cluster [19, 20]. The numbers of enumerated rooted digraphs of PW clusters $H_2O^+(H_2O)_{n-1}$ (n = 2–8) are given in [19]. The number of rooted digraphs



H-bonded O–H	Z–Z	Shared proton oscillation in Zundel core
	E–DA	Eigen O–H H-bonded to donor-acceptor water
	E–A	Eigen O–H H-bonded to acceptor water
	Z–W	Zundel O–H H-bonded to water
	DA–A	O–H of donor-acceptor water H-bonded to acceptor water
	DA–DA	O–H of donor-acceptor water H-bonded to donor-acceptor water
free O–H	P	free O–H of Eigen and Zundel ion core
	D1A2	free O–H of single-donor-double-acceptor water
	D1A1	free O–H of single-donor-single-acceptor water
	A(s)	free O–H of acceptor water (symmetric stretching)
	A(a)	free O–H of acceptor water (anti-symmetric stretching)

Fig. 13: O–H topological types in protonated water clusters. Local minimum structures of representative small protonated water clusters containing two Zundel and three Eigen isomers are shown. In their digraph representations, the O–H topological types are illustrated. Each of small vertex with undirected edge in digraphs represents free O–H. The O–H topological types contain six H-boned O–H and four free O–H. Free O–H of acceptor water A has two stretching modes: symmetric stretching A(s) and anti-symmetric stretching A(a).

for PW octamer ($n=8$) is 896 063: this is the maximum number of its topologically distinct structures. It is not trivial to search full potential energy surface of PW octamer using a reliable level of calculation. Making use of different computational methods [41, 42], 134 topologically distinct local minima of PW octamers were obtained [43] at the theoretical level of MP2/aug-cc-pVDZ. The lowest energy structure is **TP1**, which is Eigen type. **TP26** is the 26th stable isomer, and the most stable isomer of the Zundel type. To gain insights into the thermodynamic stability of topologically distinct isomers, temperature-dependent populations of the optimized local minimum structures were estimated on the basis of the molecular partition functions [43] using the left method of Fig. 5. At a temperature lower than ~140 K, **TP1** is dominant. In the range between ~150

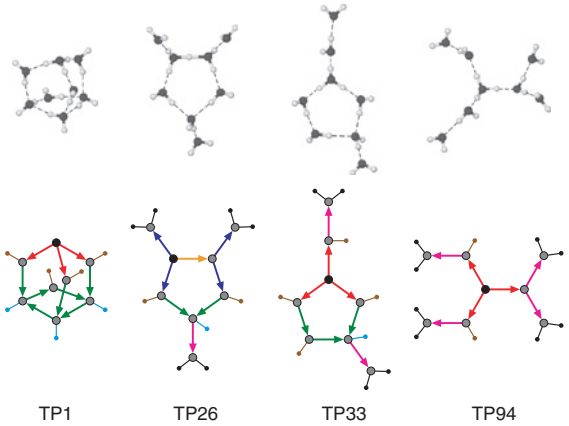


Fig. 14: Optimized structures and digraph representations of representative protonated water octamers. The O–H topological types in each structure are indicated in the corresponding rooted digraph.

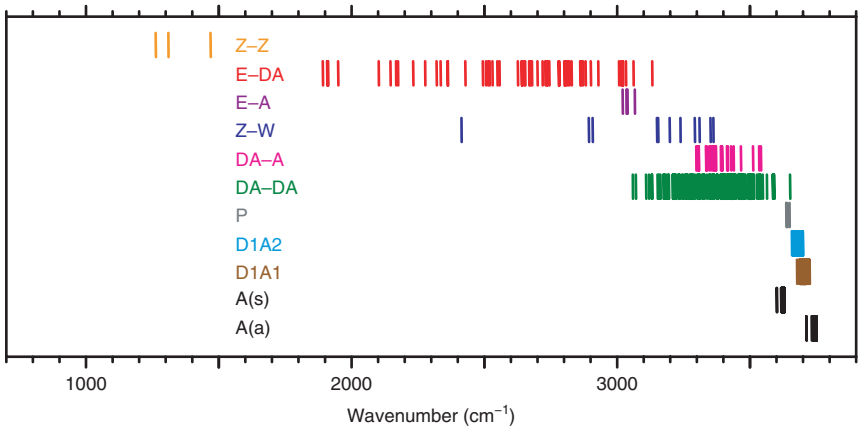


Fig. 15: Harmonic frequencies of O–H stretching modes of protonated water octamers. The wavenumber was calculated at the MP2/aug-cc-pVDZ level of theory and scaled by 0.955.

and ~250 K, TP26 (Zundel isomer) is dominant being followed by **TP33**. At higher temperature (>270 K), a chain structure **TP94** is dominant. The optimized structures and the corresponding digraphs of these four PW octamers are shown in Fig. 14. The O–H bonds are drawn with coloring same as in Fig. 13.

Vibrational frequency of O–H topological type

All O–H vibrational frequencies of 25 selected isomers of PW octamer are plotted in Fig. 15. The O–H vibrational frequencies spread from ~1200 cm^{-1} to ~3800 cm^{-1}: they can be classified according to the topological type of the O–H bond [43]. The vibrational frequency for free O–H of acceptor water molecule A is further classified into symmetric stretching A(s) and anti-symmetric stretching A(a) according to the vibration modes (Fig. 13).

If the O–H topological type (Fig. 13) is the same, the corresponding O–H vibrational frequency is similar even in different isomers. The O–H topological types can be used to distinguish spectral signatures of O–H stretching modes. Assign the O–H bond the topological type, and you can make a rough estimate about the vibrational frequency.

Calculated IR spectra and assignment of coexisting isomers

The calculated harmonic IR spectra of the representative structures, **TP1**, **TP26**, **TP33** and **TP94** in the range of 2000–3800 cm^{-1} are given in [43]. The range of 3600–3800 cm^{-1} corresponds to the free O–H stretching modes of the topological types including A(s) (~3630 cm^{-1}), D1A2 (~3670 cm^{-1}), D1A1 (~3700 cm^{-1}) and A(a) (~3750 cm^{-1}). The number of the bands of **TP26** and **TP33** is four in this range and it corresponds with the number of the bands of the observed spectra of PW octamer [44, 45]. Since there is a spectral signature of Zundel core (the shared proton oscillation) in the observed spectra [44], it is certain that **TP26** exists. The coexistence of either **TP33** or **TP1**, or both of them, with **TP26** does not conflict with the observed spectra [44].

Conclusion

HB pattern is used to categorize the structures of water cluster. Digraph is used to represent HB pattern. The important point is that the maximum number of HB patterns for water cluster is known: thus, we can use the

HB pattern as a means to divide the configuration space of water cluster with the same number of constituent water molecules into the topologically distinct structures. The constituent water molecules can be classified into the local HB types according to the way how the water molecule participates in the HB network. The dipole moment of a water molecule is enhanced because of the polarization by hydrogen bonding with surrounding water molecules. The influence is rationalized by the local HB type of each water molecule. In protonated water cluster, all the O–H bonds are classified into 10 topological types, utilizing the combination of the local HB types of the contributing water molecules in the O–H bond. The O–H topological types can be used to distinguish spectral signatures of O–H stretching modes.

In this article, we showed that HB patterns of water clusters and local HB types of water molecules in water clusters play important roles to grasp essential qualities from large amounts of computational data. This concept to use the topological aspect of the HB network can be applied to other systems than water clusters.

Acknowledgment: This mini-review is based on the research which has been performed at the Quantum Chemistry Research Group of Hiroshima University. We acknowledge all of the previous and present students in the Group for their hard work and helpful discussion.

Article note: A special collection of invited papers by recipients of the IUPAC Distinguished Women in Chemistry and Chemical Engineering Awards.

References

[1] A. Cayley. *Philos. Mag.* **47**, 444 (1874).
[2] D. H. Rouvray. *Chem. Soc. Rev.* **3**, 355 (1974).
[3] H. Wiener. *J. Am. Chem. Soc.* **69**, 17 (1947).
[4] H. Hosoya. *Bull. Chem. Soc. Jpn.* **44**, 2332 (1971).
[5] H. Hosoya. *Internet Electron. J. Mol. Des.* **1**, 428 (2002).
[6] K. Ohno, T. Shimoaka, N. Akai, Y. Katsumoto. *J. Phys. Chem. A* **112**, 7342 (2008).
[7] S. Fujita. *Bull. Chem. Soc. Jpn.* **81**, 193 (2008).
[8] S. Fujita. *Bull. Chem. Soc. Jpn.* **83**, 1 (2010).
[9] S. Iwata, D. Akase, M. Aida, S. S. Xantheas. *Phys. Chem. Chem. Phys.* **18**, 19746 (2016).
[10] J. C. Howard, G. S. Tschumper. *WIREs Comput. Mol. Sci.* **4**, 199 (2014).
[11] M. Aida, D. Akase, H. Doi, T. Yoshida. in *Practical Aspects of Computational Chemistry II: An Overview of the Last Two Decades and Current Trends*, J. Leszczynski, M. K. Shukla (Eds.), pp. 49–68, Springer, Dordrecht (2012).
[12] T. P. Radhakrishnan, W. C. Herndon. *J. Phys. Chem.* **95**, 10609 (1991).
[13] S. McDonald, L. Ojamäe, S. J. Singer. *J. Phys. Chem. A* **102**, 2824 (1998).
[14] J. L. Kuo, J. V. Coe, S. J. Singer, Y. Band, L. Ojamäe. *J. Chem. Phys.* **114**, 2527 (2001).
[15] T. Miyake, M. Aida. *Chem. Phys. Lett.* **363**, 106 (2002).
[16] T. Miyake, M. Aida. *Internet Electron. J. Mol. Des.* **2**, 24 (2003).
[17] D. Vukičević, T. Grubeša, A. Graovac. *Chem. Phys. Lett.* **416**, 212 (2005).
[18] G. Brinkmann. *J. Math. Chem.* **46**, 1112 (2009).
[19] M. Jieli, T. Miyake, M. Aida. *Bull. Chem. Soc. Jpn.* **80**, 2131 (2007).
[20] M. Jieli, M. Aida. *J. Phys. Chem. A* **113**, 1586 (2009).
[21] T. Miyake, M. Aida. *Chem. Phys. Lett.* **427**, 215 (2006).
[22] D. Akase, M. Aida. *J. Phys. Chem. A* **118**, 7911 (2014).
[23] C. J. Burnham, S. S. Xantheas. *J. Chem. Phys.* **116**, 1500 (2002).
[24] S. A. Clough, Y. Beers, G. P. Klein, L. S. Rothman. *J. Chem. Phys.* **59**, 2254 (1973).
[25] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller. *J. Chem. Phys.* **21**, 1087 (1953).
[26] F. N. Keutsch, J. D. Cruzan, R. J. Saykally. *Chem. Rev.* **103**, 2533 (2003).
[27] M. B. Day, K. N. Kirschner, G. C. Shields. *J. Phys. Chem. A* **109**, 6773 (2005).
[28] 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, T. Keith, 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, D. J. Fox. Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford, CT (2013).

[29] Y. S. Badyal, M.-L. Saboungi, D. L. Price, S. D. Shastri, D. R. Haeffner, A. K. Soper. *J. Chem. Phys.* **112**, 9206 (2000).

[30] S. S. Xantheas, T. H. Dunning, Jr. *J. Chem. Phys.* **99**, 8774 (1993).

[31] J. K. Gregory, D. C. Clary, K. Liu, M. G. Brown, R. J. Saykally. *Science* **275**, 814 (1997).

[32] P. L. Silvestrelli, M. Parrinello. *Phys. Rev. Lett.* **82**, 3308 (1999).

[33] L. D. Site, A. Alavi, R. M. Lynden-Bell. *Mol. Phys.* **96**, 1683 (1999).

[34] E. R. Batista, S. S. Xantheas, H. Jónsson. *J. Chem. Phys.* **111**, 6011 (1999).

[35] Y. Tu, A. Laaksonen. *Chem. Phys. Lett.* **329**, 283 (2000).

[36] S. W. Rick. *J. Chem. Phys.* **114**, 2276 (2001).

[37] D. D. Kemp, M. S. Gordon. *J. Phys. Chem. A* **112**, 4885 (2008).

[38] M. L. Antipova, D. L. Gurina, V. E. Petrenko. *Russ. J. Phys. Chem. A* **87**, 449 (2013).

[39] I. Bakó, I. Mayer. *J. Phys. Chem. A* **120**, 4408 (2016).

[40] R. Kumar, J. R. Schmidt, J. L. Skinner. *J. Chem. Phys.* **126**, 204107 (2007).

[41] Y. Luo, S. Maeda, K. Ohno. *J. Phys. Chem. A* **111**, 10732 (2007).

[42] H. Teramae, T. Ishimoto, U. Nagashima. *Theor. Chem. Acc.* **130**, 671 (2011).

[43] D. Akase, H. Teramae, M. Aida. *Chem. Phys. Lett.* **618**, 51 (2015).

[44] J. M. Headrick, E. G. Diken, R. S. Walters, N. I. Hammer, R. A. Christie, J. Cui, E. M. Myshakin, M. A. Duncan, M. A. Johnson, K. D. Jordan. *Science* **308**, 1765 (2005).

[45] K. Mizuse, A. Fujii. *J. Phys. Chem. A* **116**, 4868 (2012).