Supplementary Information for

Nonlinear rotational spectroscopy reveals many-body interactions in water molecules

Yaqing Zhang, Jiaojian Shi, Xian Li, Stephen L. Coy, Robert W. Field, and Keith A. Nelson*

*Corresponding author. Email: kanelson@mit.edu.

This PDF file includes:

Simulation details

Coordinate system

Before we discuss any simulation details, it is beneficial to clarify the relevant coordinate systems (**Fig. S5**). The excitation THz field is defined in the space-fixed frame (X, Y, Z). For a linearly polarized THz field, the polarization vector is assigned along the $+Z$ axis without any loss of generality.

Hamiltonian

A quantum mechanical rotational system is governed by the rotational Hamiltonian $H(t)$, which consists of a field independent Hamiltonian H_0, and a field-molecule interaction term $H_1(t)$,

$$H(t) = H_0 + H_1(t) \tag{S.1}$$

where $H(t)$ is the total rotational Hamiltonian.

The field-free term is

$$H_0 = AJ_a^2 + BJ_b^2 + CJ_c^2 \tag{S.2}$$

where A, B, and C are three rotational constants inversely proportional to the three components of moment of inertia along a-, b- and c- axes, respectively (J_a, J_b, and J_c are the three components of the total angular momentum J).

If we choose the following coordinate identifications:

$$a \leftrightarrow z, b \leftrightarrow x, c \leftrightarrow y \tag{S.3}$$

Accordingly,

$$H_0 = AJ_z^2 + BJ_x^2 + CJ_y^2 \tag{S.4}$$

The matrix elements, for example in the basis set of the symmetric top

$$\langle JKM | H_0 | J'K'M' \rangle \tag{S.5}$$

have been evaluated in textbooks(1). The other part of the total Hamiltonian is dictated by the field-dipole interaction,

$$H_1(t) = -\boldsymbol{\mu} \cdot \boldsymbol{E}_{THz}(t) \tag{S.6}$$

which can be expressed in the language of the irreducible spherical tensor operator as follows(1),

$$H_1(t) = -\sum_{p=-1}^{1} (-1)^p T_p^{(1)}(\mu) T_{-p}^{(1)}(E_{THz}(t)) \tag{S.7}$$

Evaluation of matrix elements

In the Cartesian coordinate system (space-fixed frame), we define the components of the THz field,

$$E_{THz}(t) = (E_X, E_Y, E_Z) = (0, 0, E_{THz}(t)) \tag{S.8}$$

2

$$T_0^1(E) = E_Z = E_{THz}(t)$$

$$T_{\pm 1}^1(E) = \frac{\mp(E_X \pm iE_Y)}{\sqrt{2}} = 0 \tag{S.9}$$

On the other hand, the tensor components of the permanent dipole vector are given in the molecule-fixed frame in a similar manner,

$$\mu = (\mu_x, \mu_y, \mu_z) = (\mu_b, \mu_c, \mu_a) \tag{S.10}$$

$$T_0^{(1)}(\mu) = \mu_z = 0$$

$$T_{\pm 1}^{(1)}(\mu) = \frac{\mp(\mu_x \pm i\mu_y)}{\sqrt{2}} = \frac{\mp\mu_b}{\sqrt{2}} \tag{S.11}$$

In order to evaluate the field-molecule interaction $H_1(t)$, a coordinate transformation is needed. If we transform the dipole moment from molecule-fixed frame into space-fixed frame, that is

$$T_0^{(1)}(\mu)|_{space} = \sum_{q=-1}^{1} D_{0q}^{(1)*} T_q^{(1)}(\mu)|_{molecule} = \frac{\mu}{\sqrt{2}}(D_{0,-1}^{(1)*} - D_{0,1}^{(1)*})$$

$$T_1^{(1)}(\mu)|_{space} = \sum_{q=-1}^{1} D_{1q}^{(1)*} T_q^{(1)}(\mu)|_{molecule} = \frac{\mu}{\sqrt{2}}(D_{1,-1}^{(1)*} - D_{1,1}^{(1)*}) \tag{S.12}$$

$$T_{-1}^{(1)}(\mu)|_{space} = \sum_{q=-1}^{1} D_{-1q}^{(1)*} T_q^{(1)}(\mu)|_{molecule} = \frac{\mu}{\sqrt{2}}(D_{-1,-1}^{(1)*} - D_{-1,1}^{(1)*})$$

then $H_1(t)$ is ready to be evaluated in that both the dipole moment and the excitation THz field are in the same frame (in this case, space-fixed frame).

The matrix element of $H_1(t)$ in the basis set of the symmetric top wavefunction $|JKM\rangle$ is given as

$$\langle JKM \mid H_1(t) \mid J'K'M' \rangle$$

$$= -\mu_b E_{THz}(t) \left\langle JKM \mid \frac{1}{\sqrt{2}}(D_{0,-1}^{(1)*} - D_{0,1}^{(1)*}) \mid J'K'M' \right\rangle$$

$$= -\frac{\mu_b E_{THz}(t)}{\sqrt{2}}(-1)^{M-K}\left[(2J+1)(2J'+1)\right]^{1/2} \times \tag{S.13}$$

$$\left\{ \begin{pmatrix} J & 1 & J' \\ -K & -1 & K' \end{pmatrix} - \begin{pmatrix} J & 1 & J' \\ -K & 1 & K' \end{pmatrix} \right\} \begin{pmatrix} J & 1 & J' \\ -M & 0 & M' \end{pmatrix}$$

where D is a rank (1) Wigner rotational matrix, and $\begin{pmatrix} \cdots \end{pmatrix}$ is a 3-j symbol(1). The nonvanishing condition is

$$\Delta M = 0; \Delta J = 0, \pm 1; \Delta K = \pm 1 \tag{S.14}$$

The following relation is used in the calculation,

$$\left\langle J\Omega M \mid D_{Pq}^{(k)*}(\omega) \mid J'\Omega'M' \right\rangle$$

$$= (-1)^{M-\Omega}\left[(2J+1)(2J'+1)\right]^{1/2} \times \begin{pmatrix} J & k & J' \\ -\Omega & q & \Omega' \end{pmatrix}\begin{pmatrix} J & k & J' \\ -M & P & M' \end{pmatrix} \tag{S.15}$$

The Wigner 3-j symbol is evaluated numerically by the Racah formula(1). Alternatively, the matrix element can also be expressed in the basis set of the asymmetric-top wavefunction $|JK_aK_cM> = |J\tau M>$. Expanding $|J\tau M>$ as a linear combination of the symmetric-top wavefunction, one obtains

$$|J'\tau'M'\rangle = \sum_{K'} a_{K'\tau'}^{J'} |J'K'M'\rangle$$

$$\langle J\tau M \mid = \sum_K \langle JKM \mid (a_{K\tau}^J)* \tag{S.16}$$

Inserting the identity operator, we find,

$$\left\langle J\tau M \mid H_1(t) \mid J'\tau'M' \right\rangle$$

$$= \left\langle JK_aK_cM \mid -\mu_b E_{THz}(t)\cos\theta_b^Z \mid J'K_a'K_c'M' \right\rangle$$

$$= \sum_{K''}\sum_{K'''} \left\langle J\tau M \mid J''K''M'' \right\rangle\left\langle J''K''M'' \mid H_1 \mid J'''K'''M''' \right\rangle\left\langle J'''K'''M''' \mid J'\tau'M' \right\rangle \tag{S.17}$$

$$= \sum_K\sum_{K'} (a_{K\tau}^J)* \left\langle JKM \mid H_1(t) \mid J'K'M' \right\rangle(a_{K'\tau'}^{J'})$$

Therefore,

$$\left\langle J\tau M \mid H_1(t) \mid J'\tau'M' \right\rangle$$

$$= \begin{pmatrix} a_{K_{(1)}\tau}^J \\ a_{K_{(2)}\tau}^J \\ \vdots \\ a_{K_{(n)}\tau}^J \end{pmatrix}^* \begin{pmatrix} \langle JK|H_1|J'K'\rangle_{1,1} & \cdots & \langle JK|H_1|J'K'\rangle_{1,n} \\ \langle JK|H_1|J'K'\rangle_{2,1} & \cdots & \langle JK|H_1|J'K'\rangle_{2,n} \\ \vdots & \ddots & \vdots \\ \langle JK|H_1|J'K'\rangle_{n,1} & \cdots & \langle JK|H_1|J'K'\rangle_{n,n} \end{pmatrix}\begin{pmatrix} a_{K_{(1)}\tau}^J \\ a_{K_{(2)}\tau}^J \\ \vdots \\ a_{K_{(n)}\tau}^J \end{pmatrix} \tag{S.18}$$

$$= (\langle JKM \mid J\tau M \rangle)^* (H_1(t)_{JKM,J'K'M'})(\langle J'K'M' \mid J'\tau'M' \rangle)$$

The nonvanishing condition is the same as above. A typical implementation of the density matrix has been illustrated (**Fig. S6**) below.

Time evolution

At the condition of thermal equilibrium, the initial density matrix is

$$\rho(0) = \frac{e^{-\beta H_0}}{Tr(e^{-\beta H_0})}$$

$$\beta = \frac{1}{k_B T} \tag{S.19}$$

4

or simply by

$$\rho(0) = \sum_i P_i |\psi_i\rangle\langle\psi_i| \tag{S.20}$$

where P_i is the statistical probability of the system in the normalized pure state $|\psi_i\rangle$,

$$\sum_i P_i = 1 \tag{S.21}$$

The system evolution is described by the well-known Liouville-von Neumann equation,

$$\frac{d\rho(t)}{dt} = \frac{-i}{\hbar}[\rho(t), H(t)] . \tag{S.22}$$

the time-dependent density matrix can be calculated by the time propagation operator,

$$\rho(t) = e^{-i\frac{H(t)}{\hbar}t}\rho(0)e^{+i\frac{H(t)}{\hbar}t} \tag{S.23}$$

Orientation calculation and THz field representation

The ensemble averaged orientation is,

$$\langle\cos\theta\rangle(t) = Tr[\rho(t)\cos\theta] \tag{S.24}$$

And the nonlinear orientation factor(2),

$$\langle\cos\theta(\tau,t)\rangle|_{NL} = \langle\cos\theta(\tau,t)\rangle|_{12} - \langle\cos\theta(\tau,t)\rangle|_1 - \langle\cos\theta(\tau,t)\rangle|_2 - \langle\cos\theta(\tau,t)\rangle|_{BG} \tag{S.25}$$

Where subscript 12 denotes both THz fields on, 1 denotes only THz field 1 on, 2 denotes only THz field 2 on, and BG denotes both THz fields are off (see **Fig. S4**). Note that the orientation angle θ here is understood as the angle between THz field polarization (along $+Z$ axis) and the molecular dipole moment (along b axis), so that $\theta = \theta_{b|molecule}^{Z|space}$, which can be further evaluated by the Wigner rotational matrix,

$$\theta = \theta_{b|molecule}^{Z|space} = \frac{1}{\sqrt{2}}(D_{0,-1}^{(1)*} - D_{0,1}^{(1)*}) \tag{S.26}$$

A THz field pulse is characterized by a Gaussian-envelope modulated sine or cosine wave with a center frequency ω_0 ,

$$E_{THz}(t) = E_1 e^{-t^2/\tau^2} \cos(\omega_0 t + \varphi_1) \tag{S.27}$$

Two pulses with a time delay are described by,

$$E_{THz}(t) = E_1 e^{-t^2/\tau^2} \cos(\omega_0 t + \varphi_1) + E_2 e^{-(t+\delta t)^2/\tau^2} \cos(\omega_0(t+\delta t) + \varphi_2) \tag{S.28}$$

For the sake of simplicity, the initial phases are set to zero, $\varphi_1 = \varphi_2 = 0$.

Watson's reduced Hamiltonian

Centrifugal distortion has to be taken into consideration in order to accurately find the positions of rotational transition lines. Non-rigidity causes significant spectral line shifts, on the order of 1% of the line frequencies, and needs to be accounted for to avoid misassignments of the spectral lines. The line positions can be calculated with around 0.1% accuracy (absolute error on the order of 1 GHz) with centrifugal distortion taken into account (**Fig. S7-S10**).

One approach to including centrifugal distortion is Watson's reduced Hamiltonian. This approach is based on a perturbative treatment to high order, and is used extensively although it inherently diverges(3). Watson's reduced Hamiltonian, up to the 6th order, is sufficient to address our concern. The total Hamiltonian is written as(4, 5),

$$H = H_r + H_d^{(4)} + H_d^{(6)} \tag{S.29}$$

where the unperturbed (2nd order) term is

$$H_r = \frac{1}{2}(B+C)J^2 + \left[A - \frac{1}{2}(B+C)\right](J_z^2 - b_p J_-^2) \tag{S.30}$$

The 4th order term is given as

$$H_d^{(4)} = -\Delta_J J^4 - \Delta_{JK} J^2 J_z^2 - \Delta_K J_z^4 - 2\delta_J J^2 J_-^2 - \delta_K (J_z^2 J_-^2 + J_-^2 J_z^2) \tag{S.31}$$

And the 6th order term,

$$\begin{aligned}
H_d^{(6)} &= H_J J^6 + H_{JK} J^4 J_z^2 + \\
&\quad H_{KJ} J^2 J_z^4 + H_K J_z^6 + 2h_J J^4 J_-^2 + h_{JK} J^2 (J_z^2 J_-^2 + J_-^2 J_z^2) + h_K (J_z^4 J_-^2 + J_-^2 J_z^4)
\end{aligned} \tag{S.32}$$

with

$$\begin{aligned}
b_p &= \frac{C-B}{2A-B-C} \\
J_-^2 &= J_x^2 - J_y^2
\end{aligned} \tag{S.33}$$

The rotational constants used in this work are listed (**Table S1**).

In order to implement the calculation of matrix elements of the angular moment operator, we follow the method in textbooks(1) using commutation rules and matrix multiplication. Matrix elements of the angular momentum operator through 6th order are evaluated and listed below (**Tables S3 and S4**).

Qualitative simulation of many-body interactions between water molecules

We model water intermolecular interactions qualitatively by treating an individual transition as a two-level quantum system and considering three nearby molecules. The three two-level systems yield together a four-level system via intermolecular interactions such as dipole-dipole interactions among the molecules (**Fig. S13**). The singly excited and doubly excited energy levels are degenerate, each consisting of three energy levels. The small energy shifts induced by many-body interactions are not shown in the figure. Combined three-level and multiple-level pictures can be interpreted in a similar way. A small energy shift of the combined energy levels is needed to break

the symmetry of the combined states so that emission signals from distinct pathways do not completely cancel each other(6-8).

In order to understand the many-body interactions among water molecules, a qualitative simulation is developed. The orientation simulation follows the time propagation mechanism described above, with a new Hamiltonian in which the interactions give rise to perturbation terms represented by off-diagonal elements (labelled as a in **Fig. S14**). Diagonalization gives a set of new eigen-energies. Generally, the coupling strength is very small ($a \ll \hbar\omega$) and a multi-level quantum system with an approximately equal energy spacing $\hbar\omega$ between neighboring, non-degenerate energy levels is created by many-body interactions. Note that the coupling elements (the assignments of a) are not necessarily the same as shown in **Fig. S14**. Different arrangements of the coupling elements produce similar energy levels.

Up to 6 combined energy levels have been included in the simulation. The transition frequency is specifically chosen as 0.558 THz (the first ortho-type transition $1_{01} \leftrightarrow 1_{10}$) in water. 2Q and 3Q peaks with respect to the 0.753 THz transition can be implemented in a similar way. The transition dipole matrix is set as that of a linear or prolate symmetric-top molecule in which only coherences between distinct J states are considered. A quantitative and accurate simulation of both diagonal and off-diagonal features may require a complete knowledge of the intermolecular interaction and the newly formed Hamiltonian.

Supplementary Text

Rotational energy levels of water molecules

The population distribution of rotational states of a linear molecule at the condition of thermal equilibrium follows the Boltzmann distribution(5) (**Fig. S11**)

$$Population \propto (2J+1)e^{-E_{JKaKc}/k_B T} \tag{S.34}$$

where the factor $(2J + 1)$ is introduced due to the M sublevel degeneracy. As a result, at or not far from ambient temperature, most of the rotational population stays in low J states. The majority of the rotational population is under $J = 10$ at room temperature, and this value is used as a cut-off number to simplify the simulation. We note that (i) the population distribution within a single J state is not uniform (**Fig. S11**), in contrast to linear or symmetric-top molecules; (ii) any two successive para (ortho) rotational states are always separated by an ortho (para) state, leading to no adjacent rotational states with the same nuclear spin configuration within one specific J quantum state.

The rotational energy levels of a quantum-mechanical rigid rotor are given by

$$H_0 = \frac{J_a^2}{2I_a} + \frac{J_b^2}{2I_b} + \frac{J_c^2}{2I_c} = AJ_a^2 + BJ_b^2 + CJ_c^2 \tag{S.35}$$

where A, B, and C ($A \geq B \geq C$) are three rotational constants, each of which is inversely proportional to $I_a, I_b,$ and I_c ($I_a \leq I_b \leq I_c$), the three components of the moment of inertia. For water, the values are listed with other related rotational constants below (**Table S1**).

For a linear polar molecule (e.g. OCS), $I_a = 0; I_b = I_c$, and $A = \infty; B = C$. The energy levels are solely determined by one rotational constant B, and no angular momentum occurs along the molecular axis (the direction of the permanent dipole moment) because a torque cannot exist along the molecular axis. A (prolate) symmetric top(2) (e.g. CH_3CN), satisfying ($I_a < I_b = I_c$, and $A > B = C$), has a slightly complicated rotational energy level structure described by introducing a new quantum number K, giving the projection of angular momentum onto the molecular axis. Fortunately, the selection rule requires that $\Delta J = \pm 1; \Delta K = 0$, thus reducing the energy level structure into multiple blocks distinguished by K values. Within each energy level block, a symmetric top resembles a linear molecule.

When one moves to a small asymmetric-top molecule, especially one with a large asymmetry (e.g. water), the three principle moments of inertia are all unequal, as are the rotational constants($I_a < I_b < I_c$, and $A > B > C$), leading to an extremely irregular rotational energy level structure(4, 5). The transition energy values are no longer multiples of a single fundamental frequency(2). There is no simple pattern describing the various absorption peaks (**Fig. S12**).

Selection rules

The rotational selection rules for the water molecule is given (for linearly polarized THz fields) by(4, 5)

$$\Delta M = 0,$$
$$\Delta J = 0, \pm 1,$$
$$\Delta K_a = \pm 1, \pm 3, ...$$
$$\Delta K_c = \pm 1, \pm 3, ...$$

(S.36)

Therefore, rotational transitions of water molecules involve not only those within one specific J state (only K quantum number is changing, Q-branch) but also between distinct J states (both J and K are changing, P-/R-branches). In addition, many weak transition lines from $\Delta K = \pm 3$ also contribute to the final spectrum.

Wavefunctions of an asymmetric top

The wavefunctions of an asymmetric top are generally given as linear combinations of symmetric-top wavefunctions,

$$| JK_aK_cM \rangle = | J\tau M \rangle = \sum_{K=-J}^{J} a_K(JK_aK_c) | JKM \rangle$$

$$\tau = K_a - K_c$$

(S.37)

Two pseudo-quantum numbers K_a, K_c, or their difference τ, are introduced to characterize the wavefunction of an asymmetric top. One can choose either $|JK_aK_cM\rangle$ or $|JKM\rangle$ as the basis set to construct the Hamiltonian matrix although the latter is only exact for symmetric tops.

Spin statistics

Water naturally exists as a mixture of two nuclear-spin species, para and ortho, which are determined by the total spin of the two hydrogen nuclei. For total spin $I = 0$, water is para, and $K_a + K_c$ is even; for total spin $I = 1$, water is ortho, and $K_a + K_c$ is odd. No transition between

ortho and para is allowed, separating the water energy levels into two parts (**Fig. S12**). Different nuclear spin statistical weights for para- and ortho- transitions strongly affect transition line strengths.

Centrifugal distortion

Because the water molecule is very light, its rotational constants are notably large compared to most others. For example, the rotational constant B of water molecule is 70 times larger than that of OCS and 47 times larger than CH_3CN. As a consequence, water molecule shows an extraordinarily large centrifugal distortion effect even when the angular momentum quantum number J is very small (e.g. $J < 5$), since the product of the rotational constant and the square of the angular momentum component (e.g. BJ_b^2) contributes to the Hamiltonian(4, 5). This non-rigid behavior shifts the absorption and emission spectra and further affects the rotational spectrum.

Mixing and collision

For our nonlinear 2D THz measurements, a gas cell was used. We discovered afterward that a small amount of the previous sample, gaseous acetonitrile, was adsorbed onto the aluminum surface of the gas cell. Although no couplings between water molecule and acetonitrile were found (a similar case has been reported where no coupling occurs between atoms of different rubidium isotopes (7)), a diagonal peak splitting around 0.557 THz is clearly observed. This is formed by the coincidence and interference of two third-order coherent processes: one in the water molecule with the $1_{01} \leftrightarrow 1_{10}$ coherence, and the other in acetonitrile with $J = 29 \leftrightarrow 30$.

At the typical relative humidity of 40% in our lab, there is about 10 Torr water vapor in the ambient air, and one can make full use of the water vapor for the linear THz absorption measurement (**Fig. S1**). Frequent collisions with nitrogen and oxygen molecules in the air broadens the linewidths to roughly twice those of pure water vapor.

Radiative interaction

One possible contribution to the two-quantum peaks is radiative interaction (9). After the main THz pulse interacts with one water molecule, the emitted THz field from this molecule will subsequently excite another water molecule, and the cascaded two-step excitation can also create two-quantum diagonal peaks. Similarly, three-quantum diagonal peaks can be generated involving radiative interactions among three water molecules or many-body interactions between two nearby molecules and radiative interactions with a third molecule. Radiative interactions cannot yield energy shifts away from simple sums of the separate transition energies, so such shifts provide clear signatures of many-body interactions.

2Q and 3Q peaks at 60℃ and 21℃

All the 2D spectra shown in the main paper were recorded at approximately 60℃ in order to provide sufficient vapor pressure to yield acceptable signal/noise ratios. We recorded spectra at room temperature (21℃) in order to confirm that the off-diagonal 2Q features in the spectra arose from intermolecular complexes(10) whose concentrations would be expected to decrease far more sharply with temperature than the water vapor pressure itself. As discussed in the main paper, the intensities of the off-diagonal 2Q peaks were observed to decrease much more sharply than the other peaks in the spectra, approximately proportional to the square of the ratio of vapor pressures as would be expected for peaks arising from molecular-pair complexes. The spectra for two groups

of 2Q peaks at 60ºC and 21ºC are shown in **Figs. S16** and **S17** respectively. A spectrum showing a diagonal 3Q peak at 60ºC is also shown in **Fig. S16**. This peak was too weak to measure at 21ºC. At both temperatures, the 2Q spectra show multiple off-diagonal features which we believe indicate the presence of multiple intermolecular geometries. Their linewidths yielded an approximate dephasing time of 100 ps, which we take as an instrumentally limited lower limit for the lifetimes of the intermolecular complexes. Due to the limited signal/noise ratios in the room-temperature spectra, some of the features observed at 60ºC are not clearly distinguished, but the main features appear to be preserved with some degree of merging among nearby peaks. The room-temperature 2D spectra are presented with two different levels set for the lowest (white) colorbar scale, with the higher level allowing clearer distinction of the main features and the lower level allowing finer observation of the peak wings and the regions between peaks.

The side peaks (d1, e1, c2, d2, d3, e3, d4, and e4) arise from different coherence pathways. Energy-ladder diagrams and possible Feynman diagrams are illustrated in **Fig. S18**. In **Fig. S18a**, the ground state, the singly excited state, and the doubly excited state are shown as $|g>$, $|e'>$, and $|e''>$, respectively. The solid lines show the energy-level positions without shifts due to many-body interactions. The 2Q pump frequency is that of the coherence $|e''><g|$. Due to the energy shifts of the different intermolecular complexes, the corresponding peaks can be shifted upward or downward with respect to the 2Q diagonal peak. Peaks can also appear with shifts in the 1Q emitted signal frequency, which depends on the third THz field interaction which can yield the 1Q coherence $|e'><g|$ or $|e''><e'|$), and which depends on the particular intermolecular geometry. Note that the shift in the 1Q frequency $|e''><e'|$) may be different from the shift Δ in the 2Q frequency $|e''><g|$) because both the $|e''>$ and $|e'>$ energies may be shifted by many-body interactions involving multiple excitations. Similar effects have been observed in 2D electronic spectra of biexcitons(11).

Estimate of molecule-pair binding energies

The ratio of off-diagonal peak intensities at temperatures $T_2 = 333$ K (60ºC) and $T_1 = 294$ K (21ºC) can be used to estimate the binding energies for the intermolecular complexes. The water vapor pressure at 60ºC is greater than that at 21ºC by a factor of $X = 8.3$. Peaks arising from water monomers should be more intense at 60ºC by a factor of X. We select the strong non-rephasing diagonal peak at ($f_{probe} = 0.75$ THz, $f_{pump} = 0.75$ THz) as an internal standard for the monomer peak intensity I_m at each temperature. Considering only the vapor pressure, the intensity I_p of an off-diagonal 2Q peak arising from a molecule-pair complex would increase by a factor of X^2 at the higher temperature, and thus the intensity ratio I_p/I_m would increase by a factor of X. However, given the weak binding energies ε_P of the complexes, the increased thermal energy at 60ºC will shift the equilibrium toward separated monomers, thereby reducing the molecule-pair peak intensities I_p. The dependence of the intensity ratio I_p/I_m on temperature can be calculated by considering the equilibrium between water monomers M and molecule-pair complexes C, i.e. $2M \rightleftharpoons C$, with equilibrium constant K_p given in terms of the vapor pressures p_M and p_C by

$$K_p = \frac{\left(\dfrac{p_C}{p^\circ}\right)}{\left(\dfrac{p_M}{p^\circ}\right)^2} = \frac{p_C p^\circ}{p^2{}_M} \tag{S.38}$$

where $p^\circ = 1$ bar. K_p can be calculated from the molecular partition functions q_M and q_C using the expression

$$K_p = e^{-\varepsilon_P/k_B T} \frac{\left(\dfrac{q_C}{V}\right)\left(\dfrac{k_B T}{p^\circ}\right)}{\left[\left(\dfrac{q_M}{V}\right)\left(\dfrac{k_B T}{p^\circ}\right)\right]^2} = e^{-\varepsilon_P/k_B T}\left(\frac{h^2}{2\pi k_B T}\right)^{3/2}\left(\frac{2}{m_M}\right)^{3/2}\frac{p^\circ}{k_B T}\left(\frac{k_B T}{h\nu}\right)^3 = e^{-\varepsilon_P/k_B T}\frac{p^\circ\left(k_B T\right)^{1/2}}{\left(\pi m_M\right)^{3/2}\nu^3}$$

$$\tag{S.39}$$

where m_M is the water monomer mass and ν is the frequency of any of three vibrational modes in which the two molecules in a complex move relative to each other. The three frequencies would not really be equal to each other, but for our purposes their values will not matter except that they are assumed to be low enough that the modes are thermally excited so the vibrational partition function for each takes the simple high-temperature limiting form $q_{vib} = k_B T/h\nu$. We have simplified the treatment greatly by assuming that the two molecules in a complex have the same internal molecular vibrational modes and the same independent rotations as the separated molecules, so all the partition functions are identical except for the masses in the translational partition functions and partition function for the three intermolecular vibrational modes that account for the translational degrees of freedom that are lost when two monomers come together to form a molecule-pair complex. We have not accounted explicitly for zero-point vibrational energies of these modes since we are seeking only a rough estimate for the binding energy E_B.

Combining the expressions for K_p gives

$$p_C = p_M^2 e^{-\varepsilon_p/k_B T}\frac{(k_B T)^{1/2}}{(\pi m_M)^{3/2}\nu^3} \tag{S.40}$$

The ratios of spectral features arising from complexes and monomers (i.e. 2Q off-diagonal peaks and any others in the spectrum) are proportional to the ratio of vapor pressures p_C/p_M. We measure these ratios at two temperatures $T_1 = 294$ K and $T_2 = 333$ K, then take their ratio. The water monomer vapor pressure p_M increases by a factor of $X = 8.3$ going from T_1 to T_2.

$$\frac{\frac{p_C}{p_M}(T_2)}{\frac{p_C}{p_M}(T_1)} = \frac{p_M(T_2)}{p_M(T_1)}\frac{e^{-\varepsilon_p/k_BT_2}}{e^{-\varepsilon_p/k_BT_1}}\left(\frac{T_2}{T_1}\right)^{1/2} = X\frac{e^{-\varepsilon_p/k_BT_2}}{e^{-\varepsilon_p/k_BT_1}}(1.06) \qquad (S.41)$$

We believe the peaks with different 2Q frequency values represent different molecule-pair complex geometries, and in principle we could evaluate the each of the peak intensities at both temperatures and determine the binding energy for each distinct geometry. Given the signal/noise limitations in the spectra and the associated uncertainties in peak intensities I_p, we have instead calculated an average value of the left-hand side of eq. (S.41) by averaging the intensities of peaks b1 (b1$_3$ − b1$_6$) and b3 (b3$_1$ − b3$_3$) in the spectra shown in **Figs. S16** and **S17** respectively. The resulting value is 5.5, which as expected is somewhat less than $X = 8.3$. From this we estimate the molar binding energy $\varepsilon_P \approx 10$ kJ/mol, smaller than the binding energy of the well-known hydrogen-bonded water dimer (12-14). The somewhat wide range of intensity ratios (using the peaks (b1$_6$ and b3$_3$) gives a ratio of 5.9, while using peaks (b1$_3$ and b3$_1$) gives a ratio of 3.9 may be due to some combination of experimental uncertainties and real variation among the binding energies associated with different complex geometries.

Field dependences and Feynman diagrams for selected spectral peaks

Feynman diagrams of many spectral peaks of water molecules can be identified and assigned based on previous experimental results and analyses(2, 15), especially the diagonal peaks and off-diagonal peaks not far away from the diagonal. Other off-diagonal peaks arise from either 2Q or 3Q coherences (See **Figs. 2** and **3**). The 2D water rotational spectra from both simulations and experimental data reveal the complexity of the 2Q and 3Q signals, because they can occur due to three distinct origins: (i) many-body interactions; (ii) radiative interactions; and (iii) nonlinear responses of individual water molecules. Since the simulation conditions are similar to those in the experiments, the much smaller 2Q and 3Q signal amplitudes (blue and black signals in **Fig. S19b**) than the experimental signals (red signals in **Fig. S19b**) reveal that weak nonlinear responses (**Fig. S22b**) from non-interacting water molecules are not contributing significantly to 2Q and 3Q features. (i) and (ii) together are estimated to have a dominant contribution (>90%) to the overall multiple-quantum signals, while (iii) has a small contribution (<10%). Signals from (iii) are overlapped and merged with the strong 2Q/3Q signals in the NR spectra, but some nonlinear responses from (iii) can be observed in the R spectra, offering another approach to estimating the contributions of distinct origins of the 2Q/3Q signals and giving the same results. In addition, both NR and R contributions from (iii) can be captured in the simulation (without many-body interactions) (**Fig. 5**).

Simulated field dependences of several selected spectral peaks are shown (**Fig. S19a**). As expected, peaks (P2-P6) show a third-order dependence on THz field strength in the moderate THz field regime. P1, overlapped with the 2Q signal, is fifth order on the THz field amplitude, because no corresponding third-order responses can be assigned to it. In particular, one cannot connect the coherences |2><1| and |1><0| by one single interaction between THz fields and water molecules in the V-type energy ladder diagram (**Fig. S20**). However, its counterpart in the R spectra (i.e. P3

in **Fig. S19**) is third order since one THz excitation process can connect $|1><2|$ and $|1><0|$ on the bra side of the Feynman diagram.

As a consequence, only two peaks in our measurement belong to fifth-order signals, namely, (a) 3Q signal (**Fig. 3a**) and (b) the fifth-order signal discussed above which is overlapped with the 2Q signal (**Fig. S22b** and P1 in **Fig. S19**). The rest of the 2D spectral peaks are all third-order signals.

Note that different excitation pathways in the Feynman diagram (for example, the two distinct excitation pathways$|0\rangle\langle0| \rightarrow |1\rangle\langle0| \rightarrow |1\rangle\langle2|$ (**b** (1) in **Fig. S20**) and $|0\rangle\langle0| \rightarrow |0\rangle\langle2| \rightarrow |1\rangle\langle2|$ (**b** (2) in **Fig. S20**) could be distinguished through the use of three pulses to temporally separate each of the field interactions temporally, revealing the first coherence frequency from the first interaction in addition to the second and third field interactions, the variable delays of which reveal the 2Q coherence(11, 16).

Overview of 2D THz spectra of water vapor

We show complete non-rephasing (NR) and rephasing (R) 2D THz spectra in **Figs. S24-27**. Simulated spectra are shown in **Figs. S24** and **S26**, and **Figs. S25** and **S27** illustrate experimental spectra.

Main features in the experiments are reproduced in the simulation. Simulated NR, R, and pump-probe (PP) features match well with those in the experiments. Some peaks are magnified by appropriate factors, for the sake of clarity. In the experimental data, spectral features from acetonitrile residue are indicated. In the simulation data, weak spectral signals involving weak water transitions (e.g. see **Figs. S12** and **S23**) appear (labelled as W1-7 in **Fig. S24** and WR1-5 in **Fig. S26**). Additionally, in order to clearly present spectral peaks within (0.9 − 1.3 THz), enlarged spectra are illustrated in **Figs. S24b** and **S26b**. Peak labels in the experimental spectra (**Figs. S25&S27**) follow those in the simulation (**Figs. S24&S26**).

Minor discrepancies between experimental (**Figs. S25&S27**) and simulated (**Figs. S24&S26**) spectra occur, due to that: (i) signal-to-noise ratio limitations in the measurement; (ii) slight difference between experimental and simulated THz field profiles; (iii) slight difference between rotational transition line intensities in water molecules between experiments and simulations. Note that these discrepancies are trivial, and do not influence present results in this work. Peak positions and their relations with water rotational states are summarized in **Table. S6&S7**. For a qualitative analysis including many-body interactions, see **Fig. S13-S15** and related descriptions.
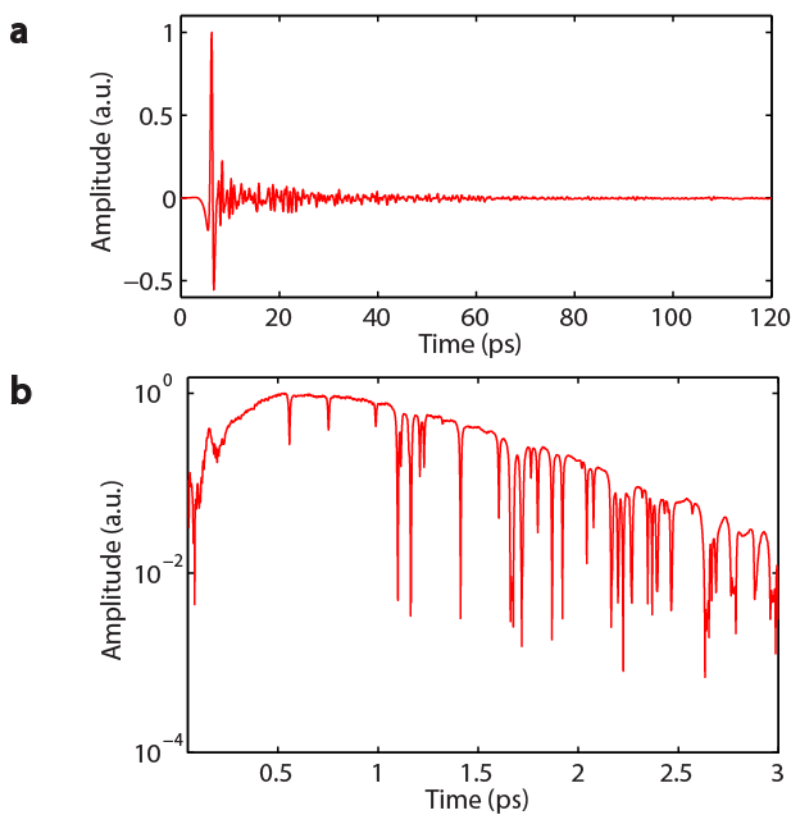
Fig. S1. Measured THz spectrum of water vapor in ambient air. **a**, The time domain signal is obtained by a typical THz time-domain spectroscopy (THz-TDS) system in which photoconductive antennae (PCA) are used for THz generation and detection as described in the text. The trailing oscillation after the main THz peak in the top panel stems from THz emission (i.e. the rotational free-induction decay) from water vapor in the ambient air. **b**, Fourier transform of the time domain signal (including both the main peak and the trailing oscillation) reveals the rotational absorption peaks.

14
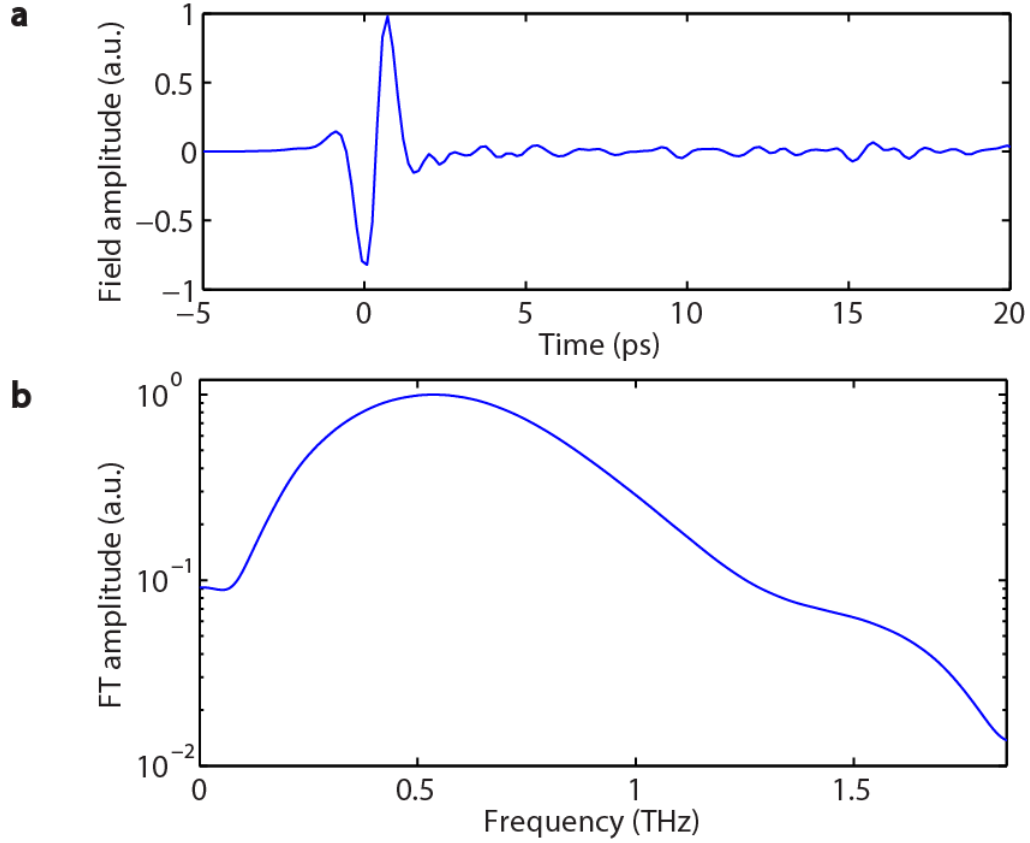
Fig. S2. Typical profile of a single-cycle THz electric field. The single-cycle THz field profile shown in this figure was generated in a lithium niobate crystal by the tilted-pulse-front technique. **a**, The time domain signal in the top panel includes the THz main peak around 0 ps and the trailing oscillation from water vapor emission. Double reflections from the gas cell window and optics in the setup appear beyond 20 ps and are excluded. **b**, FT spectrum from top panel, in which only the main THz peak is taken into account (zero-padding was added after the first large THz cycle.
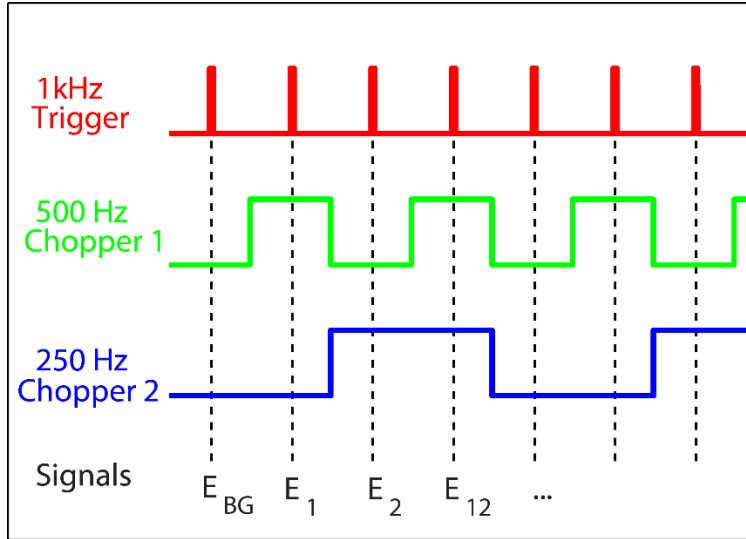
Fig. S3. Mechanism of differential chopping detection. The laser is run at a 1 kHz repetition rate (red). Chopper 1 (green) is operated at a 500 Hz frequency while the other chopper 2 (blue) is at 250 Hz. The signal sequence shown here consists of a background signal (E_{BG}), an "only THz 1 on" signal (E_1), an "only THz 2 on" signal (E_2), and a "both THz 1 and 2 on" signal (E_{12}). The nonlinear signal can be obtained by subtraction: $E_{NL} = E_{12} - E_1 - E_2$.
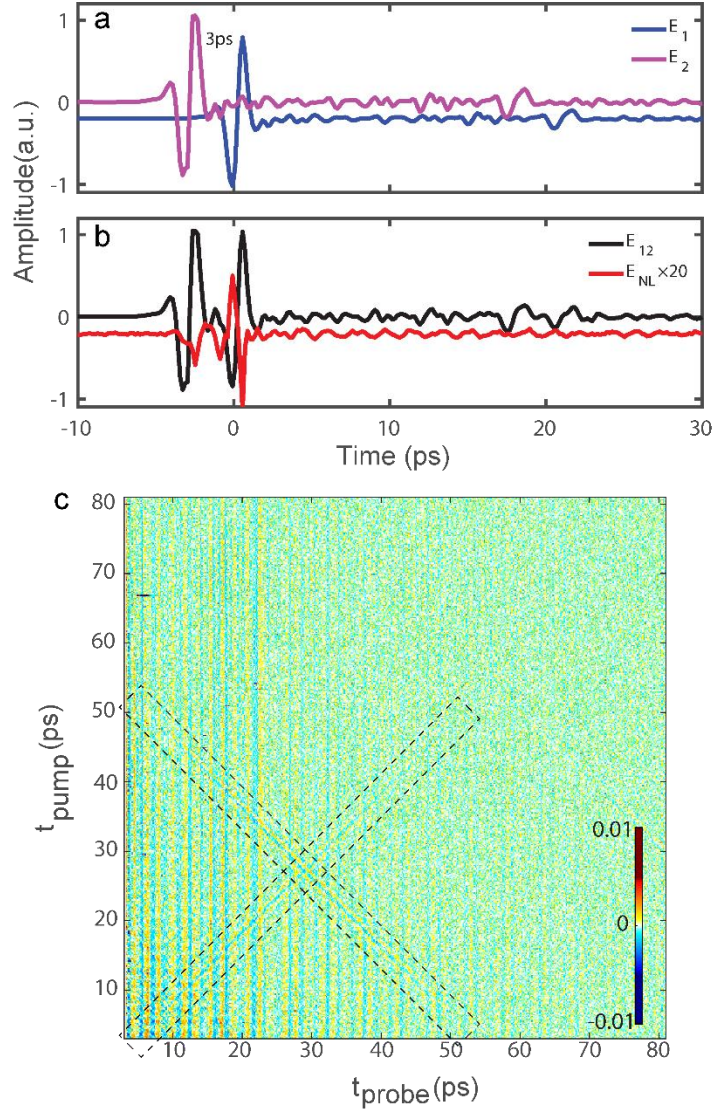
Fig. S4. Examples of linear, nonlinear, and 2D signals in the experiment. **a**, THz time-domain signals with only THz 1 (blue) and only THz 2 (magenta) on, respectively. **b**, THz signals when both THz 1 and 2 are on. The nonlinear THz signal (red) is obtained by subtraction of both E_1 and E_2 from E_{12}, namely $E_{NL} = E_{12} - E_1 - E_2$. The background noise can also be subtracted when necessary. The nonlinear signal is magnified by a factor of 20 for the sake of clarity. E_1 and E_{NL} are vertically displaced by the same value of 0.2 with respect to E_2 and E_{12}, respectively. Time delay between the two THz pulses is 3 ps. **c**, Measured 2D time-domain signal as function of the pump time and the probe time. Both pump and probe time extend from 3 ps to 83 ps, with a step size of 0.16 ps. Tilted lines in the dashed black boxes are from acetonitrile residues, which can be separated from water signals after Fourier transformation (see **Figs. 2a** and **d**). Strong THz features near 0 ps (see a and b) are removed from this 2D time-domain signal in order to highlight weaker oscillatory signal components afterwards.
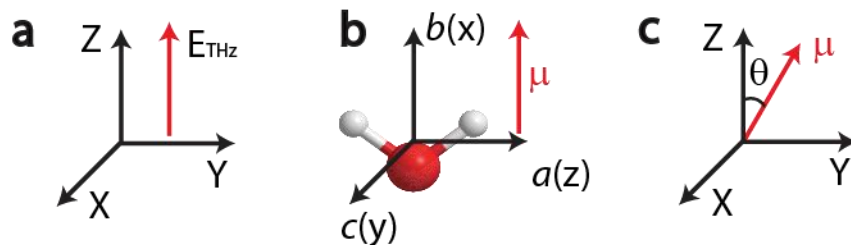
Fig. S5. Coordinate system assignment. **a**, THz field is polarized along the Z direction in the space-fixed frame (X, Y, Z). **b**, Ball-stick model of a water molecule with its principle axes (a, b, c). The permanent dipole moment of water is along its principle b-axis. **c**, Orientation angle is identified as the intersection angle between the THz field polarization vector (along Z-axis in the space-fixed frame) and the dipole moment vector (along b-axis in the molecule-fixed frame). The identification of the principle axes (a, b, c) with molecule-fixed coordinates (x, y, z) is indicated in **b**.

			Col#	1	2	3	4	5	6	7	8	9	...
			J	0	1	1	1	2	2	2	2	2	...
			K	0	-1	0	1	-2	-1	0	1	2	...
Row#	J	K											
1	0	0		*									
2	1	-1			*		X						
3	1	0				*							
4	1	1			X		*						
5	2	-2						*		X			
6	2	-1							*		X		
7	2	0						X		*		X	
8	2	1							X		*		
9	2	2								X		*	
⋮	⋮	⋮											⋱

Fig. S6. Density matrix format of the rotational Hamiltonian. Each column and row holds one specific rotational state denoted as $|JK>$ (M quantum number is omitted for simplicity). Diagonal matrix elements (red star) are the population states. Off-diagonal matrix elements (green cross) are the coherence states. Matrix elements belonging to the same J quantum number are encompassed by a blue square. When implemented in the eigen-states of an asymmetric top $|JK_aK_c>$ or $|J\tau>$, the density matrix basically remains the same structure except that K in the figure is replaced by $\tau = K_a - K_c$. The M quantum number is omitted for simplicity.
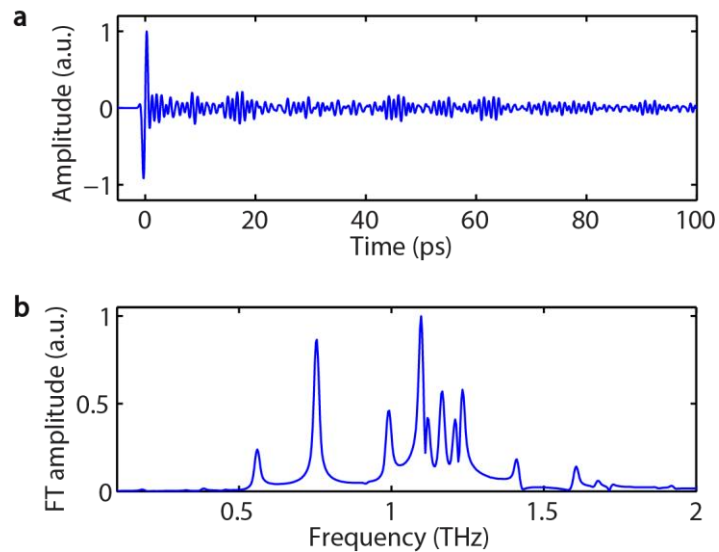
Fig. S7. Simulation of linear THz spectroscopy of water molecules. **a**, Simulated THz free-induction decay from water molecules. The maximum J quantum number was set at 15. M states considered here are those with $|M| \leq 4$. Temperature $T = 100$ K. The THz field strength was set at 300 kV/cm. **b**, FT spectrum from the time domain signal.
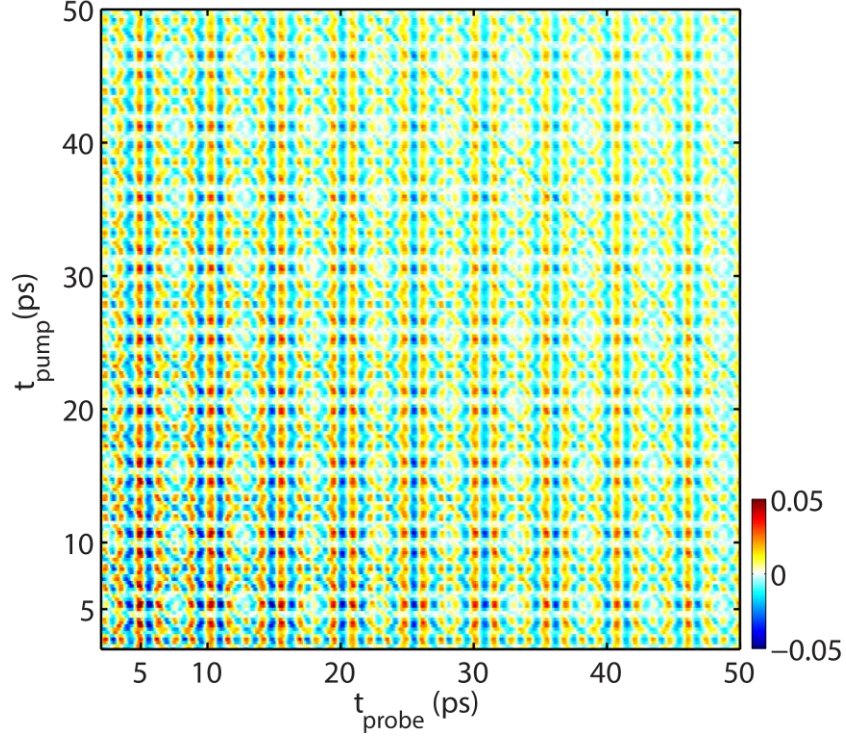
Fig. S8. Simulated 2D THz signals in the time domain. The maximum J quantum number was set at 15. M states considered were those with $|M| \leq 4$. Temperature was set at 100 K. Both THz field strengths were set at 300 kV/cm. The step size for the inter-pulse pump axis was 0.25 ps, and that for the probe axis was 0.1 ps. Both pump and probe time extend from 3 ps to 50 ps. Strong THz features near 0 ps are removed from this 2D time-domain signal in order to highlight weaker oscillatory signal components afterwards. The similar features along both time scales correspond to diagonal peaks in the 2D spectrum. The pronounced periodicity of 1.3 ps (along both axes) corresponds to the strongest diagonal peaks in the spectra, at 0.75 THz (apparent in the measured spectra of **Fig. 2** and the simulated spectra of **Fig. 4**). The features that start at $t_{\text{probe}} \approx 0$ at any time along the t_{pump} axis, i.e. signals that begin immediately after the two pulses, correspond to non-rephasing signals. The signals (clearest for the 1.3 ps periodic feature) that are at their approximate maxima along the diagonal correspond to rephasing (photon echo) signals, i.e. they reach their maxima (neglecting dephasing) at a time delay t_{probe} equal to the delay t_{pump} between the THz pump pulses.
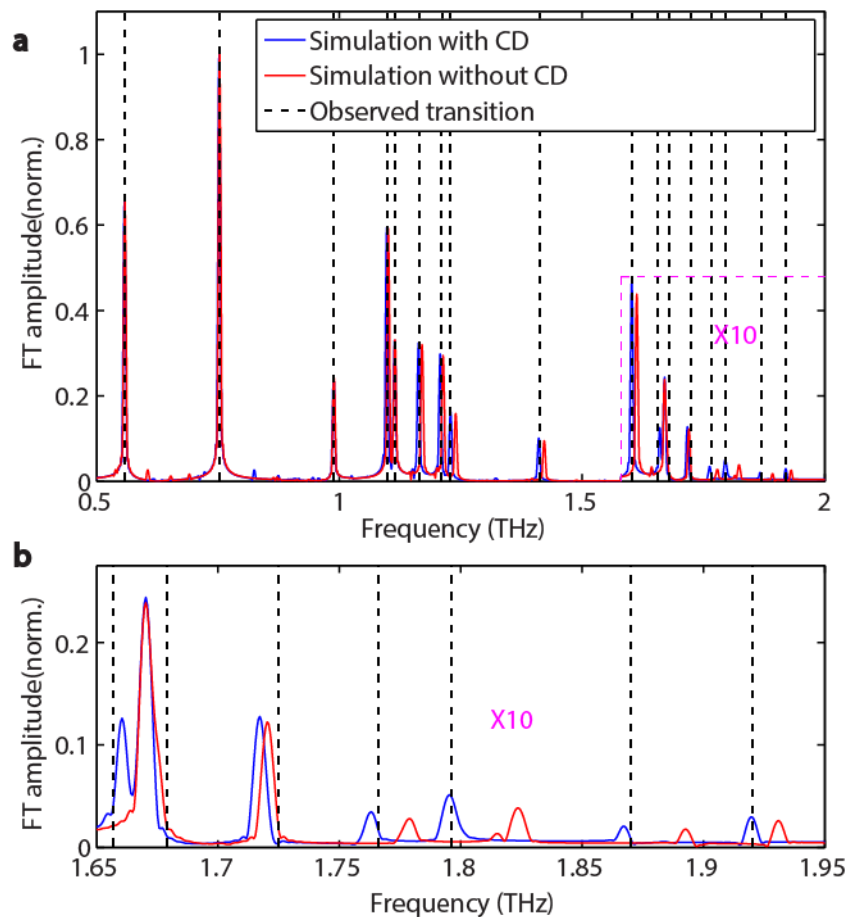
Fig. S9. Centrifugal distortion effect on transition line positions. **a**, Comparison between calculated line positions with and without centrifugal distortion (CD). **b**, Enlargement of the top panel in the region (1.65-1.95 THz). Deviations as large as 10 GHz are observed between experimentally observed transition frequencies and those calculated when no centrifugal distortion is taken into account (red line). Watson's reduced Hamiltonian has been employed to account for centrifugal distortion so that the deviation is reduced to around 1 GHz (blue line).
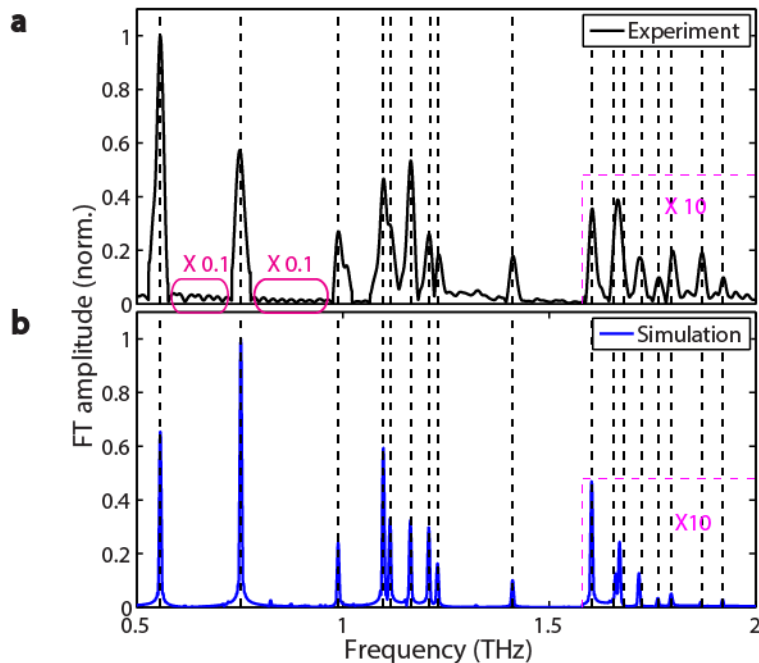
Fig. S10. Experimental and simulated water THz spectra. **a**, Measured THz spectrum of water vapor at 60ºC pumped by a THz pulse generated in a lithium niobate crystal. **b**, Simulated water spectrum using Watson's reduced Hamiltonian. The transitions above 1.5 THz are enlarged by a factor of 10 in both panels. Dashed black lines indicate the positions of observed rotational transition lines of water vapor in ambient air by a photoconductive antenna detection. Large oscillations of **a** in the frequency intervals (0.558 THz, 0.753 THz) and (0.753 THz, 0.989 THz) are from double reflections in the setup and reduced by a factor of 0.1, causing a slight distortion at the edges of peaks at 0.558 THz, 0.753 THz ,and 0.989 THz.
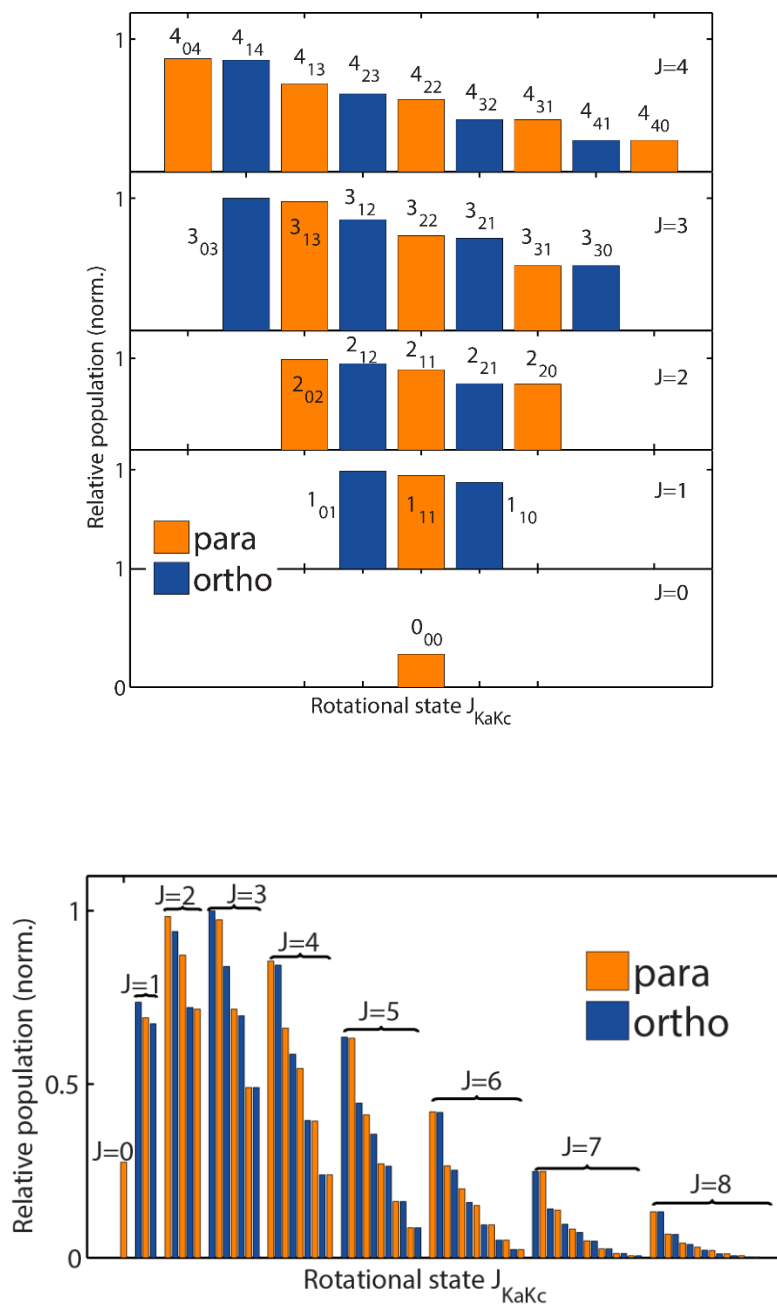
Fig. S11. Calculated population distribution of water molecules at room temperature. (Top) Relative population in each J (0-4) state is indicated. Orange and blue denote para and ortho water rotational states, respectively. The rotational states $J_{K_aK_c}$ are indicated. (Bottom) Relative population in each state up to $J = 8$. The largest populations occur at $J = 2$ and $J = 3$ when treating all rotational states sharing the same J quantum number as a whole.
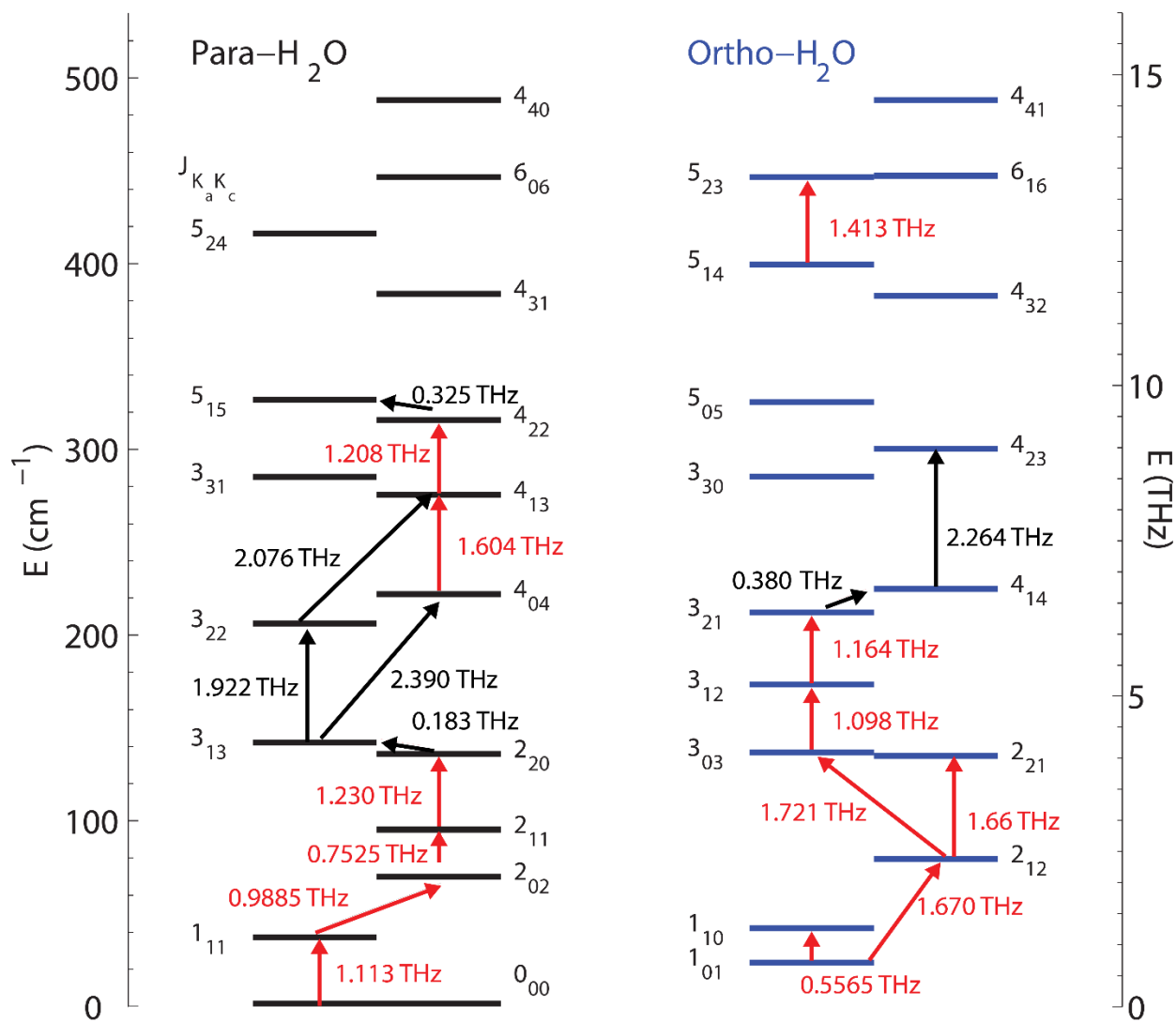
Fig. S12. Energy-level diagram of water rotational states. Rotational energy levels are divided into two parts: para-type (left side, in black) where the total nuclear spin of the two hydrogen atoms is 0, and ortho-type (right side, in blue) where the total nuclear spin is 1. Red arrows indicate allowed transitions between 0.5 THz and 1.8 THz observed in this work. Black arrows indicate other possible transitions either lower than 0.5 THz or higher than 1.8 THz. Two transitions below 2 THz, with frequencies of 1.798 THz ($6_{15} \leftrightarrow 6_{24}$ and $7_{25} \leftrightarrow 7_{34}$) and 1.871 THz ($5_{23} \leftrightarrow 5_{32}$)) are associated with higher-energy rotational states than those shown here.
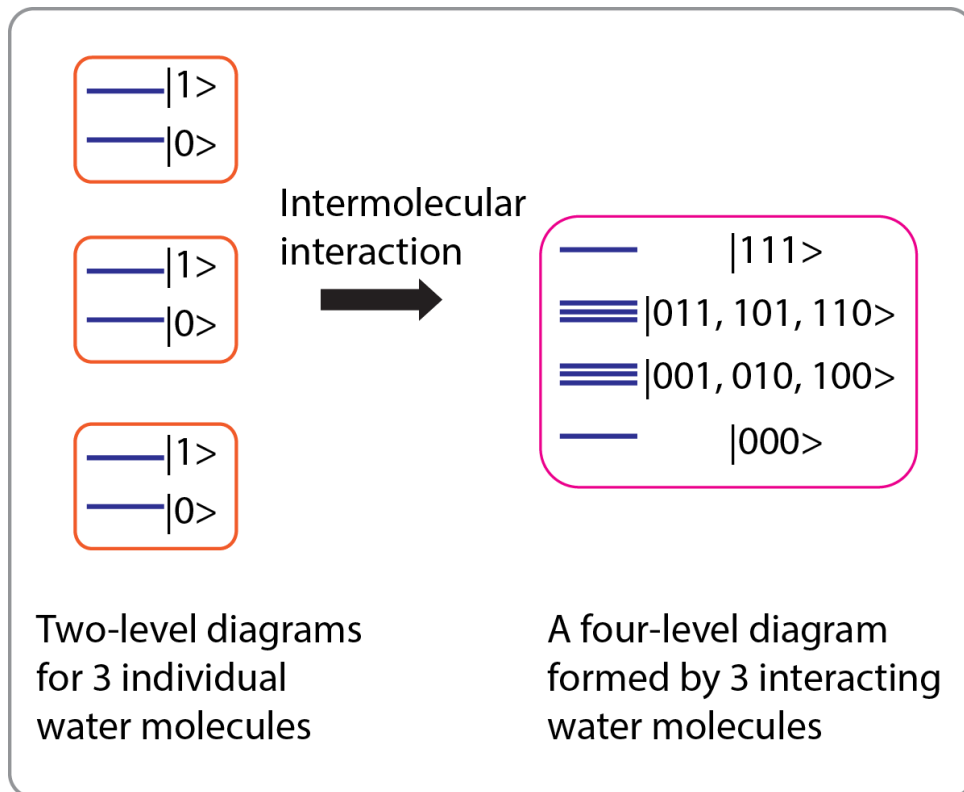
Fig. S13. Diagram of the formation of a four-level system in water molecules. Three two-level quantum systems of three individual water molecules generate a four-level quantum system via intermolecular interactions such as dipole-dipole interactions among water molecules. The singly excited and doubly excited energy levels are three-fold degenerate in the absence of interactions but non-degenerate when interactions are taken into account.

$$H = \begin{bmatrix} 0 \\ & \hbar\omega & a & a \\ & a & \hbar\omega & a \\ & a & a & \hbar\omega \\ & & & & 2\hbar\omega & a & a \\ & & & & a & 2\hbar\omega & a \\ & & & & a & a & 2\hbar\omega \\ & & & & & & & 3\hbar\omega \end{bmatrix} \quad \begin{matrix} |000> \\ |001> \\ |010> \\ |100> \\ |011> \\ |101> \\ |110> \\ |111> \end{matrix}$$

Diagonalization

$$V^{-1}HV = \begin{bmatrix} 0 \\ & \hbar\omega - a \\ & & \hbar\omega - a \\ & & & \hbar\omega + 2a \\ & & & & 2\hbar\omega - a \\ & & & & & 2\hbar\omega - a \\ & & & & & & 2\hbar\omega + 2a \\ & & & & & & & 3\hbar\omega \end{bmatrix}$$

Fig. S14. Diagram of Hamiltonian under many-body interactions. (Top) The rotational Hamiltonian under many-body interactions is subject to the impact of perturbation terms, acting as off-diagonal elements (labelled as *a*). (Bottom) Hamiltonian diagonalization gives eigen-energy levels. Note that the coupling elements are not necessarily the same as shown in this figure. Different coupling elements will produce similar splittings among the levels that are three-fold degenerate in the absence of interactions.

Fig. S15. Simulated 2Q and 3Q spectrum with many-body interactions**.** The weak interactions in the simulation are sufficient to allow 2Q and 3Q signals to appear but not enough to induce significant energy shifts, leading to only diagonal peaks. The pump-probe (PP), 1Q, 2Q (magnified by 10), and 3Q (magnified by 50) signals are illustrated. See **Fig. 3** for observed 2Q and 3Q features in the experiment.
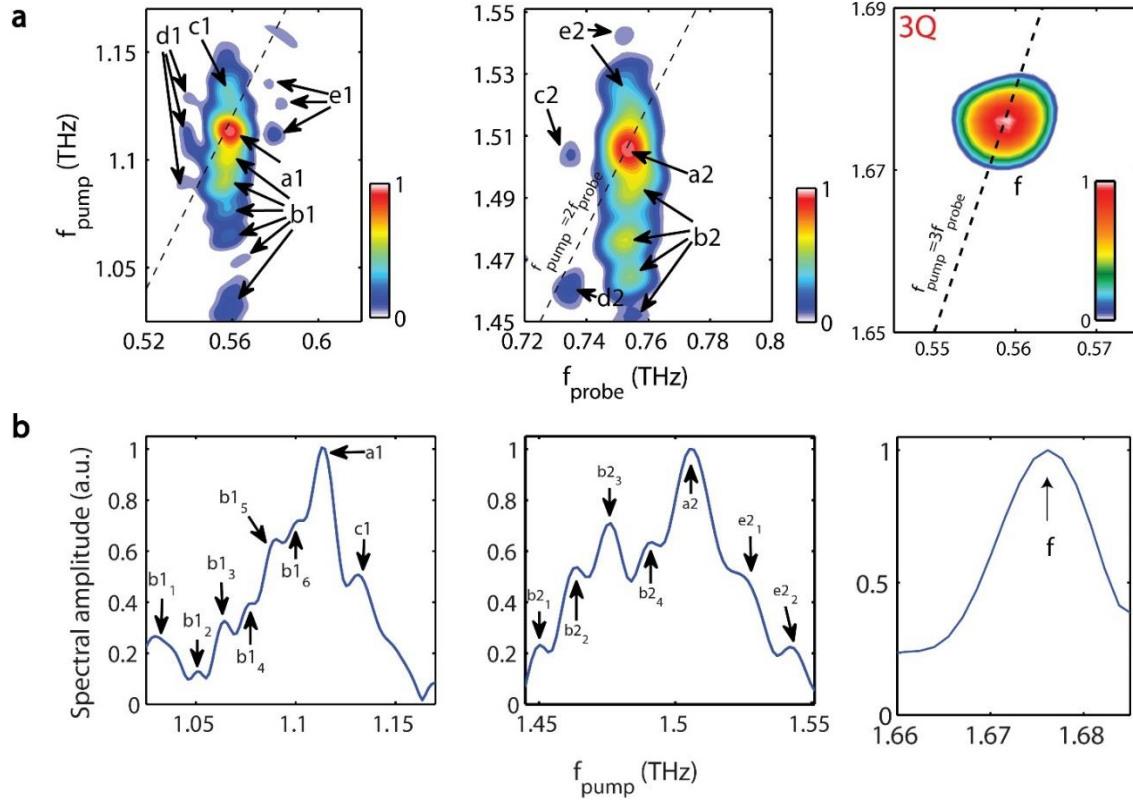
Fig. S16. Off-diagonal 2Q peak linewidths indicate multiple complex geometries and lifetimes. The 2Q diagonal, off-diagonal peaks and 3Q peaks from Fig. 3a of the main paper are replotted in **a**, and their spectral slices along the f_{pump}-axis are shown in **b**. Peak positions (uncertainties ±2 GHz) and selected peak intensity ratios are indicated in the table below. The estimated linewidths for the main peaks (a1, a2, and f) are 18 GHz, 28 GHz, and 14 GHz, giving dephasing times of 56 ps, 36 ps, and 71 ps, respectively. The 2Q off-diagonal peaks (b1$_{1-6}$, b2$_{1-4}$) have an averaged linewidth of ~10 GHz, yielding a dephasing time of ~100 ps which we believe is an instrumentally limited lower limit to the lifetimes of the metastable water complexes. No 3Q off-diagonal features are observed in the present measurements. The averaged intensities of the off-diagonal signals were used to compare with diagonal signals in Figs. 4d and 4e of the main paper. The off-diagonal 2Q spectral fine structure indicates multiple metastable complex geometries.

2Q: f_{pump} (f_{probe} = 0.558 THz)	2Q: f_{pump} (f_{probe} = 0.753 THz)	3Q: f_{pump} (f_{probe} = 0.558 THz)	Peak intensities (vs 1Q peak at (0.753 THz, 0.753 THz))
b1$_1$: 1.029 THz	b2$_1$: 1.450 THz	f: 1.676 THz	
b1$_2$: 1.051 THz	b2$_2$: 1.464 THz		
b1$_3$: 1.064 THz	b2$_3$: 1.476THz		b1$_3$: 0.23
b1$_4$: 1.078 THz	b2$_4$: 1.492 THz		b1$_4$: 0.27
b1$_5$: 1.090 THz	a2: 1.505 THz		b1$_5$: 0.43
b1$_6$: 1.102 THz	e2$_1$: 1.525 THz		b1$_6$: 0.48
a1: 1.113 THz	e2$_2$: 1.542 THz		
c1: 1.131 THz			
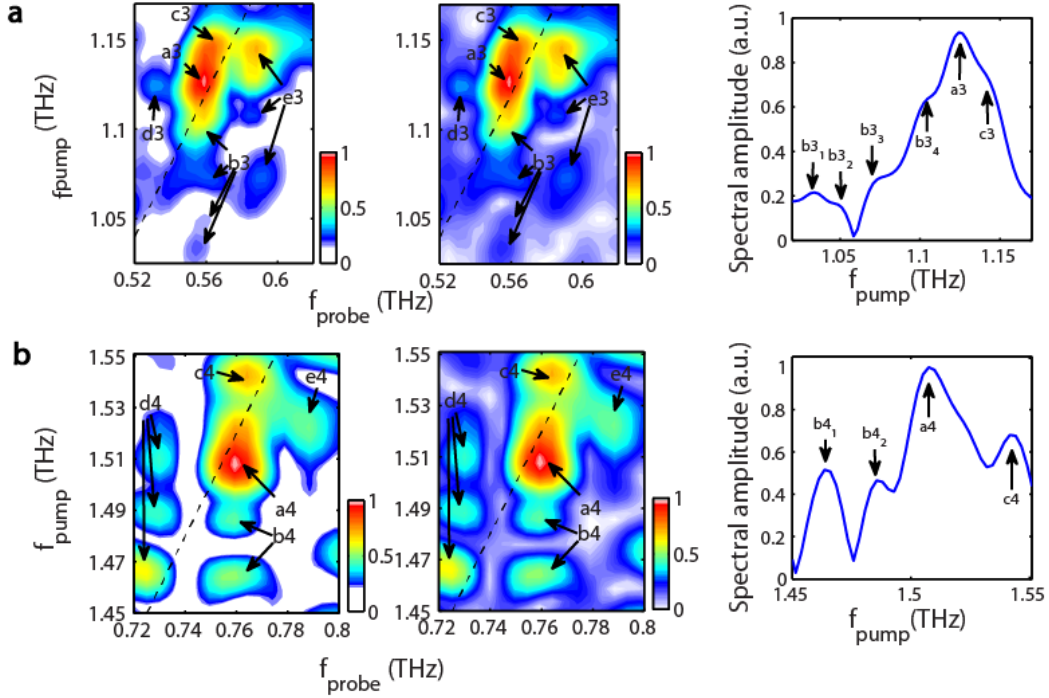
Fig. S17. Exemplar 2Q spectra and slices of water vapor at room temperature. The 2Q peaks and spectral slice near f_{probe}=0.558 THz are shown in **a** (Left: Background with intensity < 12% of the 2Q diagonal peak intensity is made white in color; Middle: Raw data with continuous colorbar scale; Right: Slice along the f_{pump} axis (f_{probe}=0.558 THz)). Similar features near f_{probe}=0.753 THz are shown in **b**. (Left: Background with intensity < 20% of the 2Q diagonal peak intensity is made white in color.) Peak positions (uncertainties ±2 GHz) are indicated in the table below. No 3Q features are observed in the present measurements at room temperature. Comparison between the room-temperature peak positions with those at 60°C (Fig. S15) shows the following features. **(i)** Some peaks remain the same: (b3$_2$,b1$_2$); (b3$_4$,b1$_6$); (b4$_1$,b2$_2$); (c4,e2$_2$). **(ii)** Some peak positions (b3$_1$,b1$_1$); (a4,a2) show slight shifts, probably within uncertainties. **(iii)** Some peaks at room temperature appear to be the spectrally merged averages of peaks that appear distinct at 60°C: (a3,a1-c1); (b3$_3$,b1$_3$-b1$_4$); (b4$_2$,b2$_3$-b2$_4$). **(iv)** Some peaks (b1$_5$, b2$_1$, e2$_1$) in the 60°C spectra are not visible at room temperature where the S/N ratio is lower. Some features (e.g. c3) are subtle and not comparable directly with counterparts at 60°C.

2Q: f_{pump} (f_{probe} = 0.558 THz)	2Q: f_{pump} (f_{probe} = 0.753 THz)	Peak intensities (vs1Q peak at (0.753 THz, 0.753 THz))
b3$_1$: 1.034 THz	b4$_1$: 1.464 THz	b3$_1$: 0.059
b3$_2$: 1.051 THz	b4$_2$: 1.486 THz	b3$_2$: 0.052
b3$_3$: 1.073 THz	a4: 1.508 THz	b3$_3$: 0.082
b3$_4$: 1.102 THz	c4: 1.542 THz	
a3: 1.124 THz		
c3: 1.144THz		
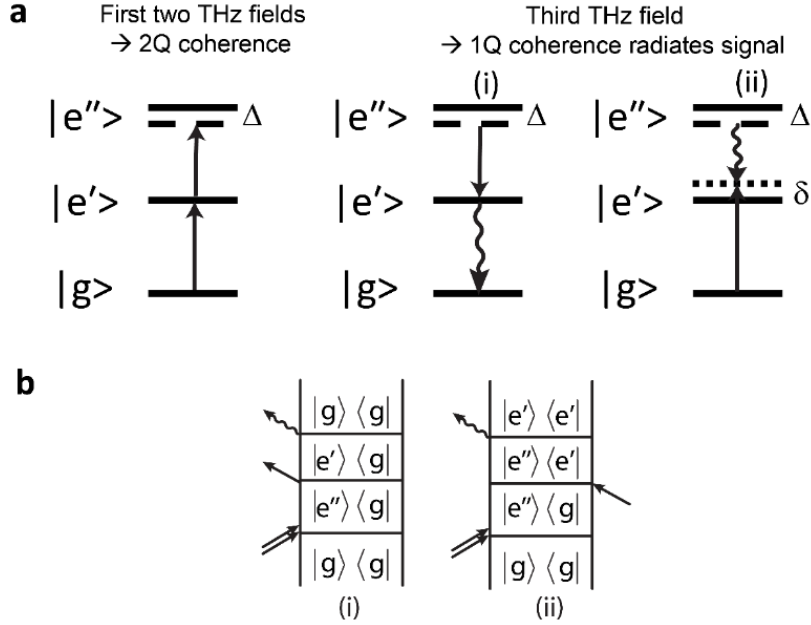
Fig. S18. Energy-ladder and Feynman diagrams of side peaks in Figs. S15 and S16. **a**, The ground state, the singly-excited state, and the doubly-excited state due to many-body interactions are shown as |g>, |e′>, and |e″>, respectively. The solid lines show the energy-level positions without many-body interactions. The first two THz fields produce the 2Q coherence, whose frequency depends on the particular intermolecular geometry. The third THz field may project |e″> downward (i) to the singly excited level |e′>, resulting in a coherence between |e′> and |g> that radiates at the usual 1Q transition frequency. Alternatively, the third THz field may promote from |g> to |e′>, resulting in a coherence between |e′> and |e″> whose frequency depends on the intermolecular geometry and whose shift from usual 1Q transition frequency may be different from the 2Q frequency shift Δ, similar to results observed in 2D electronic spectra of biexcitons(11). **b**, The 2Q pump frequency of an off-diagonal peak stems from the coherence |e″><g|. Due to the energy shifts, peaks can be shifted upward or downward with respect to the 2Q diagonal peak. Peaks can also emerge with shifts in the 1Q signal frequency, because the emitted frequency depends on the specific pathway (i) or (ii) determined by the third THz field interaction, yielding a final emitting coherence |e′><g| or |e″><e′| respectively, the latter at a frequency that also depends on the particular intermolecular geometry. The observed off-diagonal peaks and their positions (uncertainties ±2 GHz) are indicated in the table below.

Side peaks	f_{probe} (THz)	f_{pump} (THz)
Peaks at 60°C:		
$d1_1$; $d1_2$; $d1_3$	0.538; 0.539; 0.541	1.090; 1.112; 1.128
$e1_1$; $e1_2$; $e1_3$	0.579; 0.582; 0.578	1.113; 1.127; 1.136
c2; d2	0.735; 0.735	1.504; 1.461
Peaks at 21°C:		
d3; e4	0.531; 0.788	1.124; 1.522
$e3_1$; $e3_2$; $e3_3$	0.590; 0.584; 0.587	1.072; 1.107; 1.144
$d4_1$; $d4_2$; $d4_3$	0.725; 0.730; 0.730	1.466; 1.488; 1.510
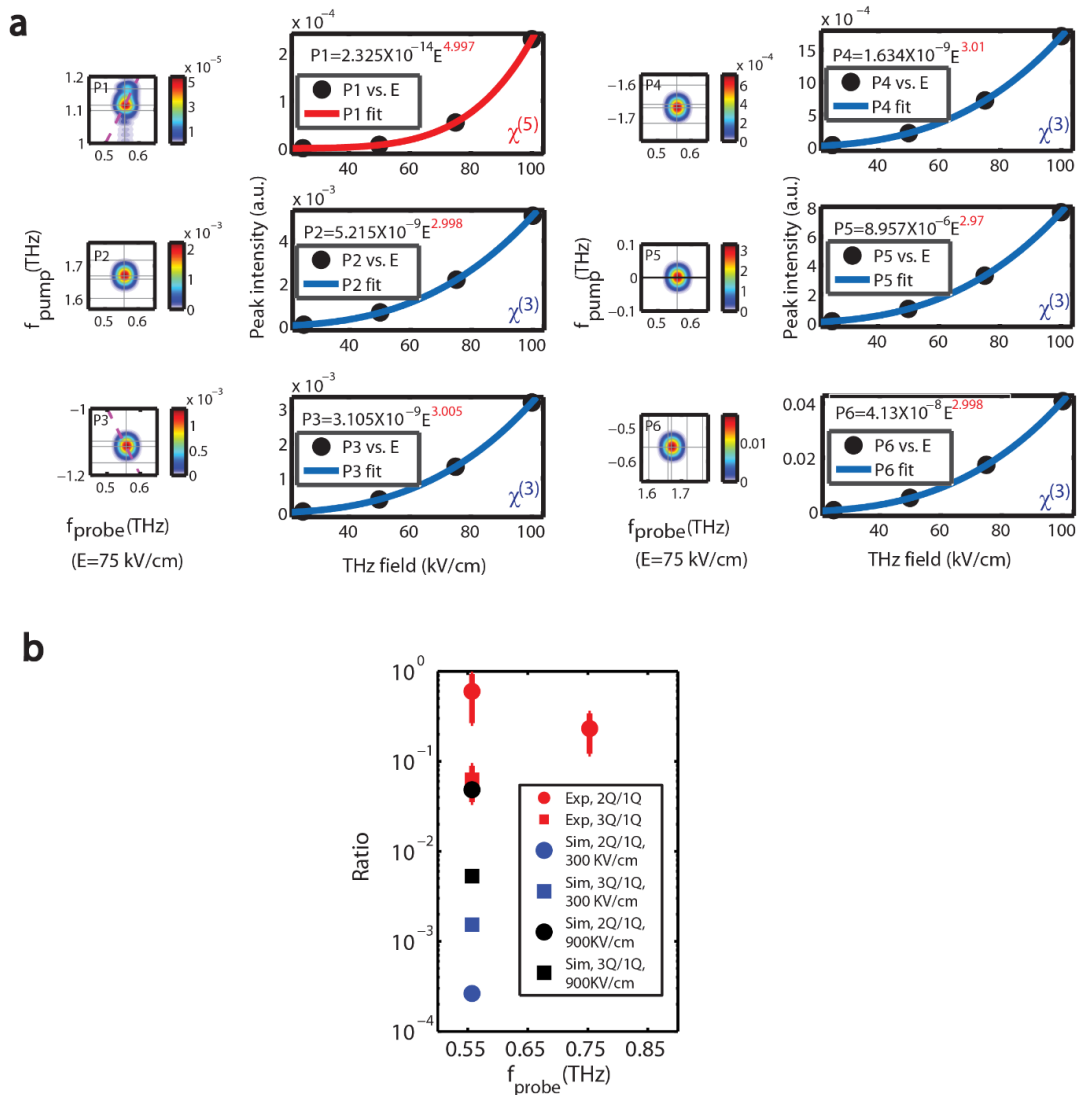
Fig. S19. Simulated field dependence of selected spectral peaks and peak ratios between experimental and simulation data. **a**, Peaks (P1-P6) are selected to show their peak intensities and field dependence fittings. In the weak to moderate THz field range, P2-P6 show third-order responses, whereas P1 is fifth order. P1 − P6 correspond to peaks labelled as NR11, NR8, R4, R6, PP1, and R13 in **Figs.S24** and **S26**. **b**, Ratios of 2Q and 3Q diagonal peaks to the corresponding 1Q diagonal peaks. Red circles and squares indicate experimental data. Points in blue and black are simulated results without many-body interactions when THz fields are 300 kV/cm and 900 kV/cm. Since the simulation conditions are similar to those in the experiments except for the absence of intermolecular interactions, the much smaller 2Q and 3Q signal amplitudes (blue and black) in the simulation than the experimental signals (red) indicate that weak nonlinear signals from non-interacting water molecules are not contributing significantly to 2Q and 3Q features.

a

$|2\rangle$

$|1\rangle$

1.670 THz

0.558 THz

$|0\rangle$

$|0\rangle = 1_{01}$ $|2\rangle = 2_{12}$

$|1\rangle = 1_{10}$

$J_{K_a K_c}$

b   P3: $\chi^{(3)}$

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	1\rangle\langle2	$ τ
$	1\rangle\langle0	$
$	0\rangle\langle0	$

(1)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	1\rangle\langle2	$ τ
$	0\rangle\langle2	$
$	0\rangle\langle0	$

(2)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	1\rangle\langle2	$ τ
$	1\rangle\langle0	$
$	1\rangle\langle1	$

(3)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	1\rangle\langle2	$ τ
$	0\rangle\langle2	$
$	2\rangle\langle2	$

(4)

c   P1: $\chi^{(5)}$

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	0\rangle\langle0	$
$	0\rangle\langle1	$
$	2\rangle\langle1	$ τ
$	2\rangle\langle0	$
$	0\rangle\langle0	$

(1)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	0\rangle\langle0	$
$	2\rangle\langle0	$
$	2\rangle\langle1	$ τ
$	2\rangle\langle0	$
$	0\rangle\langle0	$

(2)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	1\rangle\langle1	$
$	0\rangle\langle1	$
$	2\rangle\langle1	$ τ
$	2\rangle\langle0	$
$	0\rangle\langle0	$

(3)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	0\rangle\langle0	$
$	0\rangle\langle1	$
$	2\rangle\langle1	$ τ
$	0\rangle\langle1	$
$	1\rangle\langle1	$

(4)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	0\rangle\langle0	$
$	2\rangle\langle0	$
$	2\rangle\langle1	$ τ
$	0\rangle\langle1	$
$	1\rangle\langle1	$

(5)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	1\rangle\langle1	$
$	0\rangle\langle1	$
$	2\rangle\langle1	$ τ
$	0\rangle\langle1	$
$	1\rangle\langle1	$

(6)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	0\rangle\langle0	$
$	0\rangle\langle1	$
$	2\rangle\langle1	$ τ
$	2\rangle\langle0	$
$	2\rangle\langle2	$

(7)

$	0\rangle\langle0	$
$	1\rangle\langle0	$ t
$	0\rangle\langle0	$
$	2\rangle\langle0	$
$	2\rangle\langle1	$ τ
$	2\rangle\langle0	$
$	2\rangle\langle2	$

(8)

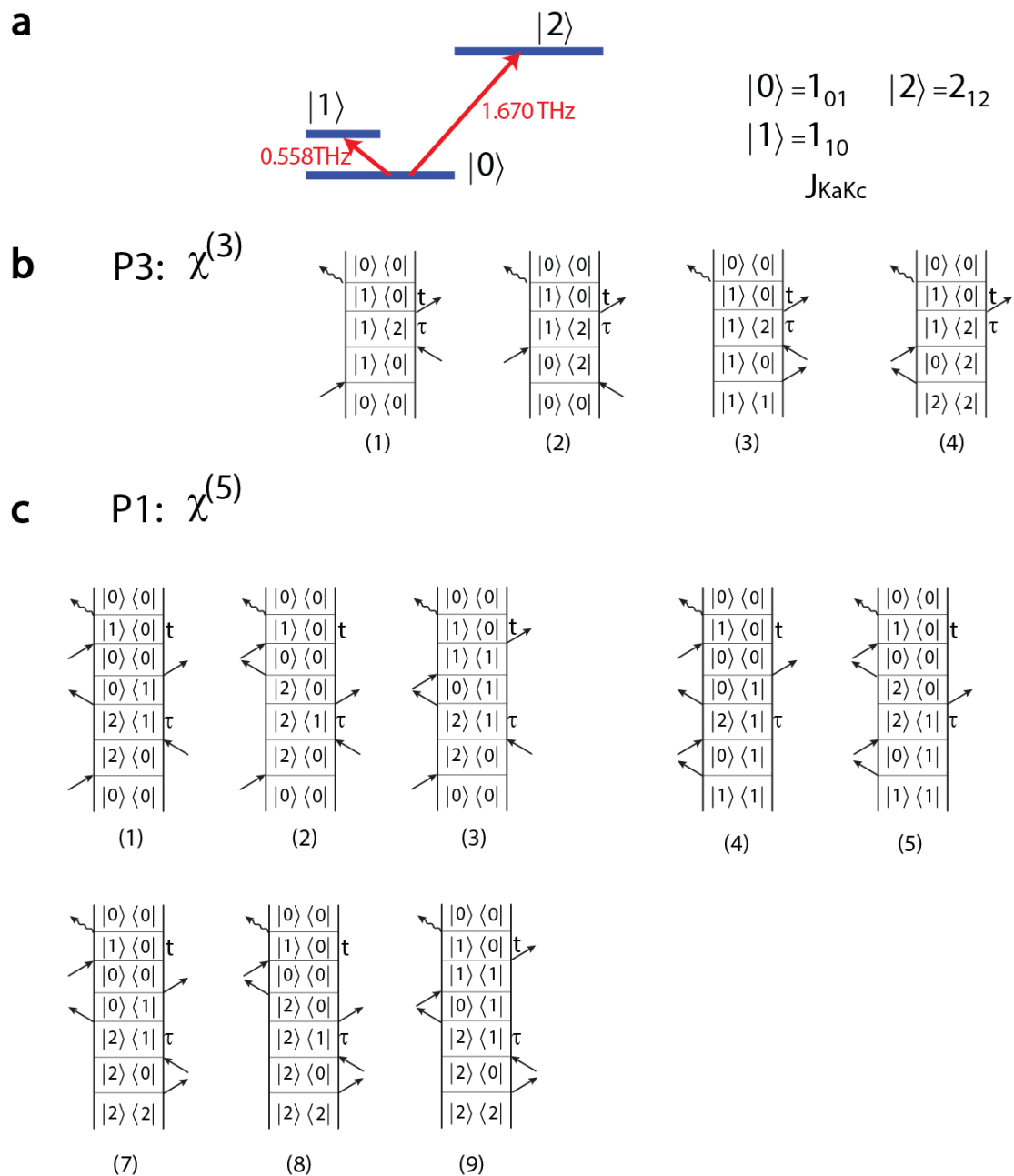| $|0\rangle\langle0|$ |
$	1\rangle\langle0	$ t
$	1\rangle\langle1	$
$	0\rangle\langle1	$
$	2\rangle\langle1	$ τ
$	2\rangle\langle0	$
$	2\rangle\langle2	$

(9)

Fig. S20. Possible Feynman diagrams for selected peaks. **a**, V-type energy ladder with defined energy level labels. The transition between rotational states |1> and |2> is forbidden. **b**, Four possible Feynman pathways for P3 (as shown in **Fig. S19**). **c**, Nine possible Feynman pathways for P1 (as shown in **Fig. S19**).
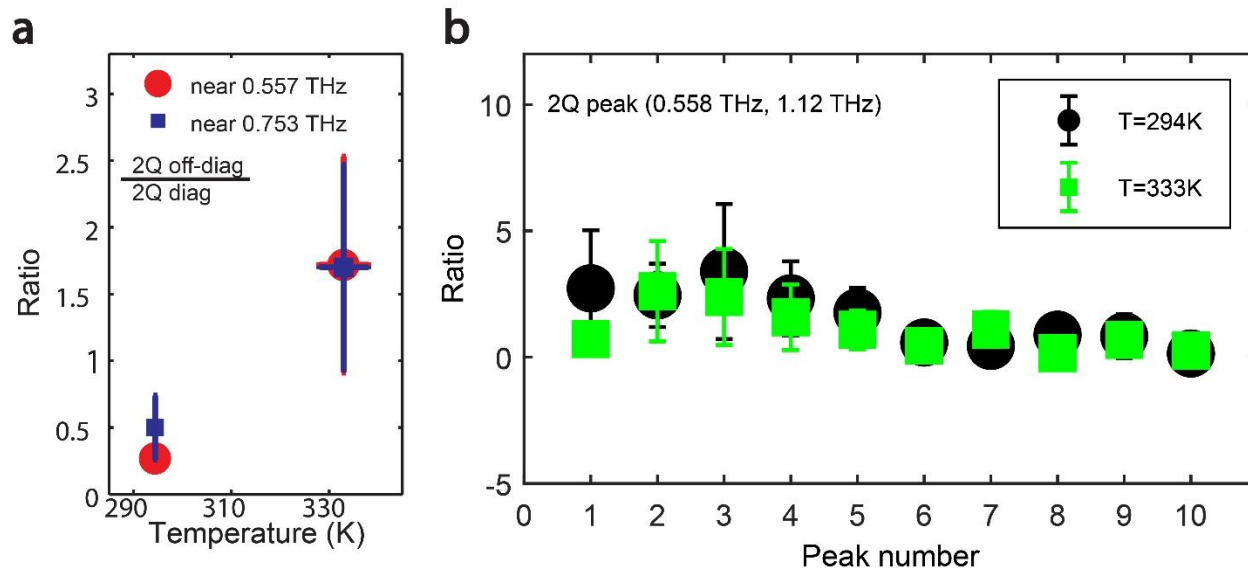
Fig. S21. Peak ratios in the experiments. a, Ratio of 2Q off-diagonal peaks to the 2Q diagonal peaks vs temperature. Red circles and blue squares are data near ($f_{probe} = 0.558$ THz, $f_{pump} = 1.115$ THz) and ($f_{probe} = 0.753$ THz, $f_{pump} = 1.506$ THz), respectively. The strong increase in the ratios with temperature suggests that 2Q off-diagonal features originate from (metastable) water complexes. **b**, Ratio of 1Q diagonal and off-diagonal rephasing and non-rephasing peak intensities to diagonal 2Q peak intensities. Black circles and green squares indicate the ratios with respect to the diagonal 2Q peak (with labels a1 and a3 in Figs. S15-S16) near (0.558 THz, 1.12 THz) at room temperature (21°C) and at 60 °C, respectively. For example, the data of Peak 1 is the ratio of the diagonal NR peak (0.989 THz, 0.989 THz) to the 2Q diagonal peak near (0.558 THz, 1.12 THz). All these peak intensities increase at the higher temperature due to the increase in water vapor pressure, but their relative intensities (given by the plotted ratios) do not change significantly, as expected for signals that all depend linearly on the water concentration. The ratios with respect to the other diagonal 2Q peak (with labels a2 and a4 in Figs. S15-S16) at (0.753 THz, 1.505 THz) yield the same conclusion. In contrast, Fig. 4d shows that the off-diagonal 2Q peak intensities increase far more than the diagonal 2Q peak intensities (and based on the present figure, far more than the NR and R peak intensities) at the higher temperature, as expected for signals from metastable molecular complexes such as weakly bound molecular pairs whose concentration scales quadratically with the water concentration. The positions of the diagonal and off-diagonal peaks selected for the plot above are as follows: Peak 1 (0.989 THz, 0.989 THz); Peak 2 (1.098 THz, 1.098 THz); Peak 3 (1.165 THz, 1.165 THz); Peak 4 (1.208 THz, 1.208 THz); Peak 5 (1.229 THz, 1.229 THz); Peak 6 (0.753 THz, 1.229 THz); Peak 7 (1.229 THz, 0.753 THz); Peak 8 (0.753 THz, -0.558 THz); Peak 9 (1.229 THz, -0.753 THz); and Peak 10 (0.753 THz, -1.229 THz), where the first (second) coordinate value is the probe (pump) frequency.
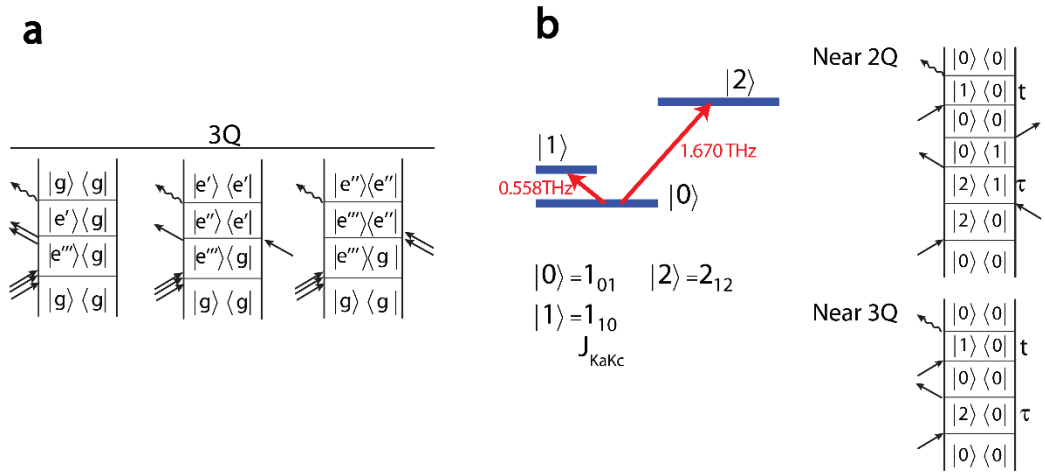
Fig. S22. Feynman diagrams for 3Q peaks and possible pathways near 2Q and 3Q peaks. **a**, The labels $|g\rangle$, $|e'\rangle$, $|e''\rangle$, and $|e'''\rangle$ denote the ground state, the singly excited state, the doubly excited state, and the triply excited state, respectively. Different diagrams are possible depending on coherence pathway selections. **b**, For the V-type energy ladder with defined energy level labels. The transition between rotational states |1> and |2> is forbidden. Besides the dominant pathways due to many-body interactions, examples for other possible contributions are illustrated here. See also **Fig. S19** & **20**.
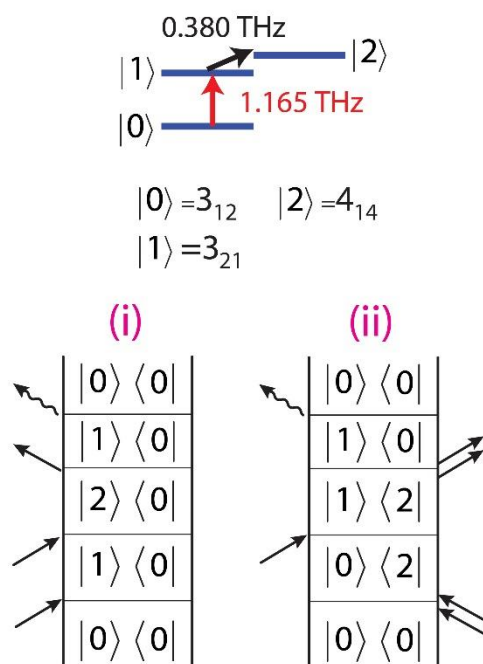
Fig. S23. Example of rotational couplings involving weak water transitions. Weak peaks in the simulation (for example, see **Fig. 4**, peak (W2) in **Fig. S24**, and peak (WR2) in **Fig. S26**) are related to the transition $3_{21} \leftrightarrow 4_{14}$ (see **Fig. S12**), the intensity of which is too weak to be observed in the present experiment. Two possible coherence pathways are shown.

Fig. S24. The complete simulated non-rephasing (NR) spectrum based on non-interacting water molecules. a, Diagonal/off-diagonal NR peaks and PP signals are shown and labelled. Some features are magnified by appropriate factors (indicated in red brackets) for clarity. For example, NR6(×5) means a non-rephasing peak magnified by a factor of 5. Weak peaks (W1-7) are nonlinear signals stemming from weak water transitions. For example, Peak W2 (f_{pump}=1.55 THz, f_{probe}=0.,38THz) involves the weak water transition of 0.38 THz. See **Fig. S23** for more details. **b**, Enlarged spectrum near 1.1 THz, which is indicated as a red dashed box in **a**. See also **Fig.2** for the measured spectrum.

Fig. S25. The complete experimental non-rephasing (NR) spectrum. NR, 2Q, 3Q and PP features are shown and labelled. Features from acetonitrile residue are indicated in a gray dashed box. Some features are magnified by appropriate factors (indicated in red brackets), for clarity. For example, NR15(×3) means a non-rephasing peak magnified by a factor of 3. Peak labels follow those in the simulation (**Fig. S24**). See also **Fig.2** for separated spectra. Background with intensity < 5% of the 2Q diagonal peak (2Q1) intensity is made white in color. For guidance, the signal-to-noise ratios for the following three peaks are as follows: 2Q1: ~40, NR1: ~30, and NR15: ~10.
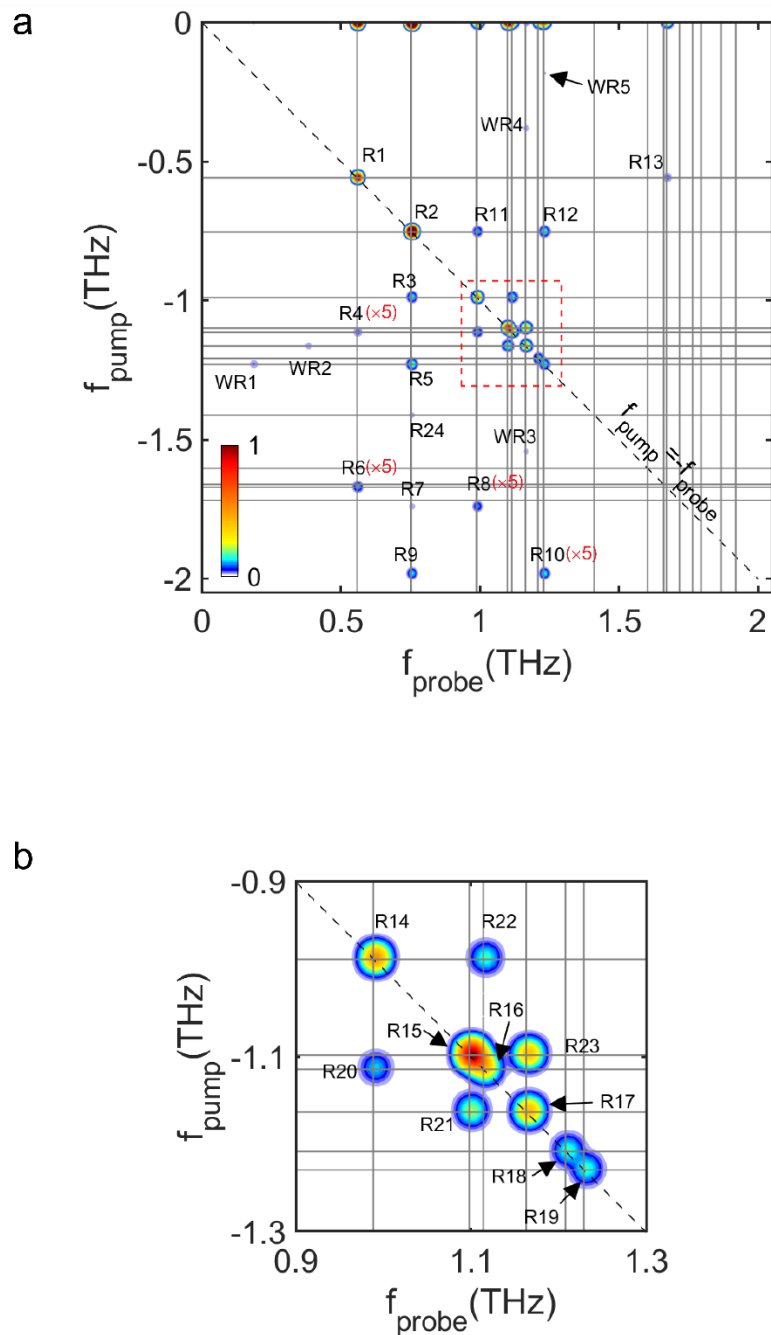
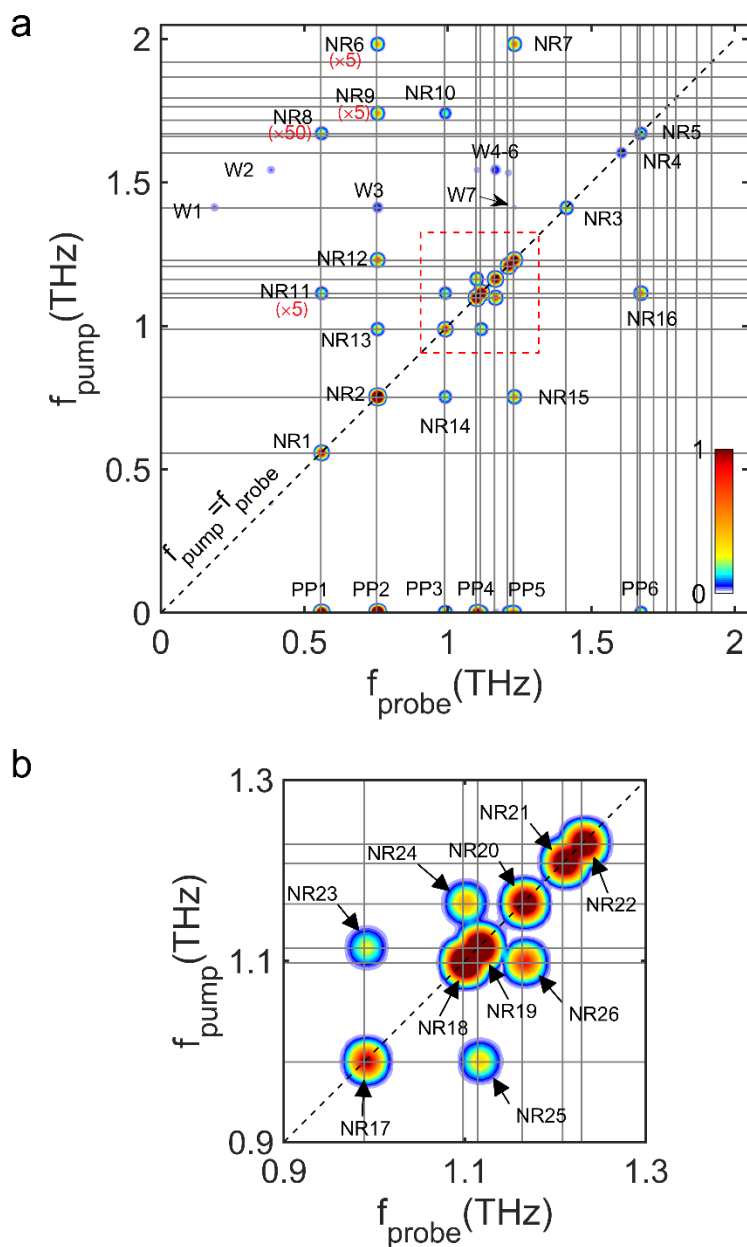**Fig. S26. The complete simulated rephasing (R) spectrum based on non-interacting water molecules. a**, R peaks and PP signals are shown and labelled. Some features are magnified by appropriate factors (indicated in red brackets), for clarity. For example, R6($\times$5) means a rephasing peak magnified by a factor of 5. Similar to weak peaks in NR spectra, weak peaks (WR1-5) are rephasing signals stemming from weak water transitions. **b**, Enlarged spectrum near 1.1 THz, which is indicated as a red dashed box in **a**. See also **Fig.2** for the measured spectrum.
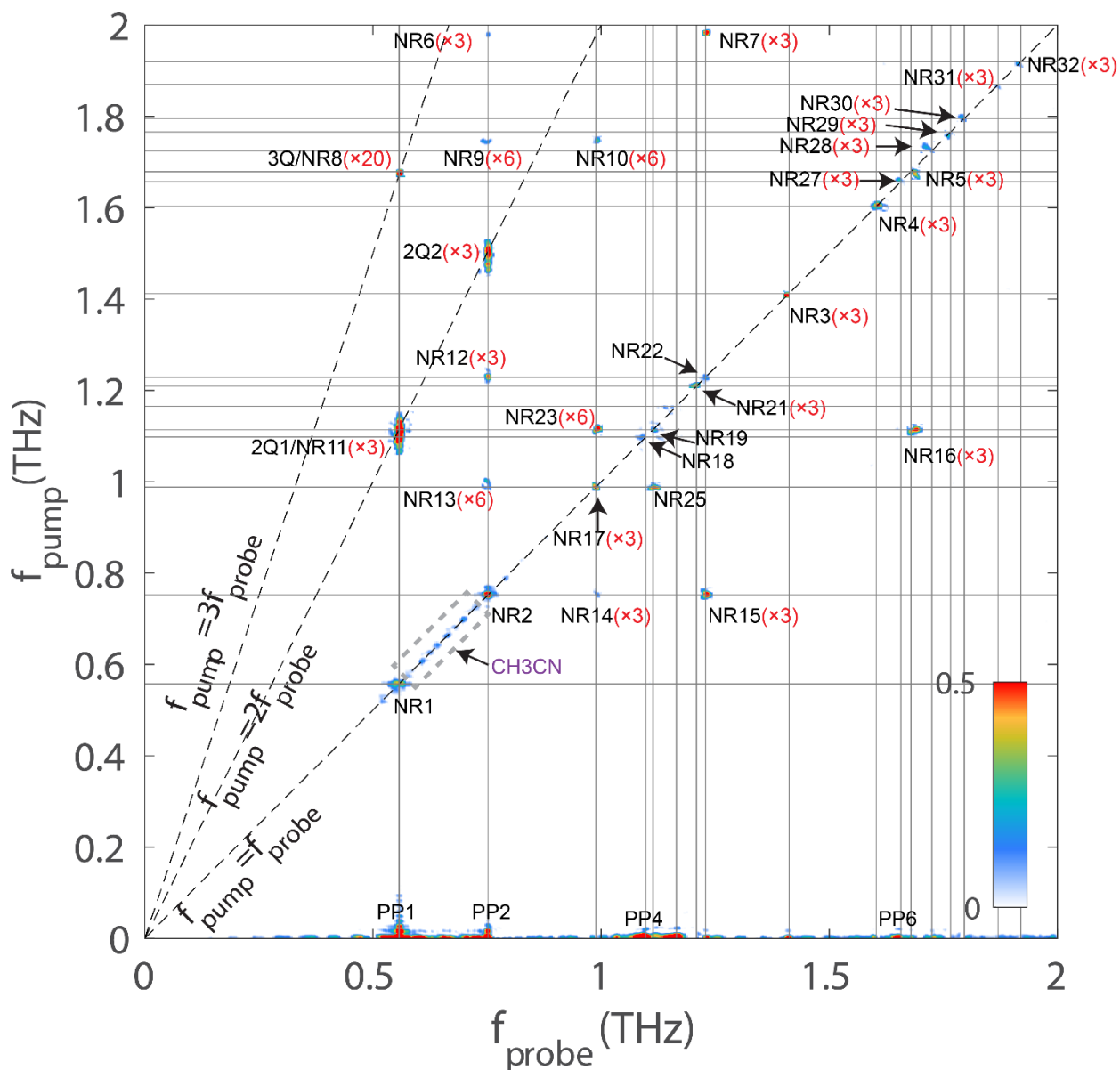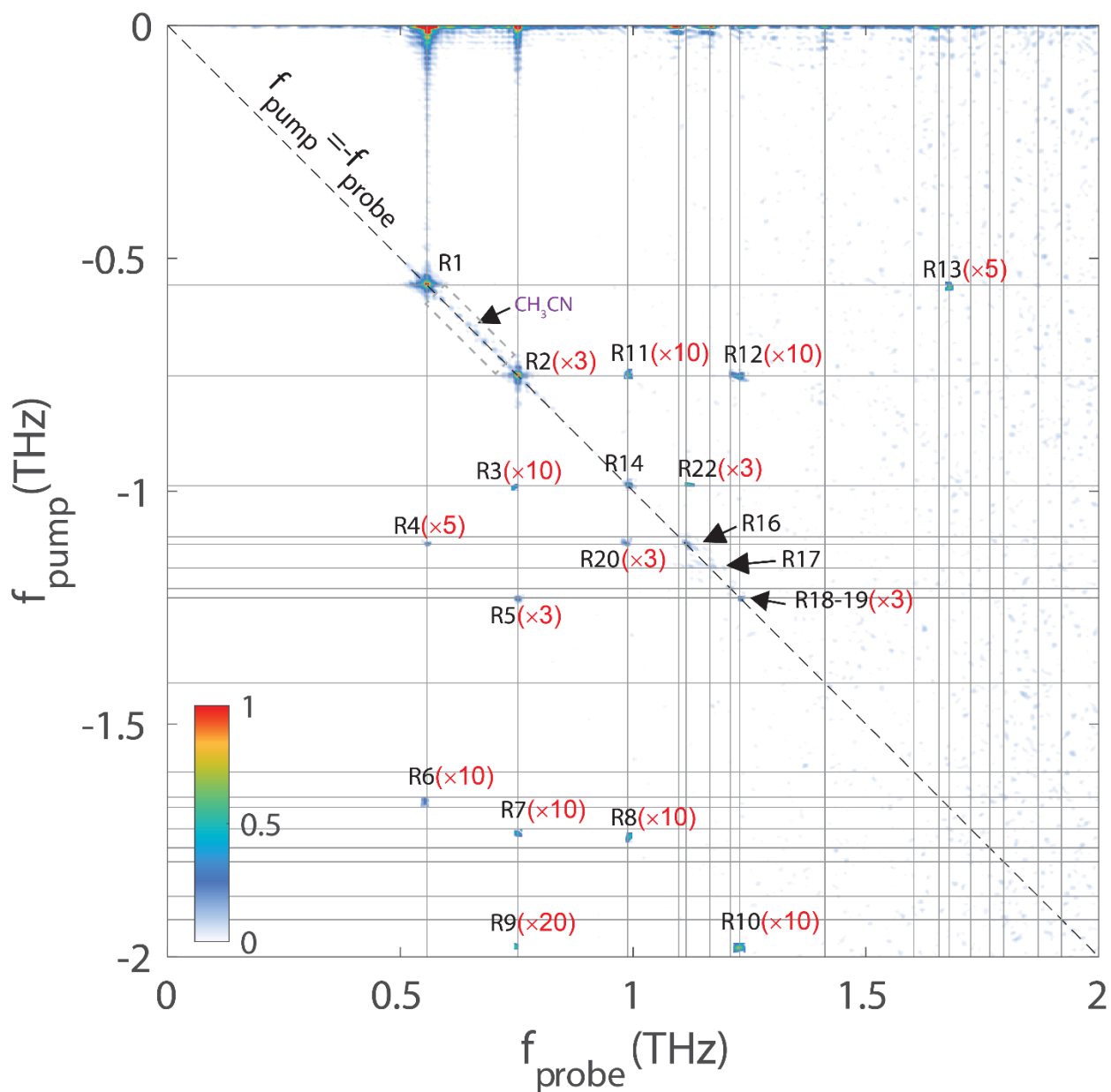
Fig. S27. The complete experimental rephasing (R) spectrum. Diagonal and off-diagonal R features are shown and labelled. Features from acetonitrile residue are indicated in a gray dashed box. Some features are magnified by appropriate factors (indicated in red brackets), for clarity. For example, R5($\times$3) means a rephasing peak magnified by a factor of 3. Peak labels follow those in the simulation (**Fig. S26**). See also **Fig.2** for separated spectra. Background with intensity $< 5\%$ of the 2Q diagonal peak (2Q1 in Fig. S25) intensity is made white in color. For guidance, the signal-to-noise ratios for the following two peaks are as follows: R1: ~30 and R13: ~3.

Rotational constants	Values (GHz)
A	835.83910
B	435.347353
C	278.139826
Δ_J	3.758280×10^{-2}
Δ_{JK}	-1.73361×10^{-1}
Δ_K	9.742470×10^{-1}
δ_J	1.521950×10^{-2}
δ_K	3.96461×10^{-2}
H	1.6284×10^{-5}
H	-5.2534×10^{-5}
H	-5.42802×10^{-4}
H	3.84436×10^{-3}
h_J	8.1611×10^{-6}
h_{JK}	-2.5607×10^{-5}
h_K	9.4784×10^{-4}

Table S1. Rotational constants of water molecules used in this work. Rotational constants of water rotational states are listed for the implementation of Watson's reduced Hamiltonian. Parameter values are taken from references (**4, 5**).

Line number	Para/Ortho type	Observed freq (THz)	Simulated freq (THz)	Obs-Sim (THz)	Transition assignment
1	Ortho	0.5574	0.5576	-0.0002	$1_{01} \longleftrightarrow 1_{10}$
2	Para	0.7525	0.7523	0.0002	$2_{02} \longleftrightarrow 2_{11}$
3	Para	0.9883	0.9885	-0.0002	$1_{11} \longleftrightarrow 2_{02}$
4	Ortho	1.098	1.098	0	$3_{03} \longleftrightarrow 3_{12}$
5	Para	1.114	1.114	0	$0_{00} \longleftrightarrow 1_{11}$
6	Ortho	1.165	1.163	0.002	$3_{12} \longleftrightarrow 3_{21}$
7	Para	1.209	1.208	0.001	$4_{13} \longleftrightarrow 4_{22}$
8	Para	1.229	1.229	0	$2_{11} \longleftrightarrow 2_{20}$
9	Ortho	1.412	1.411	0.001	$5_{14} \longleftrightarrow 5_{23}$
10	Para	1.603	1.602	0.001	$4_{04} \longleftrightarrow 4_{13}$
11	Ortho	1.657	1.660	-0.003	$2_{12} \longleftrightarrow 2_{21}$
12	Ortho	1.679	1.670	0.009	$1_{01} \longleftrightarrow 2_{12}$
13	Ortho	1.725	1.717	0.008	$2_{12} \longleftrightarrow 3_{03}$
14	Para	1.766	1.764	0.002	$6_{24} \longleftrightarrow 6_{33}$
15	Ortho Para	1.796	1.795	0.001	$7_{25} \longleftrightarrow 7_{34}$ $6_{15} \longleftrightarrow 6_{24}$
16	Ortho	1.870	1.867	0.003	$5_{23} \longleftrightarrow 5_{32}$
17	Para	1.920	1.920	0	$3_{13} \longleftrightarrow 3_{22}$

Table S2. Pure rotational transition lines of water vapor. Both observed and simulated pure rotational transition lines from the present work are listed. The typical deviation between observation and simulation is approximately 1 GHz.

Matrix elements	Results (given J, K, M)
$\langle JKM\lvert J_x\rvert J, K\pm 1, M\rangle$	$\dfrac{1}{2}[J(J+1)-K(K\pm 1)]^{1/2}$
$\langle JKM\lvert J_y\rvert J, K\pm 1, M\rangle$	$\mp\dfrac{i}{2}[J(J+1)-K(K\pm 1)]^{1/2}$
$\langle JKM\lvert J_z\rvert JKM\rangle$	K
$\langle JKM\lvert \boldsymbol{J}^2\rvert JKM\rangle$	$J(J+1)$
$\langle JKM\lvert J_x^2\rvert JKM\rangle$	$\dfrac{1}{2}\left[J(J+1)-K^2\right]$
$\langle JKM\lvert J_y^2\rvert JKM\rangle$	$\dfrac{1}{2}\left[J(J+1)-K^2\right]$
$\langle JKM\lvert J_z^2\rvert JKM\rangle$	K^2
$\langle JKM\lvert J_-^2\rvert J, K\pm 2, M\rangle$	$\dfrac{1}{2}[J(J+1)-K(K\pm 1)]^{1/2}[J(J+1)-(K\pm 1)(K\pm 2)]^{1/2}$
$\langle JKM\lvert \boldsymbol{J}^4\rvert JKM\rangle$	$[J(J+1)]^2$
$\langle JKM\lvert \boldsymbol{J}^2 J_z^2\rvert JKM\rangle$	$J(J+1)K^2$
$\langle JKM\lvert J_z^4\rvert JKM\rangle$	K^4
$\langle JKM\lvert \boldsymbol{J}^2 J_-^2\rvert J, K\pm 2, M\rangle$	$J(J+1)\cdot\dfrac{1}{2}[J(J+1)-K(K\pm 1)]^{1/2}[J(J+1)-(K\pm 1)(K\pm 2)]^{1/2}$
$\langle JKM\lvert J_z^2 J_-^2\rvert J, K\pm 2, M\rangle$	$K^2\cdot\dfrac{1}{2}[J(J+1)-K(K\pm 1)]^{1/2}[J(J+1)-(K\pm 1)(K\pm 2)]^{1/2}$
$\langle JKM\lvert J_-^2 J_z^2\rvert J, K\pm 2, M\rangle$	$(K\pm 2)^2\cdot\dfrac{1}{2}[J(J+1)-K(K\pm 1)]^{1/2}[J(J+1)-(K\pm 1)(K\pm 2)]^{1/2}$

Table S3. Calculated matrix elements of the rotational Hamiltonian up to the 4th order. Matrix elements of the angular momentum operator up to 4th order employed in this work are evaluated and listed here (**1**).

Matrix elements	Results (given J, K, M)
$\langle JKM \vert \mathbf{J}^6 \vert JKM \rangle$	$[J(J+1)]^3$
$\langle JKM \vert \mathbf{J}^4 J_z^2 \vert JKM \rangle$	$[J(J+1)]^2 K^2$
$\langle JKM \vert \mathbf{J}^2 J_z^4 \vert JKM \rangle$	$[J(J+1)]K^4$
$\langle JKM \vert J_z^6 \vert JKM \rangle$	K^6
$\langle JKM \vert \mathbf{J}^4 J_-^2 \vert J, K \pm 2, M \rangle$	$[J(J+1)]^2 \cdot \frac{1}{2}[J(J+1) - K(K \pm 1)]^{1/2}[J(J+1) - (K \pm 1)(K \pm 2)]^{1/2}$
$\langle JKM \vert \mathbf{J}^2 J_z^2 J_-^2 \vert J, K \pm 2, M \rangle$	$J(J+1)K^2 \cdot \frac{1}{2}[J(J+1) - K(K \pm 1)]^{1/2}[J(J+1) - (K \pm 1)(K \pm 2)]^{1/2}$
$\langle JKM \vert \mathbf{J}^2 J_-^2 J_z^2 \vert J, K \pm 2, M \rangle$	$J(J+1)(K \pm 2)^2$ $\cdot \frac{1}{2}[J(J+1) - K(K \pm 1)]^{1/2}[J(J+1) - (K \pm 1)(K \pm 2)]^{1/2}$
$\langle JKM \vert J_z^4 J_-^2 \vert J, K \pm 2, M \rangle$	$K^4 \cdot \frac{1}{2}[J(J+1) - K(K \pm 1)]^{1/2}[J(J+1) - (K \pm 1)(K \pm 2)]^{1/2}$
$\left\langle JKM \vert J_-^2 J_z^4 \vert J, K \pm 2, M \right\rangle$	$(K \pm 2)^4 \cdot \frac{1}{2}[J(J+1) - K(K \pm 1)]^{1/2}[J(J+1) - (K \pm 1)(K \pm 2)]^{1/2}$

Table S4. Calculated matrix elements of the rotational Hamiltonian of the 6th order. Matrix elements of the angular momentum operator of 6th order are evaluated and listed (**1**).

Exp. #	T(°C)	Position (f_{probe} (THz))	2Q diagonal peak (label (intensity))	2Q off-diagonal peak (label (intensity))	THz field strength(kV/cm)
1	21	~ 0.56	a3(1.0)	$b3_1$(0.23); $b3_2$(0.16); $b3_3$(0.27)	E1=400, E2=400
		~ 0.75	a4(1.0)	$b4_1$(0.52); $b4_2$(0.46)	
2	21	~ 0.56	a3(1.0)	$b3_2$(0.20); $b3_3$(0.29); $b3_4$(0.28)	E1=300, E2=350
		~ 0.75	a2(1.0)	$b2_3$(0.43); $b2_4$(0.28);	
3	21	~ 0.56	a3(1.0)	$b3_2$(0.28); $b3_3$(0.44)	E1=300, E2=350
		~ 0.75	a4(1.0)	$b4_1$(0.35); $b4_2$(0.96)	
4	60	~ 0.56	a3(1.0)	$b1_3$(0.60); $b1_4$(0.73); $b1_5$(1.1);	E1=440, E2=290
		~ 0.75	a2(1.0)	$b2_2$(0.98); $b2_3$(1.3); $b2_4$(1.2)	
5	60	~ 0.56	a1(1.0)	$b1_4$(1.3); $b1_5$(1.8)	E1=300, E2=350
		~ 0.75	a2(1.0)	$b2_2$(2.6); $b2_3$(2.8)	
6	60	~ 0.56	a1(1.0)	$b1_3$(2.8); $b1_4$(2.5); $b1_5$(2.4); $b1_6$(2.2)	E1=300, E2=350
		~ 0.75	a2(1.0)	$b2_3$(1.0); $b2_4$(1.9)	

Table S5. Data for calculating the ratio of 2Q off-diagonal to 2Q diagonal peak intensities. Data points from six experiments are used to calculate the ratio (Fig. 4d). Experiments (1 – 3) are data at room temperature, and experiments (4 – 6) are at 60°C. 2Q diagonal peak intensities are normalized to 1 (column 4). 2Q off-diagonal peak intensities are listed in column 5. The two THz field strengths (E1 and E2) used for each experiment are listed in the last column. Peak intensities (column 5) are normalized by E2 and the square of E1, and converted to the values when both THz strengths are 400 kV/cm. For example, the last data of the 2Q off-diagonal peak ($b2_4$) in column 5 is the product of its original value and the factor $1/(300^2 \times 350) \times (400^2 \times 400)$, where E1 = 300 kV/cm and E2=350 kV/cm are the THz fields used in the experiment. In addition, a gas-cell window transmission of 82.3% is taken into account to correct the THz field strength that reached the water molecules (because, for example, Exp. #2 and 3 measure the data from ambient air without the gas cell, whereas Exp. #5 and 6 are for water vapor in the gas cell). The averaged intensity ratio of 2Q off-diagonal peaks to 1Q diagonal peaks increases by a factor of 5.5 as the temperature rises from room temperature (21°C) to 60°C, while the water vapor pressure increases by a factor of 8.3.

Peak labels	Positions (f_{probe}, f_{pump}) (THz)	Related rotational states (J_{KaKc}) and peak origins
NR1	(0.558, 0.558)	1_{01}, 1_{10}
NR2	(0.753, 0.753)	2_{02}, 2_{11}
NR3	(1.41, 1.41)	5_{14}, 5_{23}
NR4	(1.60, 1.60)	4_{04}, 4_{13}
NR5	(1.67, 1.67)	2_{12}, 2_{21}; 1_{01}, 2_{12}
NR6	(0.753, 1.98)	2_{02}, 2_{11}, 2_{20}
NR7	(1.23, 1.98)	2_{02}, 2_{11}, 2_{20}
NR8/3Q	near (0.558, 1.67)	Many-body interaction; Radiative coupling; 1_{01}, 1_{10}, 2_{12}
NR9	(0.753, 1.74)	1_{11}, 2_{02}, 2_{11}
NR10	(0.989, 1.74)	1_{11}, 2_{02}, 2_{11}
NR11/2Q1	near (0.558, 1.12)	Many-body interaction; Radiative coupling; 1_{01}, 1_{10}, 2_{12}
NR12	(0.753, 1.23)	2_{02}, 2_{11}, 2_{20}
NR13	(0.753, 0.989)	1_{11}, 2_{02}, 2_{11}
NR14	(0.989, 0.753)	1_{11}, 2_{02}, 2_{11}
NR15	(1.23, 0.753)	2_{02}, 2_{11}, 2_{20}
NR16	(1.67, 1.12)	1_{01}, 1_{10}, 2_{12}
NR17	(0.989, 0.989)	1_{11}, 2_{02}
NR18	(1.10, 1.10)	3_{03}, 3_{12}
NR19	(1.11, 1.11)	0_{00}, 1_{11}
NR20	(1.16, 1.16)	3_{12}, 3_{21}
NR21	(1.21, 1.21)	4_{13}, 4_{22}
NR22	(1.23, 1.23)	2_{11}, 2_{20}
NR23	(0.989, 1.12)	0_{00}, 1_{11}, 2_{02}
NR24	(1.10, 1.16)	3_{03}, 3_{12}, 3_{21}
NR25	(1.12, 0.989)	0_{00}, 1_{11}, 2_{02}
NR26	(1.16, 1.10)	3_{03}, 3_{12}, 3_{21}
NR27	(1.66, 1.66)	2_{12}, 2_{21}
NR28	(~1.72, ~1.72)	2_{12}, 3_{03}
NR29	(1.77, 1.77)	6_{24}, 6_{33}
NR30	(1.80, 1.80)	7_{25}, 7_{34}; 6_{15}, 6_{24}
NR31	(1.87, 1.87)	5_{23}, 5_{32}
NR32	(1.92, 1.92)	3_{13}, 3_{22}
2Q2	near (0.753, 1.51)	Many-body interaction; Radiative coupling
PP1; PP2	(0.558, 0); (0.753, 0)	1_{01}, 1_{10}; 2_{02}, 2_{11}
PP3; PP4	(0.989, 0); (1.10, 0)	1_{11}, 2_{02}; 3_{03}, 3_{12}
PP5;	(1.23, 0)	2_{11}, 2_{20}
PP6	(1.67, 0)	2_{12}, 2_{21}; 1_{01}, 2_{12}
W1	(0.185, 1.41)	2_{11}, 2_{20}, 3_{13}
W2	(0.380, 1.54)	3_{12}, 3_{21}, 4_{14}
W3	(0.753, 1.41)	2_{02}, 2_{11}, 2_{20}, 3_{13}
W4	(1.10, 1.54)	3_{03}, 3_{12}, 3_{21}, 4_{14}
W5	(1.16, 1.54)	3_{12}, 3_{21}, 4_{14}
W6	(1.21, 1.53)	4_{13}, 4_{22}, 5_{15}
W7	(1.23, 1.41)	2_{11}, 2_{20}, 3_{13}

Table S6. Peak positions and related rotational states for NR spectra in Fig.S24&S25.

Peak labels	Positions (f_{probe}, f_{pump}) (THz)	Related rotational states (J_{KaKc}) and peak origins
R1	(0.558, -0.558)	$1_{01}, 1_{10}$
R2	(0.753, -0.753)	$2_{02}, 2_{11}$
R3	(0.753, -0.989)	$1_{11}, 2_{02}, 2_{11}$
R4	(0.558, -1.11)	$1_{01}, 1_{10}, 2_{12}$
R5	(0.753, -1.23)	$2_{02}, 2_{11}, 2_{20}$
R6	(0.558, -1.67)	$1_{01}, 1_{10}, 2_{12}$
R7	(0.753, -1.74)	$1_{11}, 2_{02}, 2_{11}$
R8	(0.989, -1.74)	$1_{11}, 2_{02}, 2_{11}$
R9	(0.753, -1.98)	$1_{11}, 2_{02}, 2_{11}$
R10	(1.23, -1.98)	$1_{11}, 2_{02}, 2_{11}$
R11	(0.989, -0.753)	$1_{11}, 2_{02}, 2_{11}$
R12	(1.23, -0.753)	$2_{02}, 2_{11}, 2_{20}$
R13	(1.67, -0.558)	$1_{01}, 1_{10}, 2_{12}$
R14	(0.989, -0.989)	$1_{11}, 2_{02}$
R15	(1.10, -1.10)	$3_{03}, 3_{12}$
R16	(1.11, -1.11)	$0_{00}, 1_{11}$
R17	(1.16, -1.16)	$3_{12}, 3_{21}$
R18	(1.21, -1.21)	$4_{13}, 4_{22}$
R19	(1.23, -1.23)	$2_{11}, 2_{20}$
R20	(0.989, -1.11)	$0_{00}, 1_{11}, 2_{02}$
R21	(1.10, -1.16)	$3_{03}, 3_{12}, 3_{21}$
R22	(1.12, -0.989)	$0_{00}, 1_{11}, 2_{02}$
R23	(1.16, -1.10)	$3_{03}, 3_{12}, 3_{21}$
WR1	(0.185, -1.23)	$2_{11}, 2_{20}, 3_{13}$
WR2	(0.380, -1.16)	$3_{12}, 3_{21}, 4_{14}$
WR3	(1.16, -1.54)	$3_{12}, 3_{21}, 4_{14}$
WR4	(1.17, -0.380)	$3_{12}, 3_{21}, 4_{14}$
WR5	(1.23, -0.182)	$2_{11}, 2_{20}, 3_{13}$

Table S7. Peak positions and related rotational states for R spectra in Fig.S26&S27.

Supplementary References

1. Zare RN (2013) *Angular momentum: understanding spatial aspects in chemistry and physics* (Wiley-Interscience).
2. Lu J*, et al.* (2016) Nonlinear two-dimensional terahertz photon echo and rotational spectroscopy in the gas phase. *Proceedings of the National Academy of Sciences* 113(42):11800-11805.
3. Bernath PF (2002) The spectroscopy of water vapour: Experiment, theory and applications. *PCCP* 4:1501-1509.
4. Gordy W & Cook RL (1984) *Microwave molecular spectra* (Wiley).
5. Townes CH & Schawlow AL (2013) *Microwave spectroscopy* (Courier Corporation).
6. Dai X*, et al.* (2012) Two-Dimensional Double-Quantum Spectra Reveal Collective Resonances in an Atomic Vapor. *Physical Review Letters* 108(19):193201.
7. Lomsadze B & Cundiff ST (2017) Frequency combs enable rapid and high-resolution multidimensional coherent spectroscopy. *Science* 357(6358):1389-1391.
8. Lomsadze B & Cundiff ST (2018) Frequency-Comb Based Double-Quantum Two-Dimensional Spectrum Identifies Collective Hyperfine Resonances in Atomic Vapor Induced by Dipole-Dipole Interactions. *Physical Review Letters* 120(23):233401.
9. Fleischer S, Field RW, & Nelson KA (2012) Commensurate Two-Quantum Coherences Induced by Time-Delayed THz Fields. *Physical Review Letters* 109(12):123603.
10. Vigasin A (2010) On the possibility to quantify contributions from true bound and metastable pairs to infrared absorption in pressurised water vapour. *Molecular Physics* 108(18):2309-2313.
11. Turner DB, Stone KW, Gundogdu K, & Nelson KA (2009) Three-dimensional electronic spectroscopy of excitons in GaAs quantum wells. *The Journal of Chemical Physics* 131(14):144510.
12. Mukhopadhyay A, Xantheas SS, & Saykally RJ (2018) The water dimer II: Theoretical investigations. *Chemical Physics Letters* 700:163-175.
13. Ruscic B (2013) Active Thermochemical Tables: Water and Water Dimer. *The Journal of Physical Chemistry A* 117(46):11940-11953.
14. Rocher-Casterline BE, Ch'ng LC, Mollner AK, & Reisler H (2011) Communication: Determination of the bond dissociation energy (D0) of the water dimer, (H2O)2, by velocity map imaging. *The Journal of Chemical Physics* 134(21):211101.
15. Lu J*, et al.* (2019) Two-Dimensional Spectroscopy at Terahertz Frequencies. *Multidimensional Time-Resolved Spectroscopy*, eds Buckup T & Léonard J (Springer International Publishing, Cham), pp 275-320.
16. Somma C, Folpini G, Reimann K, Woerner M, & Elsaesser T (2016) Two-Phonon Quantum Coherences in Indium Antimonide Studied by Nonlinear Two-Dimensional Terahertz Spectroscopy. *Physical Review Letters* 116(17):177401.