Supplementary Information for: Understanding and Tracking the Excess Proton in Ab Initio Simulations; Insights from IR Spectra

Chenghan Li[†], Jessica M. J. Swanson[*‡]

†Department of Chemistry, Chicago Center for Theoretical Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, United States
‡Department of Chemistry, Biological Chemistry Program, and Center for Cell and Genome Science, The University of Utah, Salt Lake City, Utah 84112, United States

Proof of donor identity independency of rCEC

In order to show rCEC is not dependent on the identity of donor atom, we first consider a two-state case (Figure S2). When state searching starts from atom I, following eq 6 and 8,

$$c_i^2 = \frac{1}{1 + c_j^2/c_i^2} = \frac{1}{1 + f_{ij}^{\mathrm{ct}}} \tag{S1a}$$

$$c_j^2 = c_j^2/c_i^2 \times c_i^2 = \frac{f_{ij}^{\mathrm{ct}}}{1 + f_{ij}^{\mathrm{ct}}} \tag{S1b}$$

While regarding J as pivot, the coefficients are

$$c_j^2 = \frac{1}{1 + c_i^2/c_j^2} = \frac{1}{1 + f_{ji}^{\mathrm{ct}}} \tag{S2a}$$

$$c_i^2 = c_i^2/c_j^2 \times c_j^2 = \frac{f_{ji}^{\mathrm{ct}}}{1 + f_{ji}^{\mathrm{ct}}} \tag{S2b}$$

Recall that the charge transfer factor from state $|i\rangle$ to state $|j\rangle$ is $f_{ij}^{\mathrm{ct}} = \exp\left(-k\delta_{IJH}\right)$ while $f_{ji}^{\mathrm{ct}} = \exp\left(-k\delta_{JIH}\right)$. This leads to the fact that

$$f_{ij}^{\mathrm{ct}} f_{ji}^{\mathrm{ct}} = 1 \tag{S3}$$

given that $\delta_{IJH} = r_{JH} - r_{IH} = -\delta_{JIH}$. Then, we have from eq S2

$$c_j^2 = \frac{1}{1 + f_{ji}^{\mathrm{ct}}} = \frac{f_{ij}^{\mathrm{ct}}}{f_{ij}^{\mathrm{ct}} + f_{ij}^{\mathrm{ct}} f_{ji}^{\mathrm{ct}}} = \frac{f_{ij}^{\mathrm{ct}}}{f_{ij}^{\mathrm{ct}} + 1} \tag{S4a}$$

$$c_i^2 = \frac{f_{ji}^{\mathrm{ct}}}{1 + f_{ji}^{\mathrm{ct}}} = \frac{f_{ij}^{\mathrm{ct}} f_{ji}^{\mathrm{ct}}}{f_{ij}^{\mathrm{ct}} + f_{ij}^{\mathrm{ct}} f_{ji}^{\mathrm{ct}}} = \frac{1}{f_{ij}^{\mathrm{ct}} + 1} \tag{S4b}$$

Thus, the coefficients starting from I (eq S1) are recovered, i.e. rCEC is invariant with donor identity change in the two-state case. The conclusion can be easily generalized to the multi-state by noticing the coefficient of any state $|i_k\rangle$ is propagated from a parent state $|i_1\rangle$ by multiplying f_{ct}'s along a hydrogen bond chain $|i_1\rangle \to |i_2\rangle \to \cdots \to |i_k\rangle$. An alternative donor index may change $|i_k\rangle$ into the parent of $|i_1\rangle$. Thus, the fact that $f_{i_k i_1}^{\mathrm{ct}} f_{i_1 i_k}^{\mathrm{ct}} = f_{i_2 i_1}^{\mathrm{ct}} f_{i_1 i_2}^{\mathrm{ct}} \cdot f_{i_3 i_2}^{\mathrm{ct}} f_{i_2 i_3}^{\mathrm{ct}} \cdots f_{i_k i_{k-1}}^{\mathrm{ct}} f_{i_{k-1} i_k}^{\mathrm{ct}} = 1$ means the relative ratio of $c_{i_1}^2$ and $c_{i_k}^2$ is independent of the donor index. Additionally, $\{c_i^2\}$'s are normalized to 1, so identical c_j^2/c_i^2 for $\forall i, j$ results in identical $\{c_i^2\}$.

Vibrational intensity of c_i^2

We also computed the vibrational intensity (vibrational density of states; VDOS) of the state population c_i^2 (Figure S3A) from

$$\mathrm{VDOS}_i(\omega) = \int \left\langle c_i^2(0) c_i^2(t) \right\rangle e^{-i\omega t} \mathrm{d}t. \tag{S5}$$

Noticeably, the 2800 cm^{-1} frequency can be seen in the fourth largest c_i^2, which includes the contributions from proton stretch modes of O_0-O_{1z} and also O_0-O_{1y}, O_{1x}-O_{2x1} and O_{1x}-O_{2x2} in a more Zundel-like configuration (Figure S3B). Noting that O_{2x1} and O_{2x2} are the second solvation shell of the H_3O core, 2800 cm^{-1} frequency of the 2nd shell in Figure 1C could be attributed to these proton stretch modes. This additionally verifies the rCEC is capturing the vibrational modes in the hydrated proton complex but the decaying excess charges hide the IR signal in CEC's spectrum.

SI Figures

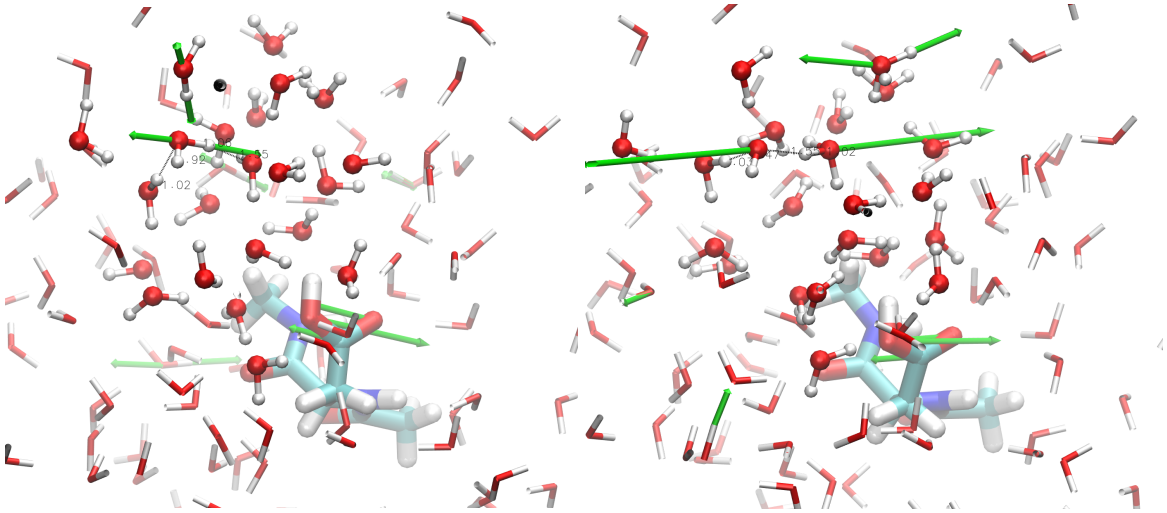

Figure S1. Snapshots extracted from 5527.5 fs and 6136.5 fs of the metadynamics simulation of aspartic acid in water showing the starting point and the resulting configuration of the decomposition procedure. The bias forces acting on atoms are shown by green arrows with a scale of 25 kJ/mol/Å^2 force per 1 Å arrow length. Forces under 25 kJ/mol/Å^2 are not shown for clarity.

Figure S2. An illustration of two state case. The proton donor oxygen is denoted as the green ball and the red ball is denoting its acceptor.

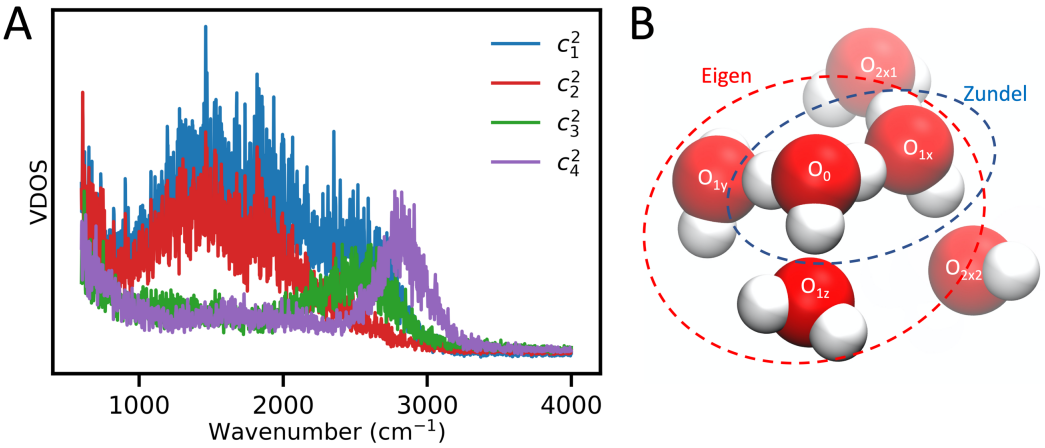

Figure S3. (A) The VDOS of the 4 largest c_i^2. (B) The oxygen notation. The oxygens with two largest c_i^2 are denoted as O_0 and O_{1x}. The other two oxygens in the first solvation shell of O_0 are denoted as O_{1y} and O_{1z}, in order of increasing δ. The two closest oxygens in the first solvation shell of O_{1x} are denoted as O_{2x1} and O_{2x2}.

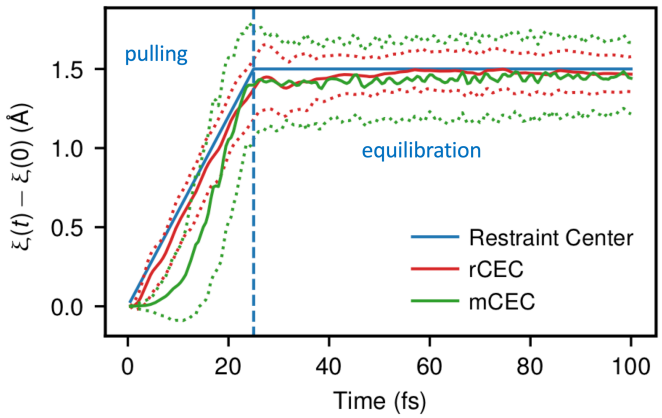

Figure S4. Time evolution of the deprotonation collective variable ξ using rCEC (red) and mCEC (green) and the restraint center (blue) in SMD runs. The average value over all SMD runs (solid line) of rCEC follows the restraint center (blue) while mCEC shows an obvious hysteresis. The standard deviation interval (dotted lines) of rCEC is much smaller than mCEC. This suggests that, when pulling the mCEC, the proton seems stickier to the amino acid, implying mCEC is missing some slow motions in the collective proton disassociation process, which may cause an overestimation of free energy barrier due to finite sampling. Also, the new CEC shows a much smaller standard deviation, revealing its robustness when handling various initial configurations.