Anharmonic Vibrational Modes of Nucleic Acid Bases Revealed by 2D IR Spectroscopy

Anharmonic Vibrational Modes of Nucleic Acid Bases Revealed by 2D IR Spectroscopy

Chunte Sam Peng, Kevin C. Jones, and Andrei Tokmakoff*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

*Corresponding Author. Telephone: 617-253-4503, Fax: 617-253-7030, tokmakof@mit.edu

Abstract

Polarization dependent two-dimensional infrared (2D IR) spectra of the purine and pyrimadine base vibrations of five nucleotide monophosphates (NMPs) were acquired in D_2O at neutral pH in the frequency range 1500 cm^{-1} –1700 cm^{-1}. The distinctive cross-peaks between the ring deformations and carbonyl stretches of NMPs indicate that these vibrational modes are highly coupled, in contrast with the traditional peak assignment which is based on a simple local mode picture such as C=O, C=N, and C=C double bond stretches. A model of multiple anharmonically coupled oscillators was employed to characterize the transition energies, vibrational anharmonicities and couplings, and transition dipole strengths and orientations. No simple or intuitive structural correlations are found to readily assign the spectral features, except in the case of guanine and cytosine which contain a single local CO stretching mode. To help interpret the nature of these vibrational modes, we performed density functional theory (DFT) calculations and found that multiple ring vibrations are coupled and delocalized over the purine and pyrimidine rings. Generally, there is close correspondence between the experimental and computational results, provided that the DFT calculations include explicit waters solvating hydrogen bonding sites. These results provide direct experimental evidence of the delocalized nature of the nucleotide base vibrations via non-perturbative fashion, and will serve as building blocks for constructing a structure-based model of DNA and RNA vibrational spectroscopy.

I. Introduction

Nucleic acids, namely DNA and RNA can adopt various secondary structures through hydrogen bond interactions between the nucleic acid bases. These base-paired secondary structures form double helices, stem-loop structures, pseudoknots, and G-quadruplexes, and their sequence dependent structural variations as well as flexibility are important components in biological processes such as DNA replication, packaging, and transcription. The ability to predict, manipulate, and monitor the atomic structures of these biomolecules is a focus in drug delivery and other bioengineering applications. In addition to being the carrier of genetic information, DNA nanotechnology has utilized artificial DNA strands to design molecular self-assemblies for a wide variety of technological purposes.[1-3] Therefore, understanding the interactions between the nucleic acid bases and the DNA/RNA secondary structural dynamics will be essential to study the interplay of these biomolecules, as well as improve the design of DNA nano-constructs.

Vibrational spectroscopy is a valuable technique to study nucleic acids because of its sensitivity to base pairing, sugar conformation, and glycosidic torsion angles. The vibrational IR and Raman spectra of nucleic acids and the general assignment of resonances have been known for decades[4-8] and have been used to structurally characterize single-stranded,[9,10] double-stranded DNA,[11-13] and various RNA structural elements.[14,15] In the mid-IR range from 800 - 1800 cm^{-1}, there are four main spectral regions of interest that reflect nucleic acid conformations. (a) Bands originated from the in-plane base vibrations, such as C=O stretch, C=N stretch, and ND_2 bending modes, are located in the 1500 –1800 cm^{-1} region. (b) The base-ribose vibrations that are sensitive to the glycosidic bond rotation have vibrational frequencies within the range of 1250 – 1500 cm^{-1}. (c) The phosphate-ribose vibrations which report on the backbone conformation lie

within the frequency range of 1000 –1250 cm^{-1}. (d) The ribose vibrations which provide reliable markers for the sugar puckering configurations can be observed in the 800 –1000 cm^{-1} region.

Even with the extensive body of work that exists, a description of the vibrational motions assigned to different resonances, and an understanding of how these are influenced by interactions in secondary structures remain elusive. Traditional spectral assignments are based on simple local modes such as C=O, C=C, and C=N double bond stretches,[4,8,16-18] whereas many computational studies suggest highly delocalized DNA vibrations.[18-22] Progress towards the study of DNA/RNA conformational dynamics is limited by the ability to accurately assign the different vibrational bands. This is because most experiments are limited to linear FTIR or Raman spectroscopy where the details of vibrational couplings are not readily accessible. On the other hand, two-dimensional infrared spectroscopy (2D IR) possesses the capability to characterize potential energy surfaces by obtaining the diagonal and off-diagonal anharmonicities and vibrational couplings through the analysis of cross-peaks. 2D IR experiments open an avenue to directly measure the vibrational couplings without the aid of isotope labeling, which can potentially alter the delocalized nature of these vibrations. Recent theoretical work on DNA bases and Watson-Crick base pairs showed that it is important to include anharmonic corrections for normal mode frequency calculations in order to reach reasonable agreement with the gas-phase IR measurement.[23] Additionally, the ability to perform polarization selective experiments allows one to obtain the relative orientations of the transition dipoles, which is a sensitive measure of the vibrational mixing underlying the anharmonic modes of the system.

2D IR spectroscopy on nucleic acids is relatively undeveloped, and much remains to be understood in order to use it as a practical tool for studying their structure and dynamics.

Krummel and Zanni reported the first 2D IR studies of Watson-Crick guanine-cytosine (G-C) base pairs and found that both electrostatics and base stacking effects are important in modeling the inter-base couplings between carbonyl stretches.[24,25] Theoretical calculations on the vibrational dynamics of DNA by Cho and coworkers investigated the vibrational basis modes and couplings of Watson-Crick base-paired DNA, as well as the effects of different DNA conformations on the exciton delocalization.[19,20,26,27] More recently, Elsaesser and coworkers have performed pump-probe and 2D IR experiments to characterize the vibrational couplings between NH and NH_2 stretching modes in DNA oligomers containing adenine-thymine (A-T) base pairs at different hydration levels[28,29] and guanine-cytosine (G-C) base pairs in chloroform.[30]

As a step toward understanding the spectral signatures of nucleotide interactions, we characterized the DNA and RNA vibrations of individual bases by acquiring polarization dependent 2D IR spectra of the purine and pyrimadine base vibrations of five nucleotide monophosphates (NMPs). We focus on the 1500 –1700 cm^{-1} spectral region that contains in-plane base vibrations. The distinctive cross-peaks between the vibrational modes of NMPs, such as ring vibrations and C=O stretches, indicate that these vibrational modes are highly coupled anharmonic bond stretching vibrations. Through the analysis and modeling of the experimental 2D IR spectra containing the IR active in-plane vibrations, we have characterized the transition energies, vibrational anharmonicities and couplings, and the magnitude and relative orientations of the transition dipoles. To help interpret the molecular origins of these vibrational modes, we also performed density functional theory (DFT) calculations. The consistency between experiment and calculations indicates that multiple ring vibrations are delocalized over the purine or pyrimidine rings.

II. Materials and Methods

Samples. Five nucleotide 5'-monophosphates (NMP): Adenosine 5'-monophosphates (AMP), guanosine 5'-monophosphate (GMP), uridine 5'-monophosphate (UMP), thymidine 5'-monophosphate (TMP), and cytidine 5'-monophosphate (CMP) were purchased from Sigma Aldrich. NMPs were chosen for the experiment because the addition of ribose to nucleic acid bases increases the water solubility (especially for adenine and guanine) and removes a number of possible tautomers. The samples were dissolved in excess D_2O (Cambridge Isotopes) and lyophilized twice for hydrogen / deuterium exchange until the absorbance of the OH stretching band (~ 3400 cm^{-1}) drops to below 5% of that of the fingerprint modes of the nucleotide monophosphate in the 6 μm region. For FTIR and 2D IR measurements, the lyophilized sample was dissolved in PBS buffer prepared in D_2O (137 mM NaCl, 150 mM KCl, 10 mM Na_2DPO_4, and 2 mM NaD_2PO_4) at pH = 6.9 (pD = 7.3). The solutions were prepared at concentrations of 20 mg/ml such that the OD's of the various fingerprint vibrations are within 0.1 to 0.3.

Spectroscopic Methods. For both FTIR and 2D IR measurements, 25 μL of sample solution was sandwiched between two CaF_2 windows that were separated by a 50 μm Teflon spacer. The experiments were carried out at room temperature. FTIR spectra were acquired using a Nicolet 380 FTIR spectrometer at 1.0 cm^{-1} resolution and averaging 64 one-second scans. 2D IR spectra were obtained as described previously.[31,32] Briefly, spectra were collected under parallel (ZZZZ) and perpendicular (ZZYY) polarization geometries at τ_2 = 150 fs. The ω_1 dimension was obtained by Fourier transformation of τ_1, which was scanned to 3.0 ps and 2.4 ps in 4 fs step until the signal decayed to 1 % level for the rephasing and non-rephasing spectra, respectively. The third-order signal was dispersed in a monochrometer to obtain the ω_3 dimension and collected using a 64 pixel MCT array detector. Both the rephasing and non-rephasing time

domain signal was zero-padded up to 65 ps before Fourier transformation. The rephasing and non-rephasing 2D spectra were added together to give the absorptive 2D IR spectra shown in this manuscript because they yield the highest resolution features. To acquire 2D spectra for different nucleotides, the IR excitation spectrum was tuned between 1610 cm^{-1} to 1700 cm^{-1}, in order to span the resonances as uniformly as possible. The FWHM spectral bandwidths were measured to be 90 fs in time and 162 cm^{-1} in energy using transient grating FROG.

Eigenstate-Based Modeling. To obtain information on the anharmonic nuclear potential and projection angles of the transition dipole moments in these systems, the 2D IR ZZZZ and ZZYY spectra were fit simultaneously with a least-squared fitting algorithm using a nonlinear response function formalism established previously.[33,34] This treats transitions between ground, singly- and doubly-excited eigenstates of a Hamiltonian for the coupled vibrations resonant with our laser pulses. The nonlinear response function derived from the sum of all possible Liouville pathways connecting eigenstates was Fourier transformed to give the calculated spectra. For each pair of vibrations, this model describes resonant transitions between six vibrational states with maximum total quantum number of 2: one ground state, two singly excited states, two doubly excited states, and one combination band. This model provides a description of the eigenenergy, vibrational anharmonicity, transition dipole strengths and projection angles between transition dipole moments.

2D lineshapes were parameterized by the generalized Kubo model,[35,36] in which the interaction of the system eigenstates with a harmonic bath are characterized by a frequency fluctuation correlation function of the form:

$$C(t) = \delta(t)/T_2 + \Delta^2 exp(-t/\tau) + \Delta_o^2 \quad (1)$$

$\delta(t)$ is the Dirac-delta function, T_2 describes the fast dynamics that determine the homogeneous linewidth, τ and Δ describe the intermediate regime, and Δ_o determines the inhomogeneous linewidth. Further details on the model are included in the Supplementary Information (SI). In our modeling, lineshape parameters for 1-2 transitions is constrained to match those of the fundamental transitions. Since the peak amplitudes acquired from the 2D spectral fitting are influenced by the pulse spectrum used in the 2D IR setup, the transition dipole strengths were obtained from fitting the experimental FTIR spectra after determining lineshape functions from the 2D IR spectra.

To extract information on the relative angles between the transition dipole moments, the orientational nonlinear response function was included by expressing the orientations of the four time-ordered transition dipole operators in a molecular frame, and then transformed into the laboratory frame through an orientational averaging.[37] Here we assumed the transition dipole moments from the zero- to one- quantum states ($\mu^{1.0}$) are parallel to those from the one- to two-quantum states ($\mu^{2.1}$) for the same vibrational coordinate. Therefore the relative cross-peak amplitudes from the polarization-selective measurement will only be sensitive to the angle between the two-coupled oscillators, Θ. Angles were determined from[34,38]

$$\cos\Theta = \sqrt{\frac{7-6\alpha}{1+2\alpha}} \qquad (2)$$

where α is the relative amplitude of the cross-peaks in the perpendicular and parallel spectra: $\alpha = A_{ZZYY}/A_{ZZZZ}$ The purine and pyrimidine ring vibrations were further assumed to be confined in the same plane for easier interpretation of the fitted angles.

For congested 2D spectra, there is great uncertainty in using direct spectral amplitudes of cross peaks for projection angle determination, given the extensive interference effects that exist

between positive and negative going resonances with varying lineshape. Different methods of amplitude determination for upper and lower cross-peaks, including amplitudes and volume integrations, led to variations in angle by as much as 40^{o}. For this reason, quoted angles were obtained solely from the relative amplitudes of peaks in absorptive 2D spectra obtained from fitting to the model. The error bars on the angle were calculated from the lower and upper bounds of the amplitude ratio between the ZZZZ and ZZYY spectra using the signal-to-noise ratio in the experimental spectra.

DFT Calculations. To help interpret the experimental results, *ab initio* density functional theory (DFT) calculations were performed using Gaussian 98.[39] The B3LYP hybrid functional was implemented with the 6-311G (d,p) basis set to determine the vibrational normal modes. A harmonic scaling factor of 0.9614 was applied to better match the calculated frequencies with the experiment.[40] To reduce the computational complexity, calculations were performed on nucleosides instead of nucleotide monophosphates. The phosphate vibrations are spatially separated and likely decoupled from the base vibrations, and fall within the 9 ~ 10 μm region, which is far outside of our 2D IR experimental spectral window. Lee and coworkers have performed similar calculations on DNA base derivatives and observed that quantum chemical calculations on methylated DNA bases was sufficient to capture the nature of the vibrations even with a smaller basis set 6-31G*.[19] Therefore, we chose adenosine, guanosine, uridine, thymidine, and cytidine over NMPs in these calculations. To investigate the solvent and hydrogen-bonding effect on the normal modes, the nucleosides were deuterated (see Figure 1 for deuteration scheme) and specific D_2O molecules were added to the proximity of carbonyl and ND groups.

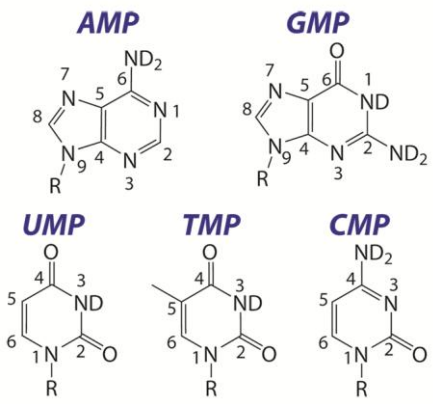


Figure 1. Structure and numbering convention of the deuterated five nucleotide 5'-monophosphates investigated. *R* represents D-ribose-5-phosphate.

III. Results

A. FTIR

The FTIR spectra of the five NMPs in D_2O in the 1500 –1700 cm^{-1} region are presented in Figure 2. In this spectral region, these bands mainly originate from in-plane double bond vibrations, such as C=O, C=N, and C=C stretch, and the ND_2 bend. The vibrational frequencies, intensities, and linewidths are sensitive to solvation and hydrogen to deuterium exchange.[16] Because an assignment convention for these modes has not been established, in this paper we will label the peaks under study with the first letter of the nucleotides, associated with a numbering scheme that goes from higher frequency to lower frequency (Figure 2).

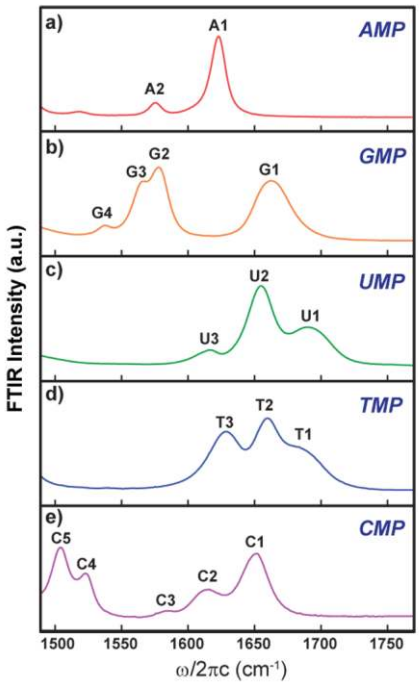


Figure 2. FTIR spectra of the five NMPs in PBS buffer at 20°C and pH 6.9. The D_2O background has been subtracted for each spectrum. The peak labels used in this paper are indicated for the various base vibrations, and their peak frequencies can be found in Table 1.

Nucleotides having carbonyl groups have IR peaks in the 1650 – 1700 cm^{-1} region. These peaks have been traditionally assigned to local C=O stretches: 1663 cm^{-1} for C^6=O in GMP, 1651 cm^{-1} for C^2=O in CMP, 1691 cm^{-1} for C^2=O and 1654 cm^{-1} for C^4=O in UMP, and 1690 cm^{-1} for C^2=O and 1661 cm^{-1} for C^4=O in TMP. These peaks in general are more intense due to the strong dipole moment of carbonyl stretches. Furthermore, they have broader linewidths (11-13 cm^{-1} FWHM) compared to the other lower frequency peaks, due to water hydrogen-bonding to the carbonyls that induce frequency shifts of these oscillators. Lower frequency modes have been assigned to be the ring vibrations consisting of C=C and C=N double bonds.[7]

Although there are similar chemical structures, the FTIR spectra of the NMPs in the spectral region of the ring vibrations are in fact quite distinct. No simple pattern exists that determines the frequency, intensity and linewidth of a particular type of vibration, leading to a

number of questions. For example, within the pyrimadines, why is there a drastic intensity variation for the two similar local C=O stretches in UMP and TMP, and why do these differ dramatically in the number and position of peaks compared to CMP? Why do the purines, AMP and GMP, have such dramatic differences in the number of resonances, their frequency, and their lineshapes?

B. 2D IR Spectra of Purines

B.1. AMP

2D IR spectra of AMP are shown in Figure 3a and 3b. The infrared spectra of AMP display one intense peak at 1625 $cm^{-1}$and a weak peak at 1578 cm^{-1}, denoted as A1 and A2 respectively, which are assigned to pyrimidine ring modes.[16,41] 2D IR resonances display positive and negative peaks that arise from the 0-1 and 1-2 transitions, respectively, so that the splitting of the doublet along the ω_3 axis gives a measure of the diagonal anharmonicity. The presence of cross-peaks indicates two anharmonically coupled vibrational modes, as would be expected for two delocalized modes of a conjugated ring system. The overall eight-peaks in the spectrum follows a classic pattern that emerges for transitions between the ground singly and doubly excited eigenstates of a pair of coupled vibrations.[42] The variation in cross-peak amplitude between parallel (ZZZZ) and perpendicular (ZZYY) polarized spectra depends on the projection angle between the two transition dipole moments. A significant enhancement of the cross-peak amplitude relative to the diagonal peaks is observed for AMP, indicating that the projection angle is close or equal to 90 degrees.[34] Both A1 and A2 and their cross-peaks appear homogeneously broadened, as seen from the diamond-like lineshapes, with diagonal linewidths (FWHM) of 11 and 9.4 cm^{-1}, respectively. The comparison of diagonal linewidth (σ), anti-

diagonal linewidth (Γ), and ellipticity, $(\sigma^2 - \Gamma^2)/(\sigma^2 + \Gamma^2)$, for the five NMPs is shown in the SI.

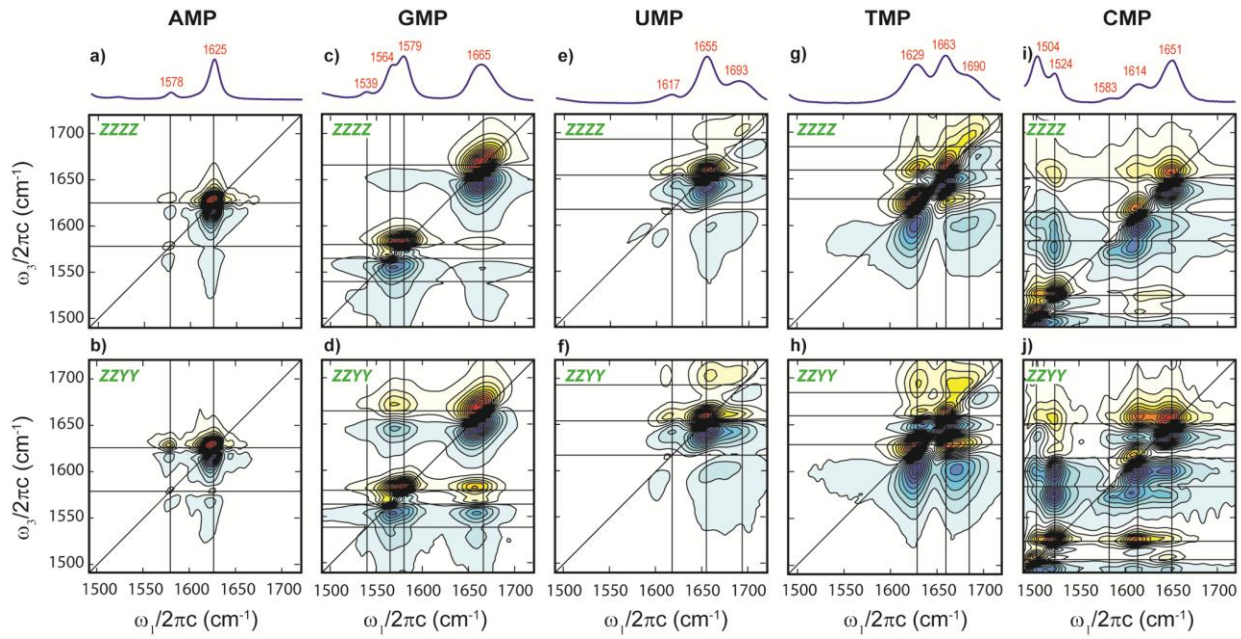


Figure 3. Experimental ZZZZ (top) and ZZYY (bottom) 2D IR spectra of the five NMPs (a,b; AMP; c,d: GMP; e,f: UMP; g,h: TMP; i,j: CMP) at 20mg/ml in PBS buffer at 20^{o}C and pH 6.9. The experimental FTIR spectra have been plotted on top of the corresponding 2D IR spectra for comparison. All contours are plotted in 7.5% intervals.

The parameters extracted from the spectral fitting of the 2D IR spectra of AMP describe the anharmonic potential surface of the two vibrational modes seen in the experiment. Calculated 2D IR spectra using these parameters are shown in Figure 4. The important parameters in describing the potential surfaces such as transition frequencies, diagonal and off-diagonal anharmonicities, as well as the angles between transition dipoles are tabulated in Table 1 and Table 2 for all nucleotides. A full list of the fitting parameters including the lineshape parameters and dipole strengths for higher order transitions is included in the SI. The extracted diagonal anharmonicities for AMP are $\Delta\omega_{A1} = 8.0$ cm^{-1} and $\Delta\omega_{A2} = 9.8$ cm^{-1}. The experimental results can

be compared with the anharmonicities obtained from DFT calculations in the gas phase by Cho and coworkers[27]: $\Delta\omega_{A1} = 7.0$ cm^{-1} and $\Delta\omega_{A2} = 0.8$ cm^{-1}. The off-diagonal anharmonicity ($\Delta\omega_{A1,A2} = 3.7$ cm^{-1}) is significantly less, suggesting that the system can be modeled as two weakly coupled normal modes. The coupling strength between the two oscillators can therefore be calculated perturbatively[43] to be 15 cm^{-1}, compared to the energy splitting of 47 cm^{-1}. From fitting the ZZZZ and ZZYY spectra simultaneously, the projection angle between μ_{A1} and μ_{A2} was determined to be 90° ± 2°.

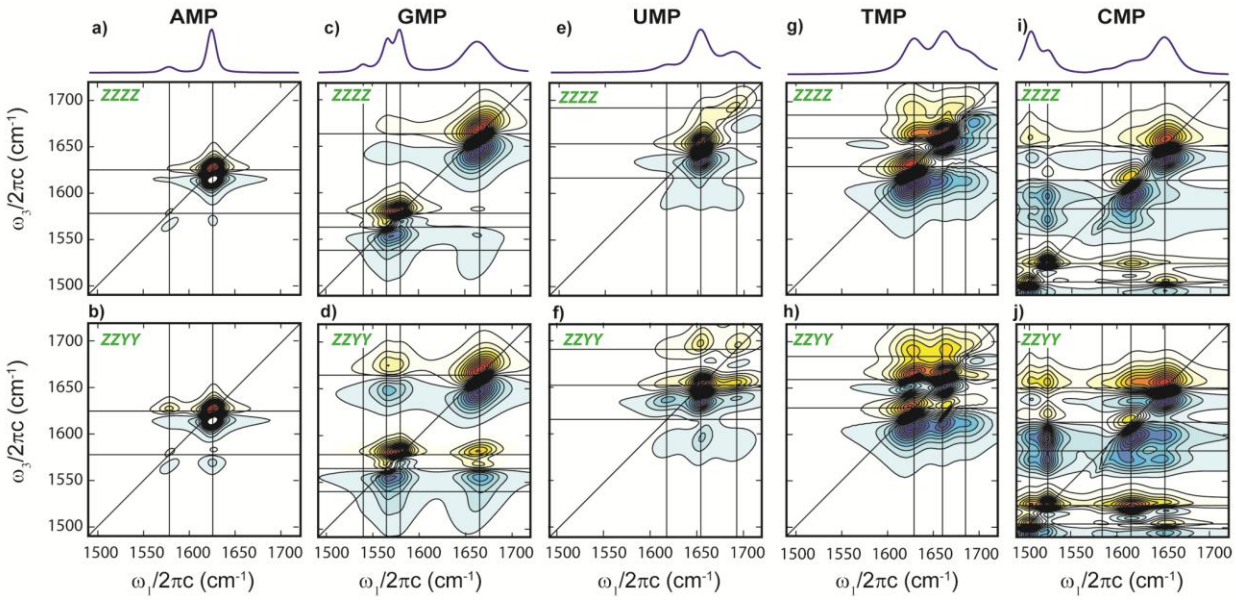


Figure 4. Simulated ZZZZ (top) and ZZYY (bottom) 2D IR spectra of the five NMPs (a,b; AMP; c,d: GMP; e,f: UMP; g,h: TMP; i,j: CMP) from 2D spectral fitting. The fitting parameters are summarized in Table 1, Table 2, and tSI. The simulated FTIR spectra using the lineshape function obtained from the 2D spectral fitting have been plotted on top of the corresponding 2D IR spectra for comparison. All contours are plotted in 7.5% intervals.

DFT calculations on deuterated adenosine in the gas phase observed two normal modes in the spectral region of interest at 1588 cm^{-1} and 1565 cm^{-1}. FTIR stick spectra for all five NMPs simulated from the calculated IR intensities are given in Figure 5. The 1588 cm^{-1} band (A1) is a pyrimidine ring mode consisting mostly of the C^4=C^5, C^5-C^6 out-of-phase stretch. More complete description of the vibration's composition can be found in Table 1 and Section IV in the SI. The 1565 cm^{-1} band (A2) contains the ring vibrations of both the imidazole and pyrimidine rings. Specifically, this highly delocalized ring deformation mode involves the C^4=C^5, C^5-C^6 in-phase stretch. The angle between the two transition dipole moment was found to be 101^o (shown in Figure 6), in good agreement with the experiment. Therefore, the qualitative

agreement of IR intensities, symmetry, and transition dipole angle with the experimental results assure us of the interpretation of these vibrational modes.

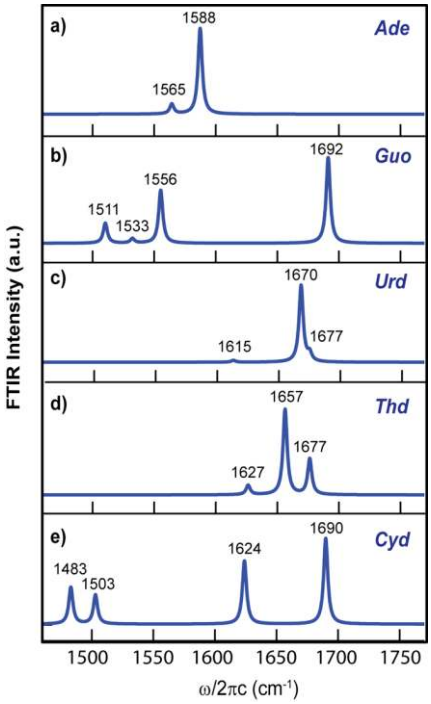


Figure 5. Calculated FTIR spectra using DFT calculated normal mode frequencies and transition dipole strengths of the five deuterated nucleosides: adenosine (a), guanosine (b), uridine (c), thymidine (d), and cytidine (e). The vibrational frequencies have been scaled by 0.9614. All peak lineshapes are generated by convoluting stick spectra with a Lorenztian function with linewidth $\Gamma = 2\ cm^{-1}$.

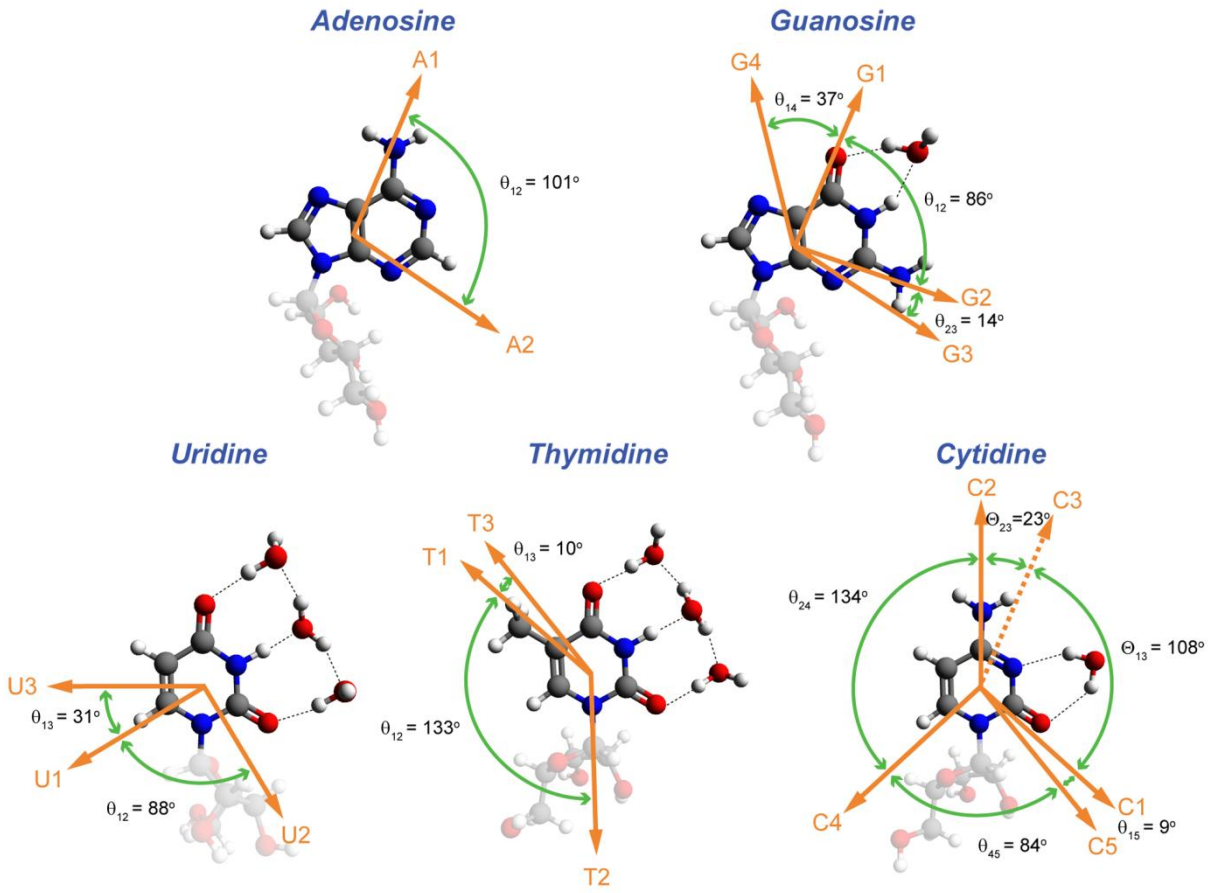


Figure 6. Transition dipole orientations (orange arrows) of the normal modes from DFT calculations on the five deuterated nucleosides. Shown here are the optimized structures with different number of explicit D_2O molecules making hydrogen-bonds to the carbonyl or ND groups on the bases. Specific normal mode displacements are shown in the SI. $\theta's$ define the calculated angles between different transition dipole moments. Only the angles between the nearest transition dipoles are shown, see Table 2 for a complete list of angles. The ribose is made less opaque for clearer viewing of the bases. Mode C3 is not seen in our DFT calculation and therefore its dipole orientation (dotted arrow) is deduced from its relative angles to other modes (Θ_{23}, Θ_{13}) based on 2D spectral fitting (Table 2).

B.2. GMP

Infrared spectra of GMP between 1500 –1800 cm^{-1} display four vibrational resonances. The intense peak at 1663 cm^{-1} (G1) is widely accepted as a local C^6=O stretch.[7,16] Its 2D lineshape is elongated along the diagonal with an unusually broad inhomogeneous linewidth of

33 cm^{-1}. The broadening of G1 is not surprising, given our understanding of how hydrogen bonding to carbonyls influences frequency shifts and line broadening. Notably, the similar broadening of the amide I band of N-methylacetamide in D_2O has been explained as containing both singly and doubly hydrogen-bonded configurations of water to the carbonyl.[44-46] Therefore, given that multiple water sites have been observed from crystal structures of DNA,[47] the inhomogeneity is most likely the result of various hydrogen-bonding configurations around the carbonyl.

FTIR shows a second peak G2 at 1578 cm^{-1}, along with a shoulder peak G3 at 1567 cm^{-1}. The presence of this shoulder peak is clearly identified by the pronounced cross-peaks between G2 and G3. The positive downhill cross-peak destructively interferes with the overtone of the G2 diagonal peak, resulting in a special peak pattern (contrasted to the eight-peak pattern for AMP) for two coupled oscillators that are separated by only a small energy gap. Additionally, the lack of relative cross-peak enhancement from the polarization-selective 2D IR spectra suggests a pair of parallel transition dipoles. On the other hand, the cross-peaks of these two modes to G1 are greatly intensified in the ZZYY spectrum, indicating a perpendicular geometry between G1 and these two vibrations. Finally, a fourth peak G4 at 1539 cm^{-1} that is barely seen in the FTIR, is detected by its cross-peaks to the higher frequency modes in the 2D spectrum. However, due to the low intensity of peak G4, the cross-peak enhancement is difficult to determine and its relative dipole orientation can only be extracted from fitting the spectra. The lineshapes of G2-G4 are homogeneous with diagonal linewidths of ~ 13 cm^{-1}, as in the case of the ring modes in AMP.

Previous assignments for these three peaks have been inconsistent: Tsuboi[4] and Mondragon-Sanchez *et al.*[10] assigned them to be local C^4=C^5 stretch and different C=N stretches; Toyama *et al.* [41] performed isotope labeling experiments to assign G2 and G3 to be pyrimidine

ring vibrations localized to N^3 and C^2, respectively, and G4 to be a delocalized vibration over the entire purine ring; while DFT calculations showed delocalized and coupled vibrations across the purine ring for these modes.[48] By fitting the GMP ZZZZ and ZZYY 2D IR spectra, the diagonal anharmonicities were obtained as: $\Delta\omega_{G1} = 11$ cm^{-1}, $\Delta\omega_{G2} = 4.0$ cm^{-1}, $\Delta\omega_{G3} = 7.8$ cm^{-1}, and $\Delta\omega_{G4} =$ 18 cm^{-1}. Krummel and Zanni measured $\Delta\omega_{G1}$ to be 14 cm^{-1} by fitting the pump-probe spectrum of dGMP.[24] The experimental values are in reasonable agreement with DFT predicted values. Wang and coworkers [23] obtained $\Delta\omega_{G1} = 19$ cm^{-1} for guanine, and Lee and Cho [20] calculated diagonal anharmonicities for C^6=O stretch (16.8 cm^{-1}) and ring vibration (13.0 cm^{-1}) for 9-methylated guanine-d3. The off-diagonal anharmonicities were found to be of nearly equal magnitude to the diagonal anharmonicities.

The relative angles between the four transition dipole moments obtained from fitting, along with comparison to DFT results, are presented in Table 2. The constraint that all four transition dipoles are in-plane was applied to match the six fitted angles into a simple geometry, remembering that solutions to Equation 2 can adopt the value of either Θ or $180° - \Theta$. It was found that μ_{G2} and μ_{G3} are close to being perpendicular with μ_{G1} ($\Theta_{G1,G2} = 89^o$, $\Theta_{G1,G3} = 111^o$), while μ_{G4} is 37^o away from μ_{G1}. The uncertainties in these angles are ~ 10^o.

DFT calculations on deuterated guanosine were performed to get a picture of the vibrations and help interpret these fitted angles. It was found that the number and location of the added water molecules in DFT calculations alters the directions of these transition dipoles, although the bond stretches involved did not change dramatically (shown in the SI). Therefore, care is needed when interpreting the 2D IR results using DFT calculations. When at least one water molecule is hydrogen-bonded to the carbonyl, the agreement between the experimentally determined angles and DFT calculated results is good, even for the projection-angles involving

the low intensity peak G4. Furthermore, from the calculation it was observed that by solvating guanosine with more D_2O molecules, the frequency of G1 redshifts while the frequencies of G2, G3, and G4 blueshift. This anti-correlated frequency shift can be rationalized through the stabilization of resonance structures which lengthen the carbonyl bond while shortening the CN bond, analogous to protein Amide I' and Amide II' modes.[49,50]

The calculated frequencies and intensities can be easily matched to the experimental results, except for the fact that the G3 mode is weaker than the G4 mode. G1 was found to contain mainly C^6=O stretch and its transition dipole aligns roughly with the CO bond, indicating a more local vibration character. G2 has larger amplitude of delocalized pyrimidine ring vibration. G3 mode is a more delocalized ring vibration that extends to the imidazole ring. G4 is also a delocalized vibration across the entire purine ring, with strong C^4=C^5 and N^7=C^8 stretches.

C. 2D IR Spectroscopy of Pyrimidines

C. 1. UMP

UMP has a more congested spectrum compared to those of the purines with three modes close in energy: one intense peak at 1654 cm^{-1} (U2), a medium peak at 1691 cm^{-1} (U1) and a weak peak at 1617 cm^{-1} (U3). Previously, U3 was attributed to the C^5=C^6 stretch where U1 and U2 have been assigned to the local C^2=O and C^4=O, respectively.[4,16] However, 2D IR spectra of UMP show couplings between all three modes, revealing some degree of delocalization. In addition, the distinct intensities and lineshapes of U1 and U2 suggest a more complicated interpretation than a pure local mode scheme. Their anti-diagonal linewidths are 16 $cm^{-1}$and 13 cm^{-1} for U1 and U2, respectively, which fall in the typical range (10-15 cm^{-1}) for homogeneous linewidths of carbonyls in aqueous solution. On the other hand, the diagonal linewidths are 32

cm^{-1} and 17 cm^{-1} for U1 and U2 respectively. As a consequence, U1 is an inhomogeneous peak whereas U2 appears homogeneous. Spectral fitting resulted in a $90^o \pm 7^o$ angle between the two vibrations, which suggests the two modes are close to being the symmetric and asymmetric C=O stretches. In fact, this observation parallels with our DFT calculation on uradine.

DFT calculations showed that U1 is the symmetric C=O stretch mixed with N^3-D bend. U2 was the asymmetric C=O stretch mixed with some C^5=C^6. Finally, U3 was C^5=C^6 stretching in-phase with C^2=O and out-of-phase with C^4=O. The mode composition of these vibrations does not change dramatically with the number of solvating water molecules; however, the vibrational frequencies and the relative dipole orientations do. Specifically, three water molecules were required to be present in the proximity of the C=O's and ND group in order to obtain an agreement with the experimentally measured relative dipole orientations.

We determined the degree of correlation between the spectral broadening of the resonances for UMP. The correlated shift of the transition frequencies $\delta\omega_i$ and $\delta\omega_j$ between two distributions is quantified by the correlation coefficient $\rho_{ij} = \left\langle \delta\omega_i \delta\omega_j \right\rangle \Big/ (\sigma_i \sigma_j)$ where σ_i and σ_j are the corresponding distribution widths.[34] Although the correlations can be expressly addressed in Kubo models of the cross-peak lineshape,[34,51] the correlation coefficient can also be more simply related to the tilt angle of the cross peak nodal line, Ψ, through $\rho_{ij} = \tan(\Psi) = \left(A_R - A_{NR}\right) / \left(A_R + A_{NR}\right)$, where A_R and A_{NR} are the cross-peak amplitudes in rephasing and non-rephasing spectra, respectively.[51,52] High degree of correlation and anti-correlation correspond to ρ_{ij} of +1 and -1, respectively, where no correlation results in $\rho_{ij} = 0$.[53,54] Based on the results shown in Table 3, U3 is well correlated with the other vibrations, whereas U1 and U2 are not. Highly correlated broadenings indicate that the

vibrational couplings are large compared to the disorder,[51] as illustrated by the cross-peak between U2 and U3 (ω_1 = 1617 cm^{-1}, ω_3 = 1654 cm^{-1}), where the tilt of the nodal line is nearly parallel to the diagonal (39^o). This is an indication that both modes have character of the conjugated C^4=O and C^5=C^6 stretches, where weakening/strengthening of one bond is correlated to the other in this conjugated polyene system. This observation again matches the vibrational modes seen in our DFT calculations and other normal mode analysis on uracil.[8,21] However, it contrasts with some *ab initio* studies which argued that U3 is the C^5=C^6 stretch while U1 and U2 are coupled C=O stretches with different participation ratio of the two carbonyls.[18,22]

C. 2. TMP

TMP spectra have three peaks with similar positions to those of UMP; however, the relative intensities change dramatically upon the methyl insertion, indicating that the TMP modes are significantly different from the UMP modes. In particular, the 1628 cm^{-1} peak (T3) is much more intense than U3, indicating a lower symmetry of the vibrational mode involving C^5=C^6 stretch. Similar to UMP, prior studies used the local mode picture and assigned the 1690 cm^{-1} peak (T1) to C^2=O stretch, 1661 cm^{-1} peak (T2) to C^4=O, and T3 to C^5=C^6 stretch.[4,8,16-18] However, Zhang *et al*. suggested that these peaks may be due to mixed vibrations.[55]

2D IR spectra of TMP reveal couplings between all three modes (≤6 cm^{-1}) that are less than the diagonal anharmonicities, which is indicative of relatively weak couplings between the modes . Spectral fitting reveals that the transition dipole moments between the carbonyl transitions T1 and T2 are 117^o (+8/–7) away from each other, which is also suggestive of a local mode assignment for the carbonyl bands. This can be compared to an angle of 133^o obtained from our DFT calculation on thymidine.

Interestingly, the TMP lineshapes are very different from those of UMP, most notably the broad overtone peak of T3 compared to U3. Although T1 is still inhomogeneous and T2 is homogeneous, their anti-diagonal linewidths are larger than those of U1 and U2 (Figure S2 in the SI). The correlation coefficients of the three pairs of TMP cross-peaks were calculated (Table 3). The result is similar to UMP in that T1 and T2 are uncorrelated while T2 and T3 are highly correlated. However, U1 and U3 are more correlated compared to T1 and T3. This effect may be explained by the insertion of the methyl group that results in the loss of C^5-H bend observed in U1, U2 and U3 modes.

Just as in our DFT calculations on uridine, three water molecules hydrogen bonded to thymidine's C=O's and ND group were needed to reach a reasonable agreement with 2D IR results. The calculations reveal that the methyl substitution decreases the degree of coupling of the two carbonyl stretches. T1 was found to be mostly a C^2=O stretch, accompanied by a small magnitude of in-phase vibration of C^4=O, with its transition dipole aligned nearly parallel to the C^2=O bond. T2 has contributions from in-phase stretching of C^4=O and C^5=C^6. Finally, T3 is mostly C^5=C^6 stretch mixed with a little C=O stretch. The normal modes and angles from the DFT calculations match the experimental results reasonably well, and suggest that the coupled vibrations in TMP are less delocalized than the ones in UMP, due to the methyl substitution.

C. 3. CMP

The spectra of CMP display five peaks in the spectral region of interest. Structurally, CMP has one C=O bond, and only one high frequency vibrational mode attributable to a CO stretch is observed at 1651 cm^{-1} (C1),[7] From the 2D IR spectra, one can see that C1 is homogeneous, in sharp contrast to the inhomogeneous lineshapes of U1 and T1. There are two

other homogeneous peaks that are close in energy at 1614 cm^{-1} (C2) and 1583 cm^{-1} (C3) with much lower intensities. Purrello and coworkers[56] assigned C2 in their UVRR spectrum of dCMP in D_2O to the C^5=C^6 stretch. Miles[57] observed that C2 and C3 disappear upon methylation on the N^3 position and dimethylation on the N^4 position, respectively. This study indicates that C2 is mixed with N^3=C^4 while C3 involves motions of the ND_2 group. The more intense cross-peaks between modes C1/C2 than C1/C3 in our 2D spectra support this peak assignment because C^2=O, N^3=C^4, and C^5=C^6 form a conjugated polyene system. Based on the enhancement of the relative amplitudes of these cross-peaks between C1, C2, and C3, the transition dipole of C1 should deviate from those of the nearly parallel C2 and C3.

Additionally, there are two more modes at 1524 cm^{-1} (C4) and 1504 cm^{-1} (C5) which also have been assigned to in-plane ring vibrations.[4] These two peaks are much more intense than C2 and C3, which can be explained by the involvement of the exocyclic ND_2 that breaks the symmetry of these ring vibrations. Interestingly, 2D IR spectra reveal that these two ring modes are coupled to the higher frequency modes C1-C3, which explain why the observed frequencies are much lower than expected for a standard C=N. In order to reduce the number of free parameters for fitting the CMP 2D spectra, the diagonal peaks C4 and C5, as well as their influence on associated cross-peaks, were set to be purely homogeneous. The anharmonicity of C1 is relatively small compared to those of C2 and C3, but comparable to those of C4 and C5. A value of 9.2 cm^{-1} for $\Delta\omega_{C1}$ is consistent with the previously reported result of 9 cm^{-1}.[24]

From our DFT calculations of harmonic vibrational frequencies on deuterated cytidine, only four vibrations were observed in the spectral window of interest, rather than the five seen experimentally. Previous DFT calculations on single-stranded poly-C also reported only one medium intensity peak in the 1616-1624 cm^{-1} region, which was assigned to an in-plane ring

vibration.[4] Since the presence of C1-C5 is experimentally established, we believe that one of the experimental peaks may be a combination band or Fermi resonance. A DFT calculated transition at 1690 cm^{-1}, corresponding to the C1 mode, has dominant C^2=O character. A second peak at 1624 cm^{-1} involves the enimine bond N^3=C^4–C^5=C^6 stretches and C–H bend. Since these contributions match our description of the C2 mode based on experimental results, we assign this peak to C2 and conclude that the DFT calculation does not find the C3 mode which requires significant ND_2 motions. Gavira and coworkers [58] observed a medium band at 1616 cm^{-1} and a very weak peak at 1582 cm^{-1} in the FTIR spectrum of cytidine 3'-monophosphate (3'-CMP) and assigned them to be the ribose C-H vibrations and in-plane base vibrations, respectively. However, this interpretation does not explain the FTIR of cytosine (data not shown) where both peaks are present. The fact that the C3 band is barely visible in the FTIR is also suggestive of a combination band.

It was found that when cytidine makes a hydrogen bond to a water molecule through the carbonyl, the relative transition dipole orientations match well with the results obtained from the 2D spectral fitting. Although the C3 mode is not described by the harmonic mode DFT calculation, it was determined from the fitting that the transition dipoles of C2 and C3 are 23^o apart and its direction is further constrained by its angles relative to other modes in a planar geometry (Figure 6).

IV. Discussion

A. Vibrational couplings and mode character

In our characterization of the IR active in-plane double bond stretching vibrations of nucleotides, we see that the 2D IR spectra and corresponding vibrational anharmonic modes are

distinct, even with the structural similarities between the purine and pyrimadine bases. Except perhaps for the case of carbonyls, there are no simple or intuitive structural correlations to the spectral features, for instance with respect to conjugation, resonance structures, and ring placement of exocyclic groups. The number of resonances between 1650 cm^{-1}–1700 cm^{-1} matches the number of carbonyls in the base, and these are inhomogeneously broadened, reflecting the sensitivity of the carbonyl frequency to hydrogen bonding to the oxygen. Comparison of the experimental results to the harmonic modes derived from DFT calculations of explicitly hydrated nucleosides is generally favorable. For the most part, significant anharmonic couplings within the bases lead to vibrational modes that are delocalized over large portions of the base.

Comparing the two purines only the G1 peak that is largely due to the C^6=O vibration is readily assigned. Due to the large energy splitting between G1 and the ring modes (> 43 cm^{-1}) and the small off-diagonal anharmonicities (< 5 cm^{-1}), it is reasonable to describe G1 as a local mode C=O stretch. The frequencies and amplitudes for the ring vibrations are difficult to predict and depend on where the exocyclic ND_2 group resides. These lower frequency (1500 cm^{-1} – 1600 cm^{-1}) ring modes in AMP, GMP, and CMP originate from the one- or two-member ring deformation coupled to the small amplitude C-ND_2 bend vibrations. It should be noted that H/D exchange redshifts the NH_2 scissorings from the 1500 cm^{-1} – 1600 cm^{-1} region to ~ 1100 cm^{-1}.

The assignment of the carbonyl double bond stretches in UMP and TMP is more complicated as the two C=O groups are more strongly coupled. Traditionally, C=O vibrations are assigned as local modes with C^2=O corresponding to U1/T1 and C^4=O to U2/T2. However, the 2D IR spectroscopy indicates that observed couplings between the carbonyl stretches are not negligible, especially for UMP. UMP has diagonal anharmonicities greater than 15 cm^{-1}, and the

off-diagonal anharmonicities vary from 8.0 cm^{-1} (U1/U2) to 25 cm^{-1} (U1/U3). The coupling constant between carbonyls in the local mode picture can be calculated perturbatively[43] to be $2V$ = 26 cm^{-1}. This coupling strength is comparable to the frequency splitting of ΔE =38 cm^{-1}. Therefore, the system is not weakly coupled, and in fact intermediate between the strong ($|2V| \gg |\Delta E|$) and weak coupling ($|2V| \ll |\Delta E|$) regimes. This finding along with the projection angle of $90^o \pm 7^o$ from the 2D spectral fitting match the results from the DFT calculations where symmetric and asymmetric C=O stretches are found. Further, the disorder (diagonal linewidth) for the carbonyls is ~30 cm^{-1}, indicating that significant variation in the mode mixing and projection angle may exist within the ensemble.

On the other hand, TMP has much smaller anharmonicities relative to UMP. With a value of $2V$ = 12 cm^{-1} between T1 and T2, compared to the 27 cm^{-1} energy splitting, a weak coupling description of the two carbonyl stretches is more appropriate. This is in accordance with an angle of $117^o \pm 2^o$ between T1 and T2, indicating a more localized C^2=O stretch for T1. However, $2V \approx 34$ cm^{-1} was obtained between T2 and T3 (ΔE = 38 cm^{-1}), showing significant mixing for the C^4=O stretch. A similar effect of couplings between the carbonyl and ring modes was also observed for CMP vibrations, indicated by the pronounced cross-peaks and calculations based on perturbation theory ($2V \approx 28$ cm^{-1}; ΔE = 37 cm^{-1}). Furthermore, the order of the frequencies of the ring vibrations involving C=C in these three pyrimidines ($\omega_{C2} < \omega_{U3} < \omega_{T3}$) is likely to be correlated with the degree of conjugation of the C=C bond in the pyrimidine ring. Therefore we conclude that the underlying in-plane base motions in UMP and TMP cannot be attributed to simple local C=O and C=C vibrations.

B. Experimental and Theoretical Limitations

By independently controlling the excitation and detection polarizations in 2D IR, one can measure the cross-peak anisotropy from which the relative dipole orientations can be obtained [34,38,59]. This is the most sensitive probe of the vibrational potential, since the relative angle between the transition dipole moments is unusually sensitive to the directionality of the bonds in the molecule, the couplings and delocalization of the vibration, and the solvent environment. In the simple case where the peaks are well-separated, the relative peak amplitude of peaks in parallel and perpendicular spectra can be obtained directly from peak heights or integrated peak volume. However, in most cases, including the spectra presented here, spectral overlap between resonances leads to interferences that greatly impair the ability to directly extract angles from spectral amplitudes. Although 2D IR spectra are linearly additive in the components, the sum over positive and negative overlapping features means that quantitative analysis of amplitudes, and therefore angles, is unreliable without knowledge of the spectral components. In the present case, it was found that angle determination based on peak volume analysis was only reliable when applied to well-isolated cross-peaks such as the pairs A1/A2, G1/G2, and G1/G3. In the case of overlapping vibrational bands, it was found that the accurate dipole projection angles obtained from spectral fitting were as much as 40° off the angles obtained from volume integration. Strategies that may help simplify such problems include the polarization-angle-scanning (PAS) 2D IR spectroscopy developed by Lee and coworkers.[60]

Our measured transition dipole orientations compare favorably with DFT calculations on the five nucleosides provided that we include explicit water molecules to hydrogen bonding sites on the base, particularly to carbonyls. Hydrogen-bonding interactions between the water and

solute strongly influence the vibrational modes, and thereby limit the ability to assign vibrations and relate them to molecular structure based solely on unsolvated DFT calculations. We find that in nucleotides, the hydrogen bonding to the water molecules can lead to vibrational frequency shifts as much as -60 cm^{-1} for carbonyl stretches, but only +10 cm^{-1} for the ring modes. However, the relative dipole orientations of the ring modes are more sensitive (dipole moments can rotate as much as 70^{o}) to the solvent environment compared to the carbonyl stretches that have more local characters such as in the case of GMP and CMP. These issues take on new meaning when considering the dynamics and disorder of the system, which can lead to sharp variations in transition dipole direction. Although *ab initio* DFT calculations are an important component of the present work, it is the combination of calculations with the 2D IR experimental constraints that lead to accurate spectral assignments and interpretation.

C. Toward quantitative models of nucleic acid secondary structure.

Although the IR spectroscopy of DNA oligomers in the region of base vibrations is sensitive to the secondary structures of DNA[8], the underlying physics required to link the molecular structure and the IR spectral features is not well understood. Current use of FTIR for investigating DNA structure is largely based on empirical findings on marker bands. For instance, it is found that G1 mode observed for free GMP at 1663 cm^{-1} is blue-shifted to 1689 cm^{-1} in Watson-Crick (WC) G-C base-pairs [61] or 1672 cm^{-1} when GMP is assembled into G-quadruplexes stabilized via Hoogsteen G-G interactions.[10] Also, upon WC duplex formation, it is also observed that G2 and G3 modes decrease in intensity.[4] A more detailed understanding of how inter-molecular interactions influence the base vibrations is needed to move beyond qualitative analysis of IR spectra.

The characterization of the individual base vibrations forms a starting point for developing structure-based spectroscopic models of nucleic acid secondary structure, and for interpreting changes in frequency, intensity, and lineshape of the base vibrations in these structures. Studies of the protein backbone amide I vibration using 2D IR spectroscopy have offered increased structural resolution and have opened up an avenue for detailed probing protein dynamics largely as a result of excitonic models that describe the coupling between local mode peptide carbonyls.[44,62] In the case of DNA and RNA, structure based models will likely take a different form, given the delocalized nature of the base vibrations and the variety of base-base couplings available. However, intermolecular interactions can be modeled as perturbations to the anharmonic vibrational modes of single bases. In the case of the CO vibrations of GC oligonucleotides, Krummel and Zanni[16] modeled the 2D IR spectra found that C=O vibrations of double helical GC pairs interact through electrostatics and base stacking, leading to delocalized modes across the base-pairs and along the helices. They also recognized that the coupling between the carbonyl and the ring modes are important for determining the inter-base coupling and the transition dipole directions.

V. Concluding Remarks

2D IR spectroscopy provides clear evidence that the in-plane ring stretching vibrations of nucleic acid bases between 1500 –1700 cm^{-1} are significantly delocalized over the purine and pyrimadine rings. Although this has been recognized in prior *ab initio* calculations, the traditional assignment and interpretation of DNA vibrations was typically based on a local mode description of particular double bond stretches. Such assignments remain reasonable for C=O vibrations of G and C. Modeling these spectra also provides a quantitative understanding of the

anharmonic vibrational potential coupling the ring modes, and the relative orientation of transition dipole moments. Such work forms the basis for constructing models for the vibrational spectroscopy of DNA and RNA secondary structures and protein-DNA interactions, in which inter-base interactions modify the frequency, intensity, mode composition, and dipole orientations of the single-base vibrations.

Acknowledgement. We thank Kim Hamad-Schifferli for providing us with a problem that motivated our first steps into the area of DNA vibrational spectroscopy. We thank Ziad Ganim for helpful discussion on 2D spectral fitting, and Andrew Horning, Carlos Baiz and Rebecca Nicodemus for assistance in DFT calculations. This work was supported by a grant from the National Science Foundation (CHE-0911107).

Supporting Information Available: Details on error analysis of the angles between transition dipoles, lineshape analysis of the experimental 2D IR spectra of the five nucleotides, detailed description of the model used to fit experimental 2D IR spectra, a full list of 2D spectra fitting parameters, snapshots of the vibrational modes investigated in this study from DFT calculations, snapshots demonstrating the influence of solvating water molecules on DFT calculations, and complete Ref. 39. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Liedl, T.; Hogberg, B.; Tytell, J.; Ingber, D. E.; Shih, W. M. *Nat Nanotechnol* **2010**, *5*, 520.

(2) Zheng, J.; Birktoft, J. J.; Chen, Y.; Wang, T.; Sha, R.; Constantinou, P. E.; Ginell, S. L.; Mao, C.; Seeman, N. C. *Nature* **2009**, *461*, 74.

(3) Liu, A.; Jones, R.; Liao, L.; Samara-Rubio, D.; Rubin, D.; Cohen, O.; Nicolaescu, R.; Paniccia, M. *Nature* **2004**, *427*, 615.

(4) Tsuboi, M. In *Applied Spectroscopy Reviews*; Brame, E. G. J., Dekker, M., Eds. New York, 1969; Vol. I, p 45.

(5) Tsuboi, M.; Benevides, J. M.; Thomas, G. J. *Proc. Jpn. Acad., Ser. B* **2009**, *85*, 83.

(6) Taillandier, E.; Liquier, J. In *Methods in Enzymology*; Lilley, D. M. J., Dahlberg, J. E., Eds.; Academic Press: 1992; Vol. 211, p 307.

(7) Banyay, M.; Sarkar, M.; Gräslund, A. *Biophys. Chem.* **2003**, *104*, 477.

(8) Liquier, J.; Taillandier, E. In *Infrared Spectroscopy of Biomolecules*; Mantsch, H. H., Chapman, D., Eds.; Wiley-Liss: 1996.

(9) Lindqvist, M.; Sarkar, M.; Winqvist, A.; Rozners, E.; Strömberg, R.; Gräslund, A. *Biochemistry* **2000**, *39*, 1693.

(10) Mondragon-Sanchez, J. A.; Liquier, J.; Shafer, R. H.; Taillandier, E. *J. Biomol. Struct. Dyn.* **2004**, *22*, 365.

(11) Urpi, L.; Ridoux, J. P.; Liquier, J.; Verdaguer, N.; Fita, I.; Subirana, J. A.; Iglesias, F.; Huynh-Dinh, T.; Igolen, J.; Taillandier, E. *Nucleic Acids Res.* **1989**, *17*, 6669.

(12) Mohammadi, S.; Klement, R.; Shchyolkina, A. K.; Liquier, J.; Jovin, T. M.; Taillandier, E. *Biochemistry* **1998**, *37*, 16529.

(13) Gousset, H.; Liquier, J.; Taillandier, E.; Sanghvi, Y.; Peoch, D. *J. Biomol. Struct. Dyn.* **1998**, *15*, 931.

(14) Abdelkafi, M.; Leulliot, N.; Baumruk, V.; Bednárová, L.; Turpin, P. Y.; Namane, A.; Gouyette, C.; Huynh-Dinh, T.; Ghomi, M. *Biochemistry* **1998**, *37*, 7878.

(15) Stancik, A. L.; Brauns, E. B. *Biochemistry* **2008**, *47*, 10834.

(16) Fodor, S. P. A.; Rava, R. P.; Hays, T. R.; Spiro, T. G. *J. Am. Chem. Soc.* **1985**, *107*, 1520.

(17) Dagneaux, C.; Liquier, J.; Taillandier, E. *Biochemistry* **1995**, *34*, 14815.

(18) Shanmugasundaram, M.; Puranik, M. *J. Raman Spectrosc.* **2009**, *40*, 1726.

(19) Lee, C.; Park, K. H.; Cho, M. *J. Chem. Phys.* **2006**, *125*.

(20) Lee, C.; Cho, M. *J. Chem. Phys.* **2006**, *125*.

(21) Nishimura, Y.; Tsuboi, M.; Kato, S.; Morokuma, K. *J. Am. Chem. Soc.* **1981**, *103*, 1354.

(22) Peticolas, W. L.; Rush, T. *J. Comput. Chem.* **1995**, *16*, 1261.

(23) Wang, G.-x.; Ma, X.-y.; Wang, J.-p. *Chin. J. Chem. Phys.* **2009**, *22*, 563.

(24) Krummel, A. T.; Mukherjee, P.; Zanni, M. T. *J. Phys. Chem. B* **2003**, *107*, 9165.

(25) Krummel, A. T.; Zanni, M. T. *J. Phys. Chem. B* **2006**, *110*, 13991.

(26) Lee, C.; Cho, M. *J. Chem. Phys.* **2007**, *126*, 145102.

(27) Lee, C.; Park, K. H.; Kim, J. A.; Hahn, S.; Cho, M. *J. Chem. Phys.* **2006**, *125*.

(28) Szyc, L.; Yang, M.; Nibbering, E. T. J.; Elsaesser, T. *Angew. Chem., Int. Ed.* **2010**, *49*, 3598.

(29) Yang, M.; Szyc, L.; Elsaesser, T. *J. Phys. Chem. B* **2011**, *115*, 1262.

(30) Yang, M.; Szyc, Ł.; Röttger, K.; Fidder, H.; Nibbering, E. T. J.; Elsaesser, T.; Temps, F. *J. Phys. Chem. B* **2011**, *115*, 5484.

(31) Chung, H. S.; Khalil, M.; Smith, A. W.; Tokmakoff, A. *Rev. Sci. Instrum.* **2007**, *78*, 063101.

(32) Jones, K. C.; Ganim, Z.; Tokmakoff, A. *J. Phys. Chem. A* **2009**, *113*, 14060.

(33) Sung, J.; Silbey, R. J. *J. Chem. Phys.* **2001**, *115*, 9266.

(34) Khalil, M.; Demirdöven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258.
(35) Schmidt, J. R.; Sundlass, N.; Skinner, J. L. *Chem. Phys. Lett.* **2003**, *378*, 559.
(36) Woutersen, S.; Pfister, R.; Hamm, P.; Mu, Y.; Kosov, D. S.; Stock, G. *J. Chem. Phys.* **2002**, *117*, 6833.
(37) Golonzka, O.; Khalil, M.; Demirdöven, N.; Tokmakoff, A. *J. Chem. Phys.* **2001**, *115*, 10814.
(38) Golonzka, O.; Tokmakoff, A. *J. Chem. Phys.* **2001**, *115*, 297.
(39) Frisch, M. J.; et al. Gaussian, Inc.: Pittsburgh PA, 1998.
(40) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.
(41) Toyama, A.; Hanada, N.; Ono, J.; Yoshimitsu, E.; Takeuchi, H. *J. Raman Spectrosc.* **1999**, *30*, 623.
(42) Khalil, M.; Tokmakoff, A. *Chem. Phys.* **2001**, *266*, 213.
(43) Hamm, P.; Lim, M.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2036.
(44) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123.
(45) DeCamp, M. F.; DeFlores, L.; McCracken, J. M.; Tokmakoff, A.; Kwac, K.; Cho, M. *J. Phys. Chem. B* **2005**, *109*, 11016.
(46) Ham, S.; Kim, J.-H.; Lee, H.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 3491.
(47) Schneider, B.; Cohen, D.; Berman, H. M. *Biopolymers* **1992**, *32*, 725.
(48) Tsuboi, M.; Takahasi, S.; Harada, I. In *Physico-Chemical Properties of Nucleic Acids*; Duchesne, J., Ed.; Academic Press Inc. : London, 1973, p 91.
(49) DeFlores, L. P.; Ganim, Z.; Ackley, S. F.; Chung, H. S.; Tokmakoff, A. *J. Phys. Chem. B* **2006**, *110*, 18973.
(50) DeFlores, L. P.; Ganim, Z.; Nicodemus, R. A.; Tokmakoff, A. *J. Am. Chem. Soc.* **2009**, *131*, 3385.
(51) Demirdoven, N.; Khalil, M.; Tokmakoff, A. *Phys. Rev. Lett.* **2002**, *89*, 237401.
(52) Roberts, S. T.; Loparo, J. J.; Tokmakoff, A. *J. Chem. Phys.* **2006**, *125*, 084502.
(53) Ge, N. H.; Zanni, M. T.; Hochstrasser, R. M. *J. Phys. Chem. A* **2002**, *106*, 962.
(54) Demirdöven, N.; Khalil, M.; Golonzka, O.; Tokmakoff, A. *J. Phys. Chem. A* **2001**, *105*, 8025.
(55) Zhang, S. L.; Michaelian, K. H.; Loppnow, G. R. *J. Phys. Chem. A* **1998**, *102*, 461.
(56) Purrello, R.; Molina, M.; Wang, Y.; Smulevich, G.; Fresco, J. R.; Spiro, T. G.; Fossella, J. *J. Am. Chem. Soc.* **1993**, *115*, 760.
(57) Miles, H. T. *Proc. Natl. Acad. Sci. U.S.A.* **1961**, *47*, 791.
(58) Gavira, J. M.; Campos, M.; Diaz, G.; Hernanz, A.; Navarro, R. *Vibrational Spectroscopy* **1997**, *15*, 1.
(59) Hochstrasser, R. M. *Chem. Phys.* **2001**, *266*, 273.
(60) Lee, K. K.; Park, K. H.; Park, S.; Jeon, S. J.; Cho, M. *J. Phys. Chem. B* **2011**, *115*, 5456.
(61) Miles, H. T.; Frazier, J. *Biochem. Biophys. Res. Commun.* **1972**, *49*, 199.
(62) Ganim, Z.; Chung, H. S.; Smith, A. W.; Deflores, L. P.; Jones, K. C.; Tokmakoff, A. *Accounts Chem. Res.* **2008**, *41*, 432.

TABLE 1: Characterization for the five deuterated NMPs in D_2O in experiments and DFT calculations. Fundamental transition energies, ω_i, diagonal anharmonicities ($\Delta\omega_{i,j}$, cm^{-1}) and the 0-1 transition dipole moments $\mu^{1,0}$.

Nucleotide	Peak	ω_i (cm^{-1}) (Fit)	$\Delta\omega_{i,i}$ (cm^{-1}) (Fit)	$\mu^{1,0}$ (Fit)[a]	ω (cm^{-1}) (Calc'd)[b]	Vibrations [c]
AMP	A1	1625	8.0	1.0	1588	ν (C^4=C^5, C^5-C^6 out-of-phase), δ (C^2-H), δ (N^6D_2), *Py*
	A2	1578	9.8	0.41	1565	ν (C^4=C^5, C^5-C^6 in-phase), ν (N^1-C^6), ν (N^3-C^4), ν (N^7=C^8), δ (C^8-H), *Py* + *Im*
GMP	G1	1665	11	1.0	1692	ν (C^6=O), δ (N^1-D), *Py*
	G2	1579	4.0	0.63	1556	ν (C^2=N^3), ν (C^6=O), δ (N^2-H), δ (N^2D_2), *Py*
	G3	1565	7.8	0.62	1533	ν (C^2=N^3-C^4=C^5), δ (C^8-H), *Py* + *Im*
	G4	1539	18	0.29	1511	ν (C^4=C^5), ν (N^7=C^8), δ (C^8-H), *Py* + *Im*
UMP	U1	1693	16	0.76	1681	ν (sym C=O), ν (C^5=C^6), δ (N^3-D), δ (C^5-H)
	U2	1655	18	1.0	1670	ν (asym C=O), ν (C^5=C^6), δ (C^5-H), δ (C^6-H), δ (N^3-D)
	U3	1617	25	0.46	1615	ν (C^5=C^6), ν (asym C=O), δ (C^5-H), δ (C^6-H), δ (N^3-D)
TMP	T1	1690	11	0.74	1677	ν (C^2=O), ν (C^4=O), δ (N^3-D)
	T2	1663	5.0	1.0	1657	ν (C^4=O), ν (C^5=C^6), δ (N^3-D), δ (C^5H_3), δ (C^6-H)
	T3	1629	10	0.94	1627	ν (C^5=C^6), ν (C^4=O), δ (C^5H_3) , δ (C^6-H)
CMP	C1	1651	9.2	1.0	1690	ν (C^2=O), ν (N^1=C^6), δ (C^6-H)
	C2	1614	15	0.55	1624	ν (N^3=C^4-C^5=C^6), ν (C^2=O), δ (C^5-H), δ (C^6-H)
	C3	1583	17	0.21	--	
	C4	1524	6.1	0.57	1503	ν (C^4-C^5), ν (N^1-C^6), δ (C^5-H), δ (C^6-H), δ (N^4D_2)
	C5	1504	8.9	0.94	1483	ν (N^3=C^4-N^4), ν (C^5=C^6), δ (C^5-H), δ (C^6-H), δ (N^4D_2)

[a] The dipole strengths reported here are from fitting the FTIR spectra using lineshape parameters from the 2D fitting. Peak amplitudes directly from 2D fitting are reported in the SI.
[b] These vibrational frequencies obtained from DFT calculations have been scaled by 0.9614.
[c] Listed here are the decomposed vibrations with larger amplitudes obtained from the DFT calculations. Snapshots of these vibrations are shown in the SI. ν : stretching; δ : bending; *Py*: pyrimidine ring vibration; *Im*: imidazole ring vibration.

TABLE 2: Couplings between base vibrations for the five deuterated NMPs in D_2O. Off-diagonal anharmonicities ($\Delta\omega_{i,j}$, cm^{-1}). $\Theta's$ are projection angles obtained from fitting the experimental 2D IR spectra, and $\theta's$ are the angles obtained from DFT calculations.

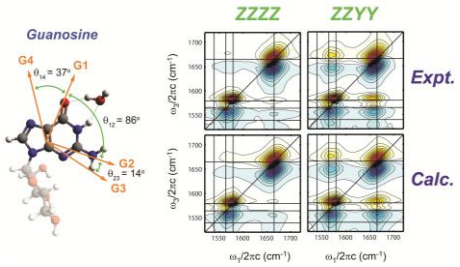
Nucleotides	Coupled Modes	$\Delta\omega_{i,j}$ (cm^{-1}) (Fit)	Θ (deg) (Fit)	θ (deg) (Calc'd)
AMP	A1 & A2	3.7	90 (±2)	101
GMP	G1 & G2	1.3	89 (+1/-10)	86
	G1 & G3	4.6	111 (+11/-12)	99
	G1 & G4	3.3	37 (+16/-37)	37
	G2 & G3	7.5	21 (±1)	14
	G2 & G4	7.0	121 (+8/-7)	121
	G3 & G4	6.8	134 (+11/-8)	134
UMP	U1 & U2	8.0	90 (±7)	88
	U1 & U3	25	35 (+5/-6)	31
	U2 & U3	15	100 (+1/-2)	101
TMP	T1 & T2	1.7	117 (±2)	133
	T1 & T3	4.9	34 (±3)	10
	T2 & T3	6.2	144 (±1)	135
CMP	C1 & C2	7.2	121 (±1)	132
	C1 & C3	7.4	108 (±2)	--
	C1 & C4	1.6	93 (+4/-3)	94
	C1 & C5	5.2	16 (+6/-16)	9
	C2 & C3	12	23 (+7/-13)	--
	C2 & C4	14	134 (±1)	134
	C2 & C5	7.6	144 (±2)[a]	142
	C3 & C4	13	152 (+5/-4)	--
	C3 & C5	8.2	124 (±2)[a]	--
	C4 & C5	1.2	71 (±1)	84

[a] The error bars for these cross-angles were determined from the negative cross-peaks because their positive cross-peaks are interfered too heavily by nearby peaks to be clearly identified. The equation relating the relative cross-peak amplitude to the dipole angle is shown in the SI.

TABLE 3: Correlation coefficients, ρ, and slope of the cross-peak nodal line, Ψ, for UMP and TMP. ρ varies from 1 for correlated to -1 for anticorrelated broadening, whereas Ψ varies from 45° to -45° over the same range.

UMP	ρ	Ψ (deg)	**TMP**	ρ	Ψ (deg)
U1 & U2	0.20	11	T1 & T2	0.29	16
U1 & U3	0.65	33	T1 & T3	0.28	15
U2 & U3	0.81	39	T2 & T3	0.98	44

TOC



Supporting Information

Anharmonic Vibrational Modes of Nucleic Acid Bases Revealed by 2D IR Spectroscopy

Chunte Sam Peng, Kevin C. Jones, and Andrei Tokmakoff*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139

*Corresponding Author. Telephone: 617-253-4503, Fax: 617-253-7030, tokmakof@mit.edu

Contents

I. Error analysis on the angles between transition dipoles

The quoted angles were the results from fitting experimental ZZZZ and ZZYY spectra simultaneously. In order to reflect the experimental S/N ratio in the error bars, we first determined the relative amplitude (α) of the cross-peaks in the ZZYY and ZZZZ spectra following Equation (2): $\cos\Theta = \sqrt{(7-6\alpha)/(1+2\alpha)}$. We then calculated the lower and upper bounds of α , from which the errors bars on the angle were determined. The errors on α was obtained through error propagation of the experimental S/N ratio with the following relation:

$$d\alpha = \alpha\sqrt{\left(\frac{dA_{\perp}}{A_{\perp}}\right)^2 + \left(\frac{dA_{\parallel}}{A_{\parallel}}\right)^2} .$$

Because Equation (2) is not a linear relation between Θ and α , for a given amplitude error, the error will be bigger and asymmetric if the angle is closer to 0^o or 180^o (see C2/C3 cross-angle for example). Initially we employed different methods to extract the cross angles directly from experimental 2D IR spectra (Figure S1), however, due to the overlapping of various peaks, these techniques resulted in angles that were as much as 40^o away from the angles accurately determined from fitting the 2D spectra.

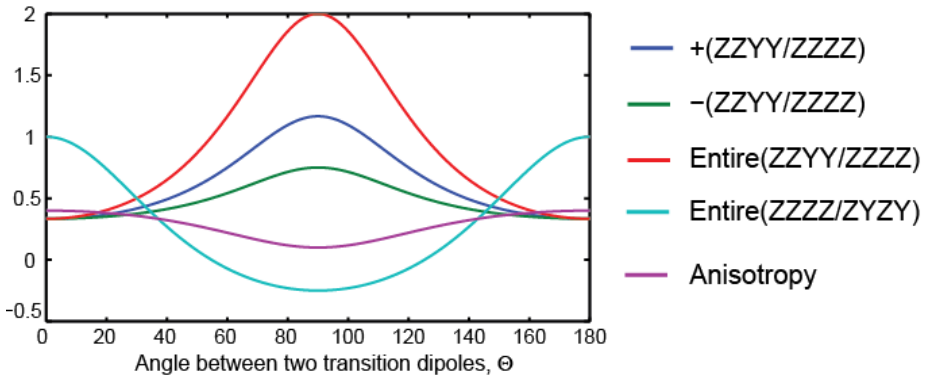


Figure S1. Cross-peak amplitude ratios as a function of the projection angle between two coupled vibrations. Five different methods are shown here: (blue) amplitude ratio A_{zzyy}/A_{zzzz} of only the absorptive cross-peak; (green) amplitude ratio A_{zzyy}/A_{zzzz} of only the emissive cross-peak; (red) integrated volume ratio A_{zzyy}/A_{zzzz} of the entire cross-peak; (cyan) integrated volume ratio A_{zzzz}/A_{zyzy} of the entire cross-peak; (purple) anisotropy determined by $(A_{zzzz}-A_{zzyy})/(A_{zzzz}+2A_{zzyy})$.

II. Lineshape analysis of the experimental 2D IR spectra of the five nucleotides

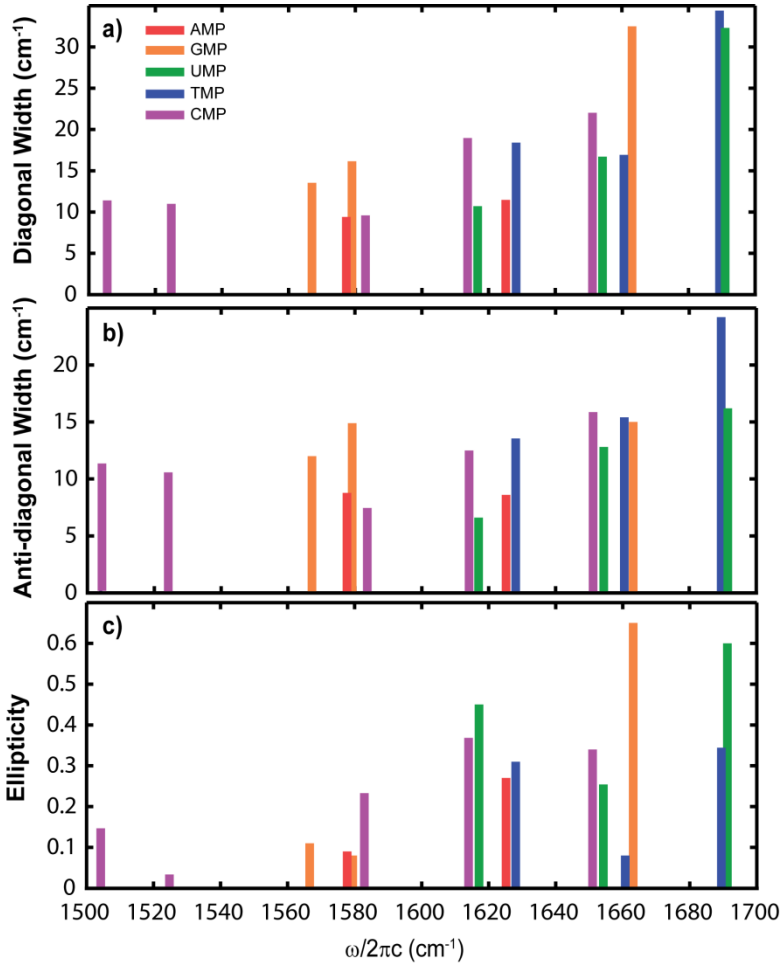


Figure S2. Lineshape analysis of the experimental ZZZZ 2D IR spectra of the five NMP's shown in Figure 3. (a) The diagonal FWHM (σ) are obtained from fitting the diagonal slices of the 2D spectra to Gaussian functions. (b) The anti-diagonal FWHM (Γ) are determined directly from the anti-diagonal slices that cut through the peak maxima of the diagonal peaks. (c) The ellipticities are calculated from the diagonal and anti-diagonal widths, $(\sigma^2 - \Gamma^2)/(\sigma^2 + \Gamma^2)$.

III. 2D IR spectra fitting model and parameters

To fit the experimental 2D IR spectra, we simulated the third-order nonlinear in the time domain and then Fourier transformed to get the spectra. The nonlinear polarization that is detected in the experiments is directly proportional to the nonlinear response function, a thrice-nested commutator of the transition dipole operators, evaluated at each light-matter interaction. The thrice-nested commutator can be expanded into eight terms, however, only four are unique since the other four are complex conjugates.

$$\overleftrightarrow{R}^{(3)}(\tau_3,\tau_2,\tau_1)=\sum_1^4 \overleftrightarrow{R}_n(\tau_3,\tau_2,\tau_1)-\overleftrightarrow{R}_n^*(\tau_3,\tau_2,\tau_1)$$

These four response function pairs differ by whether sequential operators act on the *bra* or *ket* side of ρ_g when enforcing the time-ordering. We make a further assumption to simplify the response functions by writing the nonlinear response function as a product of an isotropic nonlinear response function, which describes the vibrational dynamics, and a tensorial orientational nonlinear response function, which describes the dipole orientational dynamics. If no preferred orientations are assumed, numerical solutions exist for the tensor components in the case of two coupled oscillators.

The exact functional form of the isotropic nonlinear response function depends on how the density matrix evolves in time under the Hamiltonian, which can be presented pictorially by double-sided Feynman diagrams.[1] As an example, for a system with coupled vibrations *a* and *b*, one of the rephasing pathways R_2 *(bra/ket/bra)* can be written as:

$$R_2(\tau_3,\tau_2,\tau_1)=\mu^{0,a}\mu^{0,b}\mu^{a,0}\mu^{b,0}\times\exp(-i\omega_{a,0}\tau_3-i\omega_{a,b}\tau_2-i\omega_{0,b}\tau_1)\times F_2(\tau_3,\tau_2,\tau_1)$$

The dephasing function, $F_n(\tau_3,\tau_2,\tau_1)$, is the exponential of a linear combination of lineshape functions *g(t)* that are derived from the energy gap correlation function.

$$g(t)=\int_0^t dt''\int_0^{t''} dt' C_{ab}(t'),\ \text{where}\ C_{ab}(t')=\left\langle \delta\omega_{ab}(t')\delta\omega_{ab}(0)\right\rangle$$

The exact forms of the dephasing functions for the four vibrational response functions involving the time evolution during different time periods have been explicitly written out.[2,3] The nonlinear

[1] Tokmakoff, A. *Time-dependent quantum mechanics and spectroscopy*, 2003; Available from: http://www.mit.edu/~tokmakof/TDQMS/index.htm.
[2] Khalil, M.; Demirdöven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258.
[3] Sung, J.; Silbey, R. J. *J. Chem. Phys.* **2001**, *115*, 9266.

response functions for all the Liouville pathways are summed and Fourier transformed to give the rephasing and non-rephasing spectra. Finally the absorptive 2D IR spectra are obtained by adding rephasing and non-rephasing spectra.

In general, the time-dependent correlated fluctuations between two transition frequencies can be described by a Gaussian stochastic model with exponentially decaying auto- ($m = n$) and cross- ($m \neq n$) correlation functions:

$$C_{mn}(t) = \rho_{mn}\Delta_{mm}\Delta_{nn}exp(-|t|/\tau_{mn})$$

where *m,n* = *a, b*. ρ_{mn} is the correlation coefficient with $\rho_{mm} = 1$. In the inhomogeneous limit, the energy gap correlation function becomes $C_{mn}(t) = \rho_{mn}\Delta_{mm}\Delta_{nn}$, and the lineshape function is $g_{mn}(t) = \rho_{mn}\Delta_{mm}\Delta_{nn}t^2/2$. In this static limit, Δ_{mm} and Δ_{nn} are proportional to the magnitudes of the distribution of the transition frequencies of the system eigenstate around their central frequency. On the other hand, the energy gap correlation function is a delta function in the homogeneous limit: $C_{mn}(t) = \delta(t)/T_2^{mn}$, and the lineshape function is $g_{mn}(t) = t/\mathrm{T}_2^{mn}$, where $\mathrm{T}_2^{mn} = \rho_{mn}\Delta_{mm}\Delta_{nn}\tau_{mn}$. In this study we make use of all three cases in constructing the correlation function (Equation 1 in the manuscript).

In the present study, our primary goal is to obtain structural parameters through peak amplitudes, and not to derive dynamics from the lineshape. Therefore, we fit all the peaks in the 2D spectra to the same form of the lineshape function:

$$g(t) = t/T_2 + \Delta^2\tau^2\left[exp(-t/\tau) - 1 + t/\tau\right] + (\Delta_o t)^2/2$$

where T_2 , Δ, τ , and Δ_o are the floating parameters. This method is very robust to measuring peak amplitudes with small error, even with a large number of adjustable parameters.

In this fitting algorithm, the values of ρ_{mn} are formally related to these fit parameters. As a simpler method of quantifying correlation, we use an alternative method. It has been well established that the degree of correlation can be quantified through lineshape metrics, such as the tilt angle of the nodal line, Ψ , which we obtain as[4,5]

$$\rho_{mn} = \tan(\Psi) = \left(A_R - A_{NR}\right)/\left(A_R + A_{NR}\right)$$

[4] Demirdoven, N.; Khalil, M.; Tokmakoff, A. *Phys. Rev. Lett.* **2002**, *89*, 237401.
[5] Roberts, S. T.; Loparo, J. J.; Tokmakoff, A. *J. Chem. Phys.* **2006**, *125*, 084502.

where A_R and A_{NR} are the cross-peak amplitudes in rephasing and non-rephasing spectra, respectively. These are the values reported in Table 3.

The following tables list all the parameters used to fit the experimental 2D IR spectra for the five NMPs. ω_i is the transition frequency; $\Delta\omega_{i,i}$ and $\Delta\omega_{i,j}$ are the diagonal and off-diagonal anharmonicities, respectively; μ_i is the transition dipole strength for the zero- to one- quantum states; $\mu_{2i,i}$ is the transition dipole strength for the one- to two- quantum states; $\mu_{i/j,i}$ is the transition dipole strengths from the one-quantum state to a combination band; T_2 , Δ, τ , and Δ_o are the lineshape parameters. It should be emphasized again that these lineshape parameters are only for the purpose of fitting the peak amplitudes, and they do not imply the real dynamics. In addition, the fundamental transition dipole strengths are influenced by the spectrum used in the 2D experiments, therefore the more accurate determination of the dipole strengths were obtained from fitting the FTIR using the lineshape parameters. However, the scaling of the dipole strengths for higher transitions remains accurate in an anisotropy measurement.

a) AMP

ω_{A1}	ω_{A2}	
1625	1578	
$\Delta\omega_{A1}$	$\Delta\omega_{A2}$	$\Delta\omega_{A1/A2}$
8.00	9.80	3.65

μ_{A1}	μ_{A2}
1.00	0.41
$\mu_{2A1,A1}$	$\mu_{2A2,A2}$
1.38	0.58
$\mu_{A1/A2,A2}$	$\mu_{A1/A2,A1}$
0.86	0.41

		T_2	Δ	τ	Δ_o
Diagonal peaks	A1	1909.6	1.7	95.0	20.8
	A2	1995.0	35.8	800.3	33.5
Cross-peaks	A1/A2	2000.1	0.2	10.4	0.0

b) GMP

ω_{G1}	ω_{G2}	ω_{G3}	ω_{G4}		
1665	1579	1565	1539		
$\Delta\omega_{G1}$	$\Delta\omega_{G2}$	$\Delta\omega_{G3}$	$\Delta\omega_{G4}$		
10.94	4.00	7.76	17.98		
$\Delta\omega_{G1/G2}$	$\Delta\omega_{G1/G3}$	$\Delta\omega_{G1/G4}$	$\Delta\omega_{G2/G3}$	$\Delta\omega_{G2/G4}$	$\Delta\omega_{G3/G4}$
1.34	4.57	3.27	7.51	6.97	6.84

μ_{G1}	μ_{G2}	μ_{G3}	μ_{G4}		
1.00	0.87	0.76	0.21		
$\mu_{2G1,G1}$	$\mu_{2G2,G2}$	$\mu_{2G3,G3}$	$\mu_{2G4,G4}$		
1.40	1.18	1.04	0.49*		
$\mu_{G1/G2,G1}$	$\mu_{G1/G3,G1}$	$\mu_{G1/G4,G1}$	$\mu_{G2/G3,G2}$	$\mu_{G2/G4,G2}$	$\mu_{G3/G4,G3}$
0.84	0.76	0.36	0.77	0.27	0.38
$\mu_{G1/G2,G2}$	$\mu_{G1/G3,G3}$	$\mu_{G1/G4,G4}$	$\mu_{G2/G3,G3}$	$\mu_{G2/G4,G4}$	$\mu_{G3/G4,G4}$
1.00	1.01	1.08	0.78	0.80	0.82

		T_2	Δ	τ	Δ_o
Diagonal peaks	G1	566.1	49.7	476.2	52.8
	G2	623.2	26.2	105.8	18.5
	G3	598.8	59.9	196.1	1.9
	G4	1645.7	122.8	967.8	46.5
Cross-peaks	G1/G2	4544.3	42.7	1929.5	12.0
	G1/G3	4363.7	5.0	1694.5	12.1
	G1/G4	1866.0	118.9	721.1	0.0
	G2/G3	5111.3	6.5	1432.3	0.0
	G2/G4	1025.5	34.0	1359.0	0.0
	G3/G4	1901.0	9.9	1994.9	0.0

* A deviation from the harmonic scaling in the transition dipole strength was observed for mode G4 by comparing the amplitudes of the diagonal absorptive and emissive peaks. A relative strength of $\mu_{2G4,G4} / \mu_{G4} = 2.3$ suggests significant electrical anharmonicity, however, this observation can also be caused by the low peak intensity of G4 that led to the insensitivity of amplitude fitting.

c) UMP

ω_{U1}	ω_{U2}	ω_{U3}
1693	1655	1617
$\Delta\omega_{U1}$	$\Delta\omega_{U2}$	$\Delta\omega_{U3}$
15.96	18.32	25.00
$\Delta\omega_{U1/U2}$	$\Delta\omega_{U1/U3}$	$\Delta\omega_{U2/U3}$
7.96	24.98	15.40

μ_{U1}	μ_{U2}	μ_{U3}			
0.76	1.01	0.46			
$\mu_{2U1,U1}$	$\mu_{2U2,U2}$	$\mu_{2U3,U3}$			
0.95	1.35	0.67			
$\mu_{U1/U2,U1}$	$\mu_{U1/U3,U1}$	$\mu_{U2/U3,U2}$	$\mu_{U1/U2,U2}$	$\mu_{U1/U3,U3}$	$\mu_{U2/U3,U3}$
0.9	0.55	0.48	0.71	0.73	1.15

		T_2	Δ	τ	Δ_o
Diagonal peaks	U1	1291.7	98.0	199.6	43.4
	U2	637.7	100.3	17.4	25.7
	U3	1773.9	93.2	643.3	26.6
Cross-peaks	U1/U2	397.3	160.2	141.8	0.0
	U1/U3	862.6	102.4	338.8	0.0
	U2/U3	1977.9	31.5	518.8	38.6

d) TMP

ω_{T1}	ω_{T2}	ω_{T3}
1690	1663	1629
$\Delta\omega_{T1}$	$\Delta\omega_{T2}$	$\Delta\omega_{T3}$
10.71	5.00	10.00
$\Delta\omega_{T1/T2}$	$\Delta\omega_{T1/T3}$	$\Delta\omega_{T2/T3}$
1.73	4.92	6.15

μ_{T1}	μ_{T2}	μ_{T3}			
0.71	1.00	0.97			
$\mu_{2T1,T1}$	$\mu_{2T2,T2}$	$\mu_{2T3,T3}$			
1.2	1.42	1.37			
$\mu_{T1/T2,T1}$	$\mu_{T1/T3,T1}$	$\mu_{T2/T3,T2}$	$\mu_{T1/T2,T2}$	$\mu_{T1/T3,T3}$	$\mu_{T2/T3,T3}$
0.80	1.07	0.95	0.77	0.60	1.04.

		T_2	Δ	τ	Δ_o
Diagonal peaks	T_1	8873.7	163.43	139.88	19.62
	T2	5510.29	334.96	16.73	32.79
	T3	10904.81	175.05	97.86	44.7
Cross-peaks	T1/T2	5920.99	31.38	20.12	0
	T1/T3	6050.92	30.01	20.14	0
	T2/T3	6227.86	30	20	47.93

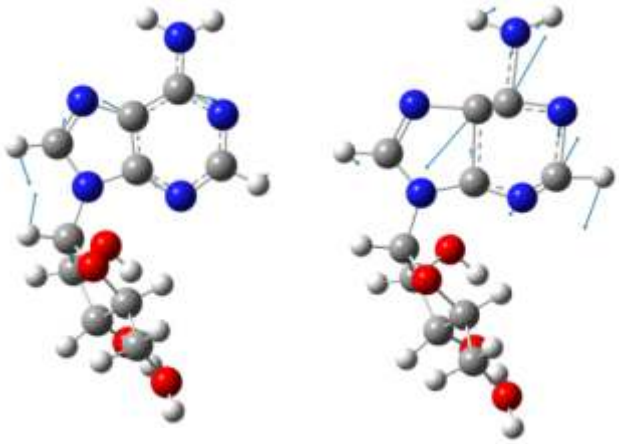
e) CMP

ω_{C1}	ω_{C2}	ω_{C3}	ω_{C4}	ω_{C4}
1651	1614	1583	1524	1504
$\Delta\omega_{C1}$	$\Delta\omega_{C2}$	$\Delta\omega_{C3}$	$\Delta\omega_{C4}$	$\Delta\omega_{C5}$
9.19	15.40	16.70	6.12	8.91
$\Delta\omega_{C1/C2}$	$\Delta\omega_{C1/C3}$	$\Delta\omega_{C1/C4}$	$\Delta\omega_{C1/C5}$	$\Delta\omega_{C2/C3}$
7.15	7.38	1.64	5.19	11.99
$\Delta\omega_{C2/C4}$	$\Delta\omega_{C2/C5}$	$\Delta\omega_{G3/G4}$	$\Delta\omega_{G3/G5}$	$\Delta\omega_{C3/C4}$
14.38	7.64	12.76	8.24	1.17

μ_{C1}	μ_{C2}	μ_{C3}	μ_{C4}	μ_{C5}					
0.85	0.62	0.19	0.63	0.61					
$\mu_{2C1,C1}$	$\mu_{2C2,C2}$	$\mu_{2C3,C3}$	$\mu_{2C4,C4}$	$\mu_{2C5,C5}$					
1.25	0.98	0.55	0.72	0.88					
$\mu_{C1/C2,C1}$	$\mu_{C1/C3,C1}$	$\mu_{C1/C4,C1}$	$\mu_{C1/C5,C1}$	$\mu_{C2/C3,C2}$	$\mu_{C2/C4,C2}$	$\mu_{C2/C5,C2}$	$\mu_{C3/C4,C3}$	$\mu_{C3/C5,C3}$	$\mu_{C4/C5,C4}$
0.61	0.24	0.64	0.59	0.40	0.62	0.40	0.30	0.70	0.65
$\mu_{C1/C2,C2}$	$\mu_{C1/C3,C3}$	$\mu_{C1/C4,C4}$	$\mu_{C1/C5,C5}$	$\mu_{C2/C3,C3}$	$\mu_{C2/C4,C4}$	$\mu_{C2/C5,C5}$	$\mu_{C3/C4,C4}$	$\mu_{C3/C5,C5}$	$\mu_{C4/C5,C5}$
0.81	0.87	0.81	0.81	0.51	0.78	0.66	0.39	0.28	0.57

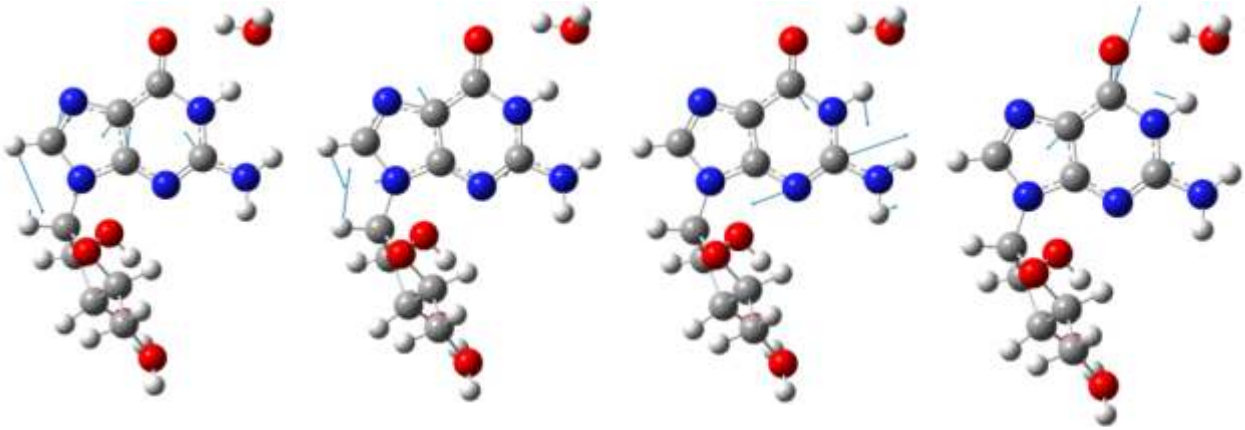
		T_2	Δ	τ	Δ_o
Diagonal peaks	C1	1584	109	204	35
	C2	2500	78	600	54
	C3	2500	17	45	46
	C4	651	--	--	--
	C5	583	--	--	--
Cross-peaks	C1/C2	2326	104	499	0
	C1/C3	862	150	143	0
	C1/C4	577	--	--	--
	C1/C5	1844	--	--	--
	C2/C3	1285	81	200	0
	C2/C4	2000	--	--	--
	C2/C5	1698	--	--	--
	C3/C4	1559	--	--	--
	C3/C5	509	--	--	--
	C4/C5	1961	--	--	--

IV. Vibrational modes investigated in this study from DFT calculations



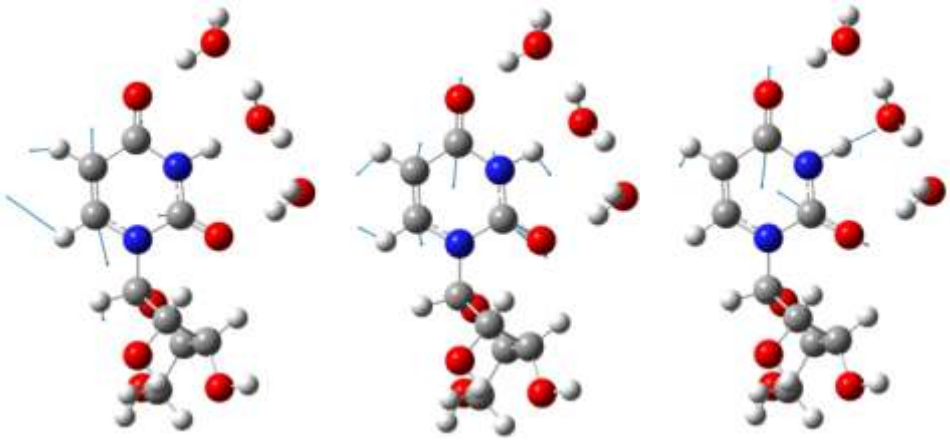
A2: 1565 cm^{-1} (0.11) A1: 1588 cm^{-1} (1.00)

Figure S3. Calculated vibrational modes A1 and A2 of deuterated adenine with the atom displacements. The frequencies have been scaled by a factor of 0.9614 and the normalized IR intensities to the more intense A2 mode are given in parentheses.



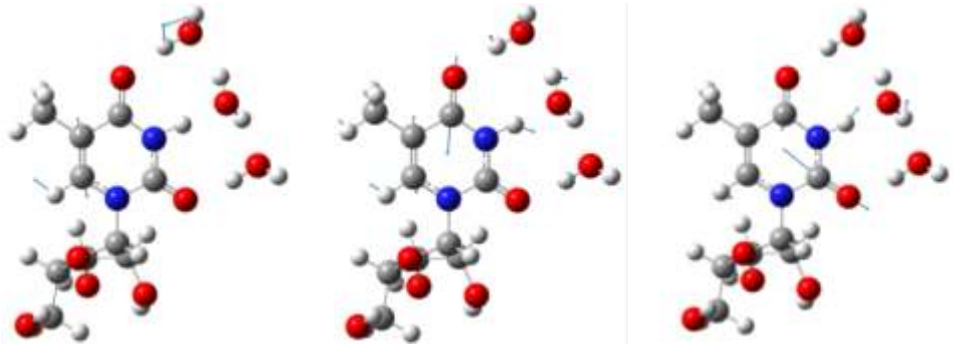
G4: 1511 cm^{-1} (0.23) G3: 1533 cm^{-1} (0.05) G2: 1556 cm^{-1} (0.62) G1: 1692 cm^{-1} (1.00)

Figure S4. Calculated vibrational modes G1 – G4 of deuterated guanosine with the atom displacements. The molecule is solvated by one D_2O molecule, which makes a hydrogen bond to the carbonyl. The frequencies have been scaled by a factor of 0.9614 and the normalized IR intensities to the most intense G1 mode are given in parentheses.



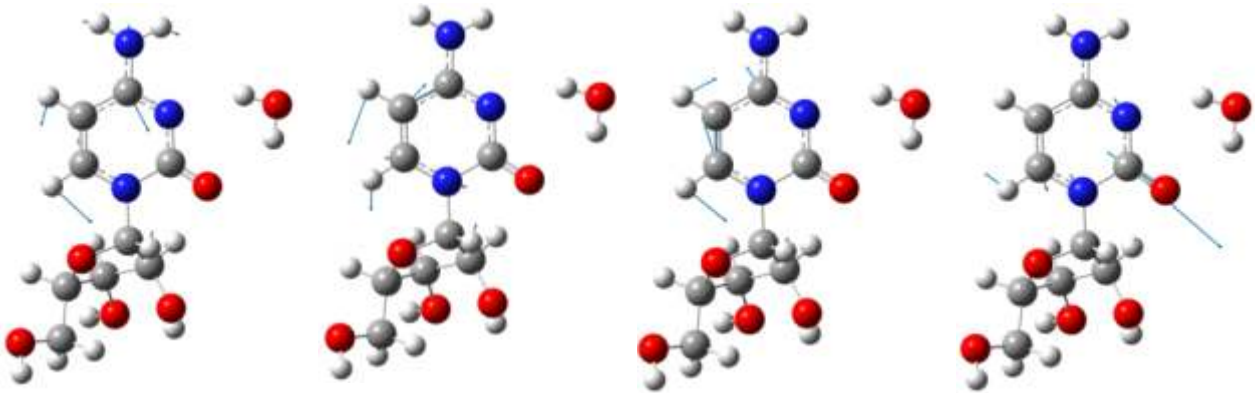
U3: 1615 cm^{-1} (0.02) U2: 1670 cm^{-1} (1.00) U1: 1681 cm^{-1} (0.10)

Figure S5. Calculated vibrational modes U1 – U3 of deuterated uridine with the atom displacements. The molecule is solvated by three D_2O molecules, which make hydrogen bonds to the carbonyls. The frequencies have been scaled by a factor of 0.9614 and the normalized IR intensities to the most intense U2 mode are given in parentheses.



T3: 1627 cm^{-1} (0.16) T2: 1657 cm^{-1} (1.00) T1: 1677 cm^{-1} (0.43)

Figure S6. Calculated vibrational modes T1 – T3 of deuterated thymidine with the atom displacements. The molecule is solvated by three D_2O molecules, which make hydrogen bonds to the carbonyls. The frequencies have been scaled by a factor of 0.9614 and the normalized IR intensities to the most intense T2 mode are given in parentheses.



C5: 1483 cm^{-1} (0.43) C4: 1503 cm^{-1} (0.34) C2: 1689 cm^{-1} (0.74) C1: 1690 cm^{-1} (1.00)

Figure S7. Calculated vibrational modes C1, C2, C4, and C5 of deuterated cytidine with the atom displacements. The molecule is solvated by one D2O molecule, which makes a hydrogen bond to the carbonyl. The frequencies have been scaled by a factor of 0.9614 and the normalized IR intensities to the most intense C1 mode are given in parentheses.

V. Influence of solvating water molecules on DFT calculations

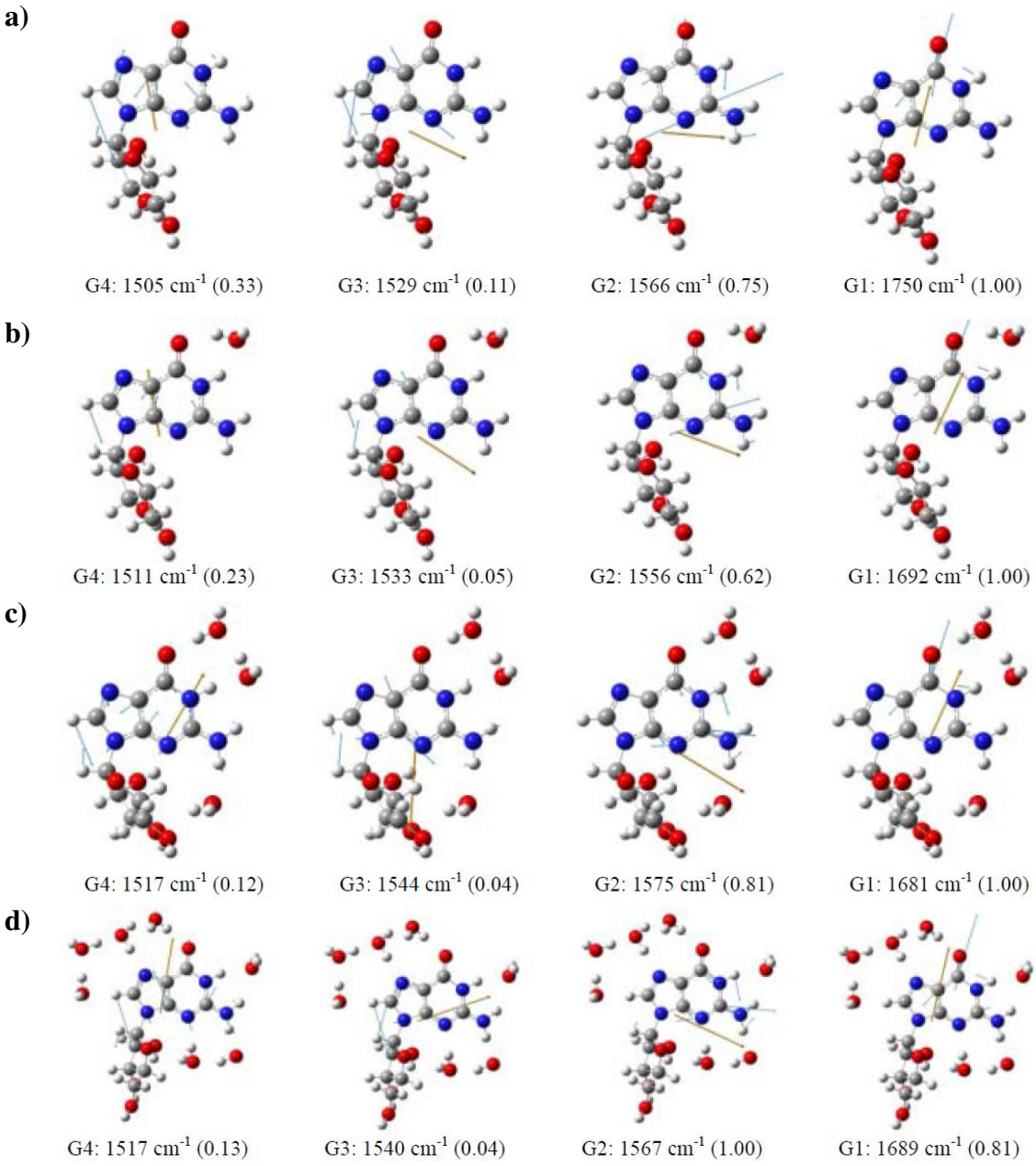


Figure S8. Vibrational normal modes G1 – G4 of guanosine as a function of the number of solvating D_2O molecules: (a) no D_2O; (b) one D_2O; (c) three D_2O; and (d) seven D_2O molecules. The frequencies have been scaled by a factor of 0.9614 and the normalized IR intensities to the most intense mode are given in parentheses.

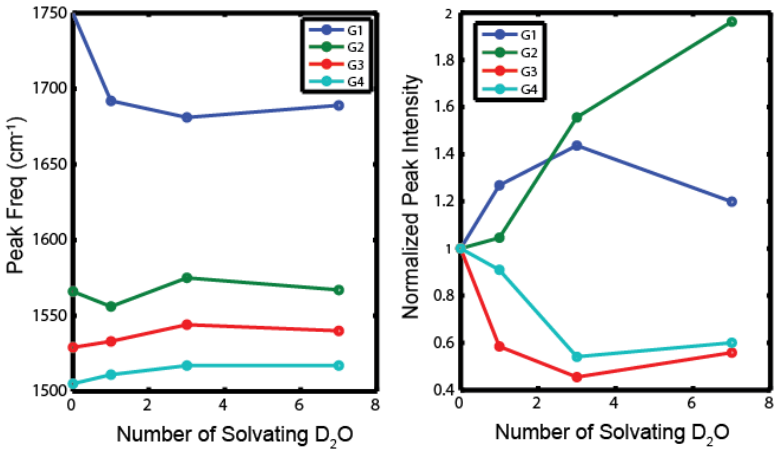


Figure S9. (Left) Peak frequencies and (right) the normalized (relative to the case without solvating water) intensity of G1 – G4 as a function of the number of solvating D_2O molecules.

VI. Complete Ref. 39

Gaussian 98, Revision A.6, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; C.Burant, J.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; G. Liu, A. L.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; G.Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A.; Gaussian, Inc.: Pittsburgh PA, 1998.