Anisotropic Dynamics of an Interfacial Enzyme Active Site Observed Using Tethered Substrate Analogs and Ultrafast 2D IR Spectroscopy

Tayler D. Hill, Sunil Basnet, Hannah H. Lepird, Blaze W. Rightnowar, and Sean D. Moran

School of Chemical and Biomolecular Sciences, Southern Illinois University Carbondale, 1245 Lincoln Drive MC 4409, Carbondale, IL 62901

KEYWORDS. Enzyme, dynamics, hydration, 2D IR spectroscopy, Fermi resonance.
ABSTRACT:

A number of recent reports have implicated ultrafast (femtosecond-picosecond) timescale motions in enzymatic activity, but relatively few experimental studies have addressed complications arising from spatially-distributed disorder, multiple substrate binding modes, or the influence of hydration dynamics on solvent-exposed active sites. Here we use ultrafast two-dimensional infrared (2D IR) spectroscopy and covalently-tethered substrate analogs to examine dynamical properties of the *Pyrococcus horikoshii* ene-reductase (PhENR) active site in two defined binding configurations. Frequency-fluctuation correlation functions of aryl-nitrile analogs reveal an end-to-end tradeoff between fast (sub-ps) and slow (>5 ps) motions. Lineshape and quantum beat analyses of Fermi resonant aryl-azide analogs demonstrate that this is an intrinsic property of the water-exposed active site. This study indicates that elements of polar pre-organization are maintained at the interface and suggests several plausible factors leading to state-selective rate enhancement and promiscuity in PhENR.
I. Introduction.

Enzyme active sites function as specialized solvent environments that allow a reacting species to sample transition state (TS) configurations orders-of-magnitude more frequently than in bulk aqueous solution.\(^1\)\(^{-5}\) Encapsulation of a reactant in a preorganized enzyme active site lowers the penalty for solvent reorganization along the reaction coordinate and stabilizes TS configurations via specific electrostatic interactions.\(^2\),\(^6\),\(^7\) In recent years, there has been increasing interest in the role of enzyme dynamics in catalysis, and although it is well-established that relatively slow (> ns) conformational motions are necessary to proceed from the reactant state (RS) to the product, the influence of fast (fs-ps) fluctuations on TS sampling remains a topic of substantial debate.\(^2\),\(^3\),\(^5\),\(^8\)-\(^22\) Depending on the system, the interplay between reactant, protein, and hydration dynamics may play an important role in enzymatic function.

Ultrafast two-dimensional infrared (2D IR) spectroscopy is a powerful probe of solvation because it combines sensitivity to electrostatics and vibrational coupling with the time resolution needed to observe rapid fluctuations in an ensemble.\(^23\) 2D IR studies of small molecules and metal complexes in solution have provided detailed insight into effects of viscosity on isomerization\(^24\) and possible relationships between short-timescale equilibrium dynamics and energy transfer in reactivity.\(^25\)-\(^27\) For enzymes and other macromolecules, vibrational probes that absorb in a ‘transparent window’ of the mid-IR spectrum are attached to amino acid side chains, reactive cofactors, or substrate/T莫斯.\(^28\)-\(^46\) Frequency fluctuation correlation functions (FFCFs) extracted from spectral diffusion measurements provide detailed descriptions of the timescales and magnitudes of local fluctuations.\(^30\)-\(^33\),\(^37\),\(^40\),\(^42\),\(^45\),\(^47\),\(^48\) Among many examples, this strategy has been used to relate active site motions to substrate selectivity in cytochrome P450\(_{cam}\)\(^33\) and to
observe underdamped scaffold vibrations in a TS-analog complex of formate dehydrogenase (FDH). Similar approaches have been used to characterize hydration dynamics at protein surfaces; for instance, in hen egg white lysozyme, low-frequency protein motions appear to be particularly susceptible to solvent slaving in agreement with several studies on the influence of hydration dynamics on conformational motion.

In systems with occluded active sites, the interpretation of the role of enzyme-as-solvent is relatively straightforward; although active site motions may be influenced by hydration dynamics at the surface, reacting species are largely shielded from dissipative fluctuations of the solvent bath. In contrast, the dynamics of active site residues and water are likely to be strongly coupled when the active site is solvent-exposed. Nevertheless, both the heterogeneity of protein surfaces and the pre-organization principle predict that such active sites will have anisotropic distributions of electric field strengths, motions, and conformational disorder over the distance of a few chemical bonds. It is reasonable to assume that where two (or more) substrate functional groups ‘participate’ in a reaction, evolution will have optimized the local characteristics of the extended (i.e., including H₂O) active site.

Here we apply 2D IR spectroscopy to measure local dynamics and disorder in the exposed active site of *Pyrococcus horikoshii* ene-reductase (PhENR). PhENR is a promiscuous redox enzyme that catalyzes the reduction of α,β-unsaturated carbonyl compounds via hydride transfer from its flavin mononucleotide (FMN) cofactor and proton transfer from a nearby side chain or H₂O. The native homodimer contains two active sites in moderately deep, pseudo-symmetric clefts lined with aromatic (Phe/Tyr/Trp) and basic (His/Arg/Lys) side chains. In a crystal structure reported by Steinkellner et al., the substrate 2-cyclohexenone (2CH) binds in two opposing
orientations where near-equal occupancy indicates that the binding modes are isoenergetic in the RS ground-state. Tentative assignments of ‘inactive’ and ‘reactive’ configurations were made based on putative proton donor-acceptor distances (DADs), but equilibrium geometries are unreliable predictors of reactivity in mobile systems. To gain insight into fluctuations about equilibrium, we constrained vibrationally isolated aryl-nitrile and fermi resonant aryl-azide substrate analogs in each orientation via covalent ligation. Time-resolved 2D IR spectra of each label reveal varying degrees of mobility and long-timescale disorder across the active site, largely irrespective of the chemical identity of the probe. This ‘dynamical anisotropy’ is an intrinsic property of the active site that has several plausible implications for catalytic activity.

II. Experimental Procedures.

A. Protein expression and purification. A codon optimized gene for PhENR containing an N-terminal His\textsubscript{10} affinity tag and a thrombin cleavage site (IDT, Coralville, IA) was sub-cloned into pET16b (Novagen, Merck, Darmstadt, Germany) and single-cysteine mutants were produced via site-directed mutagenesis with C113S combined with either H124C or H155C substitutions. BL21(DE3)pLysS cells transformed with the plasmid were grown to OD\textsubscript{600} = 1.0 at 37 °C, with shaking, in LB broth supplemented with 100 μg/mL ampicillin and 30 μg/mL chloramphenicol. Expression was induced with 1 mM IPTG and continued overnight (~16 hours) at 37 °C. Cells were pelleted, resuspended in phosphate buffer (20 mM sodium phosphate, 100 mM NaCl, pH 7.5), and lysed using a French Press. Insoluble cell debris was removed by centrifugation (25,000 x g, 1 h), and the enzyme was purified using immobilized metal (Co\textsuperscript{2+}) affinity chromatography on an AKTA Start FPLC (GE, Boston, MA). Following step gradient elution using phosphate
buffer containing 250 mM imidazole, the pure protein was exchanged back to 20 mM sodium phosphate, 100 mM NaCl, pH 7.5 on a GE HiPrep 26/10 desalting column.

**B. Vibrational Label Synthesis and Enzyme Labeling.** 4-cyano-N-phenylmaleimide (4CN-M) and 4-azido-N-phenylmaleimide (4Az-M) were synthesized from commercially available precursors using the method of Cava et al. The labels were attached to single-cysteine PhENR mutants (His124Cys or His155Cys) by adding a fivefold molar excess from a N,N-dimethylacetamide (NNDMA) stock to buffered enzyme solutions and mixing gently at room temperature for 16 hours in the dark. Insoluble unreacted labels were removed by centrifugation and size exclusion chromatography. The incorporation of the label was monitored by UV-vis spectroscopy of the FMN cofactor and labeling efficiencies were determined using Ellman’s test. A 4Az-PEG2000 model compound was prepared from 4Az-M and tetrapodal thiol-PEG2000 (Nanocs Inc., New York, NY) under similar conditions.

**C. Modeling and Simulations.** Active site and surface charge distributions were calculated via the APBS web server (https://server.poissonboltzmann.org) using experimental conditions (pH 7.5, 20 °C) and default parameters. Docking of the 4CN-M labels into the PhENR active site was performed using the Molecular Operating Environment (MOE) software package (CCG Inc., Montreal, Quebec). The p-hydroxybenzaldehyde inhibitor was deleted from the structure of the PhENR homodimer (PDB ID: 3ZOC) and the active site was defined by the FMN, one histidine residue (His124 or His155), the mutant cysteine in the opposing position (Cys155 or Cys124), Tyr4, Arg5, Tyr8, Trp30, Phe36, Arg49, and Trp158. Initial docking was performed with the Alpha PMI placement method and London dG scoring for 30 poses. Refinement was performed using the Induced Fit method with GBV/WSA dG scoring for 10 poses. Selected cysteine-reactive
orientations are shown in Figure S2. For MD simulations, the R and S enantiomers of 4-cyano- and 4-azido-N-phenylsuccinidylcysteine were built and inserted in place of either His124 or His155. Each model structure was protonated using 3D protonation default settings at pH 7.5. The substituted phenyl group of each model was positioned near the coordinates of the docked labels (and 2CH) with three-point restrained minimization steps and the structure was relaxed with a final unrestrained minimization to an essential RMS gradient of 0.001. Minimization steps used a generalized Born solvation model with a cutoff of 10 Å. The systems were solvated with rigid water and 0.1 M NaCl in a periodic 75 Å box and MD simulations were performed using the AMBER10:EHT force field (MOE)\textsuperscript{66, 67} in AMBER22 running particle mesh Ewald molecular dynamics (PMEMD).\textsuperscript{68} The structure was equilibrated to 293 K in the NVT ensemble and then to 293 K and 100 kPa in the NPT ensemble in successive 1 ns steps. Unrestrained 25 ns runs (2.0 fs steps) were performed at 293 K and 100 kPa. All simulations were performed on a workstation equipped with a 32-core AMD RYZEN 5950 processor and a Nvidia GeForce RTX 3070 Ti GPU.

**D. FTIR spectroscopy.** Unlabeled and labeled PhENR variants in 20 mM sodium phosphate buffer (pH 7.5) with 100 mM NaCl were concentrated to ~6 mM in centrifugal concentrators. Protein solutions were placed between two 2 mm CaF\textsubscript{2} windows separated by 50 μm PTFE spacers in demountable transmission cells, yielding a nominal pathlength of 50 μm. A total of 256 scans were collected at room temperature using a JASCO 6800 FTIR spectrometer (JASCO, Inc., Easton, MD) equipped with a single-channel, liquid nitrogen cooled mercury cadmium telluride (MCT) detector. Buffer backgrounds were subtracted and baseline correction was performed by low-order polynomial fitting in MATLAB (MathWorks, Natick, MA). Spectra of free (maleimide) labels and 4Az-PEG\textsubscript{2000} were collected using 100 mM and 10% (w/v) solutions, respectively.
E. 2D IR spectroscopy. For 2D IR spectroscopy, 1.75 mJ pulses from an Uptek Solutions (Bohemia, NY) Phidia Ti:sapphire amplifier, centered at 790 nm with a width of 100 fs, were directed into a TOPAS-PRIME optical parametric amplifier equipped with an AgGaS$_2$ difference frequency mixer (Light Conversion, Vilnius, Lithuania). Approximately 30 µJ mid-IR pulses, centered at 4500 nm (nitriles) or 4750 nm (azides), were directed into a PhaseTech Spectroscopy (Madison, WI) 2DQuick Array pulse-shaping spectrometer equipped with a 128 x 128-pixel MCT array detector. Samples were placed in demountable liquid cells as described above and data was collected at ambient conditions (20 °C). The pump beam was shaped with a Ge acousto-optic modulator (AOM) to produce pump pulse pairs. A total of 107 delays in 24 fs steps were acquired for each spectrum with a rotating frame frequency offset from features of interest by ~200 - 400 cm$^{-1}$ to ensure adequate sampling of the FIDs. Pump-probe waiting times (T) were adjusted by translation of computer controlled linear delay stage, and the pump-probe signal of an external standard in N,N-dimethylacetamide (NNDMA) was used to set $T_w = 0$. Four-frame phase cycling was used to reduce scatter. Signals were dispersed onto the detector using a Princeton Instruments Acton SpectraPro 2150 spectrograph and depending on signal strength, between 8 and 64 bins of 50 scans were averaged to achieve optimal signal:noise ratios. Time domain data were cosine apodized, zero-padded, averaged, and Fourier transformed in MATLAB to generate the 2D IR spectra. For nitrile spectra, probe frequencies were calibrated using propene tricarbonitrile (2,201 cm$^{-1}$ and 2,216 cm$^{-1}$) and 2-cyano-NNDMA (2,259 cm$^{-1}$). For azide spectra, probe frequencies were calibrated using 4-azidomethylbenzonitrile (2106 cm$^{-1}$) and hexylthiocyanate (2160 cm$^{-1}$).
III. Results and Discussion.

A. System Design and Characterization. The native PhENR homodimer contains two identical active sites that bind substrates in close proximity to a flavin mononucleotide (FMN) cofactor. A crystal structure reported by Steinkellner et al. (Figure 1a) showed that the substrate 2-cyclohexenone (2CH) binds in two opposing orientations separated by a ~180° in-plane rotation relative to the FMN. The active site is (broadly) positively charged and the bound substrate is substantially solvent-exposed (Figure 1b). In two-binding modes, the activating C=O group can form H-bonds with either His124 or His155 and the reactive C=C group lies within the typical π-stacking distance of the oxidized FMN. Although the FMN(N5) – 2CH(C3) hydride transfer distances are similar for both orientations, the equilibrium proton transfer distances (likely via Tyr4) differ. This fact was used to assign putative ‘inactive’ and ‘reactive’ states (Figure 1c), albeit without consideration of structural disorder.

Mutation of either of the two histidine residues to cysteine provides a convenient synthetic handle for ligating substrate analogs in defined orientations within the active site (Figure S1). We designed two thiol-reactive labels, 4-cyano-N-phenylmaleimide (4CN-M) and 4-azido-N-phenylmaleimide (4Az-M) for our study. Initial characterization of the labels indicates that the C≡N stretch in 4CN-M is effectively isolated from other vibrational modes, but that 4Az-M undergoes accidental Fermi resonance between the azide antisymmetric stretch and lower-lying ‘ring modes.’ Thus, the label pair acts as a dual probe of local and distributed dynamical phenomena. Label characterization is described in the Supplementary Materials and important details are described in the following Sections.
Figure 1. Labeling of the PhENR active site. (A) Structure of the PhENR homodimer bound to FMN (yellow) and 2CH (magenta) from PDB ID: 3ZOG. (B) Surface charge of the homodimer (left) and detail of the active site generated using the APBS webserver. Negative and positive charges are rendered in red and blue, respectively. (C) Design of vibrational labels (right) to mimic 2CH (left). Proposed reactive and inactive states of 2CH are distinguished by the orientation of the C=O group (arrows). (D) Representative models of 4CN-Cys labeled PhENR from MD simulations. Selected parameters from the simulations are shown in Figure S2. (D) Left: FMN UV-vis difference spectra of 4CN-Cys124 (solid red) and 4Az-Cys124 (dashed red) PhENR. Right: FMN UV-vis difference spectra of 4CN-Cys155 (solid blue) and 4Az-Cys124 (dashed blue) PhENR. The difference spectrum of 0.7 eq. p-hydroxybenzaldehyde bound PhENR (grey, from Figure S3) is overlaid for reference.

We tested the likely substrate analog geometries by modelling the nitrile labels. First, we docked 4CN-M to His124Cys and His155Cys PhENR and found that substrate-like binding is predicted in each mutant (Figure S1). We then replaced the cysteine and docked label with a non-natural 4-cyano-N-succinimidylicysteine (4CN-Cys) side chain (Figure 1c). The variants generated are denoted 4CN-Cys124 and 4CN-Cys155 PhENR. Representative models of the 4CN-
Cys \textit{Ph}ENR conformations are shown in Figure 1d. The H-bond accepting C≡N group and the phenyl ring mimic the C=O and cyclohexene ring of 2CH, respectively (Figure 1c). Based on structural analogy,\textsuperscript{62} the Cys124 variant is an ‘inactive state’ analog and the Cys155 variant is a ‘reactive state’ analog, but we note that the combination of conjugated probes and oxidized FMN cofactors actually represents a mixture of reactant-like and product-like states.

We then assessed the conformational flexibilities of the 4CN-Cys labels using classical molecular dynamics (MD) simulations. Restrained minimization was used to set the initial label positions to the docked/substrate coordinates. The systems were parameterized and equilibrated using the AMBER10:EHT force field and explicit H\textsubscript{2}O, and dynamics were monitored over 25 ns trajectories (Figure S2). The EHT (Extended Huckel Theory) approach assigns parameters for non-natural components and has been found to be robust for drugs that have similar functionalities to our labels.\textsuperscript{66, 67} Because the thiol-maleimide reaction generates a stereocenter on the linker, identical simulations were performed for the R and S enantiomers of each adduct. Results of the MD simulations show that linker flexibility allows the cyanophenyl groups to adopt similar stable conformations irrespective of stereoisomerism, and that various geometric parameters (Table S1) are in good agreement with the equilibrium positions of 2CH. The primary exceptions are the propensity of the S enantiomers to adopt bimodal distributions of inter-ring dihedral angles and a larger C≡N solvent exposed surface area (SASA) in 4CN-Cys124 \textit{Ph}ENR (27 ± 13\%) than in 4CN-Cys155 \textit{Ph}ENR (10.6 ± 6.3\%). Because \textit{Ph}ENR has moderate-to-high stereoselectivity for various substrates,\textsuperscript{62} it is reasonable to assume that the labeling reactions are biased toward one of two enantiomers but detailed analysis of the distribution(s) is outside the scope of this work and the comparison is provided for reference.
Samples of the 4CN-Cys and 4-azido-N-phenylsuccinimidylcysteine (4Az-Cys) labeled variants were prepared under ambient conditions and labeling efficiencies of ~70% were determined using Ellman’s test of free thiols. Upon adduct formation, orientation-dependent shifts in the FMN UV-vis absorbance were observed (Figure 1e). In the 4CN-Cys124 and 4Az-Cys124, the ~4 nm red shift is similar to that observed upon binding of p-hydroxybenzaldehyde (Figure S3), which is a competitive inhibitor and close structural analog of our labels. In contrast, the 4CN-Cys155 and 4Az-Cys155 adducts produced a ~2 nm blue shift, suggesting opposing perturbations of the FMN’s π-system. Importantly, the high degree of similarity between the UV-vis difference spectra (labeled minus unlabeled) for the 4CN-Cys and 4Az-Cys variants indicates that substitution of the vibrational probe does not have a large effect on the association of the label with the FMN in the active site.

Finally, we examined the integrity of the PhENR fold at the high concentrations required to detect 2D IR spectra of weak absorbers. Wild-type PhENR, unlabeled single-cysteine mutants, and labeled adducts were all soluble up to 6-7 mM without stabilizing agents. The FTIR spectrum of the concentrated unlabeled protein has an amide I maximum at 1634 cm\(^{-1}\) indicative of non-amyloid intramolecular β-sheets (Figure S4a). Upon dilution back to 50 μM, size exclusion chromatograms detected at 220 nm (peptide) and 450 nm (FMN) showed single major components with molecular weights of 38-41 kDa, consistent with the native homodimer (Figure S4b). Overall, substrate-like label association in native-like PhENR is likely to occur at experimental conditions.

**B. Spectroscopy of 4CN-Cys PhENR variants.** The Fourier transform infrared (FTIR) spectra of the 4CN-Cys124 and 4CN-Cys155 variants are both characterized by broad C≡N stretch
absorbances near 2230 cm\(^{-1}\) that fit to single pseudo-Voigt functions (Figure 2a). Despite the possibility of stereoisomerism and conformational disorder,\(^7\) there was no unique solution to multicomponent fits, so any effects of conformational heterogeneity fall within the inhomogeneous widths. The 4CN-Cys155 label is slightly (\(\omega_{crt} \approx 2229\) cm\(^{-1}\)) red-shifted and narrower (FWHM = 16.2 ± 0.1 cm\(^{-1}\)) compared to the 4CN-Cys124 label (\(\omega_{crt} \approx 2230\) cm\(^{-1}\); FWHM = 17.0 ± 0.1 cm\(^{-1}\)). These ~1 cm\(^{-1}\) differences are small compared to solvent-dependent shifts in the free 4CN-M label (Figure S5a), so we conclude that the mean polar environments about the C≡N probe in the labeled PhENR variants are similar. The red-shift and narrowing in 4CN-Cys155 are consistent with the proximity of the nitrile to basic residues (Arg49/His124) and reduced SASA predicted by MD simulations (Figure S2). However, this structural interpretation is incomplete and dynamical measurements are needed to understand the origins of the lineshapes.

For each 4CN-Cys labeled PhENR variant, we collected absorptive 2D IR spectra at twelve pump-probe waiting times (\(T_w\)) between 0.2 ps and 5.0 ps. Representative spectra are shown in Figure 2b. Here, peaks pairs along the diagonal correspond to \(v(01)\) and \(v(12)\) transitions of the C≡N stretch and broadening along the diagonal indicates substantial inhomogeneity in the active site compared to neat solvents (\textit{cf.} Figure S5b). The \(v(01)\) lifetimes (\(T_1\)) of 4CN-Cys124 (6.0 ± 0.4 ps) and 4CN-Cys155 (3.8 ± 0.3 ps) variants (Figure S6) are somewhat shorter than the ~7.2 ps lifetime of free 4CN-M in NNDMA (Figure S5b). The reduction in \(T_1\) upon rotation from 4CN-Cys124 to 4CN-Cys155 position is due to an increase in the efficiency of relaxation in the latter, which suggests that intermolecular (e.g., H-bonding) and/or intramolecular coupling to bath modes is orientation-dependent.\(^8\)
Figure 2. Infrared spectroscopy of 4CN-Cys labeled PhENR. (A) FTIR spectra of 4CN-Cys124 (red) and 4CN-Cys155 (blue) variants. (B) 2D IR spectra of 4CN-Cys124 (left panels) and 4CN-Cys155 (right panels) PhENR at selected $T_w$ (panel insets). Centerline points are shown as white circles and linear fits are shown as white lines. (C) CLS data for 4CN-Cys124 (red) and 4CN-Cys155 (blue). Lines are fits to the model described in the text and error bars represent standard errors from four trials.
Inspection of the 2D IR spectra in Figure 2b immediately shows that both 4CN-Cys variants undergo spectral diffusion with increasing $T_w$, indicated by increasing symmetry of the peak pairs. This effect is noticeably larger in 4CN-Cys155 PhENR. The $\nu(12)$ features show some evidence of frequency-dependent anharmonicity across the inhomogeneous ensemble, which is expected because different environments alter the C≡N potential, but may also reflect the existence of overlapping but unresolved subpopulations. The $\nu(01)$ features are also distorted, supporting the same interpretation; alternatively, these could be ascribed to background noise. One convenient way to quantify spectral diffusion of the ground state is to measure the centerline slope (CLS) of the $\nu(01)$ feature as a function of $T_w$. Because the low signal-to-noise ratios and interference between the $\nu(01)$ and $\nu(12)$ features prevent direct extraction, we fit horizontal slices (parallel to $\omega_{\text{probe}}$) within the diagonal FWHMs to sets of pseudo-Voigt functions and used the $\omega_{\text{ctr}}$ of each negative-going component to reconstruct the centerlines. The centerlines extracted from within the diagonal FWHMs are linear, so we use this measure to capture the behaviors of the majority of each ensemble.

For each variant, CLS data from four independent trials show a rapid (sub-ps) exponential decay followed by a static offset that persists beyond the timescales of the experiments (Figure 2c). In both cases, fits to the trial function $A = A_1 e^{-T_w/\tau_1} + A_s$ represent the trends, but we note that deviations of the means from the fits appear to be periodic and are consistent with a minor process with a frequency in the range of 10-20 cm$^{-1}$. These oscillations could arise from scatter, time-dependent background noise, or more interestingly, periodic oscillations of the system(s). Alternatively, they could reflect differential relaxation of two populations or weak quantum beats of two strongly-overlapping but weakly coupled modes; the latter effect is minor here and is
expanded upon in Sections III.D and III.E. Due to their small amplitudes and difficulties in fitting, they are not considered further and a ‘reduced’ form of the FFCF is calculated.

The decay of the CLS is proportional to a normalized FFCF that accounts for inhomogeneous dephasing and is typically modeled as a sum of Kubo terms, \( \sum \Delta \omega_i^2 e^{-t/\tau_i} \), where \( \Delta \omega_i \) and \( \tau_i \) are the ranges of frequencies sampled and the time constants for each dynamical process. The time constants are identical to those extracted from the CLS fits, and the static offset represents a ‘process’ with \( \tau \gg 5 \) ps. Values for \( \Delta \omega_i \) were estimated using the amplitudes \( (A_i) \) of the CLS fit via \( \Delta \omega_i = \sqrt{A_i/\sum_i A_i} \times FWHM/2\sqrt{2\ln(2)} \). The offset of CLS\((T_w = 0)\) from the theoretical value of 1.0 indicates Lorentzian contributions \( (\Gamma) \) to the lineshapes associated with homogeneous dephasing times \( (T_2) \) via \( \Gamma = 1/\pi T_2 \). Assuming that slow rotation of the macromolecular system minimizes orientational dephasing effects, \( 1/T_2 = 1/T_2^* + 1/2T_1 \), where \( T_2^* \) and \( T_1 \) are the pure dephasing time and the population relaxation time, respectively.

To calculate the full FFCFs of each variant, we proceeded using a method adapted from Ma et al. Briefly, combining pure dephasing and Kubo terms gives the generalized relaxation-free form of the FFCF, \( C(t) \), shown in Equation 1. Integration produces the lineshape function \( g(t) \) shown in Equation 2, and the real part of the Fourier transform of \( g(t) \) and a relaxation \( (T_1) \) factor returns the linear response (Equation 3).

\[
C(t) = \frac{2\delta(t)}{T_2^*} + \Delta \omega_1^2 e^{-t/\tau_1} + \Delta \omega_2^2 \quad [1]
\]

\[
g(t) = \int_0^t \int_0^{t'} d\tau' d\tau'' C(\tau'') \quad [2]
\]
\[ I(\omega) \propto \Re \int_{0}^{\infty} e^{i(\omega - \omega_{ct})t} e^{-g(t)} e^{-t/2T_{1}} dt \]  

[3]

Table 1. Apparent FFCF parameters for 4CN-Cys PhENR.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>4CN-Cys124</th>
<th>4CN-Cys155</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\omega_{ct}) (cm(^{-1}))</td>
<td>2230.2 ± 0.1</td>
<td>2228.9 ± 0.1</td>
</tr>
<tr>
<td>FWHM (cm(^{-1}))</td>
<td>17.0 ± 0.1</td>
<td>16.2 ± 0.1</td>
</tr>
<tr>
<td>(\Delta\omega_{0}) (cm(^{-1}))</td>
<td>6.2 ± 0.2</td>
<td>4.7 ± 0.3</td>
</tr>
<tr>
<td>(\Delta\omega_{1}) (cm(^{-1}))</td>
<td>3.7 ± 0.3</td>
<td>5.0 ± 0.3</td>
</tr>
<tr>
<td>(\tau_{1}) (ps)</td>
<td>0.5 ± 0.2</td>
<td>0.5 ± 0.2</td>
</tr>
<tr>
<td>(T_{1}) (ps)</td>
<td>6.0 ± 0.4</td>
<td>3.8 ± 0.3</td>
</tr>
<tr>
<td>(T_{2}) (ps)</td>
<td>3.0 ± 0.1</td>
<td>1.5 ± 0.1</td>
</tr>
<tr>
<td>(T_{2}^{*}) (ps)</td>
<td>2.4 ± 0.1</td>
<td>1.2 ± 0.1</td>
</tr>
<tr>
<td>(\Gamma) (cm(^{-1}))</td>
<td>4.5 ± 0.1</td>
<td>8.6 ± 0.5</td>
</tr>
<tr>
<td>(W_G) (cm(^{-1}))</td>
<td>14.7 ± 0.5</td>
<td>11.0 ± 0.7</td>
</tr>
</tbody>
</table>

Since \(\Delta\omega_{1}, \Delta\omega_{S}, \tau_{1}, \text{ and } T_{1}\) are already known, we obtained initial values for \(T_{2}^{*}\) by fitting the experimental FTIR spectra to simulated linear responses from Equation 3 (Figure S7).\(^{39, 47}\)

Because the time-frequency products \((\Delta\omega_{1}\tau_{1})\) of the fast decays fall within the fast fluctuation limit,\(^{23, 30}\) they also contribute to \(T_{2}^{*}\). Voigt fits to simulated linear responses from the isolated exponential terms in Equation 1 were used to partition the Gaussian and Lorentzian contributions from the fast decays. The resulting parameters of the apparent FFCFs are reported in Table 1.

Caution should be used in interpreting these results because the total CLS of overlapping components (e.g., rotamers or stereoisomers) may not approach zero even after extensive spectral diffusion.\(^{83}\) This phenomenon could obscure more complex dynamics so we refer to the values in Table 1 as parameters of ‘apparent’ FFCFs. Thus, the reported \(\Delta\omega_{S}\) values should be treated as upper limits of persistent inhomogeneity. Likewise, the difference in the magnitudes of the fast
components ($\Delta \omega_1$) could result from variations in overlap; an increase in the deviation of sub-population frequencies, appearing as inhomogeneity, would suppress the exponential decay. However, we note that evidence for this is limited to the large apparent $\Delta \omega_3$ values and distortions in the 2D IR spectra. The latter are more severe in 4CN-Cys155 PhENR where the $\Delta \omega_3$ is relatively small so we conclude that the differences in the FFCFs reflect real difference in local environmental dynamics. To the extent that conformational and stereoisomerism alter the FFCFs, they are related to heterogeneous binding and the implications of this possible phenomenon are discussed in Section IV.

Given these caveats, the picture that emerges is a “tradeoff” between slow (> 5 ps) and fast (< 1 ps) fluctuations upon rotation of the label from 4CN-Cys124 to 4CN-Cys155. To a first approximation, the $\Delta \omega_3$ values suggest that the 4CN-Cys124 label samples a rougher energy landscape than does the 4CN-Cys155 label. The origins of variations in the fast exponential components are less clear since most environmental motions are effectively ‘frozen’ on sub-picosecond timescales. One possibility is differences in the dynamics of H-bonds between the C≡N probes and the protein or the solvent. Although the current MD simulations are not detailed enough to resolve this effect, it is plausible based on the SASAs and the locations of the C≡N groups in the active site (Figure S2). Another possibility is that the fast component originates from intramolecular dynamics of the label. For instance, rotation about the inter-ring torsion ($\omega_{\text{tor}} \approx 50 – 60 \text{ cm}^{-1}$) or distal interactions with the phenyl-imido group could modulate the C≡N frequency via conjugation effects. Either of the latter effects are still consistent with variable constraints on the label imposed by the active site environment, and it is the locations of the controlling interactions that are uncertain.
Figure 3. Characterization of 4Az-M and 4Az-PEG2000. (A) Left: FTIR spectrum of 4Az-M in NNDMA. Right: Pump-probe decay of 4Az-M in NNDMA at the 2126 cm⁻¹ bleach. The solid line is a fit to Equation 4 in the text. (B) Absorptive 2D IR spectra of 4Az-M at selected T_w (panel insets). The phase twist at the half-period (T_w = 0.4 ps) is characteristic of a temporal offset between rephasing and non-rephasing signals (see Figure S9 for details). (C) FTIR spectra of 4Az-PEG2000 in neat solvents (NNDMA, N,N-dimethylacetamide; MeOH, methanol; THF, tetrahydrofuran; iPrOH, isopropanol; EtOH, ethanol; BuOH, 1-butanol; NMP, N-methylpyrrolidone; H₂O, and water). (D) Solvent-dependent relationship between high- and low-frequency peaks (R_I) and calculated unperturbed frequency separations (Δ0) for 4Az-PEG2000 obtained from solutions to Equations 5 and 6. Lines are drawn to illustrate the trends and do not represent fits. All relevant values are available in Table S2.

C. Fermi resonance in N₃ labels. A more important question regarding the interpretation of the 4CN-Cys FFCFs is the passivity of the C≡N probe itself. In a highly mobile environment, the polar and H-bonding properties of the nitrile group may dictate the formation of protein/H₂O contacts. An attractive strategy is to reference the C≡N dynamics to another probe, but there are no other unique absorbers within the bandwidth of our laser pulses. However, the minimally perturbative replacement of the C≡N group with an azide (N₃) group generates accidental Fermi resonances (FRs) between the azide antisymmetric stretch and overtones or combinations of aryl ring modes. 72-76 This leads to peak splitting, the appearance of cross-peaks in 2D IR spectra,
quantum beats (QBs) in time-resolved spectra that complicate dynamical measurements. For this reason, attempts to suppress FRs by isotope substitution and regioisomerism have been made, with mixed results. In this study, FRs are actually advantageous; while coupling delocalizes the effective probe to the aryl ring, unequal mode mixing may retain enough local information that the lineshapes of individual features report interactions of specific functional groups. Before applying this strategy in an enzyme, it is necessary to understand the behavior of the free label.

Our 4Az-M label (Figure 3A,B) shows typical signatures of FR including a doublet of peaks (2095 cm\(^{-1}\), 2129 cm\(^{-1}\)) in the FTIR spectrum, cross-peaks in the 2D IR spectrum, and a clear QB in dispersed pump-probe spectra monitored at the 2129 cm\(^{-1}\) bleach. Fitting the pump-probe trace to Equation 4 yields a beat period (\(T_{d3}\)) of 0.9 ps (\(\omega_{d3} = 37\) cm\(^{-1}\)), which is in good agreement with the 34 cm\(^{-1}\) splitting of the FTIR doublet. Selectively pumped 2D IR spectra (Figure S8) show no cross-peaks in unpumped regions, eliminating single coherence transfer (CT) pathways as the origin of any resolved signal; we expect double CT pathways (if present) to be minor as well. The cross-peaks beat with the same period, and the phase twist near the half-period is a signature of quasi-absorptive behavior.23 Detailed analysis using the method of Myers et al. (Figure S9) shows that diagonal and cross-peaks beat in the non-rephasing and rephasing spectra, respectively, as expected for coherently-coupled vibrations evolving during the \(T_w\).

\[
S_d(T_w) = S_{d1}e^{-T_w/\tau_{d1}} + S_{d2}e^{-T_w/\tau_{d2}} + S_{d3}e^{-T_w/\tau_{d3}} \cos(2\pi T_w/T_{d3} + \phi_{d3})
\]

[4]
Table 2. Pump-probe decay parameters for 4Az-M and 4Az-Cys labels

<table>
<thead>
<tr>
<th>Param.</th>
<th>$S_{1d}$ (a.u.)</th>
<th>$\tau_{1d}$ (ps)</th>
<th>$S_{2d}$ (a.u.)</th>
<th>$\tau_{2d}$ (ps)</th>
<th>$S_{3d}$ (a.u.)</th>
<th>$\tau_{3d}$ (ps)</th>
<th>$T_{3d}$ (ps)</th>
<th>$\omega_{3d}$ (cm$^{-1}$)</th>
<th>$\phi_{3d}$ (rad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4Az-M</td>
<td>0.48</td>
<td>$0.49 \pm 0.07$</td>
<td>0.18</td>
<td>7.24 ± 2.54</td>
<td>0.34</td>
<td>0.38 ± 0.03</td>
<td>0.89 ± 0.03</td>
<td>37.3 ± 1.4</td>
<td>0.20 ± 0.16</td>
</tr>
<tr>
<td>4Az-Cys124</td>
<td>0.55</td>
<td>$0.44 \pm 0.04$</td>
<td>0.18</td>
<td>2.35 ± 0.14</td>
<td>0.27</td>
<td>0.28 ± 0.03</td>
<td>0.79 ± 0.05</td>
<td>42.2 ± 2.8</td>
<td>0.12 ± 0.24</td>
</tr>
<tr>
<td>4Az-Cys155</td>
<td>0.63</td>
<td>$0.27 \pm 0.07$</td>
<td>0.28</td>
<td>1.69 ± 0.17</td>
<td>0.09</td>
<td>0.46 ± 0.17</td>
<td>0.83 ± 0.12</td>
<td>40.1 ± 5.6</td>
<td>0.28 ± 0.82</td>
</tr>
</tbody>
</table>

Fit values for pump-probe data at the FTIR maxima (Figure 4) given by Equation 4. Errors are the standard errors of nonlinear least-squares fits. Fits are reported for data where $T_w \geq 0.2$ ps to avoid non-resonant artifacts and the instrument response. The oscillation frequency ($\omega_{3d}$) was calculated from the oscillation period ($T_{3d}$).

These characteristics are largely maintained upon ligation to a non-protein thiol scaffold (4Az-PEG$^{2000}$). We analyzed the FTIR spectra of 4Az-PEG$^{2000}$ across a solvent series (Figure 3c) using the two-oscillator model described by Equations 5 and 6.\textsuperscript{89} Here, $\Delta$ is the observed frequency separation, $W$ is the anharmonic coupling constant, and $\Delta_0$ is the unperturbed frequency separation. For all solvent systems, $W \approx 12$-13 cm$^{-1}$ (Table S2) and a plot of peak intensity ratios ($R_I$) vs. $\Delta_0$ shows the characteristic ‘v-shape’ associated with frequency crossover in FR.\textsuperscript{89} No further interpretation of solvent (Kamlet-Taft) parameters\textsuperscript{72,89} was possible because the flexible PEG scaffold acts as a cosolvent. Overall, the large coupling constants and intense features provide a large sensitivity range and these results establish the 4Az-thiol adducts as sensitive reporters of solvation in mixed environments.

$$R_I = \frac{\Delta + \sqrt{\Delta^2 - 4W^2}}{\Delta - \sqrt{\Delta^2 - 4W^2}} \quad [5]$$
\[ \Delta = \sqrt{\Delta_0^2 + 4W^2} \] [6]

**D. Spectroscopy of 4Az-Cys PhENR variants.** The FTIR spectra of 4Az-Cys124 and 4Az-Cys155 PhENR demonstrate that, indeed, the orientation of the label influences FRs. Both spectra are dominated by a high-frequency signal at ~2125 cm\(^{-1}\) but have additional intensity near 2100 cm\(^{-1}\) (Figure 4a). The low-frequency peak is suppressed in 4Az-Cys155 PhENR. Dispersed pump probe traces of both variants, monitored at the high-frequency v(01) bleach (Figure 4b), undergo rapid decays (within ~ 3 ps), likely due to the fact that FR provides efficient channels for relaxation. In both systems, the traces contain oscillatory components, and fits to Equation 4 reveal periods of \(T_{d3} \approx 0.8\) ps (Table 2). The amplitudes of these beats are smaller than in 4Az-M and are reduced in 4Az-Cys155 PhENR (vs. 4Az-Cys124 PhENR), consistent with the relative peak intensities in the FTIR spectra. Comparing the labeled enzymes alone, the reduction in beat amplitude in 4Az-Cys155 PhENR is offset by larger amplitudes and shorter lifetimes of the two exponential terms in Equation 4.

The FTIR spectra of the variants were fit to pairs of pseudo-Voigt functions to extract the amplitudes, FWHMs, and center frequencies of high-frequency (H) and low-frequency (L) components (Table 3). This two-component model is an estimate and each function could include multiple near-lying features. However, we note that the intense high-frequency peak behaves similarly to the isolated C≡N absorbances (Figure 2a). Although observed transitions in FR have mixed character, the similarity to the C≡N analogs suggests that the ~2125 cm\(^{-1}\) peaks are dominated by the N\(_3\) antisymmetric stretch. For clarity, we refer to the high-frequency feature as the ‘N\(_3\) fundamental’ and the low-frequency feature as the ‘combination band’ (CB).
**Figure 4. Linear spectra of 4Az-Cys PhENR variants.** Top panels: FTIR spectra of 4Az-Cys124 (left, red) and 4Az-Cys155 (right, blue) PhENR. Each spectrum was fit to a pair of pseudo-Voigt functions (dotted lines) and the total fits are shown as solid lines. Bottom panels: Pump-probe decays of 4Az-Cys124 (left, red) and 4Az-Cys155 (right, blue) PhENR monitored at the 2125 cm\(^{-1}\) v(01) bleach. Data (circles) are fit to Equation 4 (solid lines), Fit parameters for the pump-probe and FTIR data are reported in Tables 2 & 3, respectively.

<table>
<thead>
<tr>
<th>Table 3. FTIR parameters for 4Az-Cys labeled PhENR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Param.</td>
</tr>
<tr>
<td>(A_H) (a.u.)</td>
</tr>
<tr>
<td>FWHM(_H) (cm(^{-1}))</td>
</tr>
<tr>
<td>(\omega_{ctr,H}) (cm(^{-1}))</td>
</tr>
<tr>
<td>(A_L) (a.u.)</td>
</tr>
<tr>
<td>FWHM(_L) (cm(^{-1}))</td>
</tr>
<tr>
<td>(\omega_{ctr,L}) (cm(^{-1}))</td>
</tr>
<tr>
<td>(R_I (A_H/A_L))</td>
</tr>
<tr>
<td>(\Delta) (cm(^{-1}))</td>
</tr>
<tr>
<td>(\Delta_0) (cm(^{-1}))</td>
</tr>
<tr>
<td>(W) (cm(^{-1}))</td>
</tr>
</tbody>
</table>

Using values from the fits, we calculated \(W\) and \(\Delta_0\) using Equations 5 & 6. Unlike 4Az-PEG\(_{2000}\), where \(W\) is virtually independent of the solvent (Table S2), rotation of the label from the 4Az-Cys124 to 4Az-Cys155 position results in an apparent reduction of \(W\) from 12.4 cm\(^{-1}\) to 9.9
cm\(^{-1}\) and near doubling of \(\Delta_0\) from 9.8 cm\(^{-1}\) to 17.4 cm\(^{-1}\). Comparing values in Table 3 to Figure 3d, 4Az-Cys124 PhENR lies on-trend at slightly higher \(R_I\) and \(\Delta_0\) than 4Az-PEG2000/H\(\text{O}\) but 4Az-Cys155 PhENR has \(R_I\) and \(\Delta_0\) far outside the range, suggesting a major change in environmental interaction. Taken individually, the modes that contribute to FR are subject to the same frequency tuning and lineshape broadening effects as isolated oscillators. These are modified by anharmonic coupling such that the linewidth of the otherwise dark CB is limited by the linewidth of its partner, and variations in \(W\) among members of the ensemble can lead to lineshape changes in both features. Keeping in mind that rotation of the label (roughly) exchanges the positions of the N\(_3\) and phenyl-imido groups, the behaviors of both the N\(_3\) fundamental and CB are similar to the effect of reorientation on the C≡N probe. Thus, the lineshapes are not entirely dictated by coupling and we expect that environmental effects play a significant role. The large change in \(\Delta_0\) stands out and suggests that the N\(_3\) stretch, ‘ring’ modes, or both, are more susceptible to frequency tuning than is the C≡N, which only varies by \(\sim1\) cm\(^{-1}\). However, conclusions drawn from the simplified two-state model – which does not explicitly address ensemble inhomogeneity – are suspect, and there are some inconsistencies between datasets. For instance, the mismatch between \(\sim27\) cm\(^{-1}\) frequency separation (\(\Delta\)) extracted from the FTIR fits is in poor agreement with the \(\sim40\) cm\(^{-1}\) prediction from the QBs. These deficiencies motivate a more detailed investigation via 2D IR.

For each of the 4Az-Cys PhENR variants, we collected twenty-one absorptive 2D IR spectra in 100 fs intervals between \(T_w=0.0\) ps and 2.0 ps., Representative spectra at zero delay (0.0 ps), the half-period (0.4 ps) and the full period (0.8 ps) of the QBs are shown in Figure 5a. For both 4Az-Cys124 and 4Az-Cys155 PhENR, the 2D IR spectra contain all major features of the 4Az-M spectra (Figure 3b) but are broadened and complicated by interferences. At \(T_w=0.0\) ps, the diagonal Fermi doublet is resolved in both variants and the N\(_3\) fundamental appears at the expected
Figure 5. 2D IR spectra of 4Az-Cys PhENR variants. (A) Spectra of 4Az-Cys124 (left panels) and 4Az-Cys155 (right panels) PhENR at $T_w = 0.0$ ps (top), 0.4 ps (middle), and 0.8 ps (bottom). Spectra are normalized to the N$_3$ fundamental near 2125 cm$^{-1}$ to show details. (B) Detail of the diagonal CB in 4Az-Cys124 (left) and 4Az-Cys155 (right) PhENR at $T_w = 0.0$ ps.

frequency of $\sim 2125$ cm$^{-1}$. Well-resolved cross-peaks appear below the diagonal and complex intensity above the diagonal results from interference between the $v(12)$ transitions and cross-peaks. The below-diagonal cross-peaks beat with the $\sim 0.8$ ps period and the intensity oscillations are noticeably larger in 4Az-Cys124 PhENR than in 4Az-Cys155 PhENR, consistent with the
pump-probe data (Table 2). The putative CB features also vary (Figure 5b); the 4Az-Cys124 CB is relatively narrow and the 4Az-Cys155 CB is severely broadened along the diagonal. This trend is consistent with the static components of the C≡N FFCFs ($\Delta \omega_s$) in the same positions (Table 1) and the FTIR fits in Table 3.

Analysis of the CBs and the below-diagonal cross-peaks provides an opportunity to reconcile the deviation between observed frequencies and quantum beats. In 4Az-Cys124 PhENR the weak CB appears at 2100 cm$^{-1}$ in agreement with the fit to the FTIR spectrum but substantially different than the predicted $\Delta$ from the QB. The 4Az-Cys155 PhENR CB is red-shifted (vs. the FTIR fit) with an apparent $\omega_{\text{ctr}}$ of 2088 cm$^{-1}$. This shift of about -10 cm$^{-1}$ vs. the fitted value (Table 3) brings $\Delta \approx 37$ cm$^{-1}$ into approximate agreement with the predicted value from the pump-probe QB (~40 cm$^{-1}$; Table 2). For both variants, fits to congested FTIR spectra resulted in inaccuracies or inconsistencies in the peak positions and/or the frequency separations.

To examine fine structure in the coherently-coupled ensemble, we analyzed the time-dependence of the below-diagonal cross-peak using a beat mapping methods adapted from Engel and co-workers and elsewhere.$^{90-92}$ Briefly, we interpolated 25 x 25 cm$^{-1}$ regions of the spectra at all 21 delays on 1 x 1 cm$^{-1}$ grids and fit the sign-inverted intensity at each grid point to Equation 4. Some constraints were used in fitting and are described in the Supplementary Materials text. To simplify discussion, the beat period ($T_{d3}$) was converted to frequency ($\omega_{d3}$). Frequency-resolved maps of the beat parameters ($S_{d3}$, $\tau_{d3}$, $\omega_{d3}$, and $\phi_{d3}$) centered -40 cm$^{-1}$ from the diagonal are presented in Figure 6. We note that this analysis of the absorptive spectra does not distinguish between rephasing and non-rephasing signals; the long averaging times required to acquire the 4Az-Cys PhENR data render separation of the components via phase control$^{87}$ impractical.
However, because the 4Az-M cross-peaks (Figure S9) only beat in the rephasing spectra, oscillations in the mapped region likely capture the rephasing behavior and effects of inhomogeneity to a good approximation.

**Figure 6. Beat maps of cross-peaks in 4Az-Cys labeled PhENR variants.** Top row: Below-diagonal cross-peak (left) and beat parameters for 4Az-Cys124 PhENR. Bottom row: Analysis of 4Az-Cys155 PhENR. Twenty-one spectra of each variant were interpolated on 1x1 cm\(^{-1}\) grids and each voxel was fit to Equation 4. Twenty-five equally spaced contours are plotted in the ranges indicated by the color bars. Dashed white lines are parallel to the diagonal but offset by -40 cm\(^{-1}\) along \(\omega_{\text{pump}}\). The four features (1-4) labeled in the amplitude map are discussed in the text. Note that the upper left and lower right corners of the maps have large errors (Figure S10) associated with interference and weak signal, respectively.

A few general caveats and guidelines are necessary to interpret the beat maps. First, separation of the oscillatory term in Equation 4 only retains signals arising from pathways with interstate coherence(s) during \(T_w\). Thus, pathways that lose coherence may contribute to the overall signal (Figure 5) but are not represented in the maps. Second, because 2D IR is an ensemble measurement, the parameters at each \(\omega_{\text{pump}}, \omega_{\text{probe}}\) reflect summed contributions that may arise from members with different microscopic parameters. Third, regions with \(\omega_{\text{probe}} \leq 2125\) cm\(^{-1}\) should be
interpreted with caution due to the possible presence of off-diagonal excited state absorption. Furthermore, fits to Equation 4 have large relative error at the lower right corners (Figure S10).

Despite these caveats, the beat maps reveal useful information about the inhomogeneous distributions of coherently coupled vibrations. To a first approximation, variations along $\omega_{\text{probe}}$ reflect ensemble behavior of the N$_3$ fundamental, variations along $\omega_{\text{pump}}$ reflect the that of the CB, and profiles parallel to the diagonal contain information about states with constant frequency separation ($\Delta$).

The amplitude ($S_{d3}$) maps largely reflect the shapes of the absorptive cross-peaks ($T_w = 0.0$ ps), confirming the analysis. Regions of maximum beat amplitude lie roughly along the -40 cm$^{-1}$ diagonal offset, consistent with the QB frequency of each variant. However, there is a clear distinction between the 4Az-Cys124 and 4Az-Cys155 maps, both of which contain two local maxima (labeled 1-4 in the panels). In 4Az-Cys124 features 1 & 2 are offset by ~10 cm$^{-1}$ along $\omega_{\text{probe}}$, and likely represent sub-states of the N$_3$ ensemble coupled to a relatively narrow distribution of CB frequencies. Possible origins of N$_3$ inhomogeneity in 4Az-Cys124 PhENR are ‘trapped’ conformational changes, stereoisomerism, and two rotamers about the phenyl-N$_3$ bond; a second unresolved CB at the blue edge of the N$_3$ signal cannot be ruled out. Motional narrowing of the ‘ring’ feature(s) is a likely origin of the narrow profile along $\omega_{\text{pump}}$. In 4Az-Cys155 PhENR, the overall beat amplitude is suppressed. Amplitude maxima (3 & 4) are oriented along $\omega_{\text{pump}}$, and likely represent two sub-states of the ‘ring’ system CB coupled to a relatively narrow subset of the N$_3$ ensemble, arising from a similar effect. The exchange of dependencies with position – and not the distance from the linker – suggests that the inhomogeneous distributions of coupled modes have more to do with the local polar environment than the properties of the linkages.
The lifetime ($\tau_{d3}$) maps indicate which features contribute to the overall dephasing rates. In 4Az-Cys124 $PhENR$, peak (1) has a relatively flat profile with $\tau_{d3} \approx 0.3$ ps, consistent with the pump-probe measurement (Table 2). Peak (2) dephases more rapidly, potentially explaining the persistence of the $\sim 2100$ cm$^{-1}$ diagonal CB. In 4Az-Cys155 $PhENR$, peak (3) has a $\tau_{d3}$ of $\sim 0.2-0.3$ ps and peak (4) has a shorter lifetime. Overall, it is the features at the high (1) and low (4) frequency edges of the cross-peaks that are subject to rapid dephasing. Comparison of the lifetime and amplitude maps shows that $S_{d3}$ and $\tau_{d3}$ are (globally) anticorrelated in both variants; i.e., pathways that give rise to intense QBs have shorter dephasing times.$^{91}$

In the frequency ($\omega_{d3}$) for both 4Az-Cys124 and 4Az-Cys155 variants, regions corresponding to the maxima in the amplitude maps have near-constant $\omega_{d3} \approx 40$ cm$^{-1}$, matching that of the QBs (Table 2). Invariance across $>10$ cm$^{-1}$ ranges of $\omega_{\text{pump}}$ and/or $\omega_{\text{probe}}$ could arise from either coupling between homogeneous ensembles or correlated inhomogeneity that preserves the frequency separation ($\Delta$) as site energies are shifted.$^{90}$ The flat $\omega_{d3}$ map of 4Az-Cys124 is consistent with either interpretation but the obvious broadening of the diagonal peaks and comparison to the C≡N analog slightly favor the latter. Larger variations in $\omega_{d3}$ across the 4Az-Cys155 $PhENR$ map suggest a substantial degree of uncorrelated inhomogeneity.$^{90}$ The vertically elongated region of near-constant ($\sim 40$ cm$^{-1}$) $\omega_{d3}$ indicates that mean QB frequencies are maintained for a narrow region of the N$_3$ frequency distribution even as the CB frequency shifts. One possible explanation for this behavior is an increased motional narrowing effect on the N$_3$ stretch in 4Az-Cys155 $PhENR$, in analogy to the effect seen in 4CN-Cys155 $PhENR$ (see Section B).

Finally, there is a clear difference in the phase relationships among the variants. In 4Az-Cys124 $PhENR$, $\phi_{d3}$ varies by approximately, $\pm \pi/3$ rad across the map. The positive phase at low $\omega_{\text{probe}}$
may come from interference with an off-diagonal combination band that is expected to beat \( \sim \pi \) radians out of phase with the signal of interest. Near the center of the cross-peak, the phase is \( \sim 0 \) rad. The phase variation across the map is only \( \sim 0.5 \) rad along lines parallel to the diagonal, so this relationship appears to be ‘locked’ for coupled modes with a constant \( \Delta \), supporting the conclusion of correlated inhomogeneity. In 4Az-Cys155 PhENR, there is also a small deviation in \( \phi_{d3} \) along the diagonal but the map is far less structured and \( \phi_{d3} \) is relatively uniform for all \((\omega_{pump}, \omega_{probe})\) coordinates. At the center of the cross-peak, \( \phi_{d3} \approx \pi/4 \) rad. Since the pulse sequences were identical for all measurements and the same frequency window is analyzed in both variants, we do not expect the phase differences between the variants to arise from the laser fields.\(^{90, 91}\) It is possible that this results from fitting error because the oscillatory term in Equation 4 is convoluted with a pair of exponential decays. However, the \( \sim \pi/4 \) phase shift may also reflect molecular-level phenomena including phase averaging within the ensemble or the accumulation of phase via interactions with the bath during the first coherence time.\(^{91}\) Both mechanisms implicate an increase in spatially-distributed disorder in 4Az-Cys155 PhENR with the difference being the timescale on which it influences coupling.

**E. Physical interpretation of combined results.** The PhENR active site is pseudo-symmetric but there are distinct distributions of polar side chains that differentiate two substrate (analog) binding modes. The UV-vis difference spectra (Figure 1d) reveal variations in electronic (likely \( \pi-\pi \)) interactions of the labels with the FMN cofactor as a function of orientation. The divergence between orientations is manifested in a ‘tradeoff’ between fast and slow FFCF components measured using aryl-nitrile labels (Section III.B). It is difficult to assign this to environmental effects due to internal flexibility and possible stereoisomerism of the label itself. To test this, we developed and characterized a Fermi resonant aryl-azide analog (Section III.C) capable of
reporting distributed perturbations to two ‘sites.’ Variations in N\textsubscript{3} fundamental and ‘ring’ CB linewidths at $T_w = 0.0$ ps roughly correspond to the magnitudes of the static components ($\Delta \omega_s$) of the C≡N FFCFs in comparable configurations. Maps of QBs within a well-resolved cross-peak show that broadening of the N\textsubscript{3} and ‘ring’ modes depend on orientation, and that the exchange of positions exchanges their behaviors. Thus, the tradeoff between static inhomogeneity and sub-ps fluctuations can explain the aryl-azide spectra, so despite the fact that FRs have intramolecular origin we conclude that the surrounding protein/solvent environment is the dominant influence on the spectra. That the exchange does not quantitatively preserve correlation/coupling between modes is likely due to divergences between the strengths of local interactions between different functional groups and the fact that positions thereof are not identical. To summarize, functional groups occupying the region proximal to the nonpolar end of the FMN (right side in Figure 1B) experience more long-lived inhomogeneity and those occupying the region near the polar end of the FMN (left side in Figure 1B) are less disordered on long timescales but undergo larger sub-ps fluctuations.

IV. Concluding Remarks.

We have used aryl-nitrile (4CN-Cys) and aryl-azide (4Az-Cys) substrate analogs to probe how the vibrational dynamics of a substrate analog depend on binding orientation in the PhENR active site. PhENR was selected as a model system because it represents an ‘extreme’ example of dynamic disorder\cite{4, 11} that functions as a convenient test of how much the protein structure encodes ultrafast dynamics at the hydration interface. Results from aryl-nitrile FFCF and aryl-azide QB analyses indicate that that despite substantial exposure to the solvent environment, the active site structure imposes a gradient of disorder, acting on multiple timescales, across the ~1 nm length of
the label(s), and likely, the reactive substrate(s). Our measurements do not cleanly partition fluctuations of the environment from intramolecular dynamics of the labels, but such a distinction is not required because any dynamical contribution to binding and/or reactivity necessarily involve coupled motions of both components.

Although it is difficult to draw conclusions about reactivity from observations of an unreactive analog in a simulated RS environment, gradients of coupled water-protein motions across active sites have been proposed to assist formation of Michaelis complexes and (potentially) modulate catalytically relevant motions therein.\(^5\) Throughout Section III, we discussed how possible linkage stereoisomerism complicates the interpretation of our results, but any such effects reflect heterogeneity in (non-catalytic) binding/ligation reactions that are likely to translate to the reactive system. The remainders of the observed frequency distributions come from time-varying configurations of each linkage sub-population. Varying fluctuations spanning sets of functional groups – ‘dynamical anisotropy’ – are related to the quality of polar pre-organization, and our results provide some clues as to how the efficiency of the configurational search in the RS ground state differs between binding modes in PhENR.

Perhaps the most notable result is the apparent disorder of the ‘ring’ system in the putative ‘reactive’ configuration (4Az-Cys155), where the phenyl group mimics the 2CH cyclohexene group. Traditional models of enzyme function usually attribute some portion of the catalytic effect to conformational and electrostatic constraints imposed on the substrate by the surrounding environment, and our observation runs counter to this prediction. However, a broad ensemble of states may be needed (in some cases) to access an RS configuration leading to a TS.\(^4\) The \(\leq 5\) ps timescale of our experiments is substantially shorter than that of turnover,\(^6\) so in the real system,
it is likely that much of the accessible landscape would be sampled. This behavior is also consistent with the promiscuity of the enzyme, where the binding site must be flexible enough to accommodate various substrates. 62 Regarding the activating C=O group of 2CH, we found that both the C≡N and N₃ groups in Cys155-linked analogs sense reduced electrostatic disorder on long timescales but an increase in fast fluctuations. Either optimization of C=O polarization or rapid sampling of the landscape could counteract a loss of efficiency due to disorder at the reactive (C=C) site. Finally, it is interesting that the 4Az-Cys155 analog shows increased decoupling and/or decorrelation of N₃/’ring’ frequencies. We speculate that this decorrelation may be relevant to catalysis because independent fluctuations may allow efficient sampling of the landscape where the fractional population of TS-like configurations in a disordered RS is small.

Overall, the combined strategy of comparing vibrationally ‘isolated’ and coupled substrate analogs in two defined geometries has revealed several features of the system that would not be clear from any single measure. It has also generated a number of physical and biochemical hypotheses that can be addressed in future experiments. For example, we ignored relatively minor FFCF oscillations in this analysis, but it is possible that these reflect underdamped motions of the protein 40, 41 or collective motions of interfacial H₂O, 93, 94 which has weak absorbances within the bandwidth of our laser pulses. Improvements in the experimental apparatus and more detailed MD simulations will likely clarify their meaning. This work also highlights a useful application of this troublesome class of vibrational labels; 73-76 the Fermi resonant aryl-azide probes provide a large amount of information about the interplay between environmental and intramolecular fluctuations. We anticipate that a combination of isotope editing, 73, 74, 81 phase/polarization control, 87, 95-97 mode-selective pumping, 86 and dual-frequency methods 98 could access additional details of (sub)ensemble structures and dynamics.
Ultimately, the validity of our conclusions will depend on the accuracy of our structural models and vibrational mode assignments. In future studies, we will address these concerns with high-resolution (e.g., NMR, X-ray) structure determination, QM/MM calculations, \(^99-101\) and develop linkers that eliminate the ambiguities stemming from stereoisomerism. Quantum chemical calculations will be especially useful in defining the contributions of frequency tuning and coupling to the 4Az-Cys spectra; whether strongly perturbed vibrational modes have projections along the reaction coordinate is of particular interest. \(^13\) We will also explore how the spatial distribution of label dynamics/disorder varies as a function of FMN redox state, mutations that influence catalytic efficiency, protein concentration, \(^102-104\) pressure and temperature conditions, \(^60, 62, 105\) and solvent composition. \(^42, 43, 55\) Together, these studies have the potential to provide a comprehensive view of multi-timescale protein-solvent motions in PhENR and to discern their relevance to catalytic activity.
ASSOCIATED CONTENT

Supplementary Materials. Detailed data analysis methods not described in the text. Figures and tables showing docking and MD results; UV-vis, FTIR, and SEC characterization of PhENR; IR characterization of 4CN-M; vibrational lifetimes of 4CN-Cys124 & 4CN-Cys155 PhENR; simulated linear spectra of 4CN-Cys124 & 4CN-Cys155 PhENR based on FFCF parameters; solvent dependent FR parameters for 4Az-PEG2000; mode-selective 2D IR spectra of 4Az-M; and rephasing and non-rephasing spectra of 4Az-M.

AUTHOR INFORMATION

Corresponding Author

* smoran@chem.siu.edu

Author Contributions


Notes

There are no conflicts to disclose.
ACKNOWLEDGMENT

This research was supported by the National Institute of General Medical Sciences (NIGMS) of the National Institutes of Health under award number R35GM119818 (to S.D.M.). This material is based on work supported by the National Science Foundation under award DGE-1545870 (to T.D.H.).
REFERENCES


38


