

Supporting Information: Breaking the Symmetry of Pyrimidine: Solvent Effects and Core-Excited State Dynamics

Sebastian Eckert,^{*,†} Vinícius Vaz da Cruz,^{*,†} Miguel Ochmann,[‡] Inga von Ahnen,[‡] Alexander Föhlisch,^{¶,†} and Nils Huse[‡]

[†]*Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Institute for Methods and Instrumentation for Synchrotron Radiation Research, 12489 Berlin, Germany*

[‡]*University of Hamburg, Institute for Nanostructure and Solid State Physics, Center for Free-Electron Laser Science, Luruper Chaussee 149, 22761 Hamburg, Germany*

[¶]*Universität Potsdam, Institut für Physik und Astronomie, 14476 Potsdam, Germany*

E-mail: sebastian.eckert@helmholtz-berlin.de; vinicius.vaz_da_cruz@helmholtz-berlin.de

Computational Details

The geometry of aqueous pyrimidine was optimized at the DFT/PBE0¹ with the def2-QZVPPD² basis set, using the D3BJ^{3,4} correction. The RIJCOSX approximation⁵ was used with the def2/J⁶ auxiliary basis. The conductor-like polarizable continuum model⁷ (CPCM) was used using the Gaussian charge scheme⁸ with a scaled Van der Waals cavity with radii from Bondi.⁹ The normal modes were computed at the same level of theory using numerical second derivatives for constructing the Hessian. All electronic structure calculations used the Orca package.¹⁰

Electronic RIXS Calculations As mentioned in the main text, the purely electronic RIXS scattering amplitudes are computed within the RSA-TD-DFT method described in detail within ref.¹¹ The ground-state Kohn-Sham orbitals used as reference were computed at the same level of theory described above, except that the smaller def2-TZVP(-f)² basis set was adopted.

RIXS simulations for the optimized geometry The linear response equations were solved with the restricted subspaces defined as $(N_c, N_o, N_u) \equiv (2, 15, 20)$, where N_c is the number of core orbitals included (the $1a_1$ and $2b_1$ N1s orbitals), N_o is the number of occupied valence orbitals included and N_u is the number of unoccupied orbitals considered. A total of 320 roots was computed comprising the manifold of valence excited state and the 20 lowest core-excited states. The electronic transition dipole moments were computed from the Orca output using Multifwn.¹²

RIXS simulations sampled from Molecular Dynamics The effect of the hydrogen bonding interactions and thermal geometrical distortions was assessed by computing also an averaged RIXS spectrum sampled from molecular dynamics simulations (see Sec. further below). From the MD snapshots we constructed a minimally solvated snapshot including pyrimidine and the first solvation shell at the nitrogen atoms, more specifically All water

molecules with a hydrogen within the distance to the first minimum in the N–H radial distribution (See Fig. 2) function were included explicitly. A CPCM layer was added around the minimally solvated cluster to account for the bulk liquid effects. The same level of theory from the previous section was used. A similar orbital truncation was used, except the subspaces were $(2, N_o, 20)$ where $N_o = 15 + 4 \times n_{H_2O}$, where each explicit water molecule adds 4 occupied orbitals to the calculation (removing the O1s orbitals).

Energy shifts A rigid shift of 11.125 eV was applied to the simulated X-ray absorption spectra of the isolated molecule and the MD-sampled structures presented in Fig. 1 and Fig. 2 of the main manuscript. The core-excited state potentials in Fig. 3a of the main manuscript were shifted by 11.179 eV with respect to the vertical point of the RSA-TD-DFT potentials to align the experimental and theoretical detuning axis.

Molecular coordinates The molecular coordinates used to define the symmetry planes and dipole moment orientations discussed in the main article result from the atom positions in Tab. 1.

Table 1: Optimized cartesian coordinates for pyrimidine with implicit solvation.

Element	x [Å]	y [Å]	z [Å]
N	0.6832	1.1868	0.0000
N	0.6832	-1.1868	0.0000
C	1.2782	0.0000	0.0000
C	-0.6484	1.1770	0.0000
C	-0.6484	-1.1770	0.0000
C	-1.3749	0.0000	0.0000
H	2.3636	0.0000	0.0000
H	-1.1452	2.1414	0.0000
H	-1.1452	-2.1414	0.0000
H	-2.4559	0.0000	0.0000

Normal Modes of Pyrimidine

Table 2 collects the normal modes computed as described above.

Table 2: Normal modes of aqueous pyrimidine computed at the DFT/PBE0 level with the def2-QZVPPD basis set, using the D3BJ correction and the CPCM model. Modes are order in ascending frequency, the classic label in terms of Wilson's^{13–15} notation is also given.

mode	irrep.	frequency (cm-1)	frequency (eV)	Wilson's notation
1	b_1	347.06	0.043	16b
2	a_2	404.54	0.050	16a
3	b_2	633.74	0.079	6b
4	a_1	691.80	0.086	6a
5	b_1	737.91	0.091	4
6	b_1	833.96	0.103	10b
7	b_1	997.11	0.124	5
8	a_2	1023.14	0.127	17a
9	a_1	1023.43	0.127	12
10	b_1	1046.20	0.130	17b
11	a_1	1091.32	0.135	1
12	b_2	1100.22	0.136	18b
13	a_1	1170.36	0.145	9a
14	b_2	1230.73	0.153	14
15	b_2	1275.59	0.158	15
16	b_2	1396.40	0.173	3
17	a_1	1447.11	0.179	19a
18	b_2	1505.56	0.187	19b
19	a_1	1632.62	0.202	8a
20	b_2	1638.06	0.203	8b
21	a_1	3175.59	0.394	20a
22	a_1	3181.37	0.394	13
23	b_2	3182.14	0.395	7b
24	a_1	3228.31	0.400	2

Core-Excited Potential Energy Scans along the Normal Modes

To identify which modes of pyrimidine are active upon core-excitation, we performed energy scans for both the ground state and the two near-degenerate core-excited states. The core-

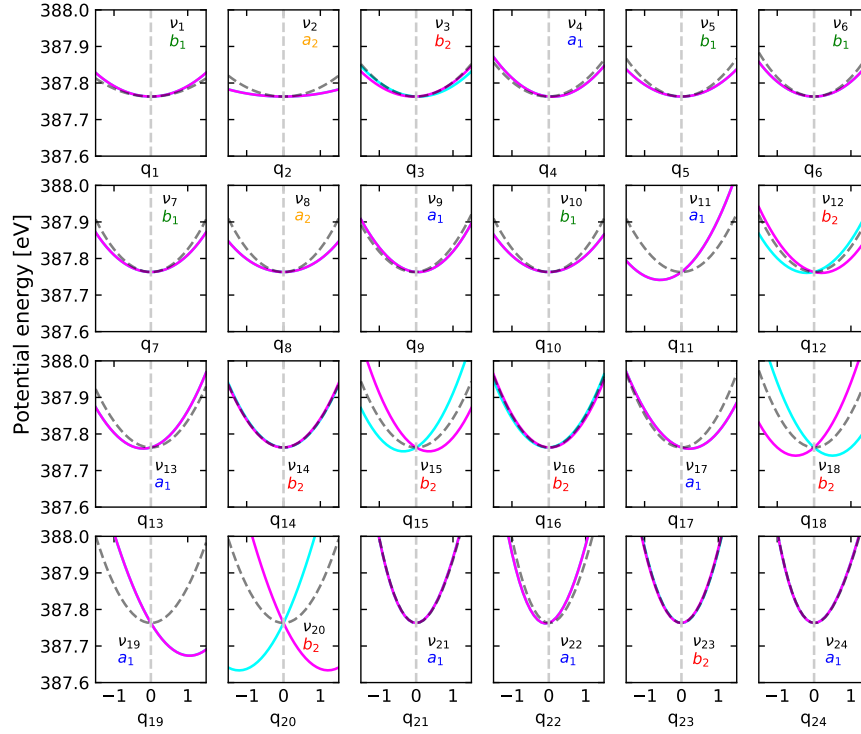


Fig. S 1: Potential energy scans along the 24 normal modes of pyrimidine for the ground state (dashed gray line, the minimum of this curve was shifted to the vertical excitation energy for comparison) and two diabatic core-excited states $|1s_{N_1}^{-1}2a_2^1\rangle$ and $|1s_{N_2}^{-1}2a_2^1\rangle$ (cyan and magenta). Modes are labelled in ascending frequency, the corresponding irreducible representation is also shown. Note that q_{ν_i} are dimensionless normal mode coordinates.

excited states were computed both at the TD-DFT level and using the Z+1 approximation. The Z+1 approximation was shown in ref.¹¹ to be more robust, therefore we adopted those results and they are shown in Fig. 1

Final RIXS cross-section

Using the definition of the scattering amplitude F_{f,ν_f} from the main text the double differential RIXS cross-section reads

$$\sigma(\omega', \omega) = r_0^2 \frac{\omega'}{\omega} \sum_{f,\nu_f} |F_{f,\nu_f}|^2 \Delta(\omega - \omega' - \omega_{f0} - (\epsilon_{\nu_f} - \epsilon_0), \Gamma_f), \quad (1)$$

where ω and ω' are the incoming and outgoing photon frequencies, respectively. $\omega_{f0} = (\epsilon_f - \epsilon_0)/\hbar$ is the transition frequency difference between the initial state and final state. And where $\Delta(x, \Gamma) = \Gamma/\pi(x^2 + \Gamma^2)$ is the normalized Lorentzian lineshape function .

Average over orientations

We need to average the cross-section over all orientations, as the molecules in solution are randomly oriented. Expressing the dot products explicitly as $\mathbf{e} \cdot \boldsymbol{\mu}_{nm} = \sum_{\alpha} e_{\alpha} \mu_{nm}^{\alpha}$, $\alpha \in \{x, y, z\}$, let us write the averaged RIXS cross-section as

$$\begin{aligned} \overline{\sigma(\omega', \omega)} &= r_0^2 \frac{\omega'}{\omega} \sum_{\alpha,\beta,\gamma,\delta} \overline{e'_{\alpha} e'_{\beta} e_{\gamma} e_{\delta}} \sum_{f,k,k'} \mu_{fk}^{\alpha} \mu_{k'f}^{\beta} \mu_{k0}^{\gamma} \mu_{0k'}^{\delta} \\ &\times \sum_{\nu_f} \langle \Psi_{k'}(\omega) | \nu_f \rangle \langle \nu_f | \Psi_k(\omega) \rangle \Delta(\omega - \omega' - \omega_{f0} - (\epsilon_{\nu_f} - \epsilon_0), \Gamma_f) \end{aligned} \quad (2)$$

following Gel'mukhanov and Ågren¹⁶ let us carry out the average by averaging over the photon polarization tensor in Eq. (2)

$$\overline{e'_{\alpha} e'_{\beta} e_{\gamma} e_{\delta}} = \frac{1}{30} \left[\delta_{\alpha\beta} \delta_{\gamma\delta} (3 + \cos^2 \chi) + (\delta_{\alpha\gamma} \delta_{\beta\delta} + \delta_{\alpha\delta} \delta_{\beta\gamma}) \frac{1}{2} (1 - 3 \cos^2 \chi) \right], \quad (3)$$

where χ is the angle between the incoming polarization and the detection direction.

Effective anti-symmetric wave packet

In our model, we take into account two vibrational modes, namely the $8a$ and $8b$ modes of pyrimidine. Although excitation into the symmetric $8a$ mode is important for describing the lineshape (specially for the elastic band) it does not contribute to the symmetry breaking. Therefore, in Fig. 3 of the main text, we use for illustration purposes an effective wave packet $|\tilde{\Psi}_k(\omega)\rangle$ where we sum over the contributions of the symmetric mode

$$|\tilde{\Psi}_k(\omega)\rangle = \sum_{\nu'_k{}^{(s)}} \langle \nu'_k{}^{(s)} | \Psi_k(\omega) \rangle, \quad (4)$$

$$= i \sum_{\nu'_k{}^{(s)}} \sum_{\nu_k} \frac{\langle \nu'_k{}^{(s)} | \nu_k \rangle \langle \nu_k | \nu_0 \rangle}{\omega - \omega_{k0} - (\epsilon_{\nu_k} - \epsilon_{\nu_0}) + i\Gamma}, \quad (5)$$

since $|\nu_k\rangle = |\nu_k^{(s)}\rangle |\nu_k^{(a)}\rangle$ and $\epsilon_{\nu_k} = \epsilon_{\nu_k^{(s)}} + \epsilon_{\nu_k^{(a)}}$, for simplicity we label the mode $8a$ as (s) and $8b$ as (a). Hence, we have $\langle \nu'_k{}^{(s)} | \nu_k \rangle = \delta_{\nu'_k{}^{(s)} \nu_k^{(s)}} |\nu_k^{(a)}\rangle$

$$|\tilde{\Psi}_k(\omega)\rangle = i \sum_{\nu'_k{}^{(s)}} \sum_{\nu_k^{(a)}, \nu_k^{(s)}} \frac{\delta_{\nu'_k{}^{(s)} \nu_k^{(s)}} \langle \nu_k^{(s)} | \nu_0^{(s)} \rangle |\nu_k^{(a)}\rangle \langle \nu_k^{(a)} | \nu_0^{(a)} \rangle}{\omega - \omega_{k0} - (\epsilon_{\nu_k^{(s)}} + \epsilon_{\nu_k^{(a)}} - \epsilon_{\nu_0^{(s)}} - \epsilon_{\nu_0^{(a)}}) + i\Gamma}, \quad (6)$$

$$= i \sum_{\nu_k^{(a)}, \nu_k^{(s)}} \frac{\langle \nu_k^{(s)} | \nu_0^{(s)} \rangle |\nu_k^{(a)}\rangle \langle \nu_k^{(a)} | \nu_0^{(a)} \rangle}{\omega - \omega_{k0} - (\epsilon_{\nu_k^{(s)}} + \epsilon_{\nu_k^{(a)}} - \epsilon_{\nu_0^{(s)}} - \epsilon_{\nu_0^{(a)}}) + i\Gamma} \quad (7)$$

$$= i \sum_{\nu_k^{(s)}} \langle \nu_k^{(s)} | \nu_0^{(s)} \rangle |\psi_{k, \nu_k^{(s)}}^{(a)}\rangle \quad (8)$$

where the partial wave packets $\psi_{k, \nu_k^{(s)}}^{(a)}$ are formally equivalent to the intermediate wave packets described in the mixed $nd+md$ RIXS model described by Vaz da Cruz et al..¹⁷

Incoming photon bandwidth and spectrometer resolution

The RIXS spectra were computed using the core-excited state lifetime broadening for nitrogen $\Gamma = 0.06$ eV. The final state broadening Γ_f was chosen as 0.05 eV to be able to distinguish the individual vibrational excitations. For comparison to the experimental spec-

tra, RIXS spectra were simulated with a 0.02eV excitation energy spacing and convoluted on the excitation energy axis with a 0.25eV Gaussian broadening to account for the finite excitation bandwidth and a 0.4eV gaussian broadening on the emission energy scale to account for the spectrometer resolution.

Self-absorption correction

The absorption and emission of the photons in the dense sample causing self-absorption effects, which mainly affect the electronically elastic emission channel, is accounted for using the expression^{18,19}

$$\tilde{\sigma}^{\text{RIXS}}(\omega, \omega') = \frac{\sigma^{\text{RIXS}}(\omega, \omega')}{1 + \frac{\sigma^{\text{XAS}}(\omega')}{\sigma^{\text{XAS}}(\omega)}}. \quad (9)$$

Physico-Chemical Picture

Let us try to give deeper insight into the symmetry breaking and detuning dependence observed. Starting from the scattering amplitudes, we have two equivalent pictures

$$\text{localized: } F_{f,\nu_f} = F_{f,\nu_f}^{(1)} + F_{f,\nu_f}^{(2)} \quad (10)$$

$$\text{delocalized: } F_{f,\nu_f} = F_{f,\nu_f}^{(1b_2)} + F_{f,\nu_f}^{(1a_1)} \quad (11)$$

ignoring the vibrational motion, both yield the same final state intensities at the ground-state minimum geometry. The two pictures are connected by the linear combinations

$$\boldsymbol{\mu}_{f,1b_2} = \frac{1}{\sqrt{2}}(\boldsymbol{\mu}_{f,1s_{N_1}} - \boldsymbol{\mu}_{f,1s_{N_2}}), \quad \boldsymbol{\mu}_{f,1a_1} = \frac{1}{\sqrt{2}}(\boldsymbol{\mu}_{f,1s_{N_1}} + \boldsymbol{\mu}_{f,1s_{N_2}}), \quad (12)$$

which yields the following relations

$$\mu_{1b_2,2a_2}^z = \sqrt{2}\mu_{1s_{N_1},2a_2}^z, \quad \mu_{7b_2,1b_2}^x = \sqrt{2}\mu_{7b_2,1s_{N_1}}^x, \quad \mu_{11a_1,1b_2}^y = \sqrt{2}\mu_{11a_1,1s_{N_1}}^y. \quad (13)$$

Lastly, we have the following symmetry relations

$$\begin{aligned}\mu_{7b_2,1s_{N_1}}^x &= -\mu_{7b_2,1s_{N_2}}^x, & \mu_{7b_2,1s_{N_1}}^y &= \mu_{7b_2,1s_{N_2}}^y, \\ \mu_{11a_1,1s_{N_1}}^x &= \mu_{11a_1,1s_{N_2}}^x, & \mu_{11a_1,1s_{N_1}}^y &= -\mu_{11a_1,1s_{N_2}}^y.\end{aligned}\tag{14}$$

Role of nuclear dynamics

To analyse the main effect of symmetry breaking, and to compare with the experiment, instead of dealing with the differential cross-section as in the main text, let us work with the integral RIXS cross-section for a given final state f

$$S_f(\omega) = \int_{-\infty}^{\infty} d\omega' \sigma_f(\omega', \omega), \quad \sigma_f = \sum_{\nu_f} |F_{f,\nu_f}|^2 \Delta(\omega - \omega' - \omega_{f0} - (\epsilon_{\nu_f} - \epsilon_0), \Gamma_f) \tag{15}$$

since F_{f,ν_f} does not depend on ω' and the Lorentzian $\Delta(x, \Gamma)$ is normalized, we obtain

$$S_f(\omega) = \sum_{\nu_f} |F_{f,\nu_f}|^2 = F_{f,\nu_f}^\dagger F_{f,\nu_f}, \tag{16}$$

where the scattering amplitude reads

$$F_{f,\nu_f} = F_{f,\nu_f}^{(1)} + F_{f,\nu_f}^{(2)}, \quad F_{f,\nu_f}^{(k)} = -i(\mathbf{e}' \cdot \boldsymbol{\mu}_{fk})(\mathbf{e} \cdot \boldsymbol{\mu}_{k0}) \times \langle \nu_f | \Psi_k(\omega) \rangle, \tag{17}$$

substituting this equation into Eq. (16), we get

$$S_f(\omega) = S_f^{(1)} + S_f^{(2)} + S_f^{(\text{int})}, \tag{18}$$

$$S_f(\omega) = \sum_{\nu_f} |F_{f,\nu_f}^{(1)}|^2 + |F_{f,\nu_f}^{(2)}|^2 + 2\text{Re}[F_{f,\nu_f}^{(1)\dagger} F_{f,\nu_f}^{(2)}], \tag{19}$$

where the first two terms are the cross-section for independent scattering via either nitrogen atom, and the last term is the interference term between the two channels. Let us analyse

the first two terms,

$$S_f^{(k)} = |F_{f,\nu_f}^{(k)}|^2 = (\mathbf{e}' \cdot \boldsymbol{\mu}_{fk})^2 (\mathbf{e} \cdot \boldsymbol{\mu}_{k0})^2 \times \sum_{\nu_f} \langle \Psi_k(\omega) | \nu_f \rangle \langle \nu_f | \Psi_k(\omega) \rangle$$

$$S_f^{(k)} = (\mathbf{e}' \cdot \boldsymbol{\mu}_{fk})^2 (\mathbf{e} \cdot \boldsymbol{\mu}_{k0})^2 \times \langle \Psi_k(\omega) | \Psi_k(\omega) \rangle \quad (20)$$

where we used $\sum_{\nu_f} |\nu_f\rangle \langle \nu_f| = 1$, also note that $\Psi_k(\omega)$ is not normalized and $\langle \Psi_k(\omega) | \Psi_k(\omega) \rangle = \sigma_{\text{vib}}^{\text{XAS}}$. Now, let us turn our attention to the interference term

$$S_f^{(\text{int})} = 2(\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_1}})(\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_1},2a_2})(\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_2}})(\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_2},2a_2})\text{Re}\langle \Psi_1(\omega) | \Psi_2(\omega) \rangle, \quad (21)$$

where $\langle \Psi_1(\omega) | \Psi_2(\omega) \rangle$ is the overlap between the two core-excited wave packets.

Symmetry breaking as a function of detuning

Using the core-excited symmetry breaking parameter, as defined in the main text

$$\zeta(\omega) = 1 - \text{Re} \frac{\langle \Psi_1(\omega) | \Psi_2(\omega) \rangle}{\langle \Psi_1(\omega) | \Psi_1(\omega) \rangle}, \quad (22)$$

noting that $\langle \Psi_1(\omega) | \Psi_2(\omega) \rangle = \langle \Psi_2(\omega) | \Psi_2(\omega) \rangle$ and introducing the vibrationally compensated integral cross-section $\tilde{S}_f = S_f / \langle \Psi_1(\omega) | \Psi_1(\omega) \rangle$ we arrive at the following expression

$$\tilde{S}_f(\omega) = (\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_1}})^2 (\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_1},2a_2})^2 + (\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_2}})^2 (\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_2},2a_2})^2 + \quad (23)$$

$$2(\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_1}})(\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_1},2a_2})(\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_2}})(\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_2},2a_2})(1 - \zeta(\omega)). \quad (24)$$

For large detuning $\Omega \ll 0$, the wave packets remain at the vertical excitation point and do not move leading to maximal overlap between the two core-excited states, so that

$$\zeta(\omega) \rightarrow 0 \quad (25)$$

$$\tilde{S}_f(\omega) = \left[(\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_1}})(\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_1},2a_2}) + (\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_2}})(\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_2},2a_2}) \right]^2 \quad (26)$$

$$\tilde{S}_f(\omega) = \left[(\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1b_2})(\mathbf{e} \cdot \boldsymbol{\mu}_{1b_2,2a_2}) + (\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1a_1})(\mathbf{e} \cdot \boldsymbol{\mu}_{1a_1,2a_2}) \right]^2 \quad (27)$$

$$\tilde{S}_f(\omega) = (\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1b_2})^2 (\mathbf{e} \cdot \boldsymbol{\mu}_{1b_2,2a_2})^2 \quad (28)$$

as it can be seen, in this limit symmetry breaking is quenched, and the localized and delocalized picture lead to the same result with partial cancellation of the emission transition dipoles, as well as the closing of the forbidden $1a_1$ scattering channel.

In the case of resonant excitation $\Omega = 0$, the wave packets are spread in opposite directions, and their overlap is minimized. In this case $0 < \zeta(\omega) \leq 1$. For a clearer qualitative picture, let us assume the extreme case so that

$$\zeta(\omega) \rightarrow 1$$

$$\tilde{S}_f(\omega) = (\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_1}})^2 (\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_1},2a_2})^2 + (\mathbf{e}' \cdot \boldsymbol{\mu}_{f,1s_{N_2}})^2 (\mathbf{e} \cdot \boldsymbol{\mu}_{1s_{N_2},2a_2})^2 \quad (29)$$

in this limit, symmetry breaking is maximum, and we obtain a result which corresponds to the sum of intensities for scattering via either nitrogen atom.

Evaluating these expressions for all the relevant final states, and using the symmetry relationships between the transition dipole moments from Eq. 14 we obtain for the two

extreme cases

$$\Omega \ll 0, \tag{30}$$

$$S_{2a_2} = 4(e'_z e_z)^2 (\mu_{1s_{N_1}, 2a_2}^z)^4 S_{\text{vib}}(\omega),$$

$$S_{7b_2} = 4(e_z \mu_{1s_{N_1}, 2a_2}^z)^2 (e'_x \mu_{7b_2, 1s_{N_1}}^x)^2 S_{\text{vib}}(\omega),$$

$$S_{11a_1} = 4(e_z \mu_{1s_{N_1}, 2a_2}^z)^2 (e'_y \mu_{11a_1, 1s_{N_1}}^y)^2 S_{\text{vib}}(\omega),$$

$$\Omega = 0, \tag{31}$$

$$S_{2a_2} = 2(e'_z e_z)^2 (\mu_{1s_{N_1}, 2a_2}^z)^4 S_{\text{vib}}(\omega),$$

$$S_{7b_2} = 2(e_z \mu_{1s_{N_1}, 2a_2}^z)^2 \left[(e'_y \mu_{7b_2, 1s_{N_1}}^y)^2 + (e'_x \mu_{7b_2, 1s_{N_1}}^x)^2 \right] S_{\text{vib}}(\omega),$$

$$S_{11a_1} = 2(e_z \mu_{1s_{N_1}, 2a_2}^z)^2 \left[(e'_x \mu_{11a_1, 1s_{N_1}}^x)^2 + (e'_y \mu_{11a_1, 1s_{N_1}}^y)^2 \right] S_{\text{vib}}(\omega).$$

Details of the Molecular Dynamics Simulations

All molecular dynamics simulations were carried out with the Gromacs package.²⁰ To model aqueous pyrimidine we adopted the accurate force field and charges derived from quantum-chemical calculations reported by Cacelli et al.,²¹ while water was modelled by the flexible single point charge model (SPC/F_w) described by Wu et al.²²

Initially, a single pyrimidine molecule was placed in a cubic box of length 3.4 nm, the box was filled with 1280 water molecules using the *gmx_solvate* tool from Gromacs. All calculations used periodic boundary conditions, the temperature was controlled by a modified Berendsen thermostat (0.1 ps time-constant) and in NPT runs the Parrinello-Rahman pressure-coupling (2 ps time-constant) was used. For the calculation of the Coulomb and van der Waals term a cut-off of 1.4 nm was employed, while long-range electrostatics were treated via the particle mesh Ewald (PME) procedure. The initially generated system was first optimized. Next, the system was equilibrated in two steps: first an NVT equilibration at T = 298 K was run for 500 ps ($\Delta t = 0.5$ fs), followed by an NPT equilibration at T = 298 K and P = 1 bar, which was also run for 500 ps ($\Delta t = 0.5$ fs). Throughout the equilibration

steps, the coordinates of the complex were constrained to the center of the simulation box. Lastly, a production NPT run was carried out for 10 ns ($\Delta t = 0.2$ fs, $T = 298$ K and $P = 1$ bar). From the final run snapshots were collect every 1 ps. They were subsequently used for analysis of the radial distribution functions and for sampling explicitly solvated configurations for the RIXS spectral calculations.

Radial distribution function

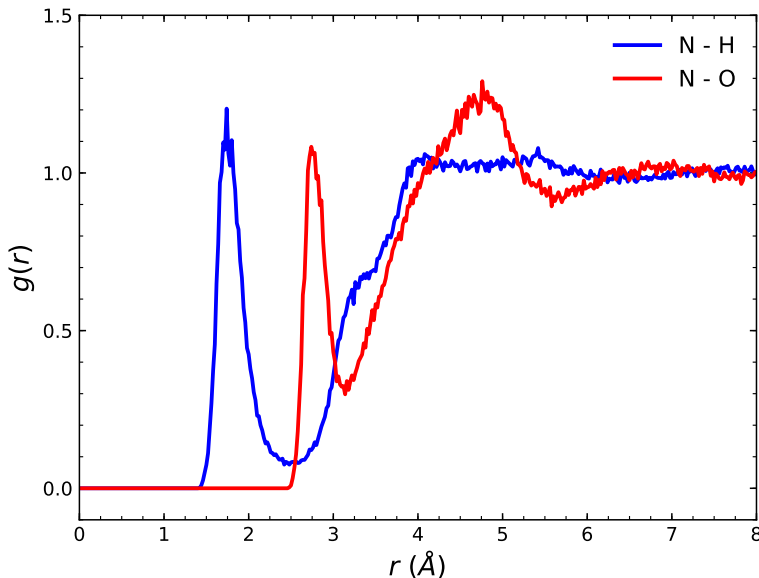


Fig. S 2: Radial distribution functions between the nitrogens of pyrimidine and the hydrogen atoms (blue line) and oxygen atoms of the solvent water molecules.

Hydrogen bond statistics

The hydrogen bonding properties of pyrimidine and water were primarily analysed via the radial distribution functions, computed with the *gmx_rdf* tool. Hydrogen bond statistics were computed with the *gmx_hbond* tool. The obtained results are shown in Tab. 3

Table 3: Hydrogen bonding analysis from the MD simulation of pyrimidine in aqueous solution. The r_{NH} distance corresponds to the maximum of the $g_{XH}(r)$ and $g_{HO}(r)$ pair-correlation functions. The average number of donated HBs (n_{HB}) from water to pyrimidine is defined via the criteria $\angle NHO < 20^\circ$ and $r_{N-O} < 3.5 \text{ \AA}$.

	r_{N-H}	$r_{N-O} \text{ (\AA)}$	$\angle NOH \text{ (}^\circ\text{)}$	n_{HB}
N ₁	1.77	2.83 ± 0.16	10.23 ± 4.82	1.04 ± 0.51
N ₂	1.77	2.83 ± 0.16	10.29 ± 4.84	1.05 ± 0.52
N ₁ + N ₂	1.77	2.83 ± 0.16	10.26 ± 4.83	2.08 ± 0.72

References

- (1) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods Without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (2) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (3) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (4) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- (5) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, Approximate and Parallel Hartree-Fock and Hybrid DFT Calculations. A Chain-of-Spheres-Algorithm for the Hartree-Fock Exchange. *Chem. Phys.* **2009**, *356*, 98 – 109.
- (6) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.
- (7) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
- (8) Lange, A. W.; Herbert, J. M. A Smooth, Nonsingular, and Faithful Discretization Scheme for Polarizable Continuum Models: The Switching/Gaussian Approach. *J. Chem. Phys.* **2010**, *133*, 244111.
- (9) Bondi, A. v. Van der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68*, 441–451.

- (10) Neese, F. The ORCA Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (11) Vaz da Cruz, V.; Eckert, S.; Föhlisch, A. TD-DFT Simulations of K-edge Resonant Inelastic X-Ray Scattering Within the Restricted Subspace Approximation. *Phys. Chem. Chem. Phys.* **2021**, *23*, 1835–1848.
- (12) Lu, T.; Chen, F. Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* **2012**, *33*, 580–592.
- (13) Wilson, E. B. The Normal Modes and Frequencies of Vibration of the Regular Plane Hexagon Model of the Benzene Molecule. *Phys. Rev.* **1934**, *45*, 706–714.
- (14) Lord, R.; Marston, A.; Miller, F. A. Infra-red and Raman Spectra of the Diazines. *Spectrochim. Acta* **1957**, *9*, 113–125.
- (15) Albert, S.; Quack, M. High Resolution Rovibrational Spectroscopy of Pyrimidine: Analysis of the B₁ Modes ν_{10b} and ν_4 and B₂ Mode ν_{6b} . *J. Mol. Spectrosc.* **2007**, *243*, 280–291.
- (16) Gel'mukhanov, F.; Ågren, H. Resonant Inelastic X-Ray Scattering With Symmetry-Selective Excitation. *Phys. Rev. A* **1994**, *49*, 4378–4389.
- (17) Vaz da Cruz, V.; Ignatova, N.; Couto, R. C.; Fedotov, D. A.; Rehn, D. R.; Savchenko, V.; Norman, P.; Ågren, H.; Polyutov, S.; Niskanen, J. et al. Nuclear Dynamics in Resonant Inelastic X-Ray Scattering and X-Ray Absorption of Methanol. *J. Chem. Phys.* **2019**, *150*, 234301.
- (18) Vaz da Cruz, V.; Ertan, E.; Couto, R. C.; Eckert, S.; Fondell, M.; Dantz, M.; Kennedy, B.; Schmitt, T.; Pietzsch, A.; Guimarães, F. F. et al. A Study of the Water Molecule Using Frequency Control Over Nuclear Dynamics in Resonant X-Ray Scattering. *Phys. Chem. Chem. Phys.* **2017**, *19*, 19573–19589.

- (19) Sun, Y.-P. Spontaneous and Stimulated X-Ray Raman Scattering. Ph.D. thesis, KTH Royal Institute of Technology, Stockholm, 2011.
- (20) Lindahl,; Abraham,; Hess,; van der Spoel, GROMACS 2019 Source code. doi:10.5281/zenodo.2424363, Zenodo, 2018.
- (21) Cacelli, I.; Ferretti, A.; Prampolini, G. Perturbative Multireference Configuration Interaction (CI-MRPT2) Calculations in a Focused Dynamical Approach: A Computational Study of Solvatochromism in Pyrimidine. *J. Phys. Chem. A* **2015**, *119*, 5250–5259.
- (22) Wu, Y.; Tepper, H. L.; Voth, G. A. Flexible Simple Point-Charge Water Model With Improved Liquid-State Properties. *J. Chem. Phys.* **2006**, *124*, 024503.