SUPPLEMENTAL INFORMATION

i-Motif DNA in isolated Hemiprotonated Cytosine Dimers, Studied by IR Spectroscopy and Theoretical Calculations

Ana Parejo Vidal[1,©], Yuika Okura[2,©], Keisuke Hirata[2,3], Vijay Madhav Miriyala[6,©], Pavel Hobza[6], Shun-ichi Ishiuchi[2,3,4], Masaaki Fujii[3,4,5], Mattanjah de Vries[1,3]

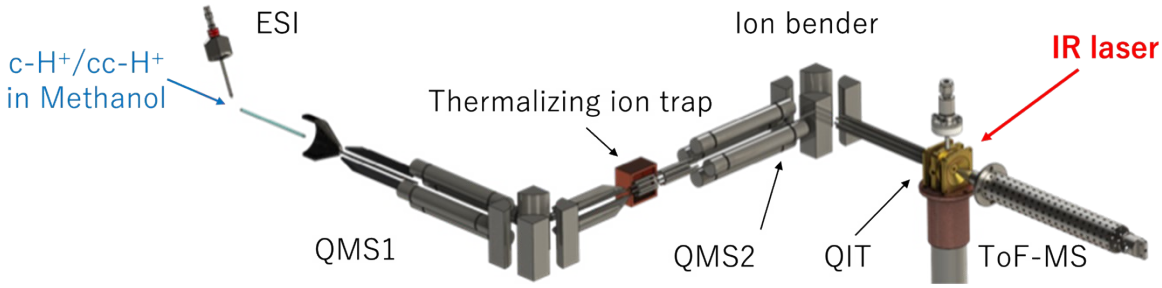

Figure S1: Experimental set up

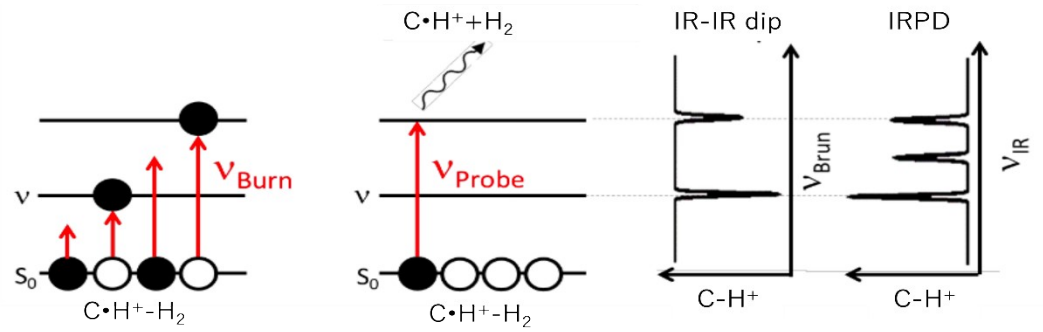

Figure S2: Schematic representation of the IR-IR double resonant spectroscopy technique.

The left panel illustrates the excitation process: the burn laser selectively excited a specific vibrational mode (ν_{Burn}), while the probe laser (ν_{Probe}) interrogates the isomer's response, isolating its spectrum. The central panel represents the ion tagging process, where a hydrogen-tagged protonated cytosine complex dissociates upon IR excitation. The right panel compares the resulting IR-IR depletion (dip) spectrum, which isolates specific isomers, with the full IRPD spectrum capturing all contributing species (ν_{IR}).

Figure S3 supports the assignment of the protonated cytosine monomer discussed in the main text. The most intense peak at 1669 cm^{-1} can be attributed to either the CN stretch and associated modes appearing at 1675 cm^{-1} in the keto form or the mode at 1701 cm^{-1} in the enol form. Additionally, the experimental IRPD peak at 1823 cm^{-1} aligns with the CO stretch in the keto form at 1862 cm^{-1}, while the peak at 1611 cm^{-1} corresponds to a mode in the enol structure at the same frequency

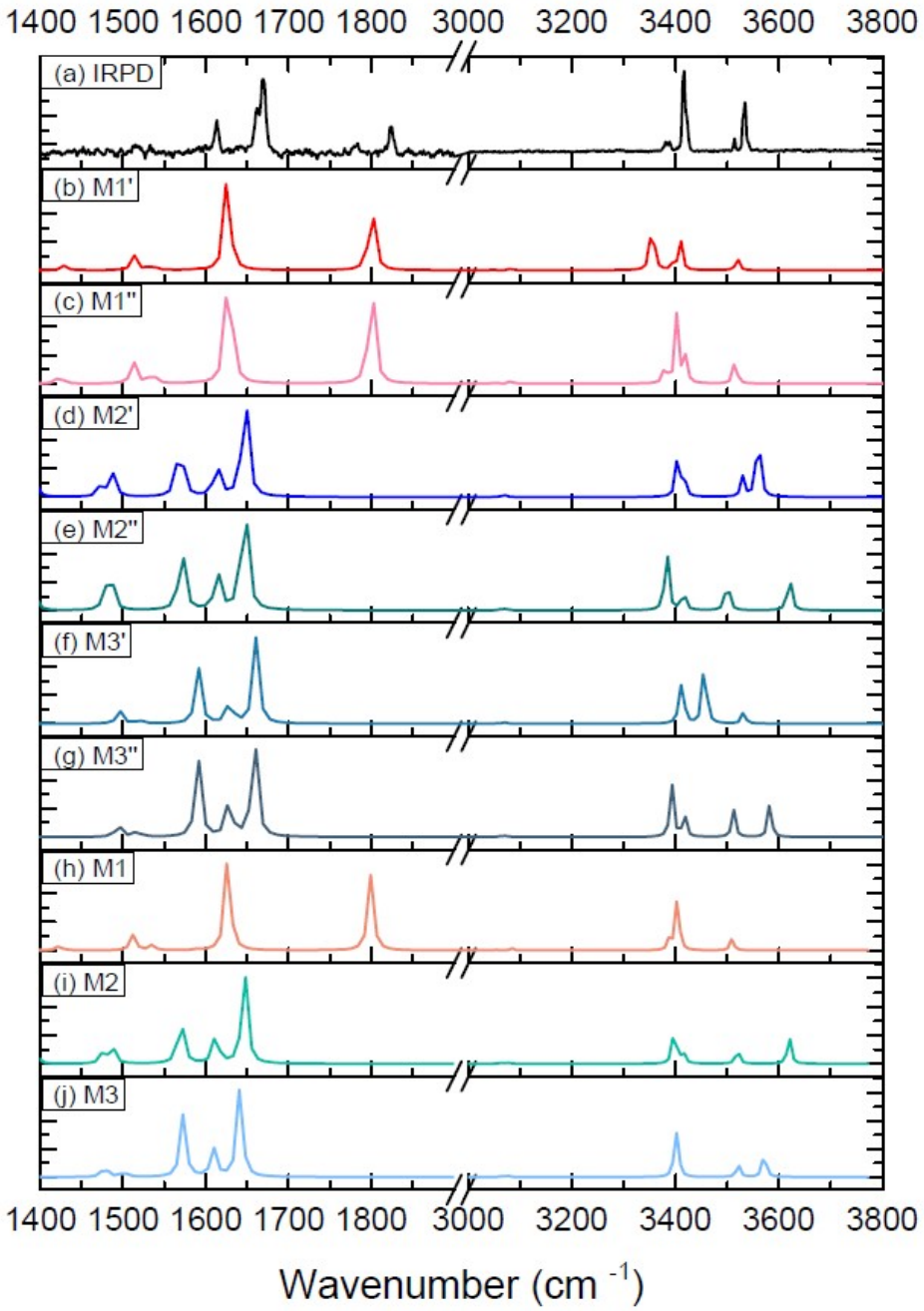

Figure S3: Comparison of the measured IRPD (black line, (a) and the calculated vibrational spectra for the H_2-tagged keto structures M1' (b), M1" (c) with the H_2 molecule located in the carbonyl or the amine group respectively, H_2-tagged enol structures M2' (d) and M2" (e) with the H_2 molecule located in the hydroxy or the amine group respectively, the untagged keto structure M1 (f), and the untagged enol structures M2 (g) and M3 (h). The computed spectra are shifted by a scaling factor of 0.956 for the 6 micron region of the IR, and 0.945 for the 3 micron region of the IR.

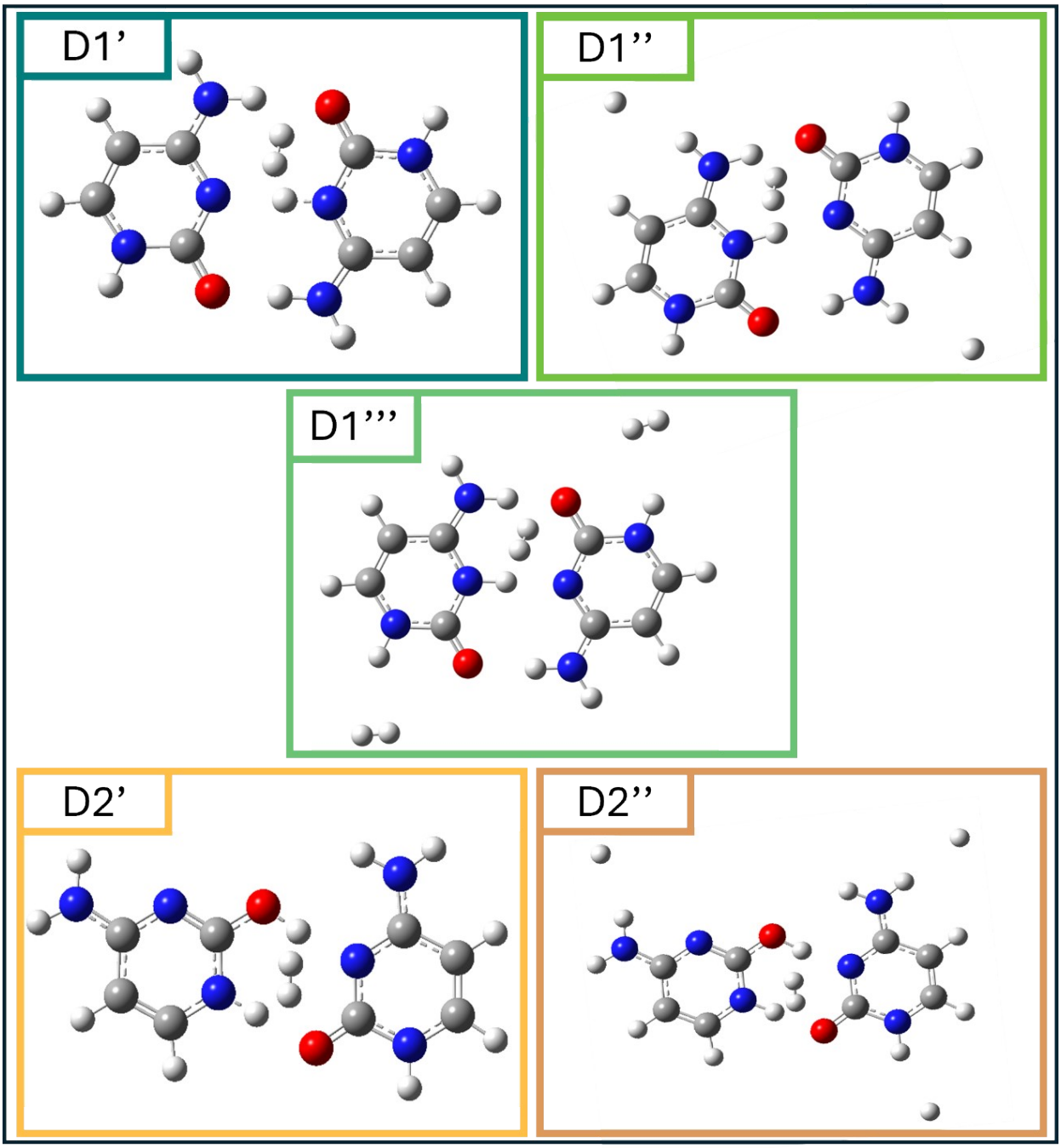

Figure S4: Optimized structures of the most stable hydrogen-tagged hemiprotonated cytosine dimers at the MP2/cc-pVTZ level. The molecular configurations include D1 and D2 tagged with either one hydrogen molecule (D1' and D2') or multiple hydrogen molecules (D1'', D1''', D2''). In D1'' and D1''', three hydrogen molecules are attached, while D2'' is tagged with four hydrogen molecules. These variations reflect distinct hydrogen bonding arrangements, highlighting the structural adaptability of D1 and D2 under tagging conditions.

Figure S5 compares the experimental IRPD spectrum with calculated spectra of untagged (D1 and D2) and hydrogen-tagged configurations (D1', D1'', D1''', D2', and D2''). The untagged structures (panels (b) and (c)) reproduce the dominant experimental features in both the 3-micron and 6-micron regions, providing a strong match to the observed spectrum. Hydrogen tagging introduces subtle shifts in key vibrational modes, particularly O-H and N-H stretches, due to interactions with the hydrogen molecules. These shifts refine peak assignments and further validate the alignment between experimental and theoretical spectra for D1 and D2, with D1 as the dominant structure. Again, from these calculations we conclude that the tagging has no significant impact on the spectra and the occurrence of multiple tags and different tagged structures does not affect the results. Notably, the ability of D1 and D2 to retain their characteristic spectral features under tagging conditions underscores their robustness as the dominant contributors to the gas-phase IRPD spectrum.

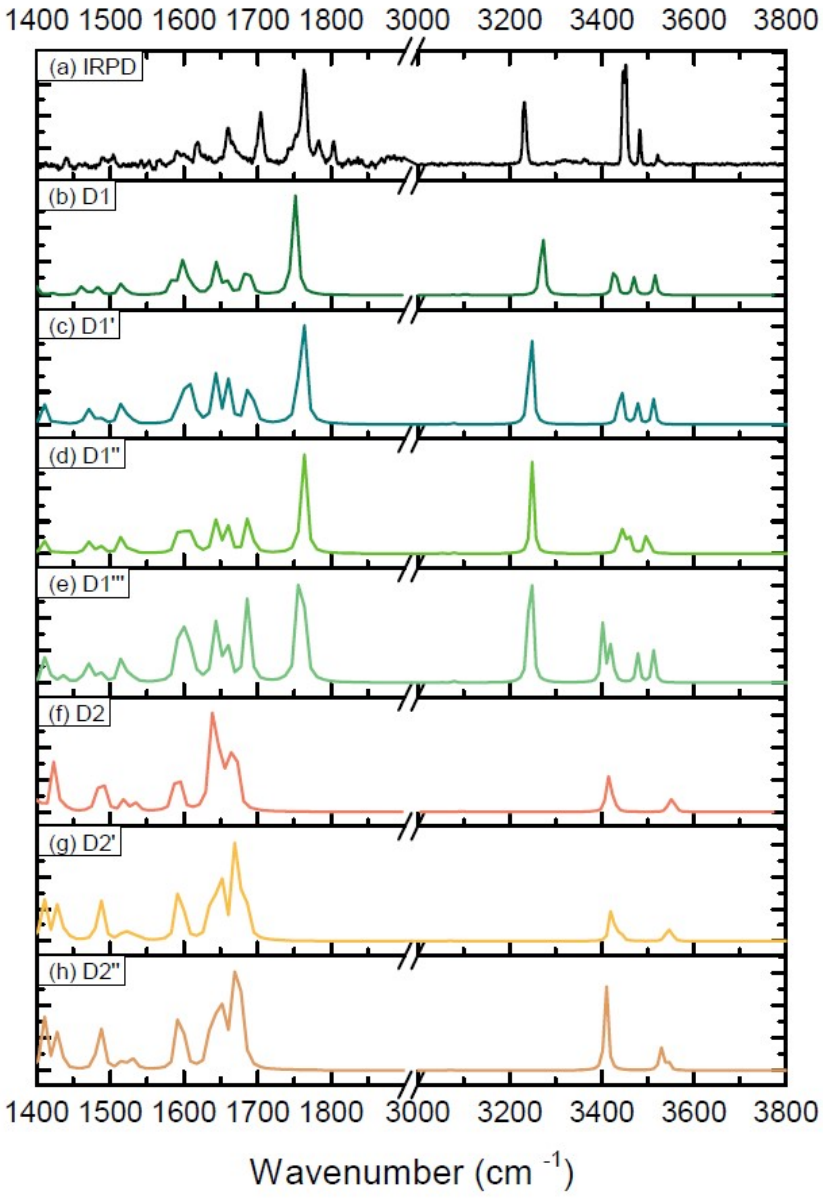

Figure S5: Comparison of the experimental IRPD spectrum (black line, panel (a)) with the calculated vibrational spectra of the two most stable hemiprotonated cytosine dimer configurations. Panels (b) and (c) represent the untagged D1 and D2 structures, respectively, while panels (d)–(g) show the H_2-tagged versions of D1 and D2, corresponding to D1', D1'', D1''', and D2'', as shown in Figure S4. Calculated spectra were scaled using a factor of 0.956 for the 6-micron region (1400–2000 cm^{-1}) and 0.945 for the 3-micron region (3000–3800 cm^{-1}). This comparison highlights the influence of hydrogen tagging on vibrational modes and the consistency of D1 and D2 with experimental data