JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Letter

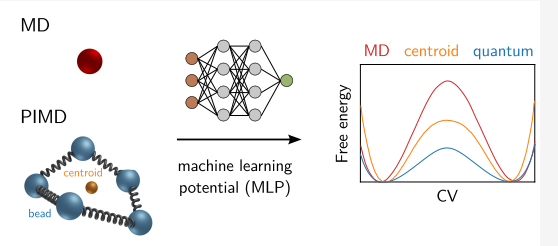
Quantum Free Energy Profiles for Molecular Proton Transfers

Aran Lamaire, Maarten Cools-Ceuppens, Massimo Bocus, Toon Verstraelen, and Veronique Van Speybroeck*

Cite This: *J. Chem. Theory Comput.* 2023, 19, 18−24 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Although many molecular dynamics simulations treat the atomic nuclei as classical particles, an adequate description of nuclear quantum effects (NQEs) is indispensable when studying proton transfer reactions. Herein, quantum free energy profiles are constructed for three typical proton transfers, which properly take NQEs into account using the path integral formalism. The computational cost of the simulations is kept tractable by deriving machine learning potentials. It is shown that the classical and quasi-classical centroid free energy profiles of the proton transfers deviate substantially from the exact quantum free energy profile.



For many molecular systems, a quantum mechanical description of the electronic potential energy surface (PES) is the preferred simulation technique to properly model the interparticle interactions. Using the appropriate thermodynamic ensemble, the properties of these systems can then be derived at realistic temperatures and pressures by means of molecular dynamics (MD) simulations. While the underlying PES is hereby quantum mechanically determined by solving the electronic structure problem, the nuclear motion on the PES is most often described classically by Newtonian dynamics. As a consequence, nuclear quantum effects (NQEs) such as zero-point energy, the quantization of energy levels, or tunnelling are completely neglected, in spite of their importance in many chemical and biological processes.[1] In particular, at low temperatures and for light-weighted atoms such as hydrogen, NQEs can significantly alter the atomic behavior. To include the quantum mechanical nature of the nuclei in MD simulations, one can use path integral MD (PIMD) simulations, which rely on Feynman's path integral formulation of quantum mechanics.[2] Notwithstanding the substantial computational overhead of *ab initio* PIMD, describing both electrons and nuclei quantum mechanically, the ever improving performance of computing resources alongside the development of machine learning potentials (MLPs) with an *ab initio* accuracy is steadily establishing the technique as a common modeling approach.[3−7]

For activated processes, both classical and quantum MD simulations require the use of enhanced sampling methods to sample all the configurations involved in the rare event of crossing the energy barrier between two (meta)stable states. For classical MD, numerous enhanced sampling protocols have been devised to calculate free energy profiles, including umbrella sampling,[8] metadynamics,[9] thermodynamic integration, and variationally enhanced sampling.[10] Apart from a few dedicated path integral analogues, such as path integral metadynamics,[11] many classical enhanced sampling techniques are also applied to PIMD using the path centroid, which contracts the beads of the path integral ring polymer to a single atom (Figure 1). However, this only provides a quasi-classical approximation to the quantum free energy profile, as the beads of the path integral ring polymer rather than the averaged path centroids are subjected to the actual quantum free energy profile, as also pointed out by refs 6 and 12. The quantum free energy profile is thus obtained by treating both the electrons and nuclei quantum mechanically and expressed as a function of bead related quantities. Nevertheless, the path centroids can

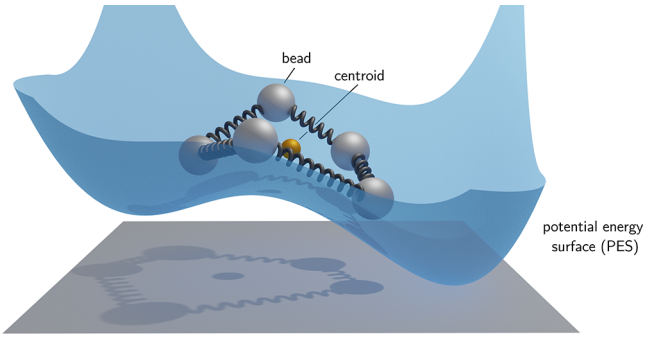


Figure 1. Schematic representation of a path integral ring polymer with five beads on a potential energy surface (PES). The path centroid is indicated in yellow.

Received: August 25, 2022
Published: December 23, 2022

still be used to bias a PIMD simulation and surmount the free energy barrier, just as for the classical counterpart, if the free energy profile of the centroid is afterward transformed to a profile for the beads.

In this work, the influence of both NQEs and a proper description of the quantum free energy profile is examined for three molecular proton transfers: (1) the concerted proton transfer in the formic acid dimer, (2) the sigmatropic proton rearrangement in pentadiene, and (3) the tautomerization of acetaldehyde to vinyl alcohol. These reactions, aside from being ubiquitous reaction types within organic synthesis, are specifically selected to illustrate the impact of the different free energy descriptions on processes with different free energy barrier shapes and heights (ranging from about 25−275 kJ/mol). Furthermore, as the importance of NQEs is known to increase with decreasing temperature,[1] a temperature range of 200−450 K is considered.

METHODS

Path Integral Molecular Dynamics (PIMD). To include the quantum mechanical nature of the atomic nuclei in a molecular dynamics (MD) simulation, one can rely on Feynman's path integral formulation of quantum mechanics. Using the Born−Oppenheimer approximation, the electronic and nuclear degrees of freedom in an MD simulation can be decoupled to obtain an electronic potential energy surface (PES) that is only parametrically dependent on the nuclear coordinates. Depending on the system that is to be studied, the PES can be described by classical force fields, machine learning potentials (MLPs), or *ab initio* solutions of the Schrödinger equation. For the nuclei, however, one generally adopts a classical treatment, sampling the classical canonical ensemble with statistical weights proportional to $e^{-\beta H}$, with $\beta = \frac{1}{k_B T}$ the inverse temperature and H the Hamiltonian of the nuclear subsystem. To sample the corresponding quantum ensemble, the Hamiltonian is replaced by the appropriate quantum mechanical operator $\hat{H}$, for which the statistical weight can be cast into the following form using a Trotter expansion[13]

$$e^{-\beta\hat{H}} \approx [e^{-\frac{\beta}{2P}\hat{V}}\, e^{-\frac{\beta}{P}\hat{T}}\, e^{-\frac{\beta}{2P}\hat{V}}]^P \tag{1}$$

with $\hat{V}$ the potential energy operator and $\hat{T}$ the kinetic energy operator. In the limit of $P \to \infty$, this expansion yields exact results. By means of the path integral technique, the quantum canonical partition function can then be transformed into a classical equivalent by using an extended phase space in which every atom i is replaced by P harmonically coupled beads, that is

$$Z_P \propto \prod_{k=1}^{P}\prod_{i=1}^{N} \int \mathrm{d}\mathbf{p}_i^{(k)} \int \mathrm{d}\mathbf{r}_i^{(k)}\, e^{-\beta H_P} \tag{2}$$

with

$$H_P = \sum_{k=1}^{P}\sum_{i=1}^{N}\left[\frac{\mathbf{p}_i^{(k)^2}}{2m_i} + \frac{1}{2}m_i\omega_P^2(\mathbf{r}_i^{(k+1)} - \mathbf{r}_i^{(k)})^2\right] + \frac{1}{P}\sum_{k=1}^{P} V(\mathbf{r}_1^{(k)}, ..., \mathbf{r}_N^{(k)}) \tag{3}$$

where $\mathbf{r}_i^{(k)}$ and $\mathbf{p}_i^{(k)}$ are respectively the position and momentum of the k-th bead of the i-th particle, m_i is the mass of particle i, $\omega_P = \frac{\sqrt{P}}{\beta\hbar}$ is the angular frequency of the harmonic nearest-neighbor interaction, and $\mathbf{r}_i^{(P+1)} = \mathbf{r}_i^{(1)}$.

Quantum Free Energy Profiles. To derive a free energy profile of a reaction using enhanced sampling techniques such as umbrella sampling, metadynamics, or variationally enhanced sampling, one starts by constructing a suitable collective variable q, described by a function $Q(\mathbf{r}^N)$, that allows to drive the reaction across the activation barrier.[14] In classical statistical physics, the free energy as a function of the variable q is then defined as

$$\begin{aligned} F_{\mathrm{cl}}(q) &= -\frac{1}{\beta}\ln\frac{\int \mathrm{d}\Gamma\, \delta(Q(\mathbf{r}^N) - q)\, e^{-\beta H}}{\int \mathrm{d}\Gamma\, e^{-\beta H}} \\ &= -\frac{1}{\beta}\ln\langle\delta(Q(\mathbf{r}^N) - q)\rangle \end{aligned} \tag{4}$$

where H denotes the classical Hamiltonian, Γ is the entire phase space, and $\mathrm{d}\Gamma = \mathrm{d}\mathbf{r}^N\, \mathrm{d}\mathbf{p}^N$. In the path integral formalism, however, every atom is replaced by a set of harmonically coupled beads, so that thermodynamic quantities are determined by averaging over the beads. Therefore, the quantum free energy is defined as[6,13]

$$F_{\mathrm{quantum}}(q) = -\frac{1}{\beta}\ln\left\langle \frac{1}{P}\sum_{k=1}^{P}\delta(Q(\mathbf{r}^{(k)^N}) - q)\right\rangle \tag{5}$$

where the ensemble average $\langle\cdot\rangle$ uses the Hamiltonian H_P of the classical isomorphism (eq 3). Although the collective variable q now depends on bead quantities, it is possible to calculate the quantum free energy profile by applying a bias in a collective variable $q_{cl}^{(c)}$ that only depends on centroid quantities, in which the beads are contracted to a single atom. The position of the centroid of atom i is for instance given by

$$\mathbf{r}_i^{(c)} = \frac{1}{P}\sum_{P=1}^{N}\mathbf{r}_i^{(k)} \tag{6}$$

The resulting free energy profile $F_{\mathrm{centroid}}(q_{cl}^{(c)})$, given by

$$F_{\mathrm{centroid}}(q_{cl}^{(c)}) = -\frac{1}{\beta}\ln\langle\delta(Q(\mathbf{r}^{(c)^N}) - q_{cl}^{(c)})\rangle \tag{7}$$

can subsequently be converted into $F_{\mathrm{quantum}}(q)$ by means of a transformation of the collective variable (see Supporting Information S1)

$$\begin{aligned} F_{\mathrm{quantum}}(q) &= -\frac{1}{\beta}\ln\left[\frac{1}{P}\sum_{k=1}^{P}\langle\delta(Q(\mathbf{r}^{(k)^N}) - q)\rangle\right] \\ &= -\frac{1}{\beta}\ln\left[\frac{1}{P}\sum_{k=1}^{P}\int \mathrm{d}q_{cl}^{(c)}\, p_b(q^{(k)}|q_{cl}^{(c)})\, e^{-\beta F_{\mathrm{centroid}}(q_{cl}^{(c)})}\right] + C \end{aligned} \tag{8}$$

where C is a negligible constant, and $p_b(q^{(k)}|q_{cl}^{(c)})$ is the conditional probability that the collective variable takes the value q for bead k in a biased simulation when the centroid collective variable has a value $q_{cl}^{(c)}$. Hence, by simply tracking the values of $q^{(k)}$ and $q_{cl}^{(c)}$ throughout the enhanced sampling simulations in the centroid collective variable $q_{cl}^{(c)}$, the conditional probability $p_b(q^{(k)}|q_{cl}^{(c)})$ can be constructed, and $F_{\mathrm{centroid}}(q_{cl}^{(c)})$ can be converted into $F_{\mathrm{quantum}}(q)$. Remark that this transformation approach can be shown to be equivalent to the method proposed in ref 6 by Tuckerman *et al.* (see Supporting Information S1).

https://doi.org/10.1021/acs.jctc.2c00874
J. Chem. Theory Comput. 2023, 19, 18−24

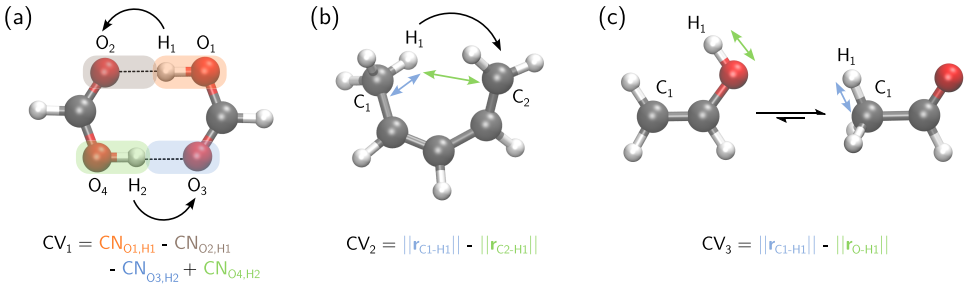


Figure 2. Molecular representation of (a) a double proton transfer in a formic acid dimer, (b) a sigmatropic proton rearrangement in pentadiene, and (c) the tautomerization of vinyl alcohol to acetaldehyde. The collective variable (CV) used to sample the proton transfer reaction is indicated below each representation, where $CN_{i,j}$ denotes a coordination number (see eq 12).

Furthermore, given that all the beads of an atom are in principle equivalent, the quantum free energy is also given by

$$F_{\text{quantum}}(q) = \frac{1}{P}\sum_{k=1}^{P} F^{(k)}(q) = -\frac{1}{P\beta}\sum_{k=1}^{P} \ln\langle\delta(Q(\mathbf{r}^{(k)^N}) - q)\rangle \tag{9}$$

which can also be obtained by biasing the ring polymer's centroid in an enhanced PIMD simulation through a similar transformation as in eq 8:

$$\begin{aligned} F_{\text{quantum}}(q) &= -\frac{1}{P\beta}\sum_{k=1}^{P} \ln\langle\delta(Q(\mathbf{r}^{(k)^N}) - q)\rangle \\ &= -\frac{1}{P\beta}\sum_{k=1}^{P} \ln\left[\int \mathrm{d}q_{cl}^{(c)}\, p_b(q^{(k)}|q_{cl}^{(c)})\, e^{-\beta F_{\text{centroid}}(q_{cl}^{(c)})}\right] + C \end{aligned} \tag{10}$$

Both formulas of the quantum free energy (eqs 8 and 10) should in principle yield the same result, as they are theoretically equivalent. Discrepancies between these different methods of calculating the quantum free energy can however arise due to the sensitivity of eq 8 to inequivalences in the sampling of the beads. In particular at lower temperatures, these bead inequivalences tend to become more prominent (see Supporting Information S10). Therefore, eq 10 is the preferred transformation approach, as it is more robust with respect to bead inequivalences. If one does not rely on centroid biased simulations, eq 9 implies that every bead should be biased separately, which would make the calculation of the quantum free energy profile P times more expensive. Strictly speaking, the equivalence of the beads also allows to bias only a single bead, although this reduces the sampling efficiency and is most likely to yield sampling difficulties, similar to the issues known for the calculation of rate constants (which can be calculated from both centroid or quantum free energy profiles).[15] More specifically, when a single bead is restrained around the top of a free energy barrier, the ring polymer possesses the freedom to orient ifself toward either side of the barrier. For a steep free energy barrier, the ring polymer is most likely to sample only one side of the barrier, which hampers a proper sampling. However, by optimizing the time constant of the PILE[16] thermostat as a function of the force constant of the umbrella bias potential, Bishop *et al.*[12] showed that it is possible to obtain converged quantum free energy profiles by biasing only a single bead. As their study only included dimers, which give rise to free energy profiles with a Lennard-Jones shape, the typical sampling difficulties near free energy peaks were not encountered. Finally, we also note that in the case of unbiased PIMD simulations the quantum free energy profile can be readily computed from eq 5 or 9 by means of a histogram of the bead collective variable.[17]

Computational Details. The proton transfer reactions studied in this work include a double proton transfer in a formic acid dimer, a sigmatropic proton rearrangement in pentadiene, and the tautomerization of acetaldehyde to vinyl alcohol (Figure 2). The free energy profile of each proton transfer was calculated using umbrella sampling in combination with a machine learning potential (MLP), which significantly reduces the computational cost of the simulations. To construct an initial classical free energy profile and to collect the necessary training data to build an MLP for each of the molecules, *ab initio* umbrella sampling simulations were performed with PSI4,[18] using a PBE0 functional[19] and a 6-311G(3df,3pd) Pople basis set.[20] From every *ab initio* umbrella simulation, the energies and forces of snapshots taken every 5 fs were used to train an equivariant MLP with our in-house developed code available at https://github.com/mcoolsce/NNFFLIB and archived on 10.5281/zenodo.6606271. The neural network architecture of the MLP is inspired by the Neural Equivariant Interatomic Potential (NequIP),[21] but the rotation order of the equivariant features is limited to $\ell = 1$ (see Supporting Information Section S2). Both classical and quantum free energy profiles were constructed with the MLP to test its stability, the performance with respect to the *ab initio* simulations, and to extract additional configurations that were not present in the initial data set. The final MLP was then obtained by training to this extended data set. More details regarding the training of the MLP are provided in Section S3 of the Supporting Information.

The test mean absolute error on the energy and the forces is limited to respectively 0.4 meV and 1.3 meV/Å for the formic acid dimer, 1.1 meV and 4.5 meV/Å for pentadiene, and 0.3 meV and 1.9 meV/Å for acetaldehyde/vinyl alcohol. A comparison between the *ab initio* and MLP free energy profiles, obtained both classically or with the inclusion of NQEs, is given in Figures S1−S6 of the Supporting Information. For the classically sampled proton transfers, the 20 ps *ab initio* umbrella sampling simulations are not fully converged, which can be easily resolved by the MLP by extending the simulation time to 250 ps, given that the MLP tremendously accelerates the evaluation of the PES (see Supporting Information Section S4). In the enhanced PIMD simulations, the spatially extended ring polymer describing the proton facilitates the sampling of the proton transfer, so that the *ab initio* free energy profiles are in excellent agreement with the MLP free energy profiles (which have an increased simulation time of 50 ps). However, from the convergence of the energy around the transition states as a function of the number of beads (Figures S7−S9), it follows that more beads are required than the 16 beads used to construct the *ab initio*

https://doi.org/10.1021/acs.jctc.2c00874
J. Chem. Theory Comput. 2023, 19, 18−24

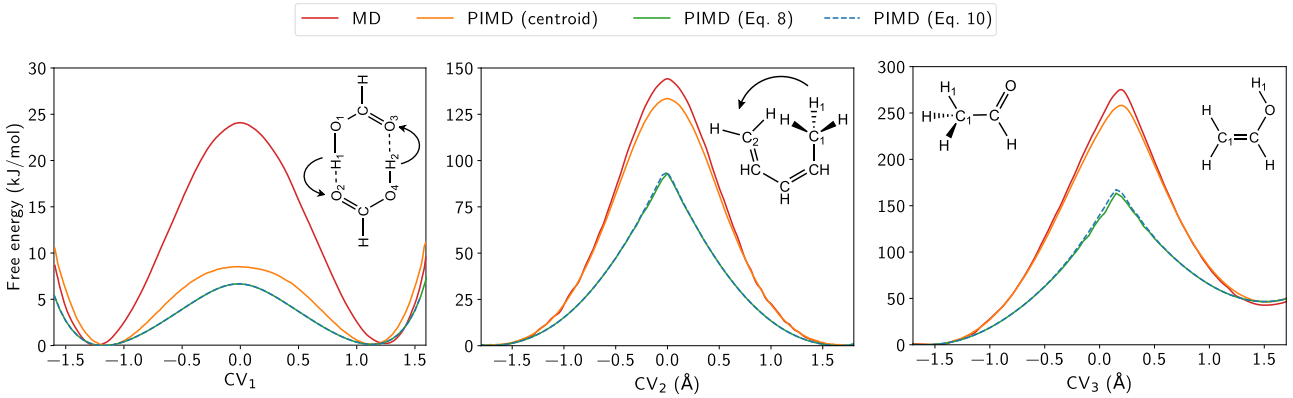


Figure 3. Free energy profiles of a double proton transfer in a formic acid dimer at 200 K (*left*), a sigmatropic proton rearrangement in pentadiene at 450 K (*middle*), and the tautomerization of acetaldehyde to vinyl alcohol at 450 K (*right*). For each proton transfer, both the classical free energy (MD) and the quantum free energy (PIMD) are reported. The PIMD (centroid) profile represents the free energy that is directly obtained when sampling along the centroid CV, before applying the appropriate transformation of eq 8 or 10.

free energy profiles. The differences between the free energy profiles derived with 16 beads or 64 beads are nevertheless limited (Figure S11).

To steer the enhanced sampling of the proton transfers using umbrella sampling, each molecule is biased according to a specific collective variable (CV). For the formic acid dimer, the collective variable for the double proton transfer is given by a difference in coordination numbers

$$\mathrm{CV}_1 = \mathrm{CN}_{\mathrm{O}_1,\mathrm{H}_1} - \mathrm{CN}_{\mathrm{O}_2,\mathrm{H}_1} - \mathrm{CN}_{\mathrm{O}_3,\mathrm{H}_2} + \mathrm{CN}_{\mathrm{O}_4,\mathrm{H}_2} \tag{11}$$

with

$$\mathrm{CN}_{i,j} = \frac{1 - \left(\frac{\mathbf{r}_{ij}}{\mathbf{r}_0}\right)^6}{1 - \left(\frac{\mathbf{r}_{ij}}{\mathbf{r}_0}\right)^{12}} \tag{12}$$

where $\mathbf{r}_{ij}$ is the interatomic distance between atoms i and j, and $\mathbf{r}_0$ = 1.4 Å. The labels of the atoms in eq 11 are indicated in Figure 2. Given that the double proton transfer in the formic acid dimer was shown to be concerted but not correlated (see Supporting Information S7),[22] a simultaneous proton transfer is assumed in the definition of CV_1. For the proton transfers in pentadiene (CV_2) and acetaldehyde (CV_3), a simple difference in interatomic distances can be used as a collective variable

$$\mathrm{CV}_2 = \left\|\mathbf{r}_{\mathrm{C}_1-\mathrm{H}_1}\right\| - \left\|\mathbf{r}_{\mathrm{C}_2-\mathrm{H}_1}\right\| \tag{13}$$

$$\mathrm{CV}_3 = \left\|\mathbf{r}_{\mathrm{C}_1-\mathrm{H}_1}\right\| - \left\|\mathbf{r}_{\mathrm{O}-\mathrm{H}_1}\right\| \tag{14}$$

for which the atomic labels are also indicated in Figure 2. All the simulations in this work relied on `PLUMED` 2.8 to implement the bias potential,[23,24] which was coupled to `Yaff`[25] to perform enhanced classical MD and to `i-PI`[26] to perform enhanced PIMD. An overview of the umbrella force constants and equilibrium values for every reaction is provided in Table S4. The temperature of the classical MD simulations was controlled by a 3-bead Nosé–Hoover chain thermostat,[27] whereas the PIMD simulations made use of a PILE-L thermostat,[16] both using a time constant of 100 fs. The classical MD and PIMD simulations used a time step of respectively 0.5 and 0.25 fs.

RESULTS AND DISCUSSION

The classical and quantum free energy profiles for a double proton transfer in a formic acid dimer, a sigmatropic proton rearrangement in pentadiene, and the tautomerization of acetaldehyde to vinyl alcohol are shown in Figure 3. These three proton transfer reactions span a classical free energy range of about 25–275 kJ/mol, which makes them ideally suited to investigate the influence of the inclusion of NQEs on the free energy. The concerted double proton transfer in the formic acid dimer has the lowest classical free energy barrier (24.1 kJ/mol at 200 K), which is reduced by a factor of about three when the atomic nuclei are treated quantum mechanically. The centroid free energy profile, which corresponds to the collective variable in which the enhanced PIMD sampling was performed, exhibits only small deviations from the actual quantum free energy profile, that is obtained after performing the transformation given by eq 8 or 10. The barrier height changes from 8.5 kJ/mol for the centroid profile to 7.3 kJ/mol for the bead profile, which is a small absolute difference (similar to the one found by Tuckerman *et al.* for hydrogen diffusion in an sII clathrate[6]), but nevertheless a reduction of 14%. Both transformations are also observed to yield the same quantum free energy profile (Figure 3), indicating a proper sampling of all PIMD beads and of the conditional probability used in the transformation.

For the proton transfer in pentadiene and acetaldehyde, by contrast, there is a striking difference between the centroid and bead profiles. By switching from a classical sampling to a quantum sampling in a centroid collective variable, the free energy barrier of the proton transfer in pentadiene is lowered by about 11 kJ/mol (from 144.2 to 133.5 kJ/mol at 450 K). After transformation to a bead collective variable, the resulting quantum free energy profile has a barrier of 93 kJ/mol, thereby reducing the barrier by an additional 40 kJ/mol, yielding a total reduction of 35% with respect to the classical description. For acetaldehyde, the reduction is even larger. The classical free energy barrier of 275.1 kJ/mol drops to 258.2 kJ/mol in a centroid description and to 163.5 kJ/mol for the actual quantum free energy profile, resulting in a diminution of 40% with respect to the classical barrier. Due to the high activation energy of the reaction, the transition state region is sampled less easily, so that a discrepancy of the order of 5 kJ/mol remains present between the two different transformation approaches of eqs 8 and 10 due to the sensitivity of eq 8 to

https://doi.org/10.1021/acs.jctc.2c00874
J. Chem. Theory Comput. 2023, 19, 18–24

inequivalences between the beads, as explained in more detail in Section S10 of the Supporting Information. At lower temperatures, the reduced thermal motion also tends to increase the inequivalences in the sampling of the beads, as shown for pentadiene in Section S10.

Each of the three proton transfer reactions clearly demonstrates the importance of an adequate quantum mechanical description of the atomic nuclei in calculating a free energy profile, in line with other proton transfers studied in the literature.[28−30] Even at a temperature of 450 K, NQEs are observed to significantly alter the classical behavior of pentadiene and acetaldehyde. The full extent of the NQEs on the free energy becomes clear when exchanging the quasi-classical centroid description for a description at the level of the bead quanta. Due to the quantum uncertainty on the position of the nuclei, mimicked by the spatial extension of the path integral ring polymer, the bead quanta experience a lower energy barrier since they are not precisely pinned to the transition state but rather envelop it (Figures 1 and 4). As the free energy barriers grow higher and are more sharply peaked, this effect is reinforced, and the dissimilarity between the centroid and the beads increases. In Figure 5, the probability density of the radius of gyration r_{gyr} and the maximal extent r_{max} of the ring polymer of the transitioning proton(s) are shown as a function of the collective variable for each molecule, with

$$r_{gyr} = \sqrt{\frac{1}{P}\sum_{k=1}^{P}\left\|\mathbf{r}_H^{(k)} - \mathbf{r}_H^{(c)}\right\|^2} \tag{15}$$

$$r_{\max} = \max_{k\in\{1,\ldots,P\}}\left\|\mathbf{r}_H^{(k)} - \mathbf{r}_H^{(c)}\right\| \tag{16}$$

Both probability densities exhibit a sparser probability distribution around the transition state, in particular for acetaldehyde, thus confirming the increased stretching of the proton's ring polymer around the transition state. Although the spatial expansion of the proton is most pronounced, also the neighboring oxygen or carbon atoms effectively increase in size, which can further aid the reaction mechanism. This is reflected in the quantum radial distribution functions (RDFs) reported in Figures S13−S16, which show a significant broadening for the peaks of the O−H and C−H atom pairs involved in the proton transfer reaction in comparison with the classical RDFs. Furthermore, the peak broadening is also observed to be more prominent near the transition state than near the free energy minima, in agreement with the proton radius of gyration.

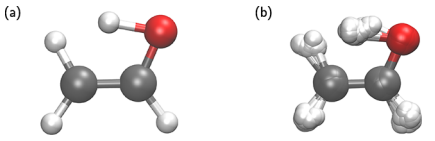


Figure 4. Centroid (a) and bead (b) representation of vinyl alcohol from an umbrella simulation with CV_3 = 0.2. In the bead representation, every atom is replaced by a ring polymer of 64 beads.

CONCLUSIONS

Notwithstanding the well-known impact of nuclear quantum effects (NQEs) on the behavior of protons and their tendency

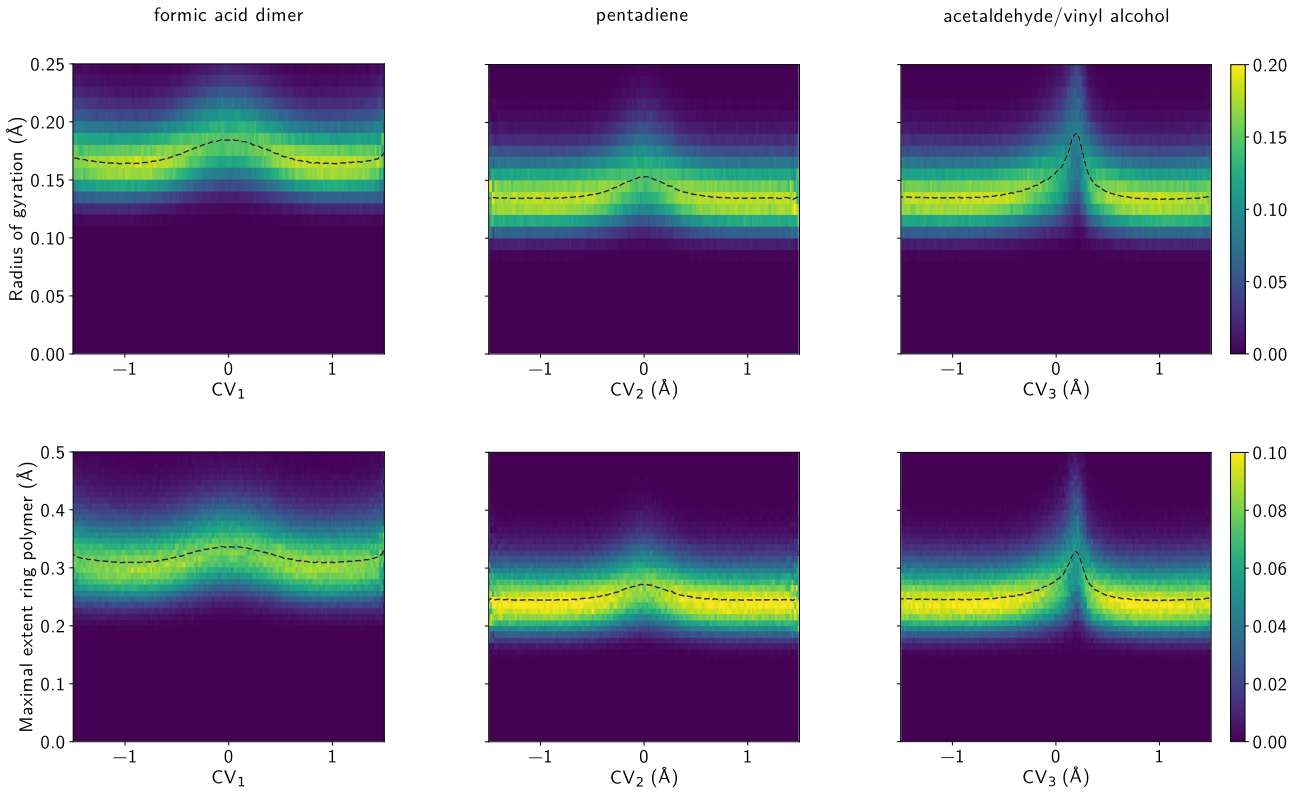


Figure 5. Conditional probability density $p(r_{gyr/max}|CV_i)$ of the radius of gyration (*top row*) and maximal ring polymer extent (*bottom row*) (eqs 15 and 16) of the transitioning proton(s) in formic acid, pentadiene, and acetaldehyde/vinyl alcohol as a function of the collective variable. The average of the probability density as a function of the collective variable is indicated by a dashed black line.

https://doi.org/10.1021/acs.jctc.2c00874
J. Chem. Theory Comput. 2023, 19, 18−24

to lower the activation energy of proton transfer reactions, this work demonstrates that also the quasi-classical approximation of the free energy, using the path integral centroid, is unable to explicitly represent the full impact of NQEs. In enhanced sampling simulations, a bias on the path integral centroid is however most convenient and computationally efficient and can provide all the required information to generate the quantum free energy profile at the level of the individual path integral beads through a postprocessing transformation. Using three molecular proton transfer reactions, spanning a classical free energy range of about 25−275 kJ/mol, the quantum free energy profiles were shown to significantly deviate from the quasi-classical centroid approximations, due to the large spatial extent of the path integral ring polymer. As the free energy barrier becomes higher and more sharply peaked, the deviation is observed to increase, as the difference between the free energy experienced by the centroid and beads is more pronounced. In order to reduce the computational cost associated with *ab initio* path integral molecular dynamics (PIMD), machine learning potentials (MLPs) were trained to accelerate the evaluation of the potential energy surface while maintaining an *ab initio* accuracy. By further exploiting the effective cost reduction of *ab initio* PIMD through the use of MLPs and by combining the MLP with PIMD acceleration techniques, NQEs can also systematically be accounted for in chemical reactions of larger molecular systems. For proton transfer reactions in particular, which are omnipresent in chemical processes, an adequate description of NQEs is indispensable to obtain reliable quantum free energy profiles.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c00874.

Mathematical details of transformation between centroid and bead free energy profiles, elaborate discussion of MLP architecture and generation of training data sets, *ab initio* validation of MLP free energy profiles, convergence of PIMD simulations, RDFs of different molecules, and temperature dependence of proton transfer in pentadiene (PDF)

MLPs of the formic acid dimer, pentadiene, and acetaldehyde/vinyl alcohol (ZIP)

■ AUTHOR INFORMATION

Corresponding Author

Veronique Van Speybroeck − *Center for Molecular Modeling (CMM), Ghent University, 9052 Zwijnaarde, Belgium;* orcid.org/0000-0003-2206-178X; Email: Veronique.VanSpeybroeck@UGent.be

Authors

Aran Lamaire − *Center for Molecular Modeling (CMM), Ghent University, 9052 Zwijnaarde, Belgium;* orcid.org/0000-0003-0093-5490

Maarten Cools-Ceuppens − *Center for Molecular Modeling (CMM), Ghent University, 9052 Zwijnaarde, Belgium;* orcid.org/0000-0002-7363-7534

Massimo Bocus − *Center for Molecular Modeling (CMM), Ghent University, 9052 Zwijnaarde, Belgium;* orcid.org/0000-0001-9474-6644

Toon Verstraelen − *Center for Molecular Modeling (CMM), Ghent University, 9052 Zwijnaarde, Belgium;* orcid.org/0000-0001-9288-5608

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c00874

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Fund for Scientific Research Flanders (FWO), Flanders Industry Innovation Moonshot (ARCLATH II, No. HBC.2021.0254), and the Research Board of Ghent University (BOF). The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO, and the Flemish Government−department EWI.

■ REFERENCES

(1) Markland, T. E.; Ceriotti, M. Nuclear quantum effects enter the mainstream. *Nat. Rev. Chem.* **2018**, *2*, 0109.

(2) Feynman, R. P. Space-Time Approach to Non-Relativistic Quantum Mechanics. *Rev. Mod. Phys.* **1948**, *20*, 367−387.

(3) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine Learning Force Fields. *Chem. Rev.* **2021**, *121*, 10142−10186.

(4) Friederich, P.; Häse, F.; Proppe, J.; Aspuru-Guzik, A. Machine-learned potentials for next-generation matter simulations. *Nat. Mater.* **2021**, *20*, 750−761.

(5) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121*, 10073−10141.

(6) Cendagorta, J. R.; Shen, H.; Bačić, Z.; Tuckerman, M. E. Enhanced Sampling Path Integral Methods Using Neural Network Potential Energy Surfaces with Application to Diffusion in Hydrogen Hydrates. *Adv. Theory Simul.* **2021**, *4*, 2000258.

(7) Li, C.; Voth, G. A. Using Machine Learning to Greatly Accelerate Path Integral Ab Initio Molecular Dynamics. *J. Chem. Theory Comput.* **2022**, *18*, 599−604.

(8) Torrie, G.; Valleau, J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(9) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.

(10) Valsson, O.; Parrinello, M. Variational Approach to Enhanced Sampling and Free Energy Calculations. *Phys. Rev. Lett.* **2014**, *113*, 090601.

(11) Quhe, R.; Nava, M.; Tiwary, P.; Parrinello, M. Path Integral Metadynamics. *J. Chem. Theory Comput.* **2015**, *11*, 1383−1388.

(12) Bishop, K. P.; Roy, P.-N. Quantum mechanical free energy profiles with post-quantization restraints: Binding free energy of the water dimer over a broad range of temperatures. *J. Chem. Phys.* **2018**, *148*, 102303.

(13) Tuckerman, M. E.*Statistical Mechanics: Theory and Molecular Simulation*; OUP Oxford: 2010.

(14) Peters, B. Reaction Coordinates and Mechanistic Hypothesis Tests. *Annu. Rev. Phys. Chem.* **2016**, *67*, 669−690.

(15) Craig, I. R.; Manolopoulos, D. E. A refined ring polymer molecular dynamics theory of chemical reaction rates. *J. Chem. Phys.* **2005**, *123*, 034102.

(16) Ceriotti, M.; Parrinello, M.; Markland, T. E.; Manolopoulos, D. E. Efficient stochastic thermostatting of path integral molecular dynamics. *J. Chem. Phys.* **2010**, *133*, 124104.

(17) Rossi, M.; Ceriotti, M.; Manolopoulos, D. E. Nuclear Quantum Effects in H^+ and OH^- Diffusion along Confined Water Wires. *J. Phys. Chem. Lett.* **2016**, *7*, 3001−3007.
(18) Smith, D. G. A.; Burns, L. A.; Simmonett, A. C.; Parrish, R. M.; Schieber, M. C.; Galvelis, R.; Kraus, P.; Kruse, H.; Di Remigio, R.; Alenaizan, A.; James, A. M.; Lehtola, S.; Misiewicz, J. P.; Scheurer, M.; Shaw, R. A.; Schriber, J. B.; Xie, Y.; Glick, Z. L.; Sirianni, D. A.; O'Brien, J. S.; Waldrop, J. M.; Kumar, A.; Hohenstein, E. G.; Pritchard, B. P.; Brooks, B. R.; Schaefer, H. F.; Sokolov, A. Y.; Patkowski, K.; DePrince, A. E.; Bozkaya, U.; King, R. A.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.
(19) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(20) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* **1984**, *80*, 3265−3269.
(21) Batzner, S.; Musaelian, A.; Sun, L.; Geiger, M.; Mailoa, J. P.; Kornbluth, M.; Molinari, N.; Smidt, T. E.; Kozinsky, B. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nat. Commun.* **2022**, *13*, 2453.
(22) Ivanov, S. D.; Grant, I. M.; Marx, D. Quantum free energy landscapes from ab initio path integral metadynamics: Double proton transfer in the formic acid dimer is concerted but not correlated. *J. Chem. Phys.* **2015**, *143*, 124304.
(23) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.
(24) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; The PLUMED Consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670−673.
(25) Verstraelen, T.; Vanduyfhuys, L.; Vandenbrande, S.*Yaff, yet another force field*. http://molmod.ugent.be/software/ (accessed 2022-02-28).
(26) Kapil, V.; Rossi, M.; Marsalek, O.; Petraglia, R.; Litman, Y.; Spura, T.; Cheng, B.; Cuzzocrea, A.; Meißner, R. H.; Wilkins, D. M.; Juda, P.; Bienvenue, S. P.; Fang, W.; Kessler, J.; Poltavsky, I.; Vandenbrande, S.; Wieme, J.; Corminboeuf, C.; Kühne, T. D.; Manolopoulos, D. E.; Markland, T. E.; Richardson, J. O.; Tkatchenko, A.; Tribello, G. A.; Van Speybroeck, V.; Ceriotti, M. i-PI 2.0: A universal force engine for advanced molecular simulations. *Comput. Phys. Commun.* **2019**, *236*, 214−223.
(27) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(28) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. On the Quantum Nature of the Shared Proton in Hydrogen Bonds. *Science* **1997**, *275*, 817−820.
(29) Cassone, G. Nuclear Quantum Effects Largely Influence Molecular Dissociation and Proton Transfer in Liquid Water under an Electric Field. *J. Phys. Chem. Lett.* **2020**, *11*, 8983−8988.
(30) Slocombe, L.; Sacchi, M.; Al-Khalili, J. An open quantum systems approach to proton tunnelling in DNA. *Commun. Phys.* **2022**, *5*, 109.

https://doi.org/10.1021/acs.jctc.2c00874
J. Chem. Theory Comput. 2023, 19, 18−24