Supporting Information for:

Fluorescence Encoded Infrared Spectroscopy: Ultrafast Vibrational Spectroscopy on Small Ensembles of Molecules in Solution

Lukas Whaley-Mayda[†‡], Samuel B. Penwell[†‡], and Andrei Tokmakoff[†*]

[†]Department of Chemistry, James Franck Institute, and Institute for Biophysical Dynamics, The University of Chicago, Illinois 60637, United States
[‡]These authors contributed equally to this work
[*]Corresponding Author: Email: tokmakoff@uchicago.edu

Contents

1 IR pulse characterization and temporal instrument response

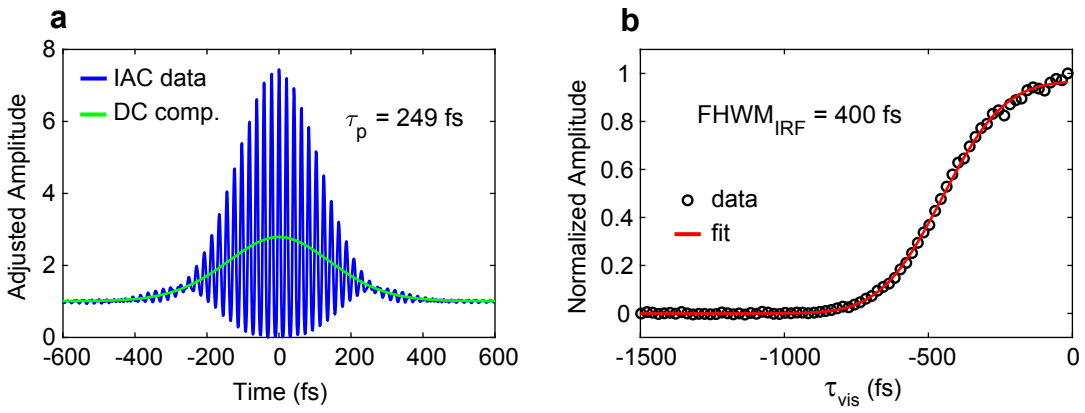

Figure S1: Pulse duration characterization. (a) Interferometric autocorrelation (blue) of the IR field at the sample position. The zero frequency component (intensity autocorrelation) is shown in green. (b) Rise of the 1-pulse FEIR signal. The red line is the fit to an error function, which produces a derivative FWHM of 400 fs.

The IR pulse duration was characterized by interferometric autocorrelation (IAC) in $AgGaS_2$ at the sample position (Figure S1a). 1 mm of CaF_2 is included in the optical path to account for transmission into the sample cell. The data is shifted and scaled so that the lowest fringe falls to zero and the long-time baseline is unity. The zero frequency band of the data is the intensity autocorrelation (IA),[1] which is extracted by a Fourier filter. We use the FWHM of the temporal pulse intensity profile as a definition of pulse duration, τ_p. Assuming a Gaussian profile, this is related to the FWHM of the IA by $\tau_p = \tau_{IA}/\sqrt{2}$. We use this relation to report the pulse duration from a Gaussian fit of the recovered IA.

The 250 fs IR pulse duration (transform limit of 130 fs) is consistent with the dispersion that results from the summed GVD of materials in the optical path. This includes the $LiGaS_2$ OPA crystal, 2 Germanium windows, a CaF_2 collimating lens, and the ZnSe beamsplitters and aspheric condenser lens. It is unclear if pulse length has a strong effect on the efficiency of FEIR signal generation provided the pulses are shorter than the vibrational lifetimes in question.

The visible pulse duration has not been directly measured at the sample. We may roughly estimate its duration by considering the rise time of the 1-pulse FEIR signal, shown in Figure S1b. The rise is fit with an error function, which represents the functional form expected by convolving Gaussian pulses with a material step response. We consider the resulting derivative FWHM, τ_{IRF}, of 400 fs to represent the duration of the temporal instrument response function (IRF). Because the FEIR signal is generated by two interactions with each field, the IRF should represent the convolution between the IR and visible temporal pulse intensity profiles, resulting in $\tau_{IRF}^2 = \tau_{p,IR}^2 + \tau_{p,vis}^2$. From this relation we find $\tau_{p,vis} = 313$ fs (transform limit > 200 fs, spectrum shown in Figure S5 is likely limited by the spectrometer resolution), however we recognize this is only a rough estimate. It is also important to recognize that employing counter-propagating pulses can wash out the time-resolution along their relative delay if signal is collected from a region with a longitudinal dimension larger than the effective pulse lengths (e.g. 30 μm for a 100 fs pulse in vacuum). In our experiment the few μm longitudinal size of the point spread function provided by the objective and confocal detection renders this effect negligible.

2 IR and visible focus characterization

The size of the IR focus at the sample was determined by measuring the transmission through a 5 μm pinhole which is raster scanned across the beam. The pinhole was mounted between a 1 mm CaF_2 window and coverslip to mimic the sample profile. Figure S2a shows an XY scan at the IR focal plane. Cross sections with Gaussian fits are shown in Figures S2c and d. The measured beam size represents the true beam size convoluted with the 5 μm pinhole. For simplicity we report the spot size as the average of the measured FWHM in the X and Y directions, yielding 10.5 μm. Figure S2b shows a Z scan along the axis of the beam, where negative Z positions are above the focus.

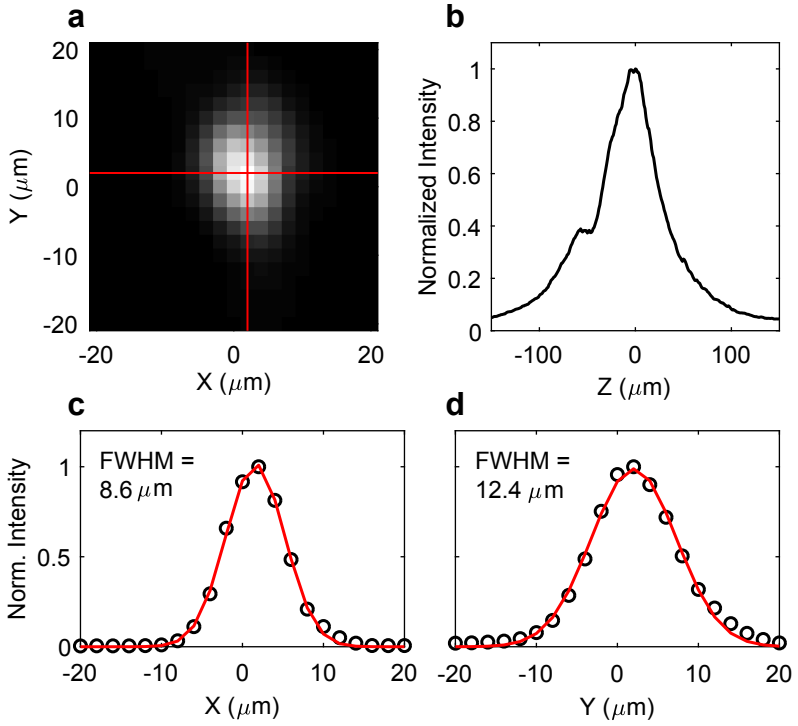

Figure S2: IR focus characterization. (a) Grayscale image of the IR intensity from a pinhole scan in the focal plane (Z = 0). The maximum and minimum intensities are set to white and black, respectively. Red lines indicate the slices plotted in (c-d). (b) Z scan at the maximum intensity XY position in (a). (c-d) Slices along the X and Y directions from the data in (a) (black circles) with Gaussian fits (red).

We determined the visible focal size by measuring the emission intensity from a 100 nm fluorescent bead as it was scanned through the focus. The bead sample was prepared by drop casting a few μL of bead solution (Fluoresbrite YO Carboxylate Microspheres 0.10 μm) on a clean coverslip. Figures S3a and b show an X cross section at the focal plane and Z scan along the optical axis, respectively. These measurements were performed without passing the fluorescence through a confocal pinhole, relying instead on the weak confocal effect from the 50 μm active area of the detector (the magnification of the objective and tube lens (f = 75 mm) produced a calculated 15 μm Airy disk at the detector). The resulting size should therefore only be taken as approximate.

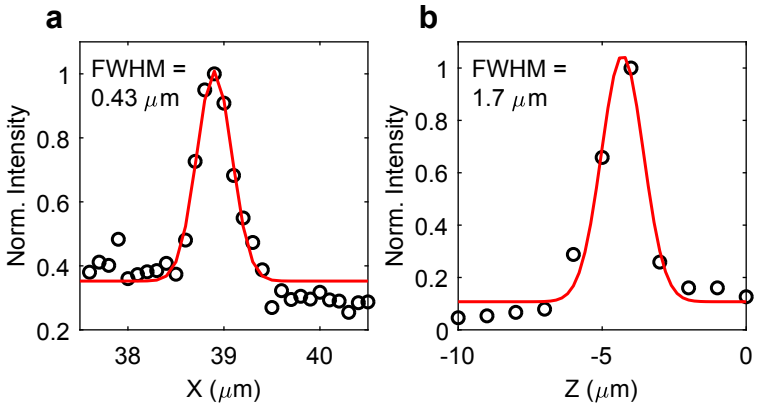

Figure S3: Visible focus characterization. (a) Bead scan across the lateral beam profile at the focus (black circles) with Gaussian fit (red). (b) Axial scan.

3 IR and visible power dependence

IR and visible power dependencies shown here for coumarin 6 in deuterated acetonitrile were measured at the peak ($\tau_{vis} = 0$) of the 1-pulse FEIR signal while chopping the IR beam at 100 Hz. The sample concentration was 40 μM. We have verified that these power dependencies do not vary with τ_{vis} delay along the transient. The IR power was adjusted by a pair of wire-grid polarizers (Specac). The IR power was measured after the polarizers and scaled to the maximum power measured at the sample. The visible power was controlled by a half-waveplate and polarizer before the doubling crystal. The reported visible powers are those measured after the spatial filter but before entering the microscope. Figure S4a shows the IR power dependence of the S_1 signal amplitude (IR_{on} - IR_{off}), as well as the background S_0 (IR_{off}). The small decrease in the background count rate as IR power is increased is likely a thermal effect due to heating of the sample. This temperature-dependent reduction in collection efficiency also affects the S_1 amplitude, which results in a slight sub-linearity to the power dependence. However, plotting the modulation ratio S_1/S_0 (Figure S4b) adjusts for this thermal effect, resulting in a linear power dependence. This thermal effect is investigated and discussed further in Section 7.

Figure S4c shows the visible power dependence with the IR pulse energy fixed at 9 nJ. Both the S_1 and S_0 amplitudes grow linearly, and the modulation ratio (Figure S4d) is constant.

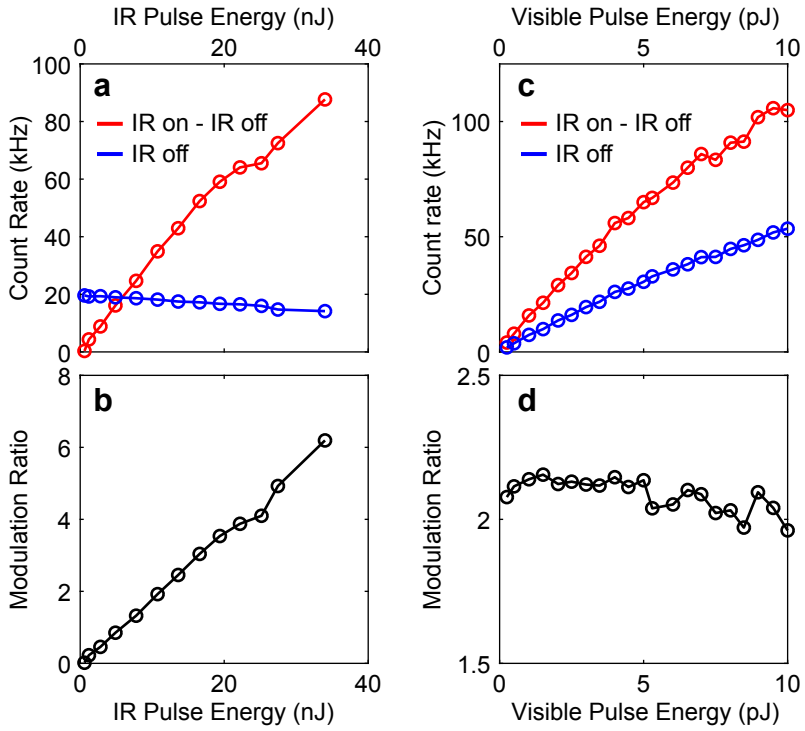

Figure S4: IR and visible power dependence of the FEIR signal. (a) S_1 (IR_{on} - IR_{off}, red) and S_0 (IR_{off}, blue) amplitude as function of IR pulse energy. (b) The modulation ratio, or quotient S_1/S_0 as a function of IR pulse energy. (c) S_1 (red) and S_0 (blue) amplitude as a function of visible pulse energy. (d) Modulation ratio as a function of visible pulse energy.

4 Molecule number characterization by FCS

For the fluorescence correlation spectroscopy (FCS) characterization of the average molecule number we used rhodamine 6G (R6G) as a proxy for coumarin 6 (C6). As shown in Figure S5a, R6G is directly on resonance with the visible encoding frequency, which allows us to perform FCS at this excitation frequency (or very close to, see Figure S5b). Using the same excitation frequency is crucial to ensure that the same optics can be used and that the effective observation volume is as close as possible to that in the FEIR experiment. Of course this comes at the cost of studying a different molecule. Our primary concern is the potential difference in the dyes' tendency to stick to interfaces, e.g. substrates in the sample cell or walls of containers, which could lead to variations in the actual local concentration. Here we proceed with the assumption that this effect is negligible in our concentration range.

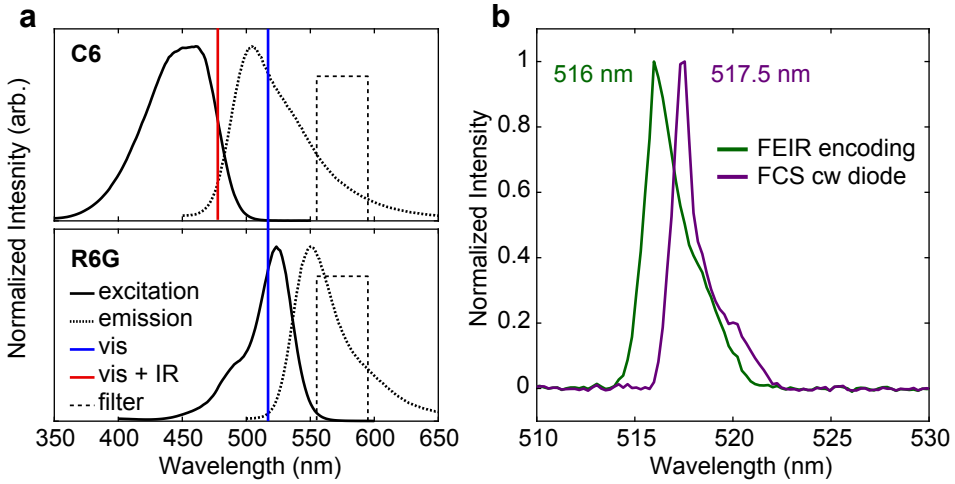

Figure S5: (a) Fluorescence excitation and emission spectra of coumarin 6 (top panel) and rhodamine 6G (bottom panel) in acetonitrile. Overlaid is the center frequency of the visible fields (FEIR encoding and cw diode, blue), the sum of FEIR encoding and IR pump frequencies (red), and the bandwidth of the fluorescence emission filter (dashed black). (b) Spectra of the FEIR encoding field (green), and the cw diode used for FCS (purple). Wavelength at the maximum intensity is indicated for each spectrum, however the line widths are likely near the resolution limit of the spectrometer (Thorlabs CCS200 fiber-coupled Czerny-Turner).

R6G solutions were prepared using the identical procedure of serial dilutions as the C6 solutions for FEIR. FCS measurements were performed in the same sample cell. A new coverslip was used for each concentration. Coverslips were cleaned by sonication in a dilute detergent solution followed by rinsing in DI and milli-Q purified water. CaF_2 windows and 50 μm teflon spacers were initially cleaned in the same way, but reused between samples after cleaning with methanol and lens tissue. The rest of the sample cell components were cleaned similarly.

Figure S6a shows the fluorescence amplitude as a function of Z position for samples between 1 and 100 nM. The coverslip-solution interface is at 0 μm, and the solution-CaF_2 interface appears between -25 and -50 μm. Effective Z distances on this plot appear scaled by the refractive index difference between layers. The edge response at the top interface (Z = 0 μm) is much sharper than the lower interface due to degradation of the focusing conditions away from the coverslip. Figure S6b shows the fluorescence amplitude at Z = -20 μm from the profiles in Figure S6a. The linear dependence of count rate on concentration down to the few nM range is an indication that our

dilution and sample preparation procedure is delivering the correct *relative* number of molecules to the observation volume.

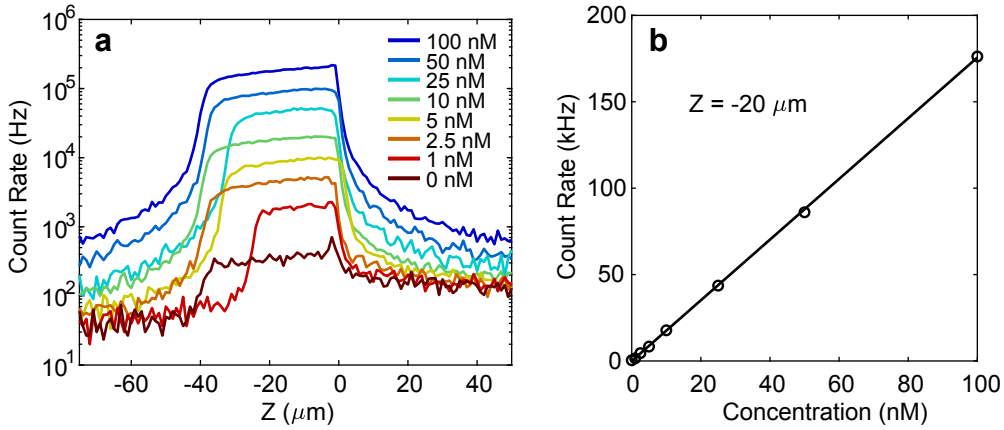

Figure S6: Depth profiles and count rates of the R6G samples. (a) Fluorescence count rate (log scale) as a function of Z position within the sample for concentrations between 0 nM (pure solvent) and 100 nM. (b) Count rate at Z = -20 μm (black circles) for the profiles in (a), with a linear fit for reference (black line).

To measure the absolute number of molecules, we use FCS as discussed in the main text. Correlation functions are measured 20 μm below the coverslip (at Z = -20 μm in Figure S6a) and averaged for 5 minutes. An example correlation function from the 1 nM R6G solution is shown in Figure S7a. Because our instrument has a single detector, the correlation functions are polluted by a large after-pulsing artifact that appears as a spike at early times ($t < 10^{-5}$ s). Figure S7b shows the baseline of the effective instrument response of the correlator produced by measuring the correlation of room light leak-through onto the detector. The after-pulsing spike has fully decayed by 10^{-5} s. For all FCS measurements we only use data after 10^{-5} s for fitting. While after-pusling artifacts are not strictly additive contributions that can be subtracted,[2] we will assume that the residual effect on the correlation amplitude after 10^{-5} s is negligible. After-pulsing artifacts are typically avoided by splitting the fluorescence onto two detectors and measuring their cross-correlation, or using a time-tagging unit to statistically filter out after-pulse events.[2]

Correlation functions are fit to the following standard FCS model function, which assumes a single species with diffusion constant D, concentration C, and a 3-dimensional Gaussian observation volume with lateral and axial $1/e^2$ radii of w_{xy} and w_z, respectively,

$$G_D(t) = \frac{1}{\langle N \rangle} \left(1 + \frac{t}{\tau_D}\right)^{-1} \left(1 + \frac{t}{\kappa^2 \tau_D}\right)^{-1/2} \tag{1}$$

$$\text{with} \quad \tau_D = \frac{w_{xy}^2}{4D}, \quad \kappa = \frac{w_z}{w_{xy}}, \quad \langle N \rangle = C\pi^{3/2} w_{xy}^2 w_z.$$

For this model, the average number of molecules $\langle N \rangle$ is defined by the concentration and effective size and shape of the observation volume. However, without assuming a known concentration or observation volume, one can treat the inverse early time amplitude (that is, $\langle N \rangle$) as a practical

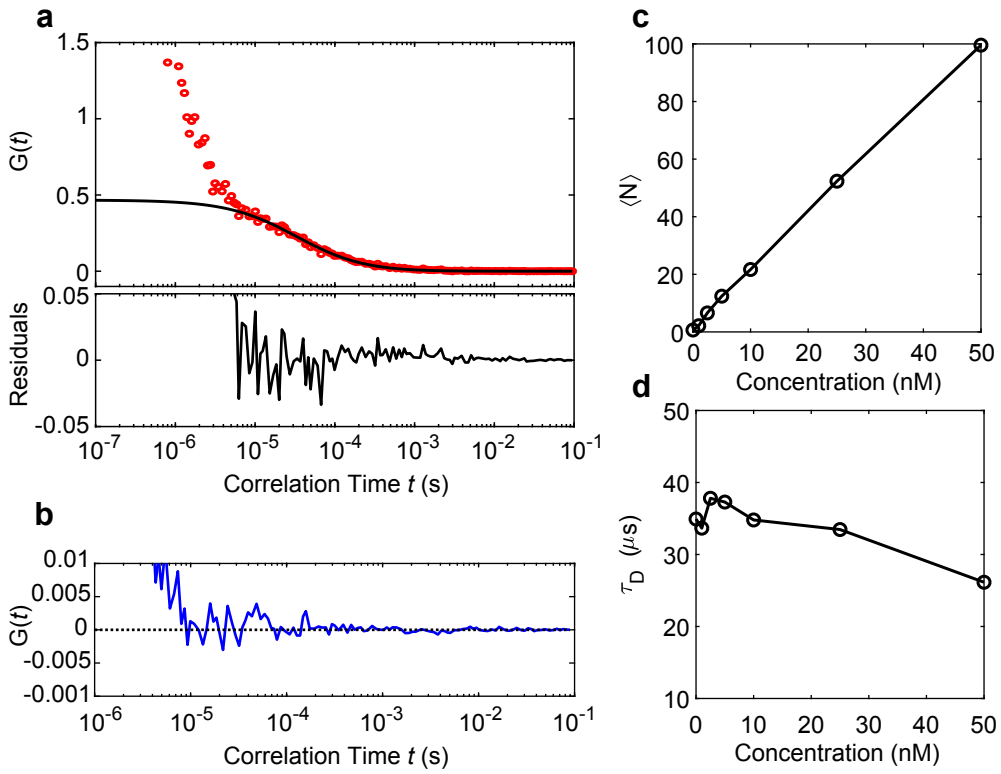

Figure S7: Correlation function fitting. (a) Correlation function from the 1 nM R6G sample (red circles) with a fit only using data with $t > 10^{-5}$ s to Eq. 1 (black). Residuals are shown in the panel below. The large amplitude spike in the data apparent for $t < 10^{-5}$ s is the detector after-pulsing artifact. (b) Detail of the correlator instrument response function near the baseline. The after-pulsing spike has negligible amplitude for $t > 10^{-5}$ s. (c-d) Average molecule number and characteristic diffusion time, respectively, from the fits as a function of concentration.

definition of the average number of molecules participating in the experiment at any given time. In general, this interpretation of the early time amplitude does not depend on the specific shape of the observation volume, provided that diffusion is the only dynamical process responsible for fluorescence fluctuations.[3] For this interpretation to be valid it is also important to ensure that the uncorrelated background count rate is small compared to the count rate from the molecules being probed. The good linearity with nearly zero y-intercept of count rate versus concentration across our dilution range (Figure S6b) suggests that the background is indeed sufficiently low. With this conceptual approach in mind, we use the model function in Eq. 1 solely to extract the early time amplitude from the measured correlation functions.

In practice, the aspect ratio κ in Eq.1 is not a sensitive parameter and is fixed at 4, the value we measure from the bead scans in Figure S3. An example fit and residuals are shown in Figure S7a. Figures S7c shows $\langle N \rangle$ as a function of concentration (same data as shown in main text Figure 3d) extracted from the fits, while Figure S7d shows the characteristic diffusion time τ_D. The diffusion time appears roughly conserved for each concentration (there should be no change), within errors produced by the quality of fitting.

5 1-pulse FEIR data fitting

The time-dependence of the 1-pulse FEIR signal shown in main text Figure 2a was analyzed by fitting to two models. The first model, shown in Figure S8a, is a simple bi-exponential decay. The data was only fit after 400 fs, which corresponds to the instrument response time determined in Section 1. This fit produces a fast time constant of 1.0 ps (relative amplitude $a_1 = 0.78$) and a slow time constant of 10.5 ps (relative amplitude $a_2 = 0.22$). The second model, shown in Figure S8b attempts to fit the entire data to a bi-exponential convolved with a Gaussian IRF. The amplitudes and time constants of the bi-exponential as well as the FWHM of the IRF are allowed to vary. This model produced slightly shorter time constants, and a slightly longer IRF width than that extracted in Section 1 (parameter values shown in the inset of Figure S8b). However, the quality of the fit is noticeably worse than the bi-exponential model after 400 fs. The important features of the data we would like to highlight here are the presence of two timescales, which we report as 1 and 10 ps, respectively. This behavior is consistent with similar measurements on coumarin 6.[4–6]

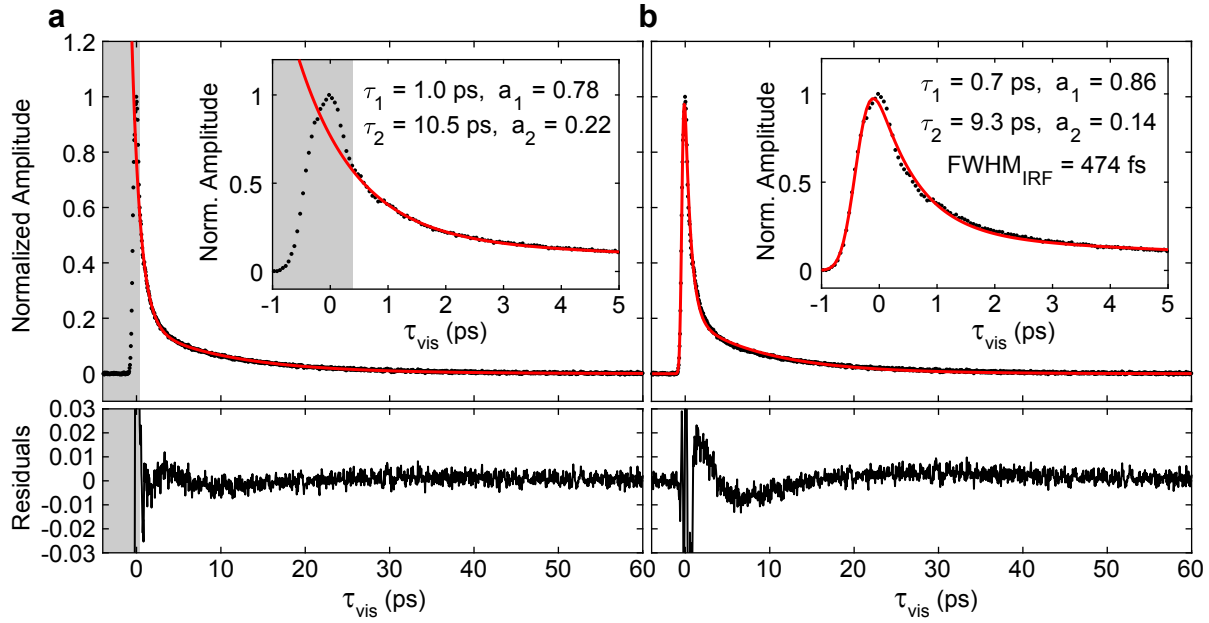

Figure S8: Fits to the 1-pulse FEIR transient. (a) Bi-exponential model. The data (black dots) are only fit after 400 fs (excluded range shaded gray), with the fit shown in red. The inset shows the first 5 ps. Time constants and relative amplitudes from the fit are listed in the inset. The lower panel shows the residuals. (b) Bi-exponential convolved with Gaussian model.

6 Continuous scanning acquisition of FEIR signals

Continuously scanning optical delays is an attractive option for collecting time domain data due to the reduction of experimental 'dead time' involved in communicating with delay stages and signal acquisition electronics. Besides decreasing the time required to sample a given number of laser shots over a range of delay values, continuous scanning can also dramatically improve signal to noise ratios if the delays are varied faster than the slow fluctuations in signal due to, for example, laser power or mechanical drift in the spectrometer. As a result, the actual speed up in acquisition time required to achieve a given signal to noise can be immense. This is particularly true in our case as the 1 MHz repetition rate is high enough to facilitate very fast scan speeds while still adequately sampling the data in real time.

Our method for fast scanning the IR delay stage (AeroTech ANT95-50L-MP) while collecting shot-to-shot data at 1 MHz has been described in Penwell *et al.*[7] Briefly, we employ the Position Synchronized Output (PSO) of the stage controller (AeroTech A3200 Npaq), which produces electronic TTL pulses synced to predefined increments (50 nm in our case) of the stage's internal encoder within a specified range of travel. The PSO waveform is used as a grid over which signal acquisition events may be referenced to in real time. How this is accomplished depends on the character of the signal channel.

In the case of the IR intensity in the dark arm of the Mach-Zehnder interferometer (MZI), a counter on the data acquisition (DAQ) card (National Instruments PCIe-6361) enumerates incoming edges of the PSO waveform while the stage moves at constant velocity. A sync waveform from the laser is then used to simultaneously trigger samples of the IR detector (Boston Electronics PVM10.6) voltage and the PSO counter value. The PSO count is then interpolated up to laser sync sampling rate to generate the position axis (see Penwell *et al* for details).

For the fluorescence channel, a second counter on the DAQ monitors the TTL output of the single photon avalanche photodiode (SPAD), i.e. counts incoming photon arrival events. The PSO counter is then used to trigger samples of the cumulative photon count. The increment in photon count between successive PSO triggers is therefore the number of photons detected within the corresponding 50 nm window of stage travel, which is sufficiently small ($\sim$0.33 fs) compared to any changes in the time domain signal at IR optical frequencies. For each repetition of the scan new photon arrivals are accumulated in their respective windows, and the data builds up as a histogram of delay bins.

The accuracy of the stage's encoder in general depends on the direction and speed of travel, so care must be taken to separately process data from each direction and use the same scan speed for the 2-pulse FEIR signal and IR reference data. Small periodic errors in the stage encoder produce spectral artifacts that appear as satellite features in FT data, analogous to 'Rowland ghosts' encountered with diffraction gratings.[8,9] These periodic errors are reproducible and may be corrected, for example by tracking the optical phase evolution of a HeNe tracer beam via the analytic signal representation of its interferogram. In our case these satellite features only appear well outside the bandwidth of our IR pulses, and we leave the positions uncorrected.

A limitation of our current aquisition procedure is that the fluorescence and IR channels cannot be detected simultaneously by our DAQ card. When performing an FT FEIR experiment, the IR reference (dark arm intensity of the MZI) is collected either immediately before or after collecting the FEIR signal. The scan speed and range of positions are kept the same. Figure S9a shows an example of time domain 2-pulse FEIR data with the corresponding IR reference (both from the

forward direction of stage travel). The phase extracted from the IR reference is used to determine the absolute timing between the IR pulse pair.[10,11] In practice, the IR reference is interpolated to the same position axis as the FEIR signal, and both are Fourier transformed using an approximate τ_{IR} time zero. The frequency domain FEIR data is divided by the complex phase factor from the IR reference and transformed back to the time domain, where the now correctly assigned negative time data is removed. A final FT of the one-sided data produces the FT FEIR spectrum as its real part.

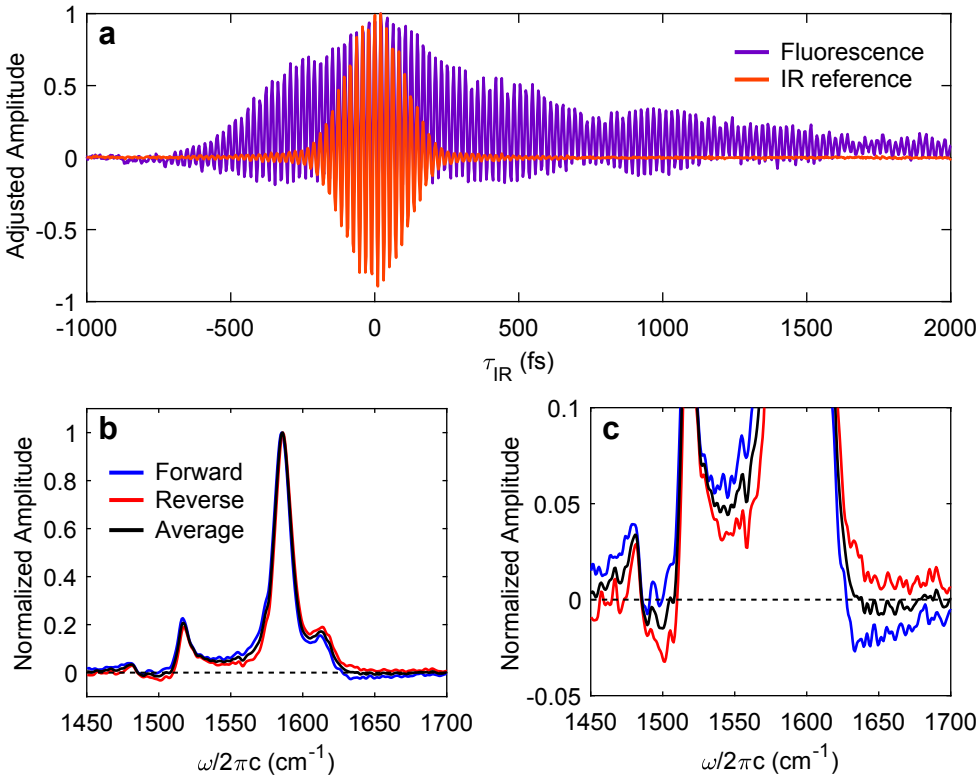

Figure S9: Referencing of FT FEIR data and residual phase error. (a) 2-pulse FEIR (purple) and IR reference (orange) data in the time domain from forward direction scans (same data as main text Figure 2b). The constant offset of both traces has been removed, and the IR reference has been multiplied by -1 to account for the π phase shift between the primary and dark outputs of the MZI. In both cases the data has been binned up to 3.3 fs steps (factor of 10 above the base encoder resolution) for ease of visualization. (b) FT FEIR spectra from the forward direction (blue), reverse direction (red), and their average (black). (c) Detail of the baseline from the spectra in (b). The black dashed line along the baseline has been added to guide the eye.

This procedure is performed separately for the forward and reverse directions of stage travel. As shown in Figures S9b and c, there are typically small residual phase errors in the FT FEIR features that vary between the two scan directions. This phase artifact is typically only visible below the 1-3% level relative to the maximum signal amplitude, and is not always oppositely signed between directions (which however roughly happens to be the case in Figure S9). We suspect this artifact may originate from or at least be exacerbated by the fact that the IR reference is not collected simultaneously with the FEIR data, allowing for instrumental drift (e.g. in the pathlength difference) between acquisitions. The character of this phase error is, however, also consistent with

a thermal effect, as discussed in Section 7. In practice, we average the forward and reverse spectra together, which often roughly cancels this effect. Nevertheless, we emphasize that the amplitude of the FT FEIR spectra shown in this work cannot be accurately interpreted below the few percent level, even if the signal to noise is significantly better.

All FEIR data shown in the main text uses a scan speed of 2 mm/s. The 2-pulse FT FEIR data uses a scan range of roughly -3 to 13 ps, which produces a base frequency resolution of 1.3 cm^{-1} before zero padding. Each scan takes roughly 1.8 seconds per direction (including all down time during stage turn around and signal processing), and typical data sets are accumulated over a few hundred scans (e.g. 239 scans per direction in Figure S9). The 50 nM FT FEIR spectrum in main text Figure 3a is the result of 351 scans/direction (25.9 minutes of total aquisition time). The 1-pulse FEIR transient shown in Figure S8 is limited to a maximum scan length of 60 ps by the presence of an unwanted IR 'ghost' pulse from a back reflection in the $LiGaS_2$ OPA crystal. This data was collected with 50 scans/direction (11.5 minute acquisition time). The forward and reverse directions were averaged together and then binned up to 50 fs steps. The 10 nM 1-pulse data in main text Figure 3b is the result of 381 scans/direction out to 13 ps (28.6 minute acquisition time).

7 Thermal signatures in FEIR data

As mentioned in Section 3, the signal and background count rates appear to be sensitive to the steady state heat load produced by absorption of the IR pulses. This can be seen more dramatically by comparing the IR power dependence of the FEIR signal measured in solvents of varying transparency. As shown in Figure S10, the power dependence is roughly linear in deuterated acetonitrile and chloroform, which have good transparency in the frequency range of the IR field. However, in ethanol, which has substantial absorption across the entire pulse spectrum, the power dependence is markedly sub-linear. This sub-linear trend does not depend on the encoding delay τ_{vis} along the transient, that is, it is not directly related to the ultrafast or picosecond response of the sample. We rationalize this behavior as a decrease in overall fluorescence collection efficiency as a function of the temperature profile in the sample. It is important to note that the glass coverslip, which is always present in our FEIR experiments, absorbs the majority of IR light incident on it (Figure S10b). However, this appears to have only a minor contribution to the thermal effect, as suggested by the approximately linear IR power dependence for the clear solvents.

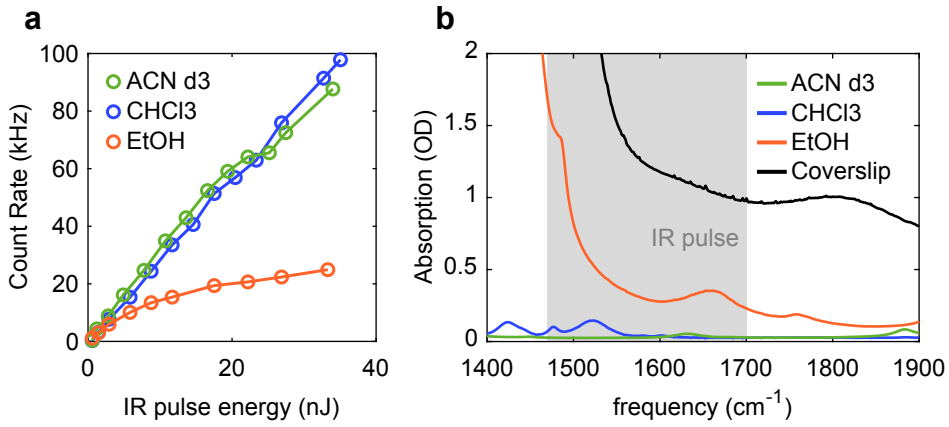

Figure S10: Solvent dependence of the thermal effect. (a) IR power dependence of the S_1 signal for coumarin 6 in deuterated acetonitrile (green), chloroform (blue), and ethanol (orange). (b) FTIR spectra of the solvents (50 μm pathlength) and the coverslip (black). The region with appreciable IR field intensity is shaded gray.

While we cannot yet assign the exact origin of this decrease in collection efficiency, we suspect it is due in part to the temperature-dependence of the sample's refractive index and/or the dye's quantum yield. To date we have not directly measured the temperature in the sample when the IR beam is present. Due to the sensitively index-dependent performance of high-NA microscope objectives, we suspect an index effect (e.g. temperature lensing) plays an important role.

Without a detailed mechanism for the thermal effect in mind, we may nevertheless characterize its influence on our FEIR measurements. As the modulation of collection efficiency appears to correlate to the amount of IR optical power absorbed by the sample, the pulse-pair Fourier transform FEIR measurements should spectrally resolve this heating response. By introducing a strong but spectrally sharp and isolated absorption into an otherwise highly transparent sample we may attempt to characterize the resulting thermal effect in a more controlled manner.

As a preliminary study we perform FT FEIR measurements on coumarin 6 in a 10:1 mixture of chloroform and acetone. The dilute acetone carbonyl stretch has a narrow, intense absorption at

1715 cm^{-1}. We tune the frequency of the IR pump to 1680 cm^{-1}, which falls in a high transparency window of chloroform, but covers the two highest frequency coumarin 6 ring modes and the acetone carbonyl stretch (shown schematically in Figure S11a). The optical densities of the dye ring modes are negligible compared to the acetone carbonyl (at typical FEIR concentrations they are below the detection sensitivity of our FTIR spectrometer). The acetone carbonyl mode should therefore act as both a thermal 'source' and 'probe'.

To help interpret these measurements we will describe a minimal model for the thermal effect. We proceed by assuming that the collection efficiency modulation is a purely bulk optical effect and that the thermal load does not influence the microscopic FEIR response. That is, we may write the total measured fluorescence in an FT FEIR experiment (encoding delay fixed at $\tau_{\mathrm{vis}} = 0$) as

$$F(\tau_{\mathrm{IR}}) = S(\tau_{\mathrm{IR}})M(\tau_{\mathrm{IR}}) \tag{2}$$

where $S(\tau_{\mathrm{IR}})$ is the FEIR response of the dye (i.e. what would be measured in the absence of any thermal effects), and $M(\tau_{\mathrm{IR}})$ is the thermal modulation. We will also assume that the magnitude of the thermal modulation depends linearly on the average IR power absorbed by the sample, so that

$$M(\tau_{\mathrm{IR}}) = 1 - \alpha P(\tau_{IR}) \tag{3}$$

where $P(\tau_{IR})$ is proportional to average IR power absorption and α is a scaling parameter that determines the strength of modulation. The thermal load most often appears to decrease the collection efficiency, so we set $\alpha > 0$. The case of $\alpha > 1$ is clearly unphysical, and for large modulations the linearity assumed in Eq. 3 most likely breaks down. The (normalized) IR power absorption as a function of interferometer delay τ_{IR} is given by

$$P(\tau_{\mathrm{IR}}) = \frac{1}{2}\Big(1 + I(\tau_{\mathrm{IR}})/I_{\max}\Big) \tag{4a}$$

$$I(\tau_{\mathrm{IR}}) = \mathrm{FT}^{-1}\Big\{S_{\mathrm{pump}}(\omega)S_{\mathrm{solv}}(\omega)\Big\} \tag{4b}$$

which is simply the interferogram of the solvent absorption spectrum $S_{\mathrm{solv}}(\omega)$ windowed by the IR pump spectrum $S_{\mathrm{pump}}(\omega)$. The notation here is meant to convey that the power absorption is constant at half its maximum for large τ_{IR} (no interference), and goes to zero for perfectly destructive interference (approximately the case at the first π out-of-phase fringes).

Implicit in Eq. 3 is the assumption that the thermal profile is at a steady state for every interferometer delay position. This means we are considering the cumulative time-averaged effect of absorbing the 1 MHz pulse train balanced by the sample's macroscopic thermal relaxation. In our first acetone experiment we will attempt to enforce steady state conditions by stepping, rather than continuously scanning, the interferometer delay stage, which allows the sample to achieve its new steady state temperature profile before signal is collected. After discussing this steady state case, we will incorporate the effects of rapidly scanning the optical delay before the temperature profile

fully relaxes.

Figure S11b shows the time domain FEIR response S, thermal modulation M, and resulting total signal F for the steady-state model. Here the FEIR response is the *total* signal and background, $S(\tau_{\mathrm{IR}}) = S_0 + S_1(\tau_{\mathrm{IR}}) + S_2 + S_{12}(\tau_{\mathrm{IR}})$, as shown for example in main text Figure 2b. For simplicity the two ring modes have been given exponential dephasing times that roughly reproduce the experimental linewidths, and the relative amplitudes of the signal and background contributions are likewise set to conform to the experimental time domain data. The acetone carbonyl absorption is also given a lorenztian lineshape.

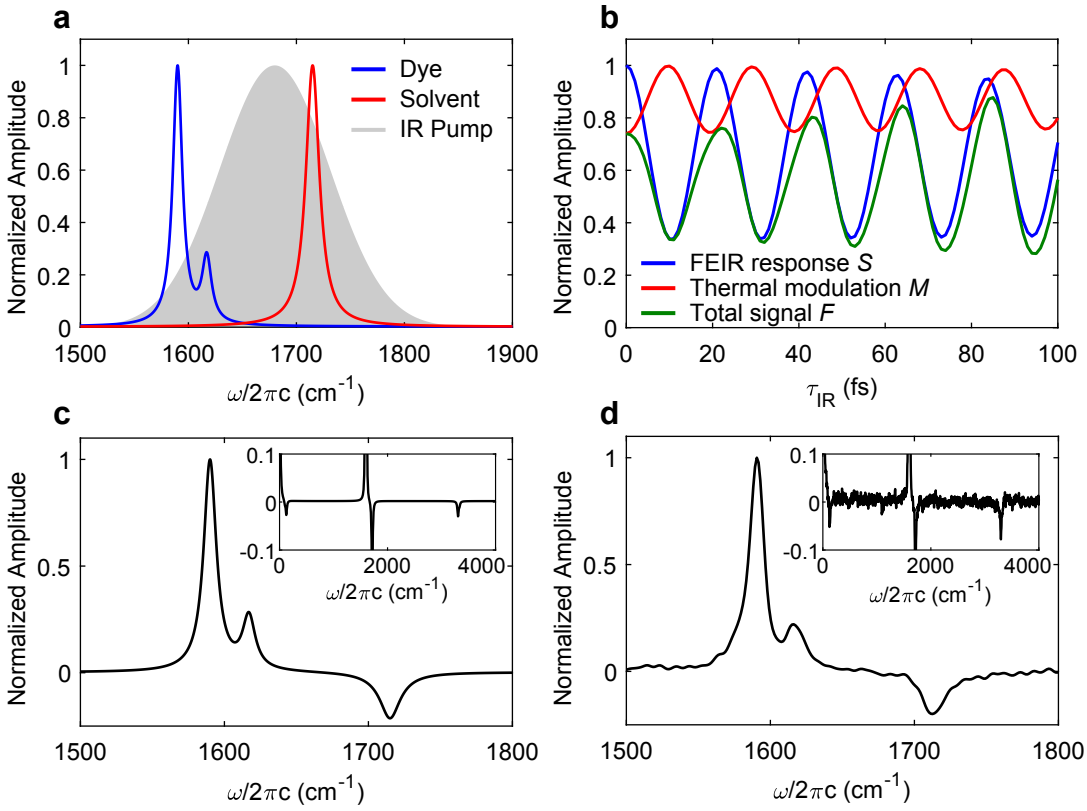

Figure S11: Steady state thermal signatures in an FT FEIR spectrum, (a - c) model and (d) data. (a) Frequency domain FEIR response of the dye ring modes (blue), absorption of the acetone carbonyl stretch in the solvent (red), and IR pulse spectrum (gray). (b) Time domain FEIR dye response S (blue), thermal modulation M from the acetone carbonyl (red), and the resulting total detected signal F (green). (c) FT FEIR spectrum calculated from F. The inset displays the baseline of the same spectrum on an extended frequency axis to show the AM sidebands. (d) Measured FT FEIR spectrum from the dilute acetone experiment, with inset displaying the same region as in (c).

The resulting model FT FEIR spectrum is shown in Figure S11c. The effect of the acetone carbonyl absorption appears as a negative feature at the same frequency (1715 cm^{-1}). This can be understood within the language of amplitude modulation (AM) in analog signal processing as being due to the 'baseband' component of the FEIR response, $S_0 + S_2 + S_1$, which has been modulated up to the frequency of M. It is negative because of the π phase shift between the modulation and the S_{12} signal (evident in the red and blue curves in Figure S11b), which physically arises because the thermal load decreases the fluorescence count rate. The modulation of the S_{12} contribution

produces AM sidebands at the sum and difference frequencies between S_{12} and M, which can be seen in the FT FEIR spectrum near 100 cm^{-1} and 3300 cm^{-1} (inset of Figure S11c). The experimental FT FEIR spectrum (Figure S11d) shows both the negative 'modulated baseband' feature as well as the AM sidebands. To get qualitative similarity between the model and data, the strength of the modulation was set to $\alpha = 0.13$.

We now describe the thermal effect for the case when FEIR data is collected by continuously scanning the interferometer stage, as described in Section 6. To do this we need to consider the finite timescale of the sample's thermal relaxation in relation to the speed at which the average power absorbed from the IR pulse-pair varies with the interferometer delay. Intuitively, one would expect that if the delay is scanned through interference fringes faster than the thermal profile can relax, then the thermal modulation will become partially washed out and lag behind the optical interference.

To capture this effect mathematically, it is convenient to describe the delay in stage displacement units, $\delta = 2c\tau_{\mathrm{IR}}$, where c is the speed of light. When the interferometer stage is continuously scanning, the displacement δ is a function of laboratory time t. We proceed by modifying Eq. 3, which is now also ultimately a function of t. Retaining the assumption of linearity, the modulation is rewritten as

$$M(t) = 1 - \alpha \int_{-\infty}^{\infty} H(t')P(t - t')dt' \tag{5}$$

where $H(t)$ is a response function that describes the sample's macroscopic thermal relaxation. For simplicity we use an exponential decay with time constant τ_R

$$H(t) = \theta(t)\frac{1}{\tau_R}e^{-t/\tau_R}, \tag{6}$$

where the Heaviside step function $\theta(t)$ has been included to enforce causality. The other quantities in Eq. 5 retain their original meaning, and Eqs. 2 and 4 still hold in their original forms after the appropriate change of variables. For example, during fast scan data acquisition the stage moves at constant velocity

$$\delta = \delta_0 + vt \tag{7}$$

where the velocity v has opposite signs during forward ($v > 0$) and reverse ($v < 0$) travel directions. The IR power absorption P is calculated from Eq. 4, transformed to lab time by $\tau_{\mathrm{IR}} = (\delta_0 + vt)/2c$, and then inserted into Eq. 5 to calculate the resulting thermal modulation M. After transforming M back to τ_{IR} units, Eq. 2 is used to find the total signal F.

The important aspect of this case is that the thermal modulation now depends on the speed and direction of stage travel. As shown in Figure S12a, the time (or pathlength) domain thermal modulations for the forward and reverse directions have reduced amplitudes and oppositely signed phase shifts with respect to the steady state modulation. Here, the thermal time constant τ_R and

the scan speed $|v|$ were chosen so that the timescale of thermal relaxation (shown by H plotted in pathlength units) is similar to the timescale of the IR optical phase evolution (shown for example by the fringes of the steady state modulation). Figure S12b shows the resulting effect on the calculated FT FEIR spectrum. The scan velocity dependent phase shift acquired by the thermal modulation manifests as a phase twist, or partially dispersive lineshape, in the negative solvent feature, while the amplitude reduction is also evident. The same effect appears in the AM sidebands (not shown).

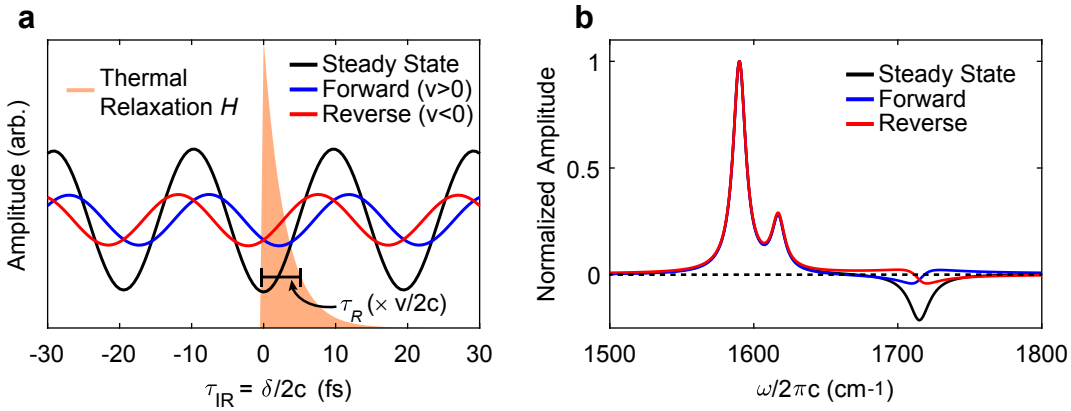

Figure S12: Model of thermal modulation with fast scanning of the IR delay. (a) Thermal amplitude modulation in the time or pathlength domain for fast scanning the delay in the forward (blue) and reverse (red) directions, compared to the steady state case of stepping the delay (black). The thermal response function H (orange) used to calculate the forward and reverse modulations is plotted in transformed units. (b) Effect on the FT FEIR spectrum. The negative solvent heating feature has reduced amplitude and acquires a phase twist with opposite signs for the forward and reverse directions.

In Figure S13 we compare FT FEIR data from the dilute acetone experiment collected alternately by stepping and fast scanning the delay. The scan direction dependent phase twist of the negative acetone feature is evident in the fast scan spectra, and is qualitatively similar to the model in Figure S12b. The distinct character of the amplitude modulation between steady state stepping, forward scans, and reverse scans can be seen through the beating it produces in the time domain.

As the speed of stage travel is increased, the rate of IR optical phase evolution will eventually surpass the thermal relaxation timescale of the sample by enough to completely wash out the interferometric structure of the thermal modulation, and the resulting uniform scaling will have no effect on the FT FEIR spectrum. On the other hand, for very slow stage motion the thermal profile in the sample will keep pace with the fringes of the IR power absorption, and the steady state behavior from Eq. 3 will be recovered. In our model we may define an intermediate point between these regimes by considering the scan speed for which the thermal relaxation time is roughly equal to a half cycle of the interferometric phase evolution, or

$$v_s = \frac{1}{\tilde{\nu}\tau_R} \tag{8}$$

where $\tilde{\nu}$ is the center frequency (in wavenumbers) of the solvent feature responsible for IR power absorption. In Figure S14 we show the scan speed dependence of the thermal signature in the

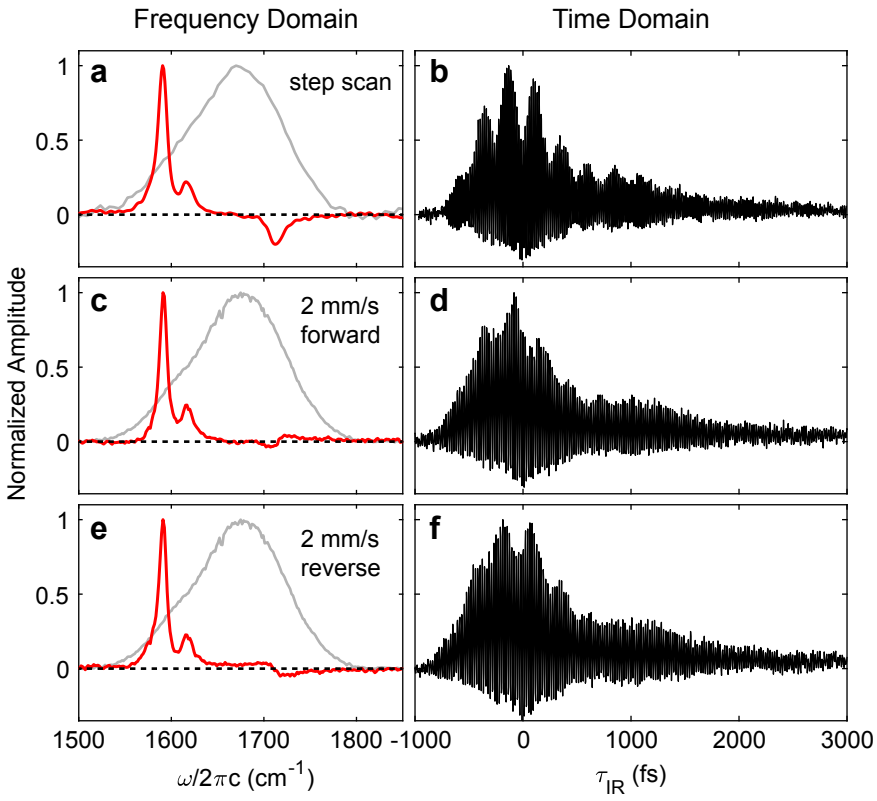

Figure S13: Thermal signatures in FT FEIR data with step and fast scan acquisition. (a - b) 2-pulse FEIR data in the frequency and time domain, respectively, for the dilute acetone experiment in the steady state case of stepping the delay (same data as Figure S11d). In the frequency domain panel the IR pulse spectrum in shown in grey, and the dashed black line along the baseline is included to guide the eye. (c - d) FEIR data from the forward direction of a fast scan with 2 mm/s stage speed. The time domain data has been shifted to remove the constant offset ($S_0 + S_2$ contributions). (e - f) Reverse direction from the same fast scan experiment. The time domain fast scan data in (d) and (f) has been binned up to 3.3 fs steps for visual comparison to the stepped data in (b).

FT FEIR spectrum over roughly three orders of magnitude for the dilute acetone experiment and model. Because the frequency of the acetone carbonyl and set of scan speeds are fixed at known values from the experiment, we may estimate the thermal time constant by adjusting its value to produce the best agreement between the set of experimental spectra and a set of model spectra with the same speed ratios between scans. This results in a thermal relaxation time constant of roughly 20 ms. At the fastest experimental scan speed of 10 mm/s, the thermal signature, although much reduced in amplitude, is still visible. This speed is near the practical upper limit of our acquisition procedure, which indicates that we will not yet be able to access the 'fast travel limit' where artifact-free FT FEIR spectra could be collected even in the presence of large thermal effects.

We conclude this section with a brief summary of our current understanding of how heating effects can influence FEIR measurements, and comment on the implications and outlook for the application of FEIR in the future. Our observations are consistent with the presence of thermally-induced amplitude modulation of the bulk fluorescence collection efficiency. This amplitude modulation is well correlated to the average optical power the sample absorbs from the IR pulse train. As

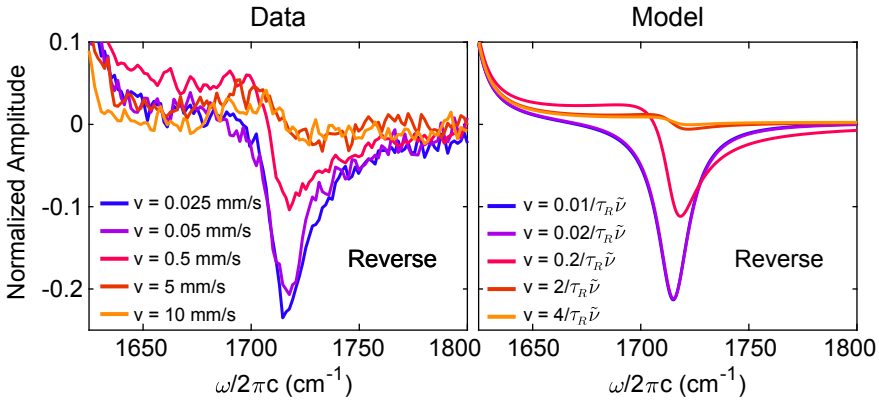

Figure S14: Scan speed dependence of the thermal effect. (a) Detail of the acetone thermal signature in experimental FT FEIR spectra for scan speeds between 25 μm/s and 10 mm/s in the reverse direction. (b) Model calculations with the same ratios between scan speeds. Speeds are reported in units of v_s as defined in Eq. 8. This set of speeds was chosen to produce the best qualitative agreement between the calculated and measured spectra, and results in the assignment of $\tau_R = 20$ ms when numerical values for the acetone absorption frequency and stage speeds are considered.

a result, if the IR absorption changes as a function of the experimentally varied optical delays, then the thermal modulation will introduce a signature, or artifact, into the measured signal. For example, FEIR experiments using an IR pulse-pair are susceptible to this artifact while the 1-pulse transient experiments are not, because average IR power absorption does not vary with the visible encoding delay. The timescale with which the thermal profile relaxes has been estimated by considering how the amplitude modulation is affected by stage velocity during continuous scanning data acquisition, and appears to be on the order of tens of milliseconds. This timescale corresponds to the sequential arrival of 10,000 IR pulses, which is consistent with our picture of 'steady state' heating due to the cumulative effect of the entire pulse train.

We rationalize our current experiment's susceptibility to thermal effects by the combination of large average IR power density and sensitive high-NA optical system for fluorescence collection. Indeed, it is interesting to compare our FEIR microscope to the instrumentation used in photothermal spectroscopy and imaging experiments (see for example Li *et al*[12]). Mitigating thermal effects will be crucial for extending the application of FEIR to condensed phase systems that lack high IR transparency, most notably to aqueous solutions. Utilizing sample flow and exploring focusing conditions that enable reductions in average IR power are possible routes forward.

References

(1) Diels, J.; Rudolph, W., *Ultrashort Laser Pulse Phenomena*, 2nd ed.; Elsevier Science: 2006.

(2) Höbel, M.; Ricka, J. Dead-time and Afterpulsing Correction in Multiphoton Timing with Nonideal Detectors. *Rev. Sci. Instr.* **1994**, *65*, 2326–2336, DOI: 10.1063/1.1144684.

(3) Widengren, J.; Mets, Ü. In *Single Molecule Detection in Solution*, Zander, C., Jorg, E., Keller, R. A., Eds.; Wiley-VCH Verlag GmbH & Co. KGaA: 2003; Chapter 3, pp 69–120.

(4) Hübner, H.; Wörner, M.; Kaiser, W.; Seilmeier, A. Subpicosecond Vibrational Relaxation of Skeletal Modes in Polyatomic Molecules. *Chem. Phys. Lett.* **1991**, *182*, 315 –320, DOI: 10.1016/0009-2614(91)80221-I.

(5) Dahinten, T.; Baier, J.; Seilmeier, A. Vibrational Energy Transfer Processes in Dye Molecules after Ultrafast Excitation of Skeletal Modes. *Chem. Phys.* **1998**, *232*, 239–245, DOI: 10.1016/S0301-0104(98)00072-X.

(6) Van Wilderen, L.; Messmer, A.; Bredenbeck, J. Mixed IR/Vis Two-Dimensional Spectroscopy: Chemical Exchange beyond the Vibrational Lifetime and Sub-Ensemble Selective Photochemistry. *Angew. Chem. Int. Ed.* **2014**, *53*, 2667–2672, DOI: 10.1002/anie.201305950.

(7) Penwell, S.; Whaley-Mayda, L.; Tokmakoff, A. Single-Stage MHz Mid-IR OPA Using $LiGaS_2$ and a Fiber Laser Pump Source. *Opt. Lett.* **2018**, *43*, 1363–1366, DOI: 10.1364/OL.43.001363.

(8) Jenkins, F.; White, H., *Fundamentals of Optics*, 4th ed.; McGraw-Hill: 1976.

(9) Gale, H. Rowland Ghosts. *Astrophys. J.* **1937**, *85*, 49–61, DOI: 10.1086/143800.

(10) DeFlores, L.; Nicodemus, R.; Tokmakoff, A. Two-Dimensional Fourier Transform Spectroscopy in the Pump-Probe Geometry. *Opt. Lett.* **2007**, *32*, 2966–2968, DOI: 10.1364/OL.32.002966.

(11) Helbing, J.; Hamm, P. Compact Implementation of Fourier Transform Two-Dimensional IR Spectroscopy without Phase Ambiguity. *J. Opt. Soc. Am. B* **2011**, *28*, 171–178, DOI: 10.1364/JOSAB.28.000171.

(12) Li, Z.; Aleshire, K.; Kuno, M.; Hartland, G. Super-Resolution Far-Field Infrared Imaging by Photothermal Heterodyne Imaging. *J. Phys. Chem. B* **2017**, *121*, 8838–8846, DOI: 10.1021/acs.jpcb.7b06065.