Supporting Information for: High-Level VSCF/VCI Calculations Decode the Vibrational Spectrum of the Aqueous Proton

Qi Yu,[*,†] William B. Carpenter,[‡] Nicholas H.C. Lewis,[‡] Andrei Tokmakoff,[‡] and Joel M. Bowman[*,†]

†*Department of Chemistry and Cherry L. Emerson Center for Scientific Computation, Emory University, Atlanta, Georgia 30322, U.S.A.*

‡*Department of Chemistry, James Frank Institute, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, Illinois 60637, U.S.A*

E-mail: qyu28@emory.edu; jmbowma@emory.edu

1. Singular Value Decomposition and Maximum Entropy Re-weighting of Experimental Spectral Components

We applied singular value decomposition (SVD) analysis to decompose the IR spectra of the acid concentration series into components relating to water and the hydrated excess proton. In SVD analysis, the concentration-dependent spectra are represented as a linear combination of n independent spectral components whose weights grow or decay as a function of concentration. A spectral series matrix A represented in terms of frequency ω and concentration c can be decomposed into three matrices:

$$A_{\omega \times c} = U_{\omega \times n} S_{n \times n} V^T_{c \times n} + E_{\omega \times c} \tag{1}$$

where $U_{\omega \times n}$ is a matrix of the two component spectra, $S_{n \times n}$ is a diagonal matrix of the singular values representing weighting of the spectral components, $V_{c \times n}$ is a matrix representing how the concentration-dependent amplitude changes of each spectral component, and $E_{\omega \times c}$ is residual noise not captured by the n components. Two components were needed to describe 99% of the spectral information. The first component, representing the water absorption spectrum, contains approximately 90% of the of the spectral information content, while a second component associated with the acid contains another 9% of spectral information and the remaining components comprising noise. The components from SVD analysis are mathematical quantities which may not represent physical absorption contributions. For instance, SVD analysis can produce negative absorption amplitudes in certain regions of the absorption spectrum.

To determine the physically relevant components, we applied a maximum entropy method to reweight the SVD components in a minimally biased way. To implement this, we introduce

a mixing matrix $T_{n\times n}$, such that

$$A_{\omega\times c} = U_{\omega\times n} S_{n\times n} T^{-1}_{n\times n} T_{n\times n} V^{T}_{c\times n}. \tag{2}$$

The reconstructed spectra are given by

$$C_{\omega\times c} = U_{\omega\times n} S_{n\times n} T^{-1}_{n\times n} \tag{3}$$

and the reconstructed concentration weights are determined as

$$a_{n\times c} = T_{n\times n} V^{T}_{c\times n}. \tag{4}$$

The elements of the mixing matrix T are chosen such that the objective function

$$F = H + \gamma P + \lambda D \tag{5}$$

is minimized. H is a term that is minimized when information entropy across the spectral components is maximized, P is a term that penalizes negative amplitudes in the spectral and concentration domains, and D is a dissimilarity term that penalizes spectral overlap between components. H is calculated as

$$H = -\sum_{\omega}\sum_{c} h_{\omega c} \ln(h_{\omega c}) \tag{6}$$

where $h_{\omega c}$ is a normalized absolute value of the first derivative of the spectrum with respect to frequency

$$h_{\omega c} = \frac{|C'_{\omega c}|}{\sum\limits_{\omega} |C'_{\omega c}|}. \tag{7}$$

The positivity function P was calculated as

$$P = \sum_n \sum_c W(a_{nc}) a_{nc}^2 + \sum_\omega \sum_c W(C_{\omega c}) C_{\omega c}^2 \tag{8}$$

where W is a weighting function for negative values

$$W(x) = \begin{cases} 0 & \text{if } x \geq 0 \\ 1 & \text{if } x < 0. \end{cases} \tag{9}$$

The dissimilarity term D was calculated as the dot product of the two spectral components normalized by their respective maximum absolute amplitudes, which was included to ensure that the second component represented the new spectral features that grew in with HCl concentration, rather than a restatement of the spectrum of 2M HCl solution. The weights of the different penalty terms can be balanced by the positive scalars γ, λ, and α. We found that the most optimal results for the HCl spectral series were obtained with $\gamma = 500$, and $\lambda = 0.005$. The global minimum of the objective function was determined with a simulated annealing algorithm.

The results of this analysis on FTIR and ATR spectra of HCl concentration series are presented in Fig. S1. For both techniques, we recorded spectra of 0-2M HCl (Figs. S1a and S1d). Because of the uncontrolled pathlengths of the FTIR spectra, the absorption at 3400 cm^{-1} in an FTIR spectrum was normalized to the value in the corresponding ATR-IR spectrum. The SVD/MEM analyses produced two positive absorption spectra, one corresponding to water absorption, and one corresponding to new absorption upon introduction of HCl (Figs. S1b and S1e). The ATR spectral components are both all positive, but there is some residual negative signal in the FTIR acid component around 3700 cm^{-1}, possibly due to over-correction by the dissimilarity term in the objection function or an effect from the variable pathlengths between FTIR signals. The ATR acid component includes some noise around 2100 cm^{-1}, likely because of the absorption by phonons in the diamond ATR prism.

We additionally include the two components resulting from the same analysis in NaCl solutions, which reveal a water component and water IR absorption perturbed by the Cl^- anions. In the ATR acid component spectrum, there is an absorption around 3450 cm^{-1} which seems to correspond to the O-H stretches perturbed by Cl^- but the correspondence is less clear in the FTIR component spectra. In Figs. S1c and S1f, we show how the two components change in amplitude with increasing acid concentration. In both ATR and FTIR spectra, the amplitude of the acid spectrum grows in linearly with acid concentration, whereas the amplitude of the water component slightly decreases as it is "diluted" with solute. The spectral components and amplitudes in the FTIR and ATR spectral series show consistent trends with each other, reinforcing the validity of the decomposition. However, the FTIR spectra contain just the absorption components of the complex index of refraction, and thus are a more direct comparison to the MULTIMODE calculations of the absorption spectra of the aqueous excess proton.

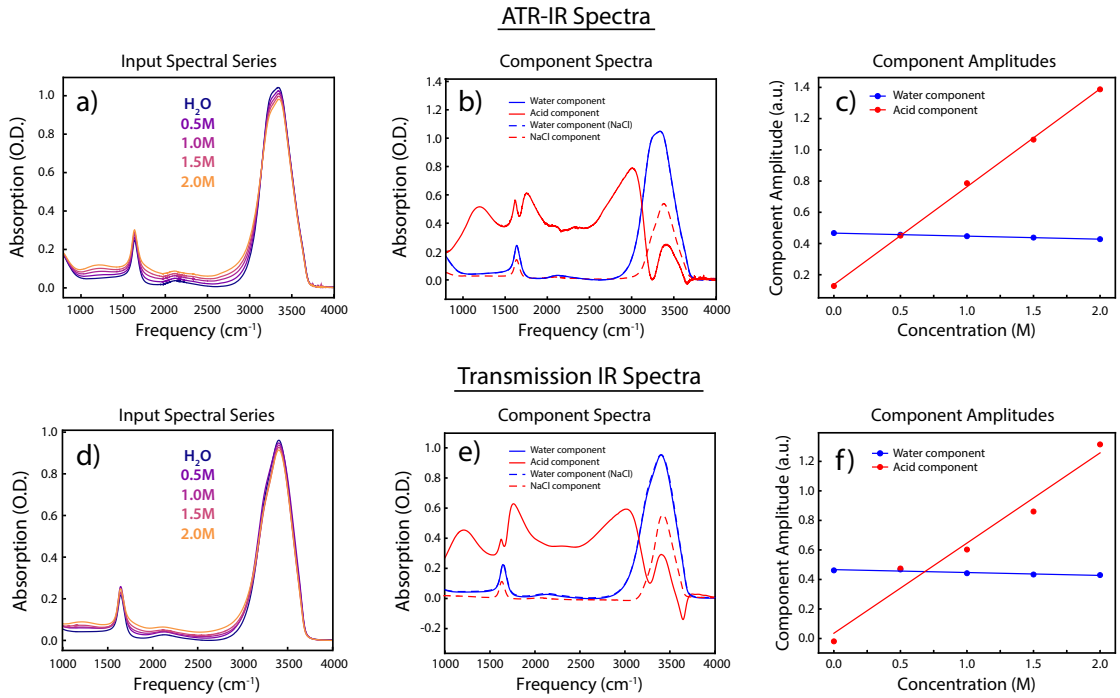

Figure S1: SVD/MEM analysis of the IR spectra of HCl in H_2O at various concentrations, measured with both FTIR and ATR-IR spectroscopy. (a) The input ATR-IR spectra of 0-2M HCl solutions, (b) the component spectra extracted from the SVD/MEM analysis, and (c) the amplitudes of the two components. (d)-(f) show the analogous curves for transmission IR spectra.

2. Calculation details

2.1 Brief description of MS-EVB simulation

The clusters were drawn from simulations of the MS-EVB 3.2 model for the excess proton in water, using the same methods and cluster configurations as in Ref. 1 The molecular dynamics simulations were performed using 256 SPC/Fw water molecules and one excess proton in the presence of a chloride ion. This reactive MD simulation was based on MS-EVB 3.2 model. The system was equilibrated in constant NVT ensemble for 1 ns with density 1.0 gm/cm^3 and room temperature 298 K. The time step was set as 0.5 fs with Nosé-Hoover chain thermostats. 1 ns constant NVE trajectories were run afterwards for generating small

clusters. The clusters are chosen with water molecules' oxygen atom within 5 Å radius from the central proton. In total, 2500 clusters were collected with time separation of 100 fs and we further chose 800 clusters randomly from the data set. The distribution of 800 clusters according to proton asymmetry coordinate (seen in Table S1) was kept the same as that in total data set. More details about the MD simulations are referred to ref 1.

Table S1: Population distribution of 800 chosen protonated water clusters from MS-EVB trajectory

δR_{OH} (Å)	Number	Population (%)	MS-EVB population(%)
0.0-0.1	184	23	23
0.1-0.2	240	30	30
0.2-0.3	248	31	31
0.3-0.4	112	14	14
0.4-	16	2	2
R_{OH} (Å)	Number	Population (%)	MS-EVB polulation(%)
1.00-1.05	82	10	11
1.05-1.10	278	35	35
1.10-1.15	290	36	36
1.15-1.20	142	18	17
1.20-	8	1	1
Zundel-like	424	53	53
Eigen-like	376	47	47

2.2 Potential energy surface and dipole moment surface

All the vibrational analysis of 800 clusters are based on accurate, *ab initio* many-body potential energy surface and dipole moment surface (PES/DMS) developed by us. Details of the PES/DMS can be referred to refs 2–4. Here, we give short description of the PES/DMS we are using. The potential energy is a many-body expansion given by

$$\begin{aligned} V =& V_h^{(1)} + \sum_i V_{w_i}^{(1)} + \sum_i V_{h,w_i}^{(2)} + \sum_{i,j} V_{w_i,w_j}^{(2)} \\ &+ \sum_{i,j,k} V_{w_i,w_j,w_k}^{(3)} + \sum_{i,j} V_{h,w_i,w_j}^{(3)} + \sum_{i,j,k} V_{h,w_i,w_j,w_k}^{(4)}. \end{aligned} \tag{10}$$

In this equation, $V_h^{(1)}$ is the potential energy surface of the monomer H_3O^+.[5] $V_{w_i}^{(1)}$, $V_{w_i,w_j}^{(2)}$ and $V_{w_i,w_j,w_k}^{(3)}$ are water 1-body, 2-body and 3-body interactions between water monomers, reported previously.[6–9] $V_{h,w_i}^{(2)}$, $V_{h,w_i,w_j}^{(3)}$ and $V_{h,w_i,w_j,w_k}^{(4)}$ are 2-body, 3-body and 4-body interactions between H_3O^+ and water monomers. The 2-body interaction is calculated from previous reported CCSD(T)/aVTZ based Zundel PES.[10] The 3-body interaction $V_{h,w_i,w_j}^{(3)}$ is fitted using Monomial Symmetrization Approach (MSA) from totally 107785 CCSD(T)/aVDZ energies with a simple correction (shift by +0.1 kcal/mol) considering Basis Set Superposition Error (BSSE). The 4-body term is a simple nonlinear fit from 1561 MP2/aVTZ data set. As to the dipole moment, similarly, this is expressed in a many-body form

$$\mu = \mu_h^{(1)} + \sum_i \mu_{w_i}^{(1)} + \sum_i \mu_{h,w_i}^{(2)} + \sum_{i,j} \mu_{w_i,w_j}^{(2)}. \tag{11}$$

This dipole moment expression includes hydronium 1-body dipole, $\mu_h^{(1)}$, water dipole moment terms $\mu_{w_i}^{(1)}$ and $\mu_{w_i,w_j}^{(2)}$ which are taken from the WHBB Dipole Moment Surface (DMS) and the hydronium-water 2-body dipole from Zundel DMS[10] As to the detailed assignment of H_3O^+ and water monomers, we applied a weighted sum of possible assignments to obtain accurate descriptions of potential energy and dipole moment.[2–4] Briefly, we first determine a "pivot" hydronium core with smallest sum of three OH nuclear distances. Other "possible" assignments of the hydronium core are also made and we apply a weight to each assignment based on the relevant OH lengths. We only consider the first solvation shell of the pivot hydronium core in order to reduce the computational cost and the total potential energy/dipole moment is a sum of all normalized assignments

$$\begin{aligned} V &= \frac{V_{pivot} + \sum_{i=1}^{3} w_i V_i}{1 + \sum_{i=1}^{3} w_i}. \\ \mu &= \frac{\mu_{pivot} + \sum_{i=1}^{3} w_i \mu_i}{1 + \sum_{i=1}^{3} w_i}, \end{aligned} \tag{12}$$

where w_i is the weight given to each assignment.

2.3 Local $H^+(H_2O)_2$ monomer analysis

From the 800 clusters drawn from MD simulations, $H^+(H_2O)_6$ clusters were selected which included the excess proton and its six closest waters. We conduct local $H^+(H_2O)_2$ monomer analysis in these clusters considering the central $H^+(H_2O)_2$ structure contributes most to the vibrational spectrum of aqueous proton. This approximation is the usual normal mode analysis of the chosen protonated water cluster but only on the central $H^+(H_2O)_2$ structure with all remaining water monomers fixed at their initial geometries.

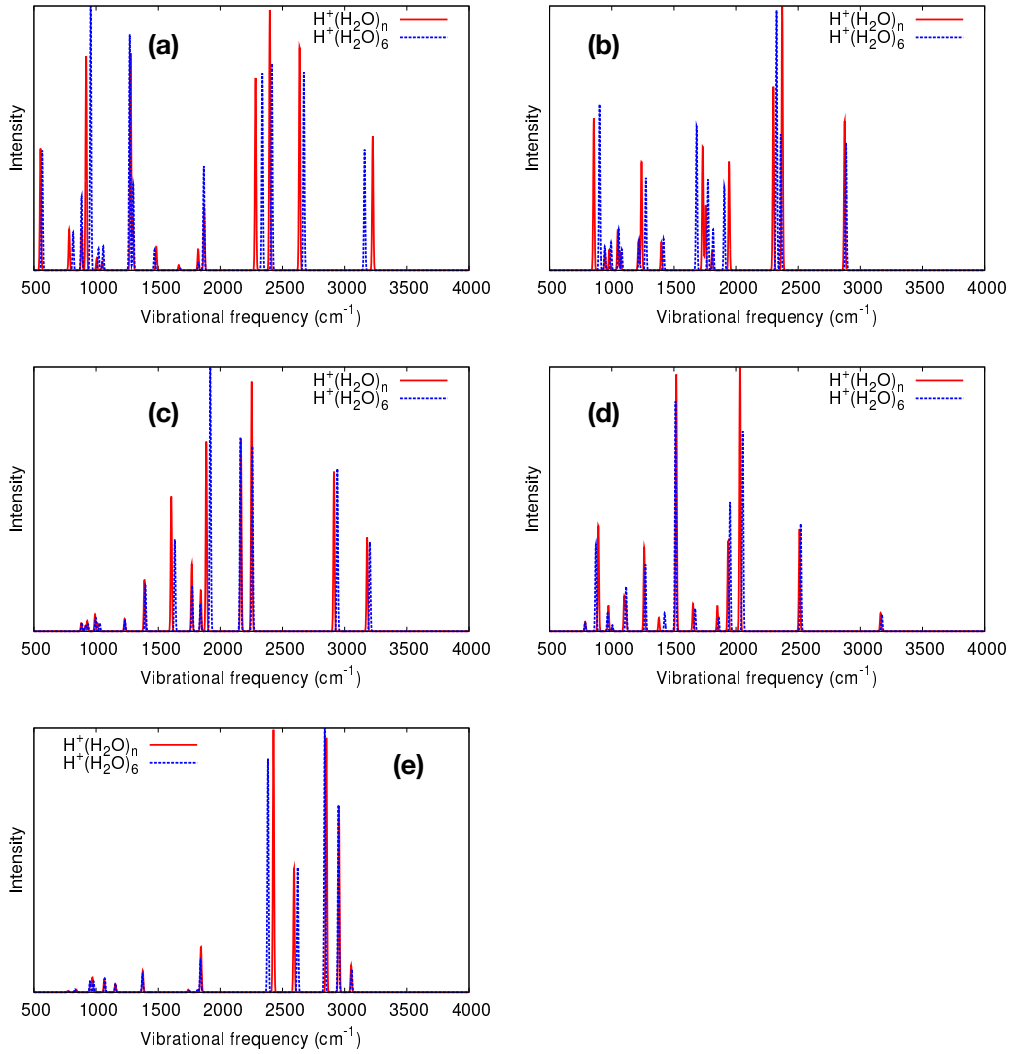

Figure S2: Examples of local monomer spectra of different clusters (a) δR_{OH} in 0-0.1 Å,(b) δR_{OH} in 0.1-0.2 Å,(c) δR_{OH} in 0.2-0.3 Å,(d) δR_{OH} in 0.3-0.4 Å. (e) $\delta R_{OH} > 0.4$ Å. Each plot includes calculation using cluster size of $H^+(H_2O)$n and $H^+(H_2O)_6$

2.4 VSCF/VCI calculations

For all 800 $H^+(H_2O)_6$ clusters, we conduct VSCF/VCI calculation using normal vectors in the local Zundel monomer analysis. For each cluster, we choose all vibrational modes from local Zundel monomer analysis with harmonic frequency larger than 700 cm^{-1}, resulting in 12-15 modes for each cluster. Here, we did not add any imaginary frequencies because it has been demonstrated that the imaginary frequencies do not contribute much to the total vibrational spectra.[1] We apply 4-mode representation of the potential (4MR) and generate the excitation space with maximum sum of singles, doubles, triples and quadruple excitations as 9, 8, 7, 6. The final size of the Hamiltonian matrix for each cluster is around 10,000 × 10,000. After obtaining the vibrational wavefunction and associated vibrational energies for all ground and excited VCI states, we calculate the quantum expectation values of OH and OO distances in the central $H^+(H_2O)_2$ structure using the ground state wavefunction. For the two isomers of gas phase $H^+(H_2O)_6$ cluster, to obtain the total spectra, we also conduct the local water monomer analysis of each isomer and then VSCF/VCI calculations of each water molecules separately. The spectra of each isomer include the contribution from central $H^+(H_2O)_2$ structure and also all remaining water molecules. All the above VSCF/VCI calculations are finished using code MULTIMODE which can conduct calculations using exact normal-coordinate Watson Hamiltonian.

3. Calculated spectra of Zundel and Eigen isomers of $H^+(H_2O)_6$

The structures of two isomers of $H^+(H_2O)_6$ are optimized at MP2/aVQZ level with NWChem6.6 suite of electronic structure codes. These strucutres are further optimized using Many-body PES and the obtained geometries are applied to the harmonic and VSCF/VCI analysis. Figure S2 and S3 show both double harmonic spectra and VSCF/VCI anharmonic spectra together with the experimental spectra. Table S2 lists peak positions in two spectra from

theory (harmonic and VSCF/VCI) and experiment. As seen in Figure S3, S4 and Table S2 , excellent agreement is reached between VSCF/VCI and experiment results. Comparing with the harmonic spectra, fully anharmonic treatment becomes necessary in capture the vibration information and VSCF/VCI works well in these cases. More discussion is provided in the main text.

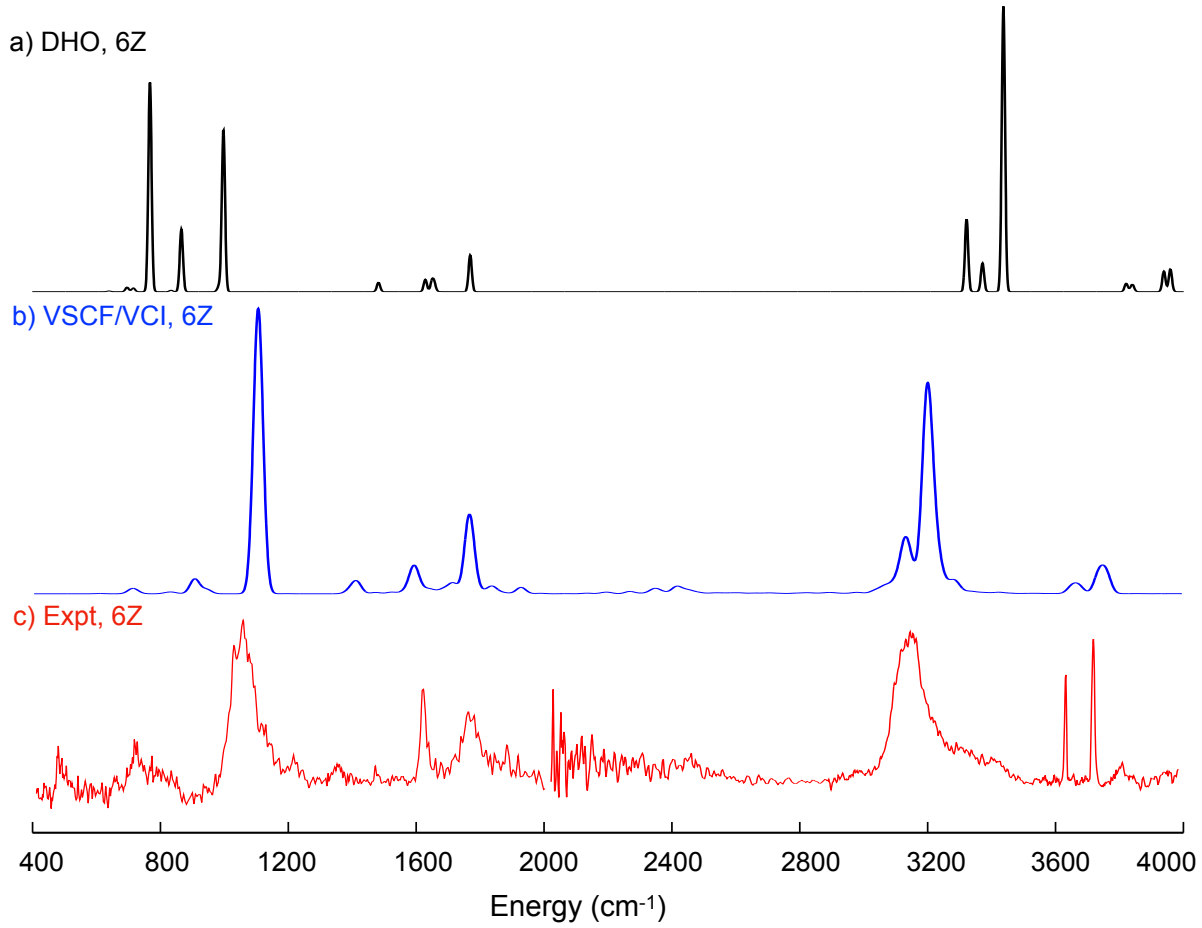

Figure S3: Calculated spectra of Zundel isomer of $H^+(H_2O)_6$

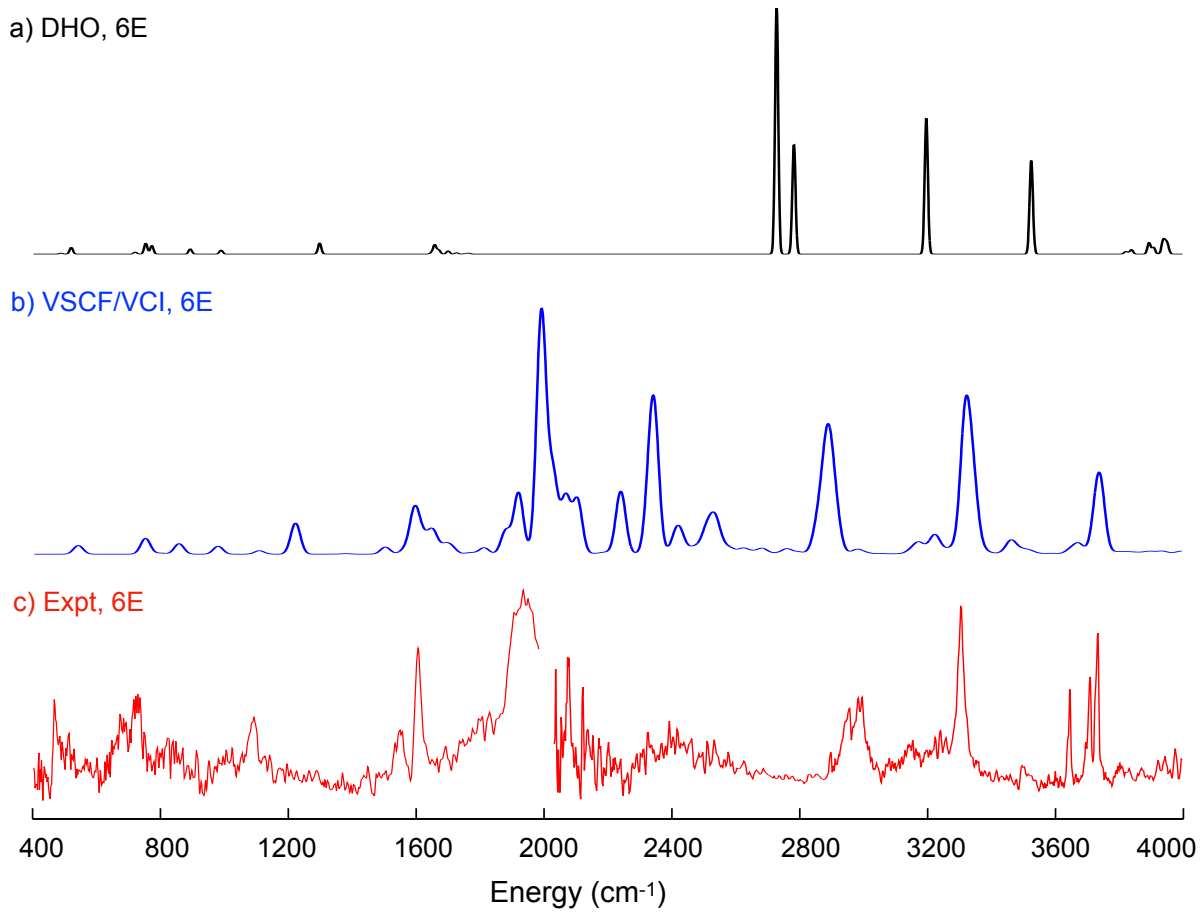

Figure S4: Calculated spectra of Eigen isomer of $H^+(H_2O)_6$

Table S2: Peak positions (cm^{-1}) of two isomers of $H^+(H_2O)_6$in the experimental spectra and calculated (harmonic and VSCF/VCI) values along with detailed assignments based on VSCF/VCI spectrum

Zundel isomer of $H^+(H_2O)_6$			
Exp. (cm^{-1})	Theory, VSCF/VCI (cm^{-1})	Theory, harmonic (cm^{-1})	Assignment
3737	3739, 3762	3938, 3958	H_2O free O-H asym-stretch
3651	3655, 3677	3844, 3856	H_2O free O-H sym-stretch
3167	3128, 3136	3353, 3371	$H_5O_2^+$ O-H stretch
1759	1765	1796	$H_5O_2^+$ shared proton bend
1618	1633, 1643	1671, 1692	H_2O bend
1050	**1102**	**964, 765**	$H_5O_2^+$ shared proton stretch
Eigen isomer of $H^+(H_2O)_6$			
Exp. (cm^{-1})	Theory, VSCF/VCI (cm^{-1})	Theory, harmonic (cm^{-1})	Assignment
3738	3744, 3750	3939, 3953	H_2O free O-H asym-stretch
3714	3730	3895	H_2O free O-H stretch
3651	3650, 3677	3825, 3840	H_2O free O-H sym-stretch
3312	3319	3525	H_2O H-bonded O-H stretch
3007	**2895**	**3197**	H_3O^+ O-H stretch
∼2425	**2345**	**2783**	H_3O^+ O-H stretch
1951	**1991**	**2728**	H_3O^+ O-H stretch
1618	1587,1619	1656,1669	H_2O bend
1097	1219	1297	H_3O^+ umbrella

bold font are frequencies of proton stretch/hydronium stretch. Experimental data are from refs. 11

4. Analysis of VSCF/VCI calculations from 800 $H^+(H_2O)_6$ clusters

4.1 VSCF/VCI spectra from some typical examples in different δR_{OH} region

Figure S5 mainly shows the VSCF/VCI spectra of 6 $H^+(H_2O)_6$ clusters which are some examples of $H^+(H_2O)_6$ in different δR_{OH} region. The six figures clearly show the shift trend of different bands in the spectra with different proton structure.

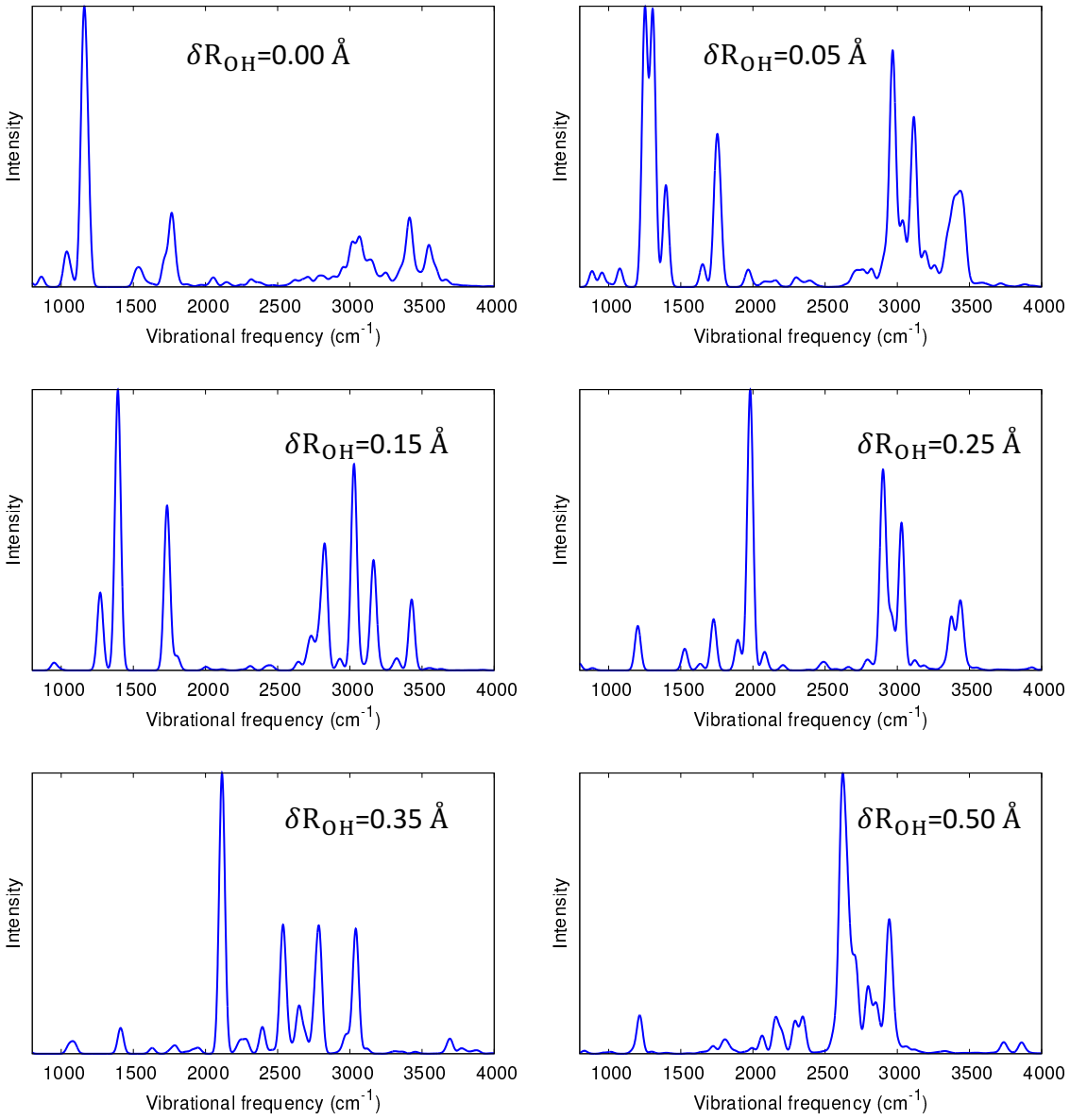

Figure S5: Decomposition of the calculated spectra of aqueous proton according to different $\langle \delta R_{OH} \rangle$ values.

Table S3: VCI coefficient of proton stretch, umbrella and bending motion at VCI state in different structures

VCI state	VCI coeff (proton stretch)	VCI coeff (umbrella)	VCI coeff (bend)
	δR_{OH}=0.00Å		
1153 cm^{-1}	0.78	0.28	0.14
1179 cm^{-1}	0.72	0.15	0.11
1517 cm^{-1}	0.16	0.47	0.24
1550 cm^{-1}	0.11	0.27	0.76
1716 cm^{-1}	0.12	0.18	0.31
1764 cm^{-1}	0.23	0.06	0.55
	δR_{OH}=0.05Å		
1251 cm^{-1}	0.48	0.44	0.13
1306 cm^{-1}	0.78	0.09	0.07
1397 cm^{-1}	0.37	0.78	0.16
1651 cm^{-1}	0.13	0.08	0.88
1738 cm^{-1}	0.10	0.04	0.85
1746 cm^{-1}	0.18	0.11	0.67
	δR_{OH}=0.15Å		
1270 cm^{-1}	0.02	0.94	0.00
1392 cm^{-1}	0.62	0.09	0.10
1731 cm^{-1}	0.08	0.00	0.85
1744 cm^{-1}	0.02	0.08	0.92
1803 cm^{-1}	0.13	0.00	0.85
	δR_{OH}=0.25Å		
1202 cm^{-1}	0.13	0.83	0.34
1527 cm^{-1}	0.11	0.23	0.85
1634 cm^{-1}	0.19	0.05	0.87
1727 cm^{-1}	0.40	0.12	0.77
1980 cm^{-1}	0.73	0.24	0.26
	δR_{OH}=0.35Å		
1411 cm^{-1}	0.38	0.71	0.06
1630 cm^{-1}	0.15	0.11	0.94
1790 cm^{-1}	0.17	0.18	0.85
1862 cm^{-1}	0.10	0.18	0.91
2114 cm^{-1}	0.64	0.48	0.30
	δR_{OH}=0.50Å		
1215 cm^{-1}	0.00	0.91	0.02
1677 cm^{-1}	0.00	0.00	0.54
1698 cm^{-1}	0.11	0.00	0.72
1723 cm^{-1}	0.10	0.18	0.91
1799 cm^{-1}	0.04	0.00	0.67
2614 cm^{-1}	0.52	0.00	0.11

4.2 Analysis of spectra from different structures and contribution from different motions

In Figure S6a), we plot the VSCF/VCI spectra of total 800 clusters, Zundel-like clusters and Eigen-like clusters which have been discussed in detail in Manuscript. From Figure 6b) to d) and also Figure 4 in the manuscript, we identify the square of VCI coefficient from different motions in the spectra.

In VSCF/VCI approach, we apply the harmonic oscillator of each normal mode as basis, so we first identify different normal modes in the central $H^+(H_2O)_2$ according to their normal vectors. For example, We identify one normal mode as proton stretch mode if one quanta along its normal vector results in the displacement of proton (any R_{OH^+}) larger than 0.01Å. The same criteria is used in identifying the flanking water stretch but on R_{OH} in flanking water monomer. We consider a normal mode to be umbrella motion if its normal vector results in change of three HOH angles of H_3O^+ structure larger than 0.8° and in the same phase. The bending motion is defined if two HOH angles of the special pair OH bond change larger than 1.0°. We also consider one normal mode as proton stretch/bend if its normal vector carries both proton stretch and bend characters.

After identifying the different normal modes included in the VSCF/VCI calculation, we record the VCI coefficient of different normal mode basis for each VCI state. The plots of square of VCI coefficient for different characters are shown in Figure S6b)-d).

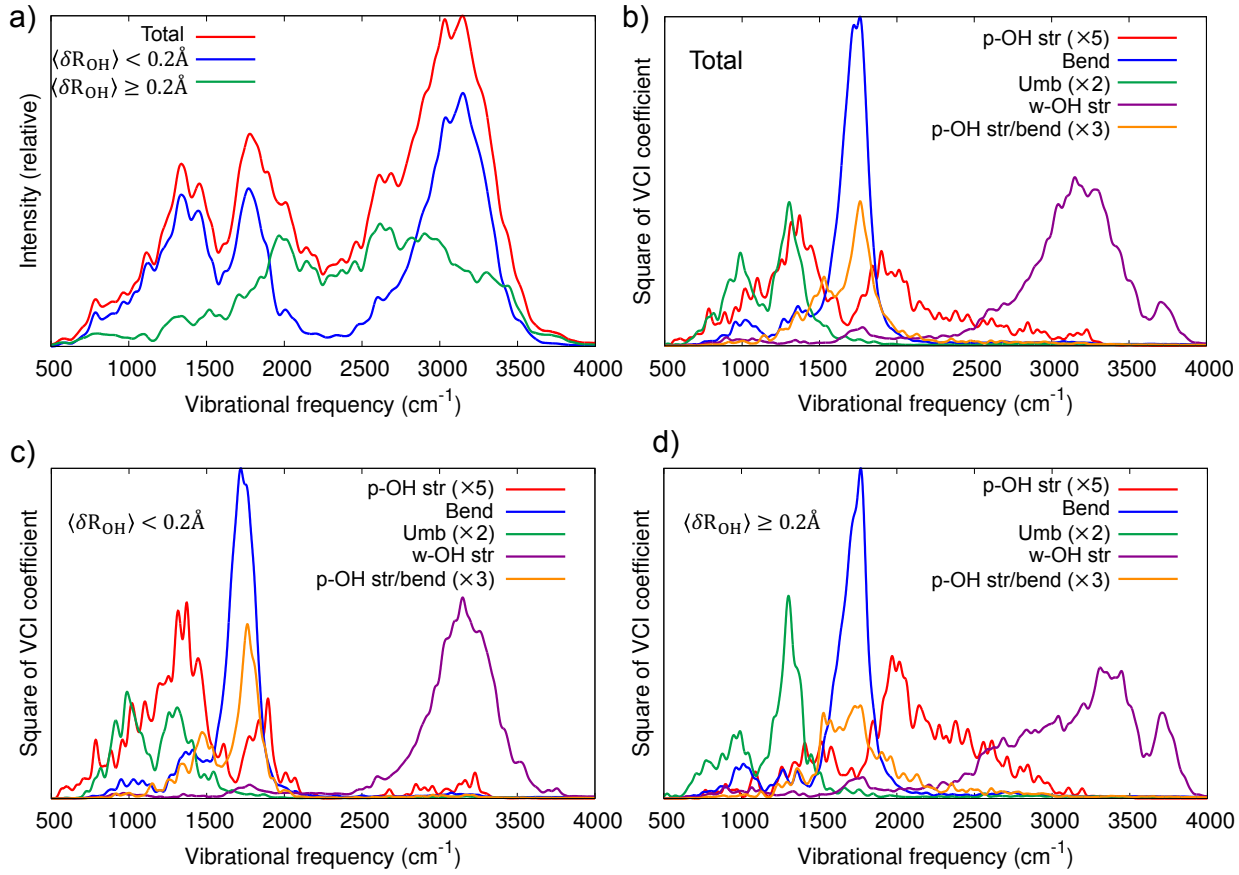

Figure S6: a) Calculated spectra of aqueous proton from all 800 clusters (red), Zundel-like clusters (blue) and Eigen-like clusters (green). b) Square of VCI coefficients for different vibrational motions in 800 clusters including p-OH str (proton stretch mode), bend, umb (umbrella mode), w-OH str (flanking water stretch mode), p-OH str/bend (mode with both proton stretch and bend characters) c)-d) Square of VCI coefficients for different vibrational motions in Zundel-like (c) and Eigen-like (d) clusters. The magnitude of proton stretch curve is amplified for 5 times and 2 times for umbrella mode, 3 times for p-OH str/bend motion

4.3 Statistical results of structural parameters from VSCF/VCI calculations

Table S4 and Figure S7-9 gives some statistics and comparisons of several important structural parameters, focusing on the initial structure's parameter and their quantum expectation values. Combining with Figure 5 in the main text, it can be concluded that the expectation value of the proton position changes most when it is treated quantum mechanically while the

oxygen atom is not influenced by nuclear effects significantly. This results in a larger fraction of configurations that would be characterized as Zundel-like, which has been observed before in other analyses that treated the proton quantum mechanically.[?]

Table S4: Population distribution of 800 chosen protonated water clusters based on VSCF/VCI calculations

$\langle \delta R_{OH} \rangle$ (Å)	Number	Population (%)
0.0-0.1	287	36
0.1-0.2	152	19
0.2-0.3	185	23
0.3-0.4	143	18
0.4-	33	4
$\langle R_{OH} \rangle$ (Å)	Number	Population (%)
1.00-1.05	55	7
1.05-1.10	246	31
1.10-1.15	220	27
1.15-1.20	243	30
1.20-	36	5
Zundel-like	439	55
Eigen-like	351	45

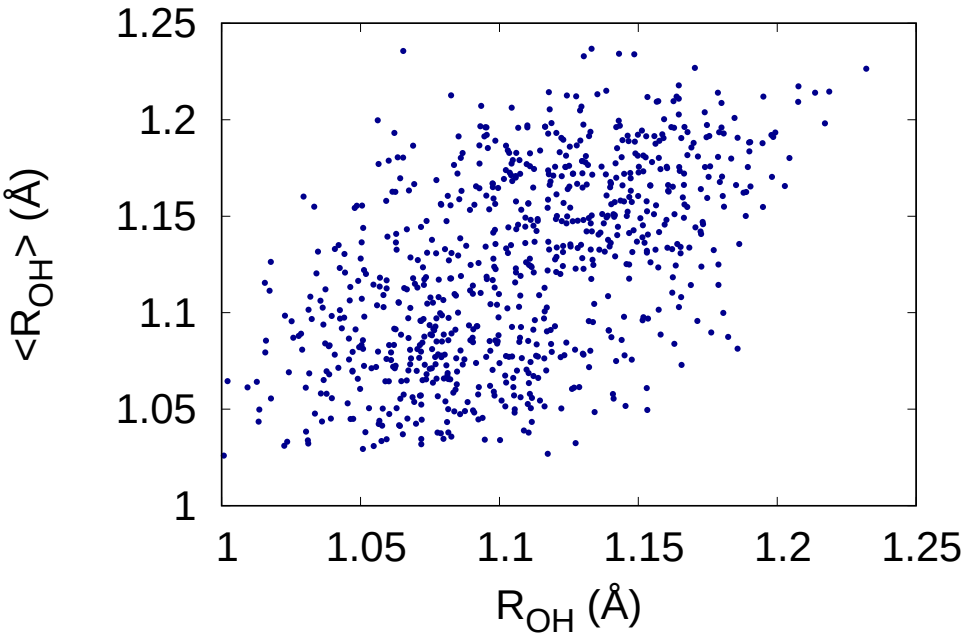

Figure S7: Scatter plot of R_{OH} in clusters and their expectation values $\langle R_{OH} \rangle$ from VSCF/VCI calculations

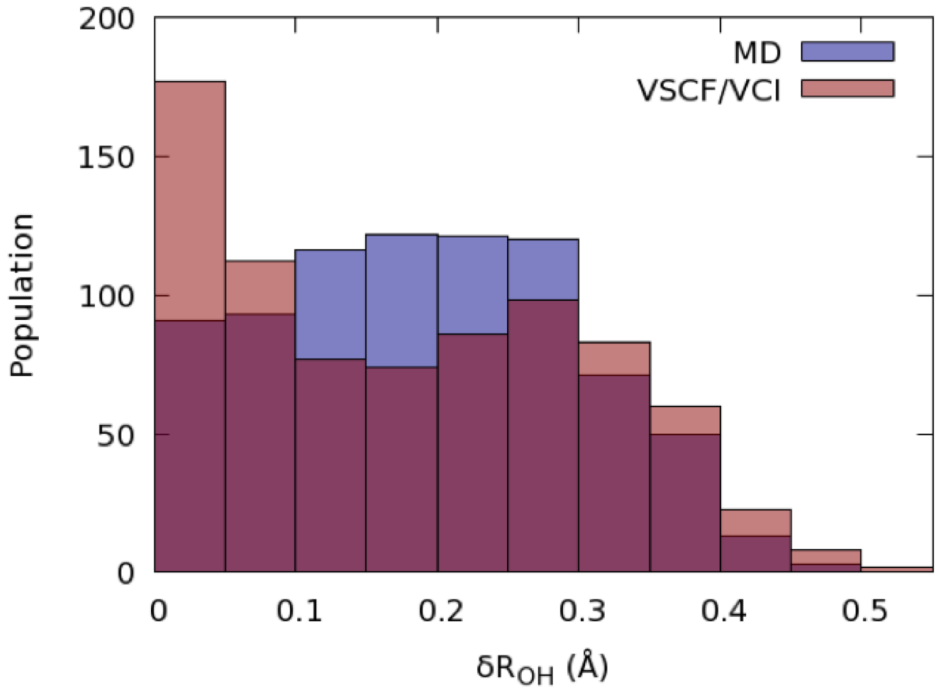

Figure S8: Distribution of δR_{OH} from 800 MS-EVB MD structures and their expectation values $\langle \delta R_{OH} \rangle$ from VSCF/VCI calculations

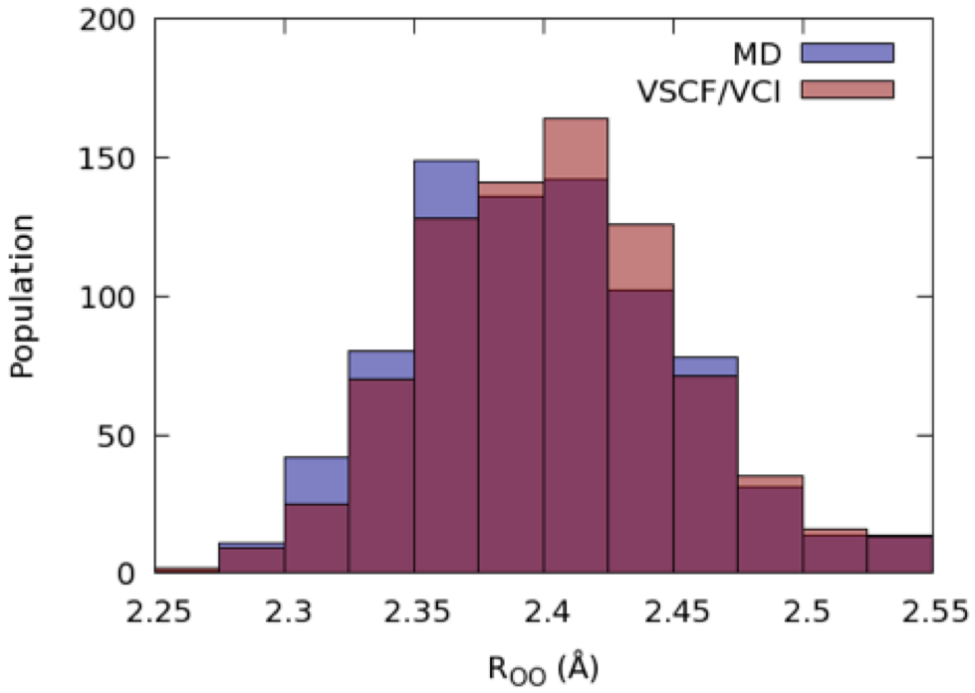

Figure S9: Distribution of R_{OO} from 800 MS-EVB MD structures and their expectation values $\langle R_{OO} \rangle$ from VSCF/VCI calculations

4.4 Correlation relationship between VSCF/VCI anharmonic proton stretch and structural parameters

Figure S10-S13 are plots that link the structural parameters with the associated proton stretch. We determine the anharmonic proton stretch in each cluster from the VSCF/VCI states with largest VCI coefficients of the proton stretch mode. Figure S10 clearly shows that R_{OO} can not provide clear relationship with proton stretch and the result does not change when we using its expectation values. Using the other parameters of R_{OH} and δR_{OH}, there still does not exist clear correlation relationship between structural parameters from initial structure and proton stretch. However, the situation changes when we consider the expectation values of these structural parameters. The correlation relationship becomes clear and strong especially when we use R_{OH} in Figure S11.

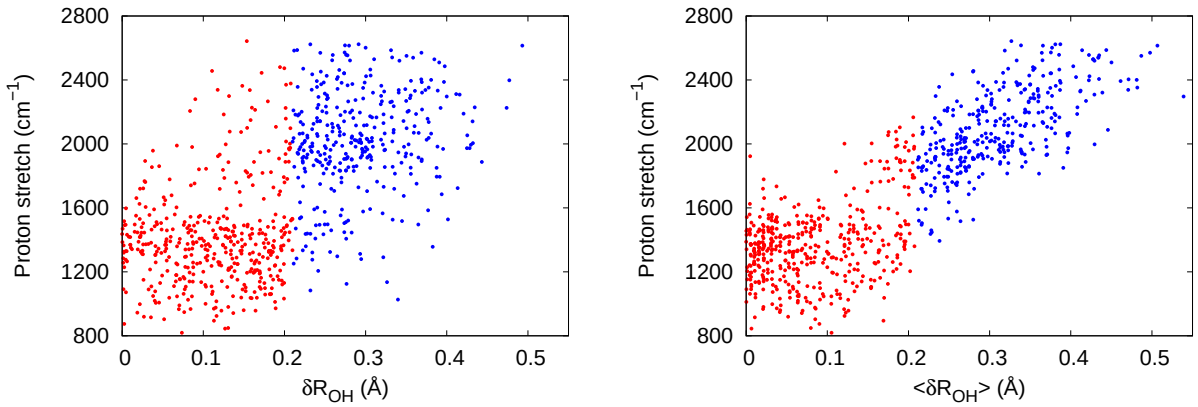

Figure S10: Relationship between proton stretch and δR_{OH}(left) and its expectation values (right). Red points are from structures with δR_{OH}<0.2Å. Blue points are from structures with $\delta R_{OH} \geq 0.2$Å.

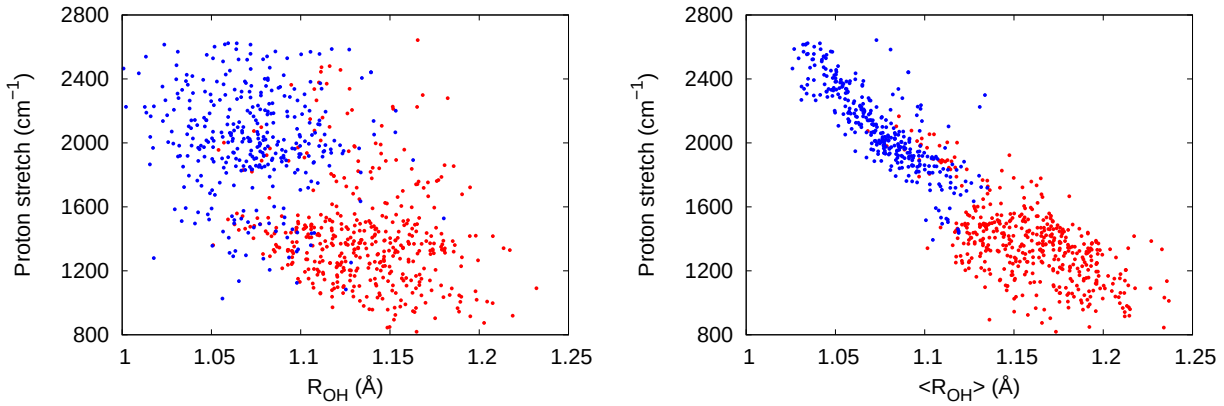

Figure S11: Relationship between proton stretch and R_{OH}(left) and its expectation values (right). Red points are from structures with δR_{OH}<0.2Å. Blue points are from structures with $\delta R_{OH} \geq 0.2$Å. Linear least square fit for red points: $\nu = a \cdot \langle R_{OH} \rangle + b$, a=-5220.1, b=7422.6, R_2=0.42, RMS=187.6 cm^{-1}. Linear least square fit for blue points: $\nu = c \cdot \langle R_{OH} \rangle + d$, c=-9018.0, b=11794.0, R_2=0.70, RMS=142.1 cm^{-1}. Linear least square fit for all 800 points: $\nu = e \cdot \langle R_{OH} \rangle + f$, e=-7941.2, f=10603.0, R_2=0.82, RMS=183.5 cm^{-1}.

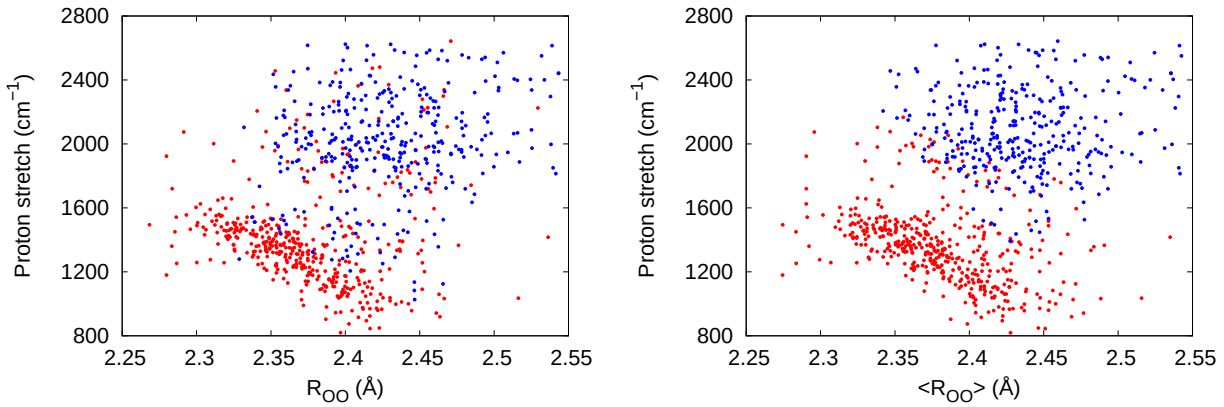

Figure S12: Relationship between proton stretch and R_{OO}(left) and its expectation values (right). Red points are from structures with δR_{OH}<0.2Å. Blue points are from structures with $\delta R_{OH} \geq 0.2$Å. Linear least square fit for red points: $\nu = a \cdot \langle R_{OO} \rangle + b$, a=-2342.0, b=6944.0, R_2=0.15, RMS=226.3 cm^{-1}.

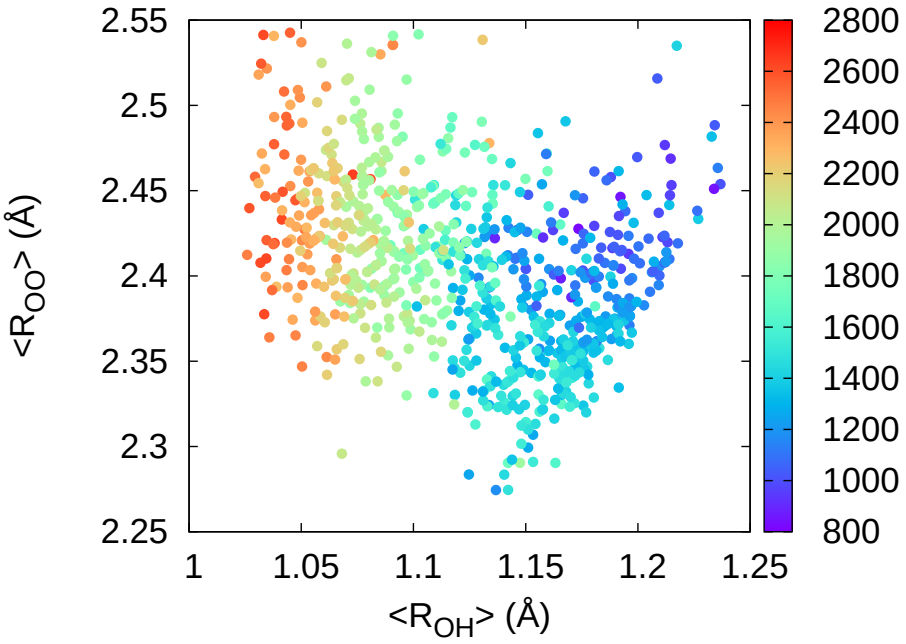

Figure S13: Relationship between $\langle R_{OO} \rangle$, $\langle R_{OH} \rangle$ and proton stretch frequency in cm^{-1} (color bar).

4.4 Decomposition of the spectra into different structures according to $\langle \delta \mathbf{R_{OH}} \rangle$ values

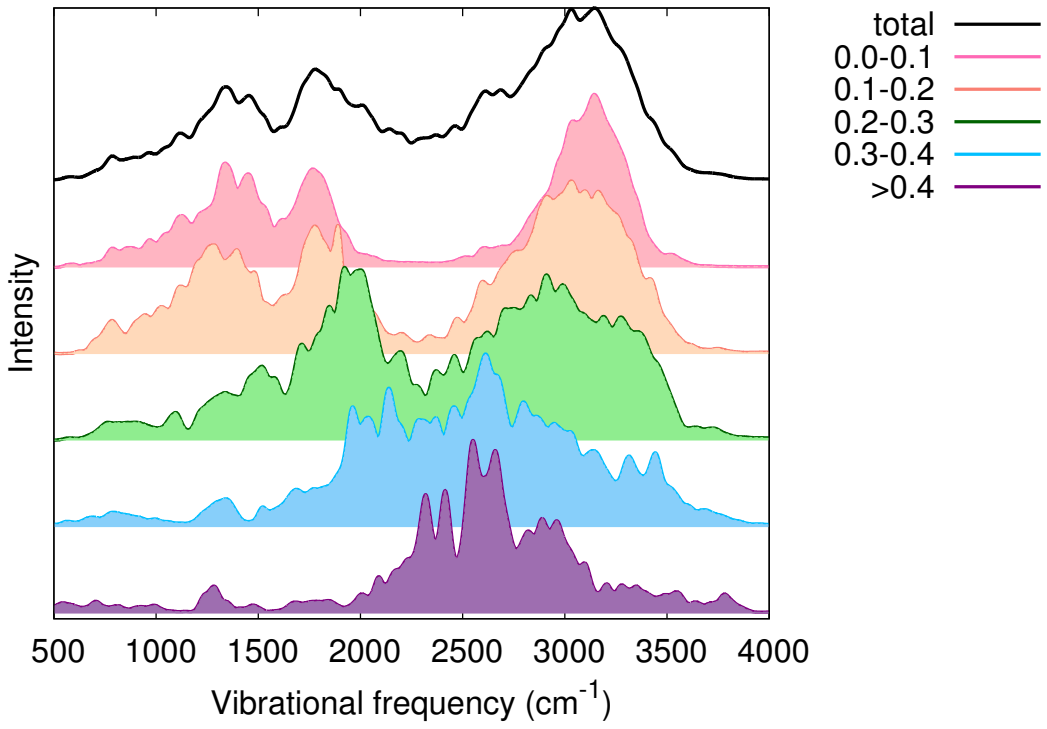

Figure S14: Decomposition of the calculated spectra of aqueous proton according to different $\langle \delta \mathrm{R_{OH}} \rangle$ values.

References

(1) Biswas, R.; Carpenter, W.; Fournier, J. A.; Voth, G. A.; Tokmakoff, A. IR Spectral Assignments for the Hydrated Excess Proton in Liquid Water. *J. Chem. Phys.* **2017**, *146*, 154507.

(2) Yu, Q.; Bowman, J. M. High-Level Quantum Calculations of the IR Spectra of the Eigen, Zundel, and Ring Isomers of $H^+(H_2O)_4$ Find a Single Match to Experiment. *J. Am. Chem. Soc.* **2017**, *139*, 10984–10987.

(3) Qu, C.; Yu, Q.; Bowman, J. M. Permutationally Invariant Potential Energy Surfaces. *Annu. Rev. Phys. Chem.* **2018**, *69*, 6.1–6.25.

(4) Heindel, J. P.; Yu, Q.; Bowman, J. M.; Xantheas, S. S. Benchmark Electronic Structure Calculations for $H_3O^+(H_2O)_n$, n=0-5 Clusters and Tests of an Existing 1,2,3-body Potential Energy Surface with a New 4-body Correction. *J. Chem. Theory Comput.* **2018**, *14*, 4553–3566.

(5) Yu, Q.; Bowman, J. M. Ab Initio Potential for $H_3O^+ \rightarrow H^+ + H_2O$: A Step to a Many-Body Representation of the Hydrated Proton? *J. Chem. Theory Comput.* **2016**, *12*, 5284.

(6) Partridge, H.; Schwenke, D. W. The Determination of an Accurate Isotope Dependent Potential Energy Surface for Water from Extensive Ab Initio Calculations and Experimental Data. *J. Chem. Phys.* **1997**, *106*, 4618.

(7) Wang, Y.; Shepler, B. C.; Braams, B. J.; Bowman, J. M. Full-dimensional, Ab Initio Potential Energy and Dipole Moment Surfaces for Water. *J. Chem. Phys.* **2009**, *131*, 054511.

(8) Wang, Y.; Huang, X.; Shepler, B. C.; Braams, B. J.; Bowman, J. M. Flexible, Ab Initio Potential, and Dipole Moment Surfaces for Water. I. Tests and Applications for Clusters up to the 22-mer. *J. Chem. Phys.* **2011**, *134*, 094509.

(9) Wang, Y.; Bowman, J. M. Ab Initio Potential and Dipole Moment Surfaces for Water. II. Local-monomer Calculations of the Infrared Spectra of Water Clusters. *J. Chem. Phys.* **2011**, *134*, 154510.

(10) Huang, X.; Braams, B. J.; Bowman, J. M. Ab initio Potential Energy and Dipole Moment Surfaces for $H_5O_2^+$. *J. Chem. Phys.* **2005**, *122*, 044308.

(11) Heine, N.; Fagiani, M. R.; Rossi, M.; Wende, T.; Berden, G.; Blum, V.; Asmis, K. R. Isomer-selective Detection of Hydrogen-bond Vibrations in the Protonated Water Hexamer. *J. A. Chem. Soc.* **2013**, *135*, 8266–8273.