TOPICAL REVIEW — 8th IUPAP International Conference on Biological Physics

Ab initio path-integral molecular dynamics and the quantum nature of hydrogen bonds*

Yexin Feng(冯页新)[1], Ji Chen(陈基)[1], Xin-Zheng Li(李新征)[2,3,†], and Enge Wang(王恩哥)[1,3]

[1] *International Center for Quantum Materials and School of Physics, Peking University, Beijing* 100871, *China*
[2] *School of Physics, Peking University, Beijing* 100871, *China*
[3] *Collaborative Innovation Center of Quantum Matter, Beijing* 100871, *China*

(Received 21 May 2015; published online 7 December 2015)

The hydrogen bond (HB) is an important type of intermolecular interaction, which is generally weak, ubiquitous, and essential to life on earth. The small mass of hydrogen means that many properties of HBs are quantum mechanical in nature. In recent years, because of the development of computer simulation methods and computational power, the influence of nuclear quantum effects (NQEs) on the structural and energetic properties of some hydrogen bonded systems has been intensively studied. Here, we present a review of these studies by focussing on the explanation of the principles underlying the simulation methods, i.e., the *ab initio* path-integral molecular dynamics. Its extension in combination with the thermodynamic integration method for the calculation of free energies will also be introduced. We use two examples to show how this influence of NQEs in realistic systems is simulated in practice.

Keywords: *ab initio* calculations, isotope, molecular dynamics, hydrogen bonds

PACS: 31.15.A–, 67.63.–r, 71.15.Pd, 82.30.Rs **DOI:** 10.1088/1674-1056/25/1/013104

1. Introduction

The hydrogen bond (HB) is a generally weak but ubiquitous intermolecular interaction. It is responsible, for example, for holding together the two strands of DNA molecules and many phases of water. As such, understanding the behaviors of HBs is lying at the core of many studies concerning the behaviors of water and biomolecules. The small mass of hydrogen (H) means that many of these problems are quantum mechanical in nature,[1–7] although all too often these nuclear quantum effects (NQEs) are neglected. These NQEs are essential for rationalizing many phenomena of te H-bonded systems, especially where quantum tunneling and zero-point motion are concerned. For example, starting from the 1950s, it has been known that replacing H by deuterium (D) in H-bonded molecular crystals with electronic structures unchanged can result in obvious variations of the lattice constants.[8] This is currently known as the Ubbelohde effect.[9] In enzyme reactions, the proton tunneling along HBs is referred to as "nature's subway". By using a neutron Compton scattering experiment, a picture concerning the impact of NQEs on the real space delocalization and vibrational properties of the proton in liquid water was also revealed.[10,11] Most recently, a cryogenic scanning tunneling microscope (STM) experiment even confirmed the isotope-dependent switching rate of water tetramer on salt between its two chirality states, which could only be rationalized by considering the NQEs.[12]

From the theoretical perspective, accurate computer simulations of these phenomena require an accurate quantum mechanical description of the finite-temperature nuclear fluctuations in real polyatomic systems. A natural choice is to construct a high-dimensional Schrödinger equation for the many-body entity of the nuclei. By solving the eigenstate wavefunctions of this Schrödinger equation, one can observe not only the statistical but also the dynamical properties of the system under investigation related to the quantum nature of the nuclei. Over the last $\sim$ 20 years, this method has been extremely successful in describing gas-phase reactions.[13–17] However, because of the notorious scaling problem in mapping the high-dimensional potential energy surface (PES) and solving the Schrödinger equation, the application of this method has been seriously limited to systems with less than $\sim$ 6 atoms. For more realistic polyatomic systems and condensed matters, a practical method is highly desired. Thanks to the development of the path-integral representation of quantum mechanics starting from the late 1940s,[18–21] a scheme practical for large polyatomic systems and condensed matters in which not only the thermal effects but also the NQEs can be accounted for statistically has been systematically presented by Feynman and Hibbs.[22] Based on such a foundation, the molecular dynamics (MD) simulation technique was combined with this path-integral representation of quantum mechanics, and a series of path-integral molecular dynam-

*Project supported by the National Natural Science Foundation of China (Grant Nos. 11275008, 91021007, and 10974012) and the China Postdoctoral Science Foundation (Grant No. 2014M550005).

†Corresponding author. E-mail: xzli@pku.edu.cn

© 2016 Chinese Physical Society and IOP Publishing Ltd

ics (PIMD) simulations were carried out in the 1980s, as reported by Chandler, Parrinello, and Berne *et al.* in Refs. [23]–[25], respectively. Analogous to these MD based simulations, the Monte–Carlo (MC) sampling technique can also be used. At roughly the same time, the properties of liquid helium, including its superfluidity, were systematically studied by Ceperley and Pollock using path-integral Monte–Carlo (PIMC) simulations.[26–29] In these PIMD/PIMC simulations, the NQEs are accounted for on the same footing as the thermal effects when the statistical properties are evaluated, as will be explained later in Section 2. Therefore, when comparing their results with the ones obtained in the MD simulations, in which the nuclei are treated classically, the NQEs can be addressed in a very clean manner.

These comparisons set up a rigorous framework for the NQEs to be investigated statistically, which is still used nowadays. However, in these early PIMD and PIMC simulations, researches often resorted to empirical potentials for descriptions of the interatomic interactions. When chemical reactions occur, e.g., a proton transfers along a HB, such a simple parametrization of the interatomic interactions may easily fail. To cope with these problems, in which the impact of the NQEs is usually more interesting, an effort toward combining the *ab initio* electronic structure calculations with the PIMD/PIMC sampling techniques has been made since the mid 1990s, first within the framework of Car–Parrinello (CP) MD (see e.g., Tuckerman, Marx, and Parrinello *et al.*'s work in Refs. [1], [30], and [31]) and later by directly using Born–Oppenheimer (BO) MD/MC.[6,7,32–34] These methods allow for bond making and breaking events, as well as the thermal and quantum nuclear fluctuations to be addressed in a seamless manner based on the forces computed "on-the-fly" as the system evolves, therefore making PIMD/PIMC simulations for problems with complicated chemical environments possible. Currently, not only different functionals within the density-functional theory (DFT)[35–37] but also many traditional quantum chemistry methods such as MP2[38,39] can be used in descriptions of the electronic structures and consequently the interatomic interactions. With such a wide range of choices for the electronic structures, one can safely rely on the results obtained from such simulations, if the accuracy of the interatomic interaction description and the sampling sufficiency are ensured. We note, however, that such an ensurance is far from being trivial.[30,35,36]

The above-mentioned recipe allows the NQEs to be described in a clean manner, if the molecular simulation time is long enough to address the issue of interest. A common problem for most molecular simulation techniques, however, is that the timescale one can afford is not long enough to allow the system to travel between different stable states, thereby hindering a direct evaluation of the free-energy differences between different meta-stable states. In such cases, to determine the relative stability of these different meta-stable states without resorting to the coarse-grained models,[40] the method one often chooses is to enhance the sampling efficiency either by imposing constraints or modifying the effective PES.[41–44] However, if it is the case that within the molecular simulation time, the system can stay within the local minima of the PES with all spatial configurations close to it explored, a comparison of the free energies associated with each local minimum directly without requiring the system to travel between different meta-stable states provides an alternative. In such cases, a method with which the free energy of each meta-stable state can be calculated is highly desired. When the influence of NQEs on these free energies needs to be quantified, a combination between the thermodynamic integration (TI) method and PIMD is a natural choice.[45,46] In the following, in addition to a general introduction to *ab initio* PIMD and its application to the studies of some H-bonded systems, a combination of the TI method with *ab initio* PIMD along with applications of this method to the free-energy calculations of some H-bonded systems will also be discussed.

The paper is organized as follows. An introduction to the *ab initio* PIMD method and its combination with TI, as well as the computational setups, are given in Section 2. In Section 3, we take a basic question in physics and chemistry, i.e., what is the impact of NQEs on the strength of HB, as an example to show how the *ab initio* PIMD method can be used to study such problems in H-bonded systems. As the second example, we quantitatively analyze the influence of NQEs on the free energy of some gas phase water clusters using a combination of *ab initio* PIMD and TI methods in Section 4. Conclusions and a short perspective are given in Section 5.

2. Methods

2.1. *Ab initio* path-integral molecular dynamics

Using the electronic structures and consequently the interatomic interactions calculated "on-the-fly", the basic idea of the *ab initio* PIMD method is to treat the parameter $\beta = 1/(k_{\mathrm{B}}T)$ associated with a finite temperature T as an imaginary time and represent quantum mechanically the density matrix of the nuclei at this finite T with an artificial polymer, through

$$\begin{aligned}
&\rho(\boldsymbol{x}_0, \boldsymbol{x}_P, \beta) \\
&= \lim_{P\to\infty}\left[\prod_{j=1}^{N}\left(\frac{m_j P}{2\beta\pi\hbar^2}\right)^{P/2}\right] \\
&\times \int_V\int_V\cdots\int_V \exp\left\{-\beta\sum_{i=1}^{P}\left[\sum_{j=1}^{N}\frac{1}{2}m_j\omega_P^2\left(\boldsymbol{x}_i^j-\boldsymbol{x}_{i-1}^j\right)^2\right.\right. \\
&\left.\left.+\frac{1}{2P}(V(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N)\right.\right.
\end{aligned}$$

$$+V(\boldsymbol{x}_{i-1}^1,\ldots,\boldsymbol{x}_{i-1}^N))\Big]\Big\}\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2,\ldots,\mathrm{d}\boldsymbol{x}_{P-1}. \quad (1)$$

Here, k_{B} is the Boltzman constant, and P is the number of replicas in the artificial polymer, which corresponds to the number of sampling points along the imaginary time path at this T. When $P=1$, the above equation reaches its classical limit, which is equivalent to an *ab initio* MD simulation. When it goes to infinite, this equation rigorously approaches its quantum limit. In practice, one needs to choose a finite P, which must be converged with respect to the physical property being investigated. m_j means the mass of the j-th nucleus, and ω_P equals $\sqrt{P}/(\beta\hbar)$. $\boldsymbol{x}_i$ means the spatial configuration of the nuclei on the i-th image (otherwise called bead) of the artificial polymer, and $\boldsymbol{x}_i^j$ is the Cartesian coordinate of the j-th atom on this image. In association with such a representation, $V(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N)$ means the potential energy of the nuclei in the i-th image of the polymer.

In Eq. (1), if we take $\boldsymbol{x}_0=\boldsymbol{x}_P$ (set the path as closed), the density matrix evolves into its diagonal part, i.e., the quantum density function. If one further integrates out the configurational space of the nuclei, the partition function of the quantum system can be obtained

$$Z^{\mathrm{Q}} = \lim_{P\to\infty}\left[\prod_{j=1}^{N}\left(\frac{m_jP}{2\beta\pi\hbar^2}\right)^{\frac{P}{2}}\right] \times\int_V\int_V\cdots\int_V \mathrm{e}^{-\beta V^{\mathrm{eff}}(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N)}\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{x}_2\cdots\mathrm{d}\boldsymbol{x}_P, \quad (2)$$

where

$$V^{\mathrm{eff}}(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N) = \sum_{i=1}^{P}\sum_{j=1}^{N}\frac{1}{2}m_j\omega_P^2\left(\boldsymbol{x}_i^j-\boldsymbol{x}_{i-1}^j\right)^2+\sum_{i=1}^{P}\frac{1}{P}V(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N). \quad (3)$$

If the partition function of the quantum system is known, one can obtain all its statistical properties. Pictorially, such a mapping of the quantum system to the classical polymer can be understood by a comparison as shown in Fig. 1 for the simplest molecule H_2. The canonical partition function of the quantum system is what we want to know. From the principles underlying the path-integral representation of the statistical mechanics as given in Ref. [22], one can construct an artificial classical polymer to simulate the density matrix of the quantum system through Eq. (1). The density function is the diagonal part of the density matrix. For the H_2 molecule, this polymer is composed of P replicas of the real system (in this case, H_2 molecule). In between the neighboring replicas, the same atom is connected by a spring interaction, whose spring constant (defined as $m_j\omega_P^2$ in Eq. (1) is determined by m_j and ω_P. Within one replica, the interatomic interaction is represented as $V(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N)$. If the PIMD simulation is based on a force field, this interaction is given by the empirical potentials. If it is *ab initio*, this interaction is calculated "on-the-fly" as the system evolves, as will be discussed later. The spring interaction and the intra-replica potential correspond to the kinetic and the potential energies of the path-integral, respectively. If the temperature is very high or the mass of the nucleus is very large, the spring interaction between the neighboring beads of this atom is very strong, and these beads will be dragged into a single point in real space. In such cases, for this nucleus, the quantum description reaches its classical limit. When the temperature is relatively low and the mass of the nucleus is small, the PIMD simulations of such a polymer will definitely present results different from those of the MD simulations (or the one with $P=1$ in Eq. (1). The NQEs will play a role in describing the properties of the system under investigation.

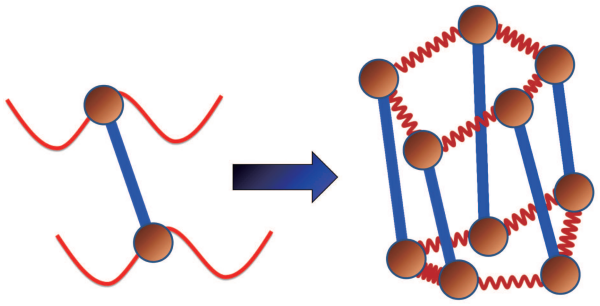


Fig. 1. Illustration of how the mapping from the canonical quantum system to a classical polymer is done in the path-integral statistical mechanics. The polymer is composed of P replicas of the real molecule. In each replica, the potential is determined by the potential of the system at the specific spatial configuration of this replica. In between the replicas, the neighboring images (beads) of the same atoms are linked by springs. The spring constant is determined by m_j and ω_P as $m_j\omega_P^2$, where $\omega_P=\sqrt{P}/(\beta\hbar)$. Therefore, the higher the temperature and the heavier the nucleus, the stronger the interaction between the beads. In the limit of $T\to\infty$ and $m_j\to\infty$, one arrives at the classical limit when all images overlap with each other. The partition function of the quantum system as shown on the left equals the configurational partition function of the polymer on the right as $P\to\infty$.

2.2. Thermodynamic integration

Using the method described above, by comparing the results obtained from the *ab initio* MD and the *ab initio* PIMD simulations, the influence of NQEs can be analyzed in a very clean manner. However, as mentioned, the molecular simulation time needs to be sufficiently long to address the problem we are interested in. When the timescale one can afford is not long enough to allow the system to travel between different meta-stable states, direct evaluation of the free-energy differences between them becomes technically difficult. One method that is often chosen is to enhance the sampling efficiency either by imposing constraints or modifying the effective PES.[41–44] But as said, if it is the case that within the molecular simulation time, the system stays within each meta-stable state with all spatial configurations close to it explored, a comparison of the free energies of the meta-stable states without requiring the system to travel between them provides an alternative. In this case, one needs a method with which the

free energy associated with each local minimum can be calculated. Here, we choose the TI method for such calculations.

The basic idea behind the TI method is that a system with unknown free energy, which we call the reference system, can be connected to a system with known free energy, which we call the real system, by varying one or more parameters. There are several choices for the reference system and the real system, depending on which property we are investigating. In the present work, we want to investigate the influence of NQEs on the free energy of the polyatomic system. Therefore, we have chosen the reference system as the one with classical nuclei, and the real system as the one with quantum nuclei. For the real system, the partition function follows Eqs. (2) and (3) and the free energy of this real system is defined as

$$F^{\mathrm{Q}} = -\frac{1}{\beta}\ln Z^{\mathrm{Q}}. \tag{4}$$

We note, however, that direct calculation of the free energy from this equation is impractical. A practical method of choice is to take a reference system whose effective potential can be expressed as

$$V^{\mathrm{eff}}(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N) = \sum_{i=1}^{P}\sum_{j=1}^{N}\frac{1}{2}m_j\omega_P^2\left(\boldsymbol{x}_i^j-\boldsymbol{x}_{i-1}^j\right)^2 + \sum_{i=1}^{P}\frac{1}{P}V(\boldsymbol{x}_{\mathrm{c}}^1,\ldots,\boldsymbol{x}_{\mathrm{c}}^N), \tag{5}$$

where $\boldsymbol{x}_{\mathrm{c}}^j$ means the Cartesian coordinate of the centroid of the j-th nucleus. For a system with effective potential given by Eq. (5), the free energy rigorously equals that of the polyatomic system with classical nuclei (for a detailed derivation, please refer to pages 191–194 in Ref. [47]). We denote the free energy of this reference system as F^{C} and assume that it is already known. Some practical routines for the calculation of this energy can be found in Refs. [48]–[52].

In between the reference system whose effective potential is given by Eq. (5) and the real system whose effective potential is given by Eq. (3), one can introduce a series of intermediate systems, whose effective potential equals

$$V^{\mathrm{eff}}(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N;\lambda) = \sum_{i=1}^{P}\sum_{j=1}^{N}\frac{1}{2}m_j\omega_P^2\left(\boldsymbol{x}_i^j-\boldsymbol{x}_{i-1}^j\right)^2 + \sum_{i=1}^{P}\frac{1}{P}\left[\lambda^2 V(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N)+(1-\lambda^2)V(\boldsymbol{x}_{\mathrm{c}}^1,\ldots,\boldsymbol{x}_{\mathrm{c}}^N)\right]. \tag{6}$$

From this definition, it is clear that the intermediate system reaches the end of the reference (real) one when $\lambda = 0$ ($\lambda = 1$). By putting Eq. (6) into the partition function as given in Eq. (2), one obtains the partition function $Z(\lambda)$ of the intermediate system. The free energy of the intermediate system is defined as

$$F(\lambda) = -\frac{1}{\beta}\ln Z(\lambda). \tag{7}$$

We note that $F(1) = F^{\mathrm{Q}}$, which is unknown, and $F(0) = F^{\mathrm{C}}$, which is known. The difference between them tells us the influence of the NQEs on the free energy of the polyatomic system. Since this $F(\lambda)$ is a continuous function of λ, this influence of the NQEs on the free energy of the polyatomic system can also be calculated as

$$\Delta F = F(1) - F(0) = \int_0^1 \mathrm{d}\lambda F'(\lambda), \tag{8}$$

where $F'(\lambda)$ means the derivative of $F(\lambda)$ with respect to λ. By putting the expression of $Z(\lambda)$ into Eq. (7) and making derivation of the free energy to λ, one obtains

$$F'(\lambda) = \left\langle \frac{2\lambda}{P}\sum_{i=1}^{P}\left[V(\boldsymbol{x}_i^1,\ldots,\boldsymbol{x}_i^N)-V(\boldsymbol{x}_{\mathrm{c}}^1,\ldots,\boldsymbol{x}_{\mathrm{c}}^N)\right]\right\rangle_{V^{\mathrm{eff}}(\lambda)}. \tag{9}$$

Here, we note that the symbol $\langle\cdots\rangle_{V^{\mathrm{eff}}(\lambda)}$ means the ensemble average of the quantity inside it generated using the effective potential as given by Eq. (6) of each λ. Thus, the influence of NQEs on the free energy of the polyatomic system can be calculated in a very clean manner.

2.3. Computational details

Two separate works will be reviewed in this manuscript, i.e., (i) the investigation of the influence of NQEs on the structural change of some H-bonded systems, and (ii) the influence of NQEs on the free energy of some water clusters. They were carried out using different codes. But in both of them, the Perdew–Burke–Ernzerhof (PBE) functional was used in the description of the electronic exchange–correlation interactions.[53] In the following, we summarize their other computational details one after the other.

For investigation of the impact of NQEs on the HB strength (using the geometric property changes as an indicator), the simulations were carried out using the CASTEP plane-wave density-functional theory (DFT) code.[54] The ultrasoft pesudopotentials were taken along with a 400 eV energy cutoff for the expansion of the wavefunction. For simulation of the water and HF clusters, a 12 Å cubic cell was used. For the organic dimers and charged water clusters, an 18 Å cubic cell was taken. For solid HF and HCl, an eight molecule supercell along with a $4\times2\times3$ Monkhorst–Pack k-point mesh was used. For the solid squaric acid, a 20-atom cell along with a $3\times3\times2$ Monkhorst–Pack k-point mesh was chosen. Time steps of 0.5–0.7 fs were used, and after thermalization, 10000 MD and 6000 PIMD steps were performed. All the MD and PIMD simulations reported for this part of the work were performed at a targeting temperature of 100 K, except in solid HCl ($N_2H_5^-$), where 50 K (300 K) was used to keep the

system stable (compared with other simulations). The simulations were carried out using the constant-volume, constant-temperature (NVT) ensemble, with the Langevin thermostat employed in controlling the temperatures. In the PIMD simulations, 16 beads were used to sample the imaginary time path-integral, along with the staging method to decouple the movement of the neighboring beads.[55] For more computational details, please refer to Ref. [7].

For the free-energy change due to NQEs in water clusters, a self-developed combination of the TI and the *ab initio* PIMD method within the Vienna *ab initio* simulation package (VASP) was used.[56,57] Projector-augmented wave potentials were employed along with a 600 eV plane wave cutoff energy for the expansion of the Kohn–Sham orbitals. The Andersen thermostat was used to control the temperature of the NVT ensemble.[58] The temperature was set as 100 K, safely with the stability of the clusters. 48 beads were used to sample the imaginary time path-integral with a time step of 0.5 fs. After thermalization, 30000 PIMD steps were used at each λ along the TI line to obtain the $F'(\lambda)$. For more computational details, please refer to Ref. [57].

3. Influence of NQEs on the structure of HBs

The first example of the *ab initio* PIMD simulations, in which the quantum nature of the nuclei plays an important role, concerns a fundamental question in physics and chemistry, i.e., what is the influence of NQEs on the strength of HBs? We start by using the structural properties, namely, the heavy-atom distances, as an indicator. Further investigations on the energetics will be discussed in Section 4. As mentioned earlier in this manuscript, it has been known for more than half a century that, in H-bonded molecular crystals, by replacing H with D, the strength of HBs and consequently the crystal structural parameters change. This phenomenon is known as the Ubbelohde effect.[8,9] The conventional Ubbelohde effect reflects an increase of the heavy-atom distance associated with the HB upon replacing H by D, which indicates that NQEs strengthen this intermolecular interaction. We note, however, that a negative Ubbelohde effect in which a decrease of the heavy-atom distance happens upon replacing H by D has also been observed in several materials.[8,9] From the perspective of molecular simulation, since the 1980s, when PIMD/PIMC becomes a conventional routine to investigate the NQEs in real polyatomic systems, a bunch of computer simulations have been carried out in a wide range of systems. A general conclusion is that the final effect is system-dependent. For example, in liquid HF, *ab initio* MD and PIMD simulations have shown that when NQEs are accounted for, the first peak of the F–F radial distribution function (RDF) sharpens and shifts to a shorter F–F distance,[59] indicating that NQEs strengthen the HB in this system. In contrast, a similar simulation for liquid water shows that the O–O RDF gets less structured when NQEs are included, indicating that NQEs weaken the HB in liquid water.[2] We note, however, that although this conclusion is probably right, it is the opposite of what has been reported in Ref. [5].

Besides this discussion concerning HBs in condensed phases, the impact of NQEs on the strength of HBs has also been widely discussed in gas-phase clusters.[59–61] For water clusters up to the hexamer, it was predicted that the NQEs weaken the HBs, as reflected by the change of the heavy-atom distances in Refs. [60] and [61]. For HF clusters, however, it was predicted that NQEs can strengthen or weaken the HBs depending on the size of the clusters.[59] For clusters smaller than tetramer, the NQEs weaken this intermolecular interaction, whereas for clusters larger than it, the strengthening of the HB is observed. In tetramer, the NQEs are negligible. With these system-dependent conclusions in mind, it is clear that a simple picture in which the influence of the NQEs on the strength of HBs and consequently the structural or energetic properties can be rationalized is highly desired. In Ref. [7], we have presented a simple picture which can be used to rationalize these different results. *Ab initio* PIMD simulations were used as the basic tool upon which the analysis was carried out. Here, using these simulations as an example, we will review this work in a way that how the *ab initio* MD and PIMD simulations are used to answer this question can be understood in a easy manner.

To be as unbiased as possible, a wide range of H-bonded systems have been chosen, including HF clusters (dimer to hexamer), H_2O clusters (dimer, pentamer, and octamer), charged, protonated, and hydroxylated water and ammonia clusters ($H_9O_5^-$, $H_9O_4^+$, $H_7O_4^-$, and $N_2H_5^-$), organic dimers (formic acid and formamide), and solids (HF, HCl, and squaric acid $C_4H_2O_4$). For each system, both *ab initio* MD simulations, in which the nuclei are classical particles, and *ab initio* PIMD simulations, in which the nuclei are quantum particles, are carried out. By comparing the difference of the statistical structures obtained from these two simulations, one can analyze the influence of NQEs in a very clean manner. The quantities we focus on in characterizing the HBs include: (i) the heavy-atom distances (denoted by X–X, where X is either O, Cl, C, N, or F), which describe the intermolecular separations; (ii) the H-bond angles (X–H$\cdots X$), which characterize the HB bending (libration); and (iii) the X–H covalent bond lengths, which describe the stretching in the H-bond donor molecules. These quantities provide an indication of H-bond strength, as will be explained later in the discussion. As the main measure of H-bond strength, we use a standard estimator based on the computed red shift (softening) in the X–H stretching frequency of the HB donor molecule. We note that there is no

perfect measure for the HB strength;[62] however, the red shift of the stretching frequency is a widely used measure [see, e.g. Refs. [63] and [64]). In Fig. 2(b), it is shown that this estimator correlates well with the computed binding energy per HB in the neutral systems we study, which is defined as the difference between the total energy of the system and the sum over its unrelaxed components.[62] The larger the red shift of the stretching frequency (measured as the ratio of the X–H stretching frequency in the H-bonded cluster to that in the free monomer), the stronger the HB.

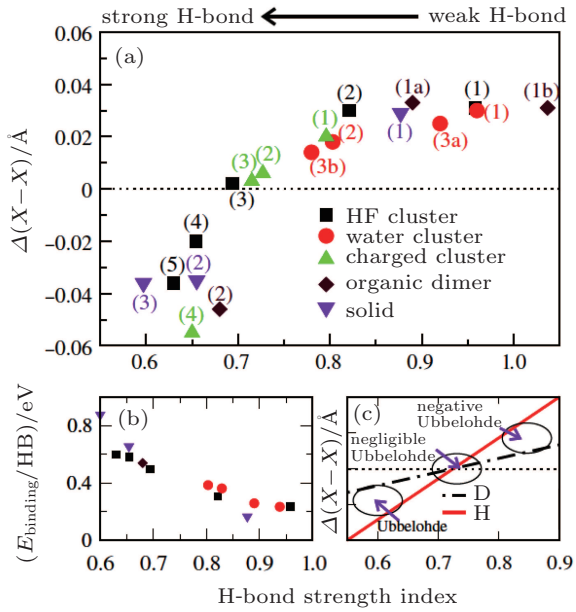


Fig. 2. (a) The differences between the shortest heavy-atom distances obtained from the PIMD and MD simulations $(X-X)^{\text{PIMD}}_{\text{average}} - (X-X)^{\text{MD}}_{\text{average}}$, denoted by $\Delta(X-X)$, are plotted as a function of the HB strength index. $\Delta(X-X)$ characterizes the impact of the NQEs on the strength of the HBs. The HB strength is defined as the ratio of the X–H stretching frequency in the H-bonded system to that in the free monomer. (b) The correlation between this HB strength index and the binding energy per HB in the neutral systems. (c) A simplified schematic illustration of the isotope (Ubbelohde) effect. We suggest that regimes of positive, negligible, and negative Ubbelohde effects exist depending on the HB strength. For the HF clusters, labels 1–5 denote the HBs in the dimer to the hexamer. For the H_2O clusters, labels 1, 2, 3a, and 3b refer to the HBs in the dimer, pentamer, and the long and short HBs in the octamer, respectively. For the charged clusters, labels 1–4 refer to $H_9O_5^-$, $H_9O_4^+$, $H_7O_4^-$, and $N_2H_5^-$. For the organic dimers, labels 1a, 1b, and 2 mean the red shifted and blue shifted HBs in the formamide and the red shifted HB in formic acid. For the solids, labels 1–3 refer to the HBs in HCl, HF, and squaric acid. The same labels also apply to Fig. 4.[7]

With the definition of the above-mentioned quantities in mind, we first look at the results for the impact of the NQEs on the strength of HBs. Upon comparing these results for the various H-bonded systems, an interesting correlation can be observed between the H-bond strength and the change in intermolecular separations, as shown in Fig. 2(a). Here, it can been seen that as the HB gets stronger, the heavy-atom separations in the PIMD simulations go from being longer than those in the MD simulations (positive $\Delta(X-X)$) to being shorter (negative $\Delta(X-X)$). In other words, the NQEs lead to longer HBs in weak H-bonded systems and shorter HBs in relatively strong H-bonded systems. We note that the HB strength increases upon going from small to large clusters and from water to HF. This trend is the main finding, and in the following we explain why it exists and discuss the implications it has for H-bonded systems in general.

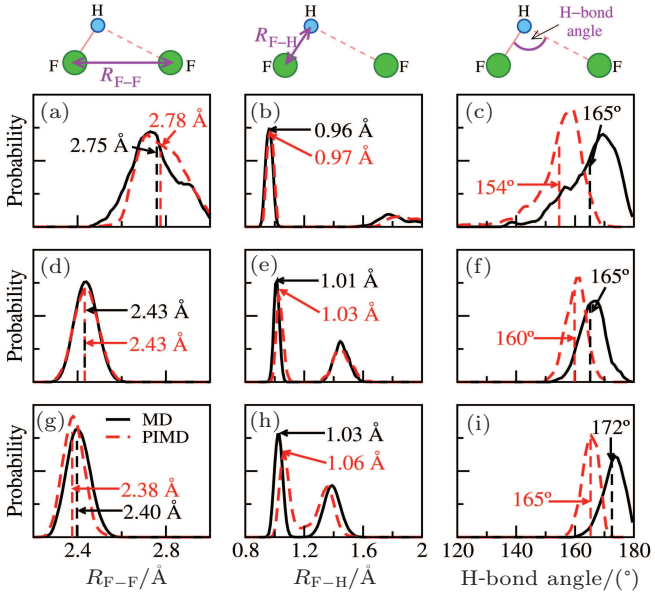


Fig. 3. HF clusters as examples for detailed analysis of the QNEs. Distributions of the F–F distances (left), the F–H bond lengths (center), and the intermolecular bending (F–H···F angle, right) from MD (solid black lines) and PIMD (dashed red lines) for a selection of systems: the HF dimer (top), the HF tetramer (middle), and the HF pentamer (bottom). The MD and PIMD averages are shown in black and red vertical dashes, respectively.[7]

To this end, we take the HF clusters (dimer, tetramer, and pentamer) as the guiding example, because upon increasing the cluster size, the HB strength increases, and the influence of the NQEs switches from a tendency to lengthen to a tendency to shorten the intermolecular separations (as seen in Ref. [65]). The results are summarized in Fig. 3, where we plot the distance and angle distributions from MD and PIMD simulations for these three clusters separately. The left column shows the final results, where we can see that, in the dimer, the averaged F–F distance is increased by including the NQEs. In the tetramer, there is no difference between the averaged quantum and classical F–F distances. While in the pentamer, the F–F distance is clearly shortened by including the NQEs. The key to understanding this variation of the heavy-atom distances is realizing that there are also related differences between MD and PIMD in the covalent F–H bond lengths (center) and H-bond angles (right). Because of anharmonic quantum fluctuations, these two geometric properties also show systematic changes. The F–H bonds are longer in the quantum (PIMD) than in the classical (MD) simulations, and this elongation becomes more pronounced as the HBs get stronger. Besides this, the HBs are more bent in the quantum than in the classical simulations, and this bending generally becomes less pronounced

as the HBs get stronger. In order to understand the influence of these variations in structure, analysis of various dimer configurations was performed. This analysis reveals that the covalent bond stretching increases the intermolecular interaction, whereas the HB bending decreases it. Taking the HF dimer as an example, a 0.04 Å increase in the F–H bond length of the donor leads to a 40 meV increase in interaction energy within the dimer; in contrast, a 21° reduction in the H-bond angle leads to a 16 meV decrease in interaction energy. In short, the F–F distance increases in the dimer as a result of a large decrease in the HB angle, but only a small increase in the covalent F-H bond length is noted upon including the NQEs. Whilst in the tetramer, the F–H stretching is sufficiently pronounced to compensate for the increase in HB bending, leaving the overall F–F distance unchanged; and in the pentamer, the F–F distance decreases because the F–H covalent bond stretching dominates over the HB bending.

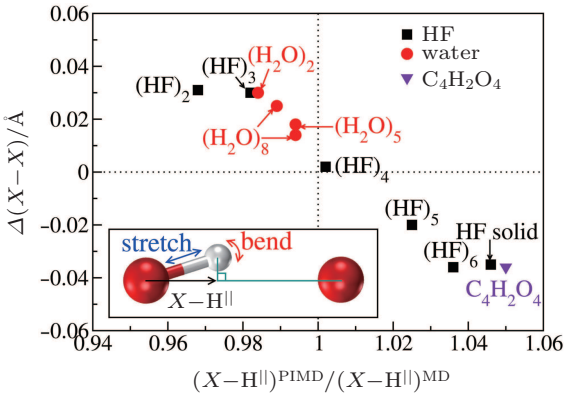


Fig. 4. Quantification of the competition between the quantum fluctuations on the stretching and bending modes. The differences in average shortest heavy-atom distances between PIMD and MD simulations ($\Delta(X-X)$, vertical axis) are plotted as a function of the ratio of the projection of the donor X–H covalent bond along the intermolecular axis from PIMD and MD simulations (horizontal axis). x larger (smaller) than 1 indicates a dominant contribution from the stretching (bending) mode when the NQEs are included. Negative values of $\Delta(X-X)$ indicate that NQEs decrease the intermolecular separation. An almost linear correlation between the two variables is observed. The inset illustrates the geometry used for projecting the donor covalent X–H bond onto the intermolecular axis. The curved red arrow represents the intermolecular bending and the straight blue arrow represents the intramolecular stretching.[7]

The above analysis provides a qualitative understanding of the trend observed. For a more rigorous examination of this picture and a quantitative description of this competition for all systems studied, we further calculate the projection (X–H$^{||}$) of the donor molecule's covalent bond along the intermolecular axis (see the inset of Fig. 4). Since X–H$^{||}$ increases upon intramolecular stretching but decreases upon intermolecular bending, this quantity itself allows the balance between stretching and bending to be evaluated to a certain extent. The influence of the NQEs is quantified by the ratio of the PIMD and MD projections, i.e., $x = (X-\mathrm{H}^{||})^{\mathrm{PIMD}}/(X-\mathrm{H}^{||})^{\mathrm{MD}}$. When this value is clearly greater than one, it indicates that when the NQEs are included, the main influence is on the stretching of the covalent bond, and when this value is clearly smaller than one, it indicates that when the NQEs are included, the main influence is on the bending of the HB. When one plots this ratio against the variations in intermolecular separations, $y = \Delta(X-X)$ (which we used to quantify the impact of the NQEs), a striking (almost linear) correlation is observed (Fig. 4). For all systems in which HB bending dominates (x clearly smaller than 1), the heavy-atom distances are longer in PIMD than in MD ($y > 0$). In cases where covalent bond stretching is dominant (x clearly larger than 1), the heavy-atom distances are shorter in PIMD than in MD ($y < 0$). With the increase of x, quantum fluctuations on the stretching mode become more dominant and the NQEs switch from weakening the hydrogen bonds to strengthening them. Thus, the overall influence of the NQEs on the HB interaction quantitatively comes down to this delicate interplay between covalent bond stretching and intermolecular bond bending. One notes that this explanation for the general case is consistent with what Manolopoulos and coworkers have elegantly shown for liquid water in Ref. [4].

4. NQEs on the free energy of water clusters

In Section 3, we have analyzed the influence of NQEs on the strength of HBs, taking the geometric changes as the indicator. A rigorous quantification of this influence, however, requires an explicit calculation of the contribution of the NQEs to the free energy of the H-bonded systems, using the method shown in Section 2. In Fig. 5, we use two examples to show how such TI simulations for the calculation of $\Delta F[n\mathrm{H_2O}]/n$ are carried out. The left panel corresponds to the results of water monomer and the right panel corresponds to those of the dimer. $F'(\lambda)$ is as defined in Eq. (9), but divided by the number of water molecules in the simulation cell. Integration of this quantity from 0 to 1 gives us the final $\Delta F[n\mathrm{H_2O}]/n$. The temperature is 100 K. After integration, $\Delta F[n\mathrm{H_2O}]/n$ equals 0.446 eV and 0.462 eV respectively in the monomer and dimer. To separate out the harmonic and anharmonic effects, we can perform a separate phonon frequency calculation for each system and integrate out the free-energy change due to NQEs at the harmonic limit through

$$\Delta F_{\mathrm{harm}} = k_{\mathrm{B}}T\int_0^{\infty} g(\omega)\ln\left[2\sinh\left(\frac{\hbar\omega}{2k_{\mathrm{B}}T}\right)\right]. \qquad (10)$$

This gives us 0.476 eV and 0.491 eV respectively for the monomer and dimer. The anharmonic effect decreases $\Delta F[n\mathrm{H_2O}]/n$ by $\sim$ 30 meV in both systems, which is totally physical and non-negligible when more delicate analysis is carried out. Numerical simulations for more systems and deeper analysis of these results are still ongoing, which we will present in Ref. [57]. Here, we just use these two examples to show how these data are generated in practice.

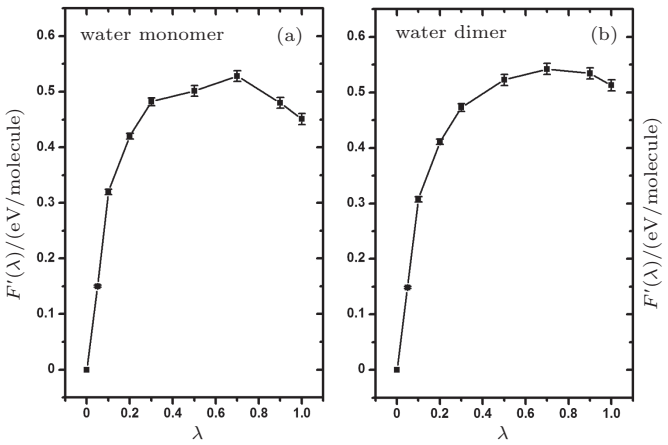


Fig. 5. Thermodynamic integration based on *ab initio* PIMD sampling for the calculation of the free-energy change due to NQEs in the water monomer and dimer. $F'(\lambda)$ is as defined in Eq. (9), but divided by the number of water molecules in the simulation cell. Integration of this quantity from 0 to 1 gives us the final $\Delta F[n\mathrm{H_2O}]/n$.[57]

5. Conclusion and perspectives

HB is generally weak, but ubiquitous and essential to life on earth. The small mass of hydrogen means that many properties of HBs are quantum mechanical in nature, although quite often their theoretical descriptions remain classical. Due to the development of computer simulation methods, this influence of NQEs on the structural and energetic properties of some hydrogen bonded systems has been systematically studied in recent years. Here we present a review of these works, by focussing on the explanation of the main methods behind the simulations, i.e., the *ab initio* PIMD method as well as its extension in combination with the thermodynamic integration method for the calculation of the free energies. As a practical example of these studies, *ab initio* PIMD simulations on the structural properties of the hydrogen bonded systems have shown that the NQEs have a tendency to strengthen the strong HBs and weaken the weak ones, which could be explained by a competition of the quantum fluctuations on the stretching and bending modes.

Acknowledgement

The computational resources were provided by the supercomputer TianHe-1A in Tianjin, China.

References

[1] Marx D, Tuckerman M E, Hutter J and Parrinello M 1999 *Nature* **397** 601
[2] Marrone J A and Car R 2008 *Phys. Rev. Lett.* **101** 017801
[3] Fanourgakis G S, Schenter G K and Xantheas S S 2006 *J. Chem. Phys.* **125** 141102
[4] Habershon T E, Markland S and Manolopoulos D E 2009 *J. Chem. Phys.* **131** 024501
[5] Chen B, Ivanov I, Klein M L and Parrinello M 2003 *Phys. Rev. Lett.* **91** 215503
[6] Li X Z, Probert M I J, Alavi A and Michaelides A 2010 *Phys. Rev. Lett.* **104** 066102
[7] Li X Z, Walker B and Michaelides A 2011 *Proc. Natl. Acad. Sci. USA* **108** 6369
[8] Matsushita E and Matsubara T 1982 *Prog. Theor. Phys.* **67** 1
[9] Ubbelohde A R and Gallagher K J 1955 *Acta. Crystallogr.* **8** 71
[10] Marrone J A, Lin L and Car R 2009 *J. Chem. Phys.* **130** 204511
[11] Burnham C J, Anick D J, Mankoo P K and Reiter G F 2008 *J. Chem. Phys.* **128** 154519
[12] Meng X Z 2015 *Nat. Phys.* **11** 235
[13] Zhang D H, Collins M A and Lee S Y 2000 *Science* **290** 961
[14] Liu K 2001 *Annu. Rev. Phys. Chem.* **52** 139
[15] Collins M A 2002 *Theo. Chem. Acc.* **108** 313
[16] Althorpe S C and Clary D C 2003 *Annu. Rev. Phys. Chem.* **54** 593
[17] Qiu M H 2006 *Science* **311** 1440
[18] Feynman R P 1949 *Phys. Rev.* **76** 769
[19] Feynman R P 1953 *Phys. Rev.* **90** 1116
[20] Feynman R P 1953 *Phys. Rev.* **91** 1291
[21] Feynman R P 1953 *Phys. Rev.* **91** 1301
[22] Feynman R P and Hibbs A R 1965 *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill Inc.)
[23] Chandler D and Wolynes P G 1981 *J. Chem. Phys.* **74** 4078
[24] Parrinello M and Rahman A 1984 *J. Chem. Phys.* **80** 860
[25] Berne B J and Thirumalai D 1986 *Annu. Rev. Phys. Chem.* **37** 401
[26] Pollock E L and Ceperley D M 1984 *Phys. Rev. B* **30** 2555
[27] Ceperley D M and Pollock E L 1986 *Phys. Rev. Lett.* **56** 351
[28] Pollock E L and Ceperley D M 1987 *Phys. Rev. B* **36** 8343
[29] Ceperley D M 1995 *Rev. Mod. Phys.* **67** 279
[30] Tuckerman M E, Marx D, Klein M L and Parrinello M 1996 *J. Chem. Phys.* **104** 5579
[31] Marx D and Parrinello M 1996 *J. Chem. Phys.* **104** 4077
[32] Morales M A, Pierleoni C, Schwegler E and Ceperley D M 2010 *Proc. Natl. Acad. Sci. USA* **107** 12799
[33] Li X Z 2013 *J. Phys.: Condens. Matter* **25** 085402
[34] Chen J 2013 *Nat. Commun.* **4** 2064
[35] Morales M A, McMahon J M, Pierleoni C and Ceperley D M 2013 *Phys. Rev. Lett.* **110** 065702
[36] Morales M A, McMahon J M, Pierleoni C and Ceperley D M 2013 *Phys. Rev. B* **87** 184107
[37] Davidson E R M, Klimes J, Alf'e D and Michaelides A 2014 *ACS Nano* **8** 9905
[38] Tachikawa M and Shiga M 2005 *J. Am. Chem. Soc.* **127** 11908
[39] Kaczmarek A, Shiga M and Marx D 2009 *J. Phys. Chem. A* **113** 1985
[40] Tozzini V 2005 *Curr. Opin. Struct. Biol.* **15** 144
[41] Bartels C and Karplus M 1998 *J. Phys. Chem. B* **102** 865
[42] Sugita Y and Okamoto Y 1999 *Chem. Phys. Lett.* **314** 141
[43] Laio A and Parrinello M 2002 *Proc. Natl. Acad. Sci. USA* **99** 12562
[44] Gao Y Q, Yang L J, Fan Y B and Shao Q 2008 *Int. Rev. Phys. Chem.* **27** 201
[45] Pérez A and von Lilienfeld O A 2011 *J. Chem. Theory Comput.* **7** 2358
[46] Habershon S and Manolopoulos D E 2011 *J. Chem. Phys.* **135** 224111
[47] Li X Z and Wang E G 2014 *Computer Simulations of Molecules and Condensed Matters: From Electronic Structures to Molecular Dynamics* (Beijing: Peking University Press)
[48] Stroobants A, Lekkerkerker H N W and Frenkel D 1987 *Phys. Rev. A* **36** 2929
[49] Frenkel D, Lekkerkerker H N W and Stroobants A 1988 *Nature* **332** 822
[50] Meijer E J, Frenkel D, LeSar R A and Ladd A J C 1990 *J. Chem. Phys.* **92** 7570
[51] Meijer E J and Frenkel D 1991 *J. Chem. Phys.* **94** 2269
[52] Eldridge M D, Madden P A and Frenkel D 1993 *Nature* **365** 35
[53] Perdew J P, Burke K and Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
[54] Clark S J 2005 *Z. Kristallogr* **220** 567
[55] Tuckerman M E, Berne B J, Martyna G J and Klein M L 1993 *J. Chem. Phys.* **99** 2796
[56] Feng Y X, Chen J, Alfè D, Li X Z, Wang E G, *et al.* 2015 *J. Chem. Phys.* **142** 064506
[57] Feng Y X *et al.* to be submitted
[58] Andersen H C 1980 *J. Chem. Phys.* **72** 2384
[59] Raugei S and Klein M L 2003 *J. Am. Chem. Soc.* **125** 8992
[60] Gregory J K and Clary D C 1996 *J. Phys. Chem.* **100** 18014
[61] Clary D C, Benoit D M and van Mourik T 2000 *Acc. Chem. Res.* **33** 441
[62] Wendler K, Thar J, Zahn S and Kirchner B 2010 *J. Phys. Chem. A* **114** 9529
[63] Xantheas S S and Dunning T H 1993 *J. Chem. Phys.* **99** 8774
[64] Cubero E, Orozco M, Hobza P and Luque F J 1999 *J. Phys. Chem. A* **103** 6394
[65] Swalina C, Wang Q, Chakraborty A and Hammes-Schiffer S 2007 *J. Phys. Chem. A* **111** 2206