




ARTICLE

<https://doi.org/10.1038/s41467-019-08926-3>

OPEN

# Protein and RNA dynamical fingerprinting

Katherine A. Niessen<sup>1</sup>, Mengyang Xu <sup>1</sup>, Deepu K. George <sup>1</sup>, Michael C. Chen <sup>2</sup>, Adrian R. Ferré-D'Amaré<sup>2</sup>, Edward H. Snell<sup>3</sup>, Vivian Cody<sup>3</sup>, James Pace<sup>3</sup>, Marius Schmidt<sup>4</sup> & Andrea G. Markelz<sup>1,3</sup>

Protein structural vibrations impact biology by steering the structure to functional intermediate states; enhancing tunneling events; and optimizing energy transfer. Strong water absorption and a broad continuous vibrational density of states have prevented optical identification of these vibrations. Recently spectroscopic signatures that change with functional state were measured using anisotropic terahertz microscopy. The technique however has complex sample positioning requirements and long measurement times, limiting access for the biomolecular community. Here we demonstrate that a simplified system increases spectroscopic structure to dynamically fingerprint biomacromolecules with a factor of 6 reduction in data acquisition time. Using this technique, polarization varying anisotropy terahertz microscopy, we show sensitivity to inhibitor binding and unique vibrational spectra for several proteins and an RNA G-quadruplex. The technique's sensitivity to anisotropic absorbance and birefringence provides rapid assessment of macromolecular dynamics that impact biology.

<sup>1</sup>Department of Physics, University at Buffalo, SUNY, Buffalo, NY, USA. <sup>2</sup>National Heart, Lung and Blood Institute, Bethesda, MD, USA. <sup>3</sup>Hauptman-Woodward Medical Research Institute & Department of Structural Biology, University at Buffalo, SUNY, Buffalo, NY, USA. <sup>4</sup>Department of Physics, University of Wisconsin, Milwaukee, WI, USA. Correspondence and requests for materials should be addressed to A.G.M. (email: [amarkelz@buffalo.edu](mailto:amarkelz@buffalo.edu))

Protein dynamics is necessary for biological function as shown 40 years ago by Austin et al.<sup>1</sup> It is still not known how thermally activated motions lead to functionally important conformational changes or how they might be tailored to optimize function. The question is relevant for the fundamental engineering of proteins for specific functions, understanding mechanisms of antibiotic resistance, and the development of new pharmaceutical strategies. Dihydrofolate reductase (DHFR) is an illustrative example of how protein dynamics are of both primary and applied importance. DHFR has been well studied by a variety of techniques and is one of the most important antitumor and antibacterial targets. Systematic NMR and crystallography measurements along with genetic analysis have demonstrated that on the millisecond time scale DHFR structural dynamics are species dependent, suggesting the evolutionary optimization of dynamics to the chemical environment of the organism<sup>2</sup>. Malaria and tuberculosis are among the diseases that are treated by antibiotics targeting DHFR in the pathogen organism. Using *Escherichia coli* DHFR (ecDHFR) it was found that the four common mutations leading to drug resistance proceed in a stepwise fashion and that the second of the four mutations, P21L, involves a residue that is not in one of the active sites, but rather is part of a dynamical loop region that allows access to the binding pocket<sup>3</sup>. The early development of this particular mutation shows that by altering the intramolecular dynamics of a critical protein, the bacteria attain a survival advantage.

The loop motion impacted by the ec DHFR P21L mutation can be associated with long range structural vibrations of the protein backbone that lie in the terahertz (THz) frequency range. While these vibrations had long been discussed theoretically, they have only recently been measured experimentally<sup>4–7</sup>. Their observation alone, however, does not provide proof of their impact on biological function. The determination of this link has been impeded by the difficulty in isolating and identifying specific motions. Typical techniques to characterize these motions have been neutron scattering, optical Kerr effect, and terahertz absorption spectroscopy<sup>5,8,9</sup>. These techniques typically measure broad single peaks that do not strongly vary with protein or functional state. The lack of distinct resonant features has prevented determination of specific vibrations and their impact on function. Recently anisotropic terahertz microscopy (ATM) was introduced, which uses the polarization dependence of an aligned array of proteins to isolate protein vibrations based on the direction of their dipole transition. ATM enabled the first observation of narrow band intramolecular protein vibrations<sup>4</sup>. The ATM spectra were found to change with inhibitor binding for the bench marking protein chicken egg white lysozyme (CEWL)<sup>10</sup>.

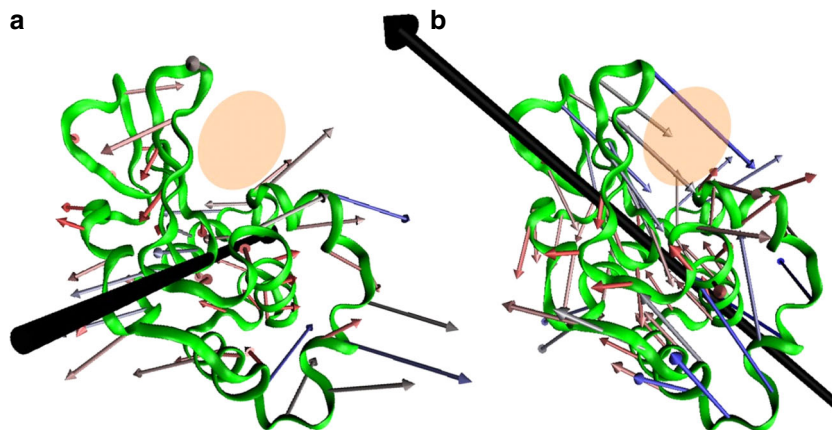
In that paper, a comparison between neutron scattering and ATM results produced the important result, confirmed by normal mode analysis, that the vibrational density of states (VDOS) is not strongly impacted by mutation or inhibitor binding, however, the directions of the vibrations are. The result was previously suggested by molecular dynamics simulations<sup>11,12</sup>. The importance of the impact of the directionality of vibrations is illustrated in Fig. 1, which shows the displacement vectors for two vibrations of CEWL with nearly equal energies. In spite of the similar energy of the vibrations, the backbone motion is very different. The vibration depicted in Fig. 1a has twisting motion while the vibration in Fig. 1b has compressive motion. CEWL catalyzes the hydrolysis of glycosidic bonds of alternating copolymers of N-acetyl-D-glucosamine and N-acetylmuramic acid in bacteria cell walls. The orange region in the molecules indicate the substrate binding site. The twisting motion in Fig. 1a does not strongly impact the access to the binding site whereas the hinging motion in Fig. 1b does. Measurements of the VDOS do not distinguish

between the two types of motions. Thus they cannot reveal if a mutation that increases activity does so because hinging vibrations were enhanced and/or twisting motions were limited. The vibrations depicted in Fig. 1 are distinguishable by light absorption. The coupling of a vibration to light is related to the change in the molecule's electric dipole with the vibrational excitation, the dipole transition. The exact evaluation of the dipole transition requires a full quantum mechanical treatment of a system with hundreds of atoms. For these large systems, typically the dipole transition is approximated by the semi-classical dipole derivative  $\partial \vec{p} / \partial \vec{q} \sim \Delta \vec{p} / \Delta \vec{q}$ , which is the change in the net molecular dipole,  $\Delta \vec{p}$ , that occurs with the vibrational atomic displacements given by the vibration vector  $\Delta \vec{q}$ . In Fig. 1, the large black arrows indicate the direction and magnitude of the dipole derivatives for the two vibrations. Only light with a component of polarization aligned along the dipole derivative can excite the vibration. Thus anisotropic absorption can provide a more detailed and specific measurement to distinguish those vibrations that impact function. However this simple picture requires either measuring a single protein, or a sample where all the proteins are aligned. To realize this macroscopic alignment, protein crystals are used. A protein crystal is a regular oriented array of proteins in an environment similar to a crowded cell with 30–70% water, most of which is mobile<sup>13</sup>. Because of weak intermolecular forces, protein dynamics are routinely measured using protein crystals<sup>14–19</sup>. As typical protein crystals are <300  $\mu\text{m}$  lateral dimensions which is less than the diffraction limited spot size for the relevant frequency range, one must use a near field approach, such as used by ATM.

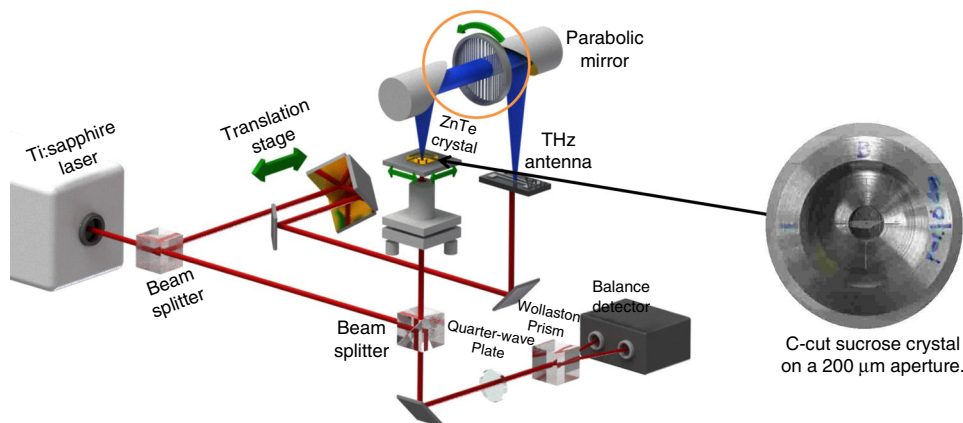
Unfortunately ATM requires the rotation of the sample in the near field: a large technological challenge. This is in part due to the requirement that while the sample's orientation relative to the THz polarization is changed, the orientation of the near field electro optic detection crystal must remain constant relative to the THz polarization to provide a constant response. Further, the imprecision of the sample centering on the rotation axis requires full imaging of the sample to ensure the same sample area is probed for each rotation. As a result, a complete 360-degree ATM data set requires 24 h. Here we introduce polarization varying anisotropy terahertz microscopy (PV-ATM) shown in Fig. 2, which rapidly determines anisotropic response. PV-ATM resolves closely spaced resonances that are indistinguishable in ATM. This enhanced spectral structure arises from sensitivity to sample birefringence, enabling higher specificity in fingerprinting of the global intramolecular vibrations. By removing the sample rotation requirement PV-ATM reduces a full 360-degree measurement to 4 h and provides open access to the sample for photo excitation and temperature control. We characterize and model PV-ATM using single crystal sucrose, and then demonstrate unique dynamical fingerprinting of bench marking biomolecules CEWL, DHFR, photoactive yellow protein (PYP) and RNA G-quadruplex<sup>20</sup>.

## Results

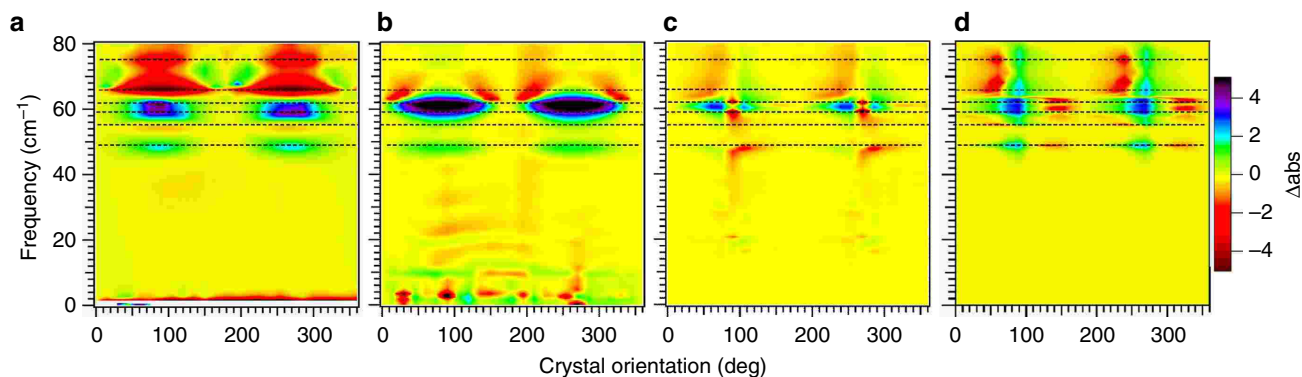
**Relationship of PV-ATM data to vibrational absorbance.** In Fig. 3 we compare surface plots of the anisotropic absorbance difference  $\Delta \text{abs}$  for *c* face polished sucrose single crystals measured by far field THz TDS, ATM, and PV-ATM. Sucrose readily forms molecular crystals, similar to protein crystals, except the size of the sucrose molecule forming the crystalline ordered array has a much smaller molecular weight, the crystal is stable with dehydration and the intermolecular contact forces are stronger. As sucrose single crystals can be grown sufficiently large to use standard far field techniques, the THz anisotropic absorption and birefringence are well established<sup>21,22</sup>. We define the anisotropic



**Fig. 1** Directionality of global intramolecular vibrations. The vector diagrams of two calculated low frequency vibrations for CEWL at **(a)**  $12.3\text{ cm}^{-1}$  ( $0.37\text{ THz}$ ) and **(b)**  $11.6\text{ cm}^{-1}$  ( $0.35\text{ THz}$ ). The orange shaded region highlights the binding region. While the two vibrations nearly overlap in energy, their motions are dramatically different. For the  $12.3\text{ cm}^{-1}$  case the largest amplitude motion is largely out of the page, with little motion around the binding site, whereas the  $11.6\text{ cm}^{-1}$  vibration moves in the page with a large compressive motion into the binding region. The black arrows indicate the light polarization direction needed to excite the vibration, demonstrating that polarization dependent absorption measurements on aligned samples readily discriminates between the two vibrations



**Fig. 2** Schematic of PV-ATM optical system. Sample rotation and X-Y imaging at each rotation is eliminated by THz polarization rotation



**Fig. 3** Relationship of PV-ATM data to vibrational absorbance for sucrose. The measured  $\Delta\text{abs}(\nu, \theta)$  using **(a)** far-field orientation dependent absorption, **(b)** near field anisotropic absorbance (ATM), and **(c)** PV-ATM. The THz radiation polarized in the  $a$ - $b$  plane and the  $0^\circ$  orientation corresponds to the THz polarization parallel to the  $a$  axis. Also shown is simulation of the PV-ATM measurement **(d)**. The simulations indicate that the increased structure in the PV-ATM spectra compared to anisotropic absorption arises from sensitivity to both the resonant anisotropic birefringence and the EO detector crystal polarization dependence

absorbance,  $\Delta\text{abs}(\omega, \theta)$ , with respect to a reference polarization direction, where

$$\Delta\text{abs}(\omega, \theta) = \text{abs}(\omega, \theta) - \text{abs}(\omega, \theta_{\text{ref}}) = -2\ln(E_t(\omega, \theta)/E_t(\omega, \theta_{\text{ref}})). \quad (1)$$

For all the sucrose measurements the THz light is normally incident to the  $a$ - $b$  plane of the crystal and  $\theta_{\text{ref}} = 0^\circ$  is taken along the  $a$  axis. The far-field measurements are shown in Fig. 3a. The positive resonances (green-blue) correspond to lattice phonons polarized along the  $b$ -axis and are consistent with previous measurements at  $49 \text{ cm}^{-1}$ ,  $59 \text{ cm}^{-1}$  and  $62 \text{ cm}^{-1}$ . In spite of the close spacing, there is a clear double peak at  $59 \text{ cm}^{-1}$  and  $62 \text{ cm}^{-1}$  in the far field measurement. The negative resonances (orange-red) correspond to phonons polarized  $90^\circ$  from the  $b$  axis. The crystal  $c$  axis partially projects onto the  $a$  axis for the monoclinic crystals so these resonances include those polarized along the  $a$  and/or  $c$  axis. The resonant frequencies at  $55$ ,  $66$ , and  $75 \text{ cm}^{-1}$  again agree with previous results. The  $\Delta\text{abs}$  features are periodic with the crystal rotation as expected for the oriented sample. Figure 3b shows the ATM results for a single  $10 \mu\text{m}$  by  $10 \mu\text{m}$  near field pixel, which do not show as much definition as the far field measurements. The resonances in Fig. 3b are coincident with those measured in the far field in Fig. 3a, including the change in polarity, however the  $59$  and  $62 \text{ cm}^{-1}$  resonances now appear as a single band, and only the  $66 \text{ cm}^{-1}$   $a/c$  axis resonance is clearly visible. Nevertheless, the ATM spectra are directly related to the anisotropic absorption and can be readily modeled using the calculated dipole transition vectors, which for proteins can be done using normal mode analysis for example<sup>4</sup>. In Fig. 3c we show the PV-ATM results. There is a dramatic change in the spectrum appearance compared to the ATM and far field. The resonant peaks and valleys now appear as inflections in the polarization angular dependence. All the resonant features in the far field data are present in the PV-ATM results. In fact, it is easier to distinguish the  $59$  and  $62 \text{ cm}^{-1}$  resonances in PV-ATM than in the far field results. For the PV-ATM results, the vibrational dipole transition direction is given by the direction of the inflection at the resonance. While the  $b$  axis resonances go from negative to positive at  $90$  degrees, the  $a/c$  axis resonances go from positive to negative. To properly model the PV-ATM data, the anisotropic refractive index of the sample as well as the polarization dependence of the electro optic detection must be included in addition to the anisotropic absorbance. Using the anisotropic absorbance and birefringent refractive index previously characterized with far field measurements, our model, shown in Fig. 3d, readily reproduces the main features measured including the signs of the inflections of the  $\Delta\text{abs}$  angular dependence. We note the PV-ATM signal is non-zero at all polarizer orientations due to a slight ellipticity of the THz polarization at the polarizer due to mirror reflections. Evidence for this is found in Supplementary Figure 2.

### PV-ATM's measurement of protein intramolecular vibrations.

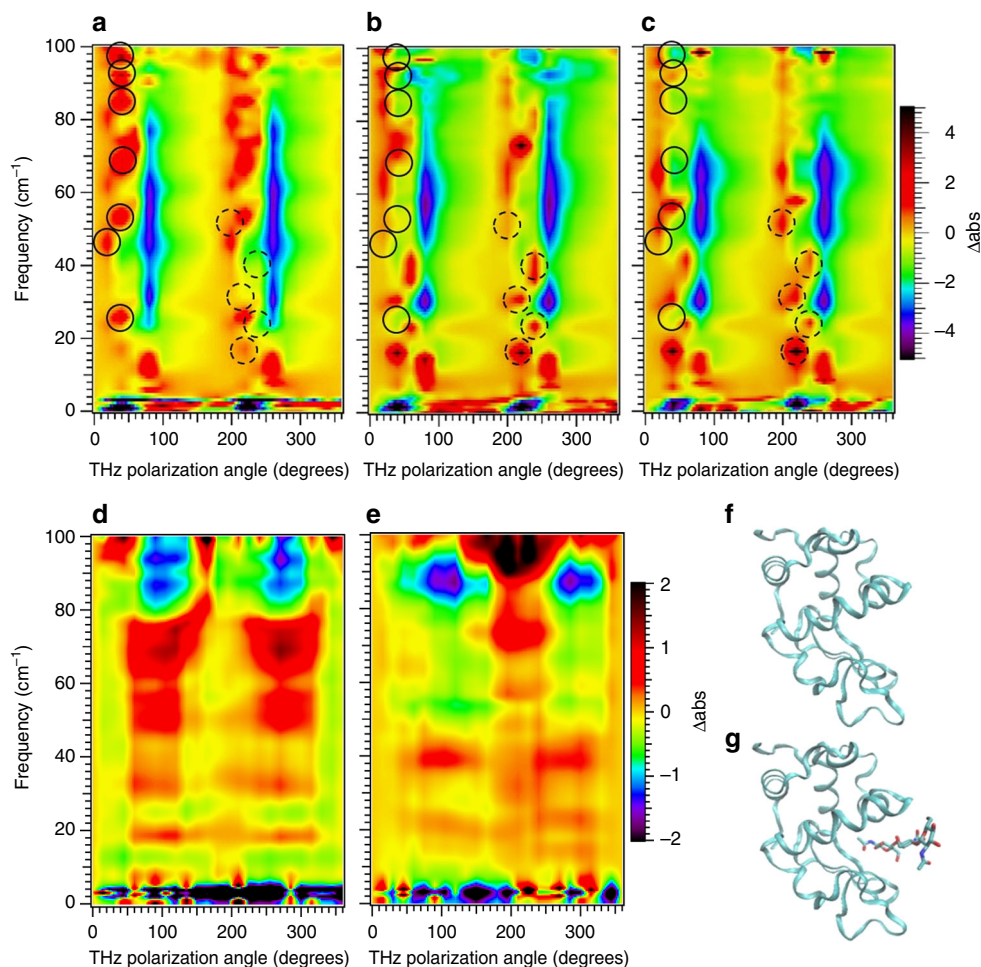
The sucrose measurements demonstrate spectroscopic reliability and sensitivity of PV-ATM. We now exploit the high spatial resolution of the near field anisotropic technique to characterize protein intramolecular vibrations. We first examine sensitivity to inhibitor binding using CEWL and the inhibitor tri acetyl glucosamine (3NAG)<sup>10</sup>. Previously it was found that inhibitor binding only slightly alters the broad featureless VDOS and isotropic absorption, while anisotropic absorption readily detects the reorientation of intramolecular vibrations. Figure 4 shows the PV-ATM results for CEWL with the THz light incident on the (110) face of tetragonal crystals, identical to the sample configuration used in the previous ATM measurements. Figure 4a

shows the free CEWL PV-ATM spectrum and Fig. 4b, c show 3NAG bound CEWL spectra for two different crystals. The PV-ATM spectrum has more distinct features than the previous ATM measurements<sup>4,10</sup>. Most likely the broad ATM spectral resonances break up into distinct features in PV-ATM due to the rapid refractive index variation accompanying closely spaced resonances, as is seen for the  $59$  and  $62 \text{ cm}^{-1}$  resonances of sucrose in Fig. 3. Comparing 4a to 4b, and 4c we see that the protein dynamics are strongly altered by the inhibitor binding as was reported previously using ATM, however PV-ATM shows considerably more spectral and angular structure. PV-ATM Resonances removed with binding are indicated by solid line circles, whereas new resonances arising with binding by dashed line circles. To ensure clarity between the two different sets of resonances, we use the  $180^\circ$  periodicity and indicate the free CEWL resonances at or near  $40^\circ$  and the inhibitor bound CEWL resonances at or near  $220^\circ$ . A comparison between Fig. 4b, c shows good crystal to crystal spectral reproducibility. Some of the differences between the two correspond to features present in the free CEWL, suggesting that they may arise from differences in inhibitor binding. For example the resonance at  $72 \text{ cm}^{-1} // 220^\circ$  is not seen in either the free CEWL or the second 3NAG-CEWL crystal. It is possible that this resonance is due to a free excess 3NAG in the first 3NAG-CEWL crystal. While we expect the underdamped intramolecular vibrations of free 3NAG at higher frequencies<sup>23–25</sup>, interstitial 3NAG bonding within the protein lattice may lead to this optically active vibration. The PV-ATM method provides a quick and sensitive determination of the presence of inhibitor binding.

The uniqueness of spectra for different biomolecules is illustrated in Fig. 5 which shows the PV-ATM spectra for photoactive yellow protein (PYP), DHFR and RNA G-quadruplex. The PV-ATM spectra for the three systems are readily distinguishable. Of particular note is the fingerprint of the quadruplex. We found the chevron structure reproducibly for several RNA G-quadruplexes with different crystal structures, but not for any protein. It is possible that this is signature of the low frequency optical response of the quadruplex structure.

### Discussion

The sucrose data in Fig. 3 shows a striking increase in structure for the PV-ATM data over the ATM, with the absorbance resonances at  $59$  and  $62 \text{ cm}^{-1}$  easily distinguished (see Supplementary Fig. 1). As mentioned earlier, modeling of the PV ATM signal shown in Fig. 3d reproduces these sharp features only when the frequency dependent sample birefringence is included. This result reveals the added structure in the case of sucrose is coming from the polarization rotation of the THz. The near field EO detection is strongly dependent on the relative THz and NIR probe polarizations. We account for this known variation by using an empty aperture reference measurement. For the THz polarizer at  $90^\circ$  the relative polarizations are such that the signal is at minimum. Thus any rotation of the THz polarization by the sample will result in an enhanced EO response. The birefringence that necessarily accompanies the anisotropic absorption lines of the sucrose provides this frequency dependent polarization rotation for the sucrose. For the biomacromolecules the additional structure over that of the ATM measurements likely arises in part from the accompanying birefringence for different resonances as well. However we cannot ignore the possibility that the increase in the EO signal could also come about from circular dichroism resonances. It has been speculated that circular dichroism can occur at these frequencies for biomolecules. The ATM technique would not have been sensitive to this effect as the relative polarization of the THz and NIR probe were held



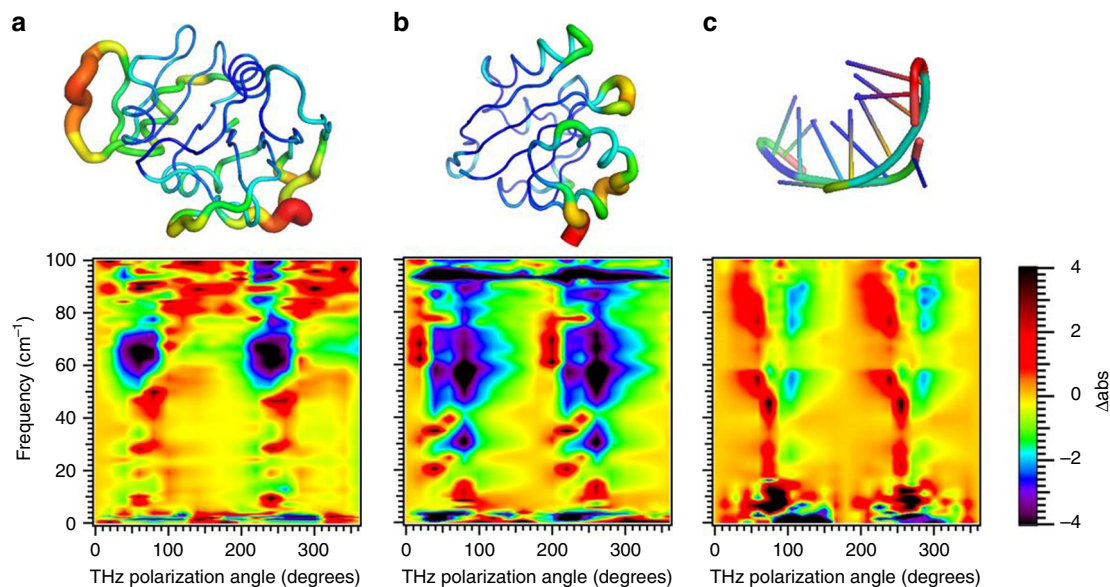
**Fig. 4** PV-ATM sensitivity to inhibitor binding. **a** CEWL; **b** CEWL-3NAG crystal 1, and **c** CEWL-3NAG crystal 2. The spectra is reproducible crystal to crystal. ATM measurements are shown for comparison (**d**) CEWL and **e** CEWL-3NAG. The simpler and faster PV-ATM technique provides spectra with increased structure over the ATM method. For the PV-ATM results solid circles indicate spectral features that are present in free CEWL and absent in inhibitor bound CEWL. Dashed circles indicate features present for inhibitor bound CEWL and absent for free CEWL. These are offset by 180° from the solid circles since all features have 180° periodicity from the regular alignment of the molecules. The large changes with inhibitor binding indicate the features arise from intramolecular vibrations and not crystal phonons. The crystal phonons are dependent on the intermolecular surface contacts and molecular mass, which change only slightly with binding as indicated by the X-ray structures for **f** free (1bwh.pdb) and for **g** inhibitor bound (1hew.pdb)

constant for the entire measurement. We are currently testing for this possibility.

Sucrose, protein and RNA crystals all consist of a regular array of molecules forming a lattice, however the origin of the THz resonances for the sucrose crystals is distinctly different from the resonances for the proteins. The sucrose anisotropic resonances seen in Fig. 3 arise from intermolecular vibrations, phonons, between sucrose molecules within the lattice. These vibrations can be visualized as net displacements of the entire molecule at a given crystal lattice site relative to the neighboring molecules. For the proteins on the other hand, the resonances arise from intramolecular vibrations: vibrations within each molecule rather than between molecules. This is apparent from the frequencies of the PV-ATM resonances for all the proteins measured. We can approximate the lattice vibrational frequency as  $\sqrt{k/m}$ , where  $m$  is the mass of the molecule and  $k$  is the effective force constant of the intermolecular bonds with neighboring molecules. The protein mass is  $\sim 1000$  times that of sucrose. While at the same time the intermolecular bond strengths for protein crystals are less than half those of sucrose crystals<sup>26</sup>. This weak intermolecular protein interactions are evident in the extreme fragility of protein crystals. The two factors together lead to protein crystal lattice

phonon frequencies of less than  $3\text{ cm}^{-1}$ <sup>27,28</sup>, a factor of 10 smaller than the observed resonances. On the other hand the PV-ATM protein resonant frequencies are consistent with normal mode analysis estimates of intramolecular vibrations<sup>29</sup>.

The CEWL binding measurements shown in Fig. 4 are an additional confirmation that the observed resonances are intramolecular vibrations. In Table 1 we show the frequency of the  $\Delta\text{abs}$  peaks observed for CEWL that are not present in the CEWL-3NAG, as well as those observed for CEWL-3NAG not present for CEWL. There are obvious large changes in the spectra with inhibitor binding. For example the peak at  $26\text{ cm}^{-1} // 40^\circ$  is not present in the 3NAG bound sample, but instead three peaks at  $16\text{ cm}^{-1} // 40^\circ$ ,  $24\text{ cm}^{-1} // 60^\circ$  and  $31\text{ cm}^{-1} // 40^\circ$  are present. The  $26\text{ cm}^{-1}$  resonance must either shift in frequency, shift in direction or no longer exist in the bound state. In all three cases we can rule out this resonance arises from the crystal lattice phonons. The maximum frequency red shift for lattice phonons is  $\leq 2\%$  with inhibitor binding due to mass loading. In addition, there would be no reorientation of the polarization of the phonon since all the CEWL and CEWL-3NAG crystals measured belong to the same tetragonal space group ( $P4_32_12$ ), and all measurements are for the same crystal face and reference



**Fig. 5** Unique dynamical fingerprints for different biomacromolecules. PV-ATM spectra for **a** *E. coli* DHFR (1dra.pdb); **b** photoactive yellow protein (2phy.pdb) and **c** RNA G-quadruplex (4xk0.pdb) rendered with Pymol B-factor putty. B-factor varies from high (red) to low (blue). While the vibrational density of states, that is the vibrational energy spectrum, for the three systems are highly similar, the actual motion for a given vibrational energy is structure specific leading to large differences in the anisotropic response

**Table 1** PV-ATM  $\Delta_{\text{abs}}$  peaks from CEWL and CEWL-3NAG crystal measurements

CEWL				CEWL-3NAG			
$\nu$	$\Delta\nu$	$\theta$	$\Delta_{\text{abs}}$	$\nu$	$\Delta\nu$	$\theta$	$\Delta_{\text{abs}}$
26	3	40°	2.78	16	4	40°	4.77/5.39
46	4	20°	1.50	24	2	60°	1.07/1.08
53	4	40°	2.36	31	3	40°	1.61/1.28
68	4	40°	2.54	41	6	60°	1.65/1.22
85	6	40°	2.84	51	3	20°	0.42/1.27
92	3	40°	3.44				
97	3	40°	3.63				

The table shows the  $\Delta_{\text{abs}}$  peak frequencies,  $\nu$  ( $\text{cm}^{-1}$ ), linewidths,  $\Delta\nu$  ( $\text{cm}^{-1}$ ), angles,  $\theta$ , and  $\Delta_{\text{abs}}$  that arise in the CEWL crystal spectrum only (left side) and the CEWL-3NAG spectra only (right side). The intensities for CEWL (CEWL-3NAG) listed for Fig. 3a (Fig. 3b, c)

polarization direction. The large frequency shifts and/or orientation changes must then arise from changes in the intramolecular vibrations.

We now turn to the possible interpretation of the resonances in Table 1. Focusing first on the  $<60 \text{ cm}^{-1}$  resonances, the CEWL resonances at  $26 \text{ cm}^{-1}$ ,  $46 \text{ cm}^{-1}$ ,  $53 \text{ cm}^{-1}$  are replaced by  $16 \text{ cm}^{-1}$ ,  $24 \text{ cm}^{-1}$ ,  $31 \text{ cm}^{-1}$ ,  $41 \text{ cm}^{-1}$ , and  $51 \text{ cm}^{-1}$  for CEWL-3NAG. A straight forward interpretation of the observed changes is that resonances that were previously associated with displacements into the substrate site are eliminated by the binding and new distinct vibrations associated with the changes in the internal couplings emerge. Referring back to Fig. 1, the vibration depicted in Fig. 1a would be unaffected by the occupancy of the substrate site, whereas the vibration in Fig. 1b has displacements that would be eliminated by binding. Then resonances at  $26 \text{ cm}^{-1}$ ,  $46 \text{ cm}^{-1}$ ,  $53 \text{ cm}^{-1}$  may all directly influence substrate access. However it is possible that increased internal coupling with substrate binding results in the  $26 \text{ cm}^{-1}$  and  $46 \text{ cm}^{-1}$  resonances blue shifting to the similarly oriented 3NAG bound  $31 \text{ cm}^{-1}$  and  $51 \text{ cm}^{-1}$  resonances. Assuming this is the case, this then isolates the  $53 \text{ cm}^{-1}$  resonance as a

candidate vibration to modulate substrate access. We now examine the  $>60 \text{ cm}^{-1}$  resonances, which appear to be absent or reoriented and decreased in intensity for CEWL-3NAG. The dramatic change in this frequency range is similar to the ATM results. It is important to reemphasize that the molecular orientation is exactly the same for the different measurements, thus the reorientation of a resonance either entirely or partially out of the plane corresponds to substantial impact on the collective atomic displacements. All resonances in this range are good candidates for examining the influence of these motions on biological function. Tests of this relevance could take the form of strong excitation at the candidate resonance frequency and monitoring ligand binding rates and/or the net catalytic rate. Challenges to this test are catalytic changes due to solvent heating with intense excitation and the relatively slow CEWL catalytic rate.

The unique PV-ATM spectra for CEWL, PYP, DHFR, and the RNA G-quadruplex are reminiscent of the small molecule optical fingerprinting in the mid infrared (MIR) range  $500\text{--}1800 \text{ cm}^{-1}$ . Specific bonds and/or neighboring bonds result in vibrational energies, which can be optically measured. The resonant absorption lines provide a signature of the specific chemical compound. The spectra can be complex, and the assignment of lines to specific motions is challenging. However, because the spectra are compound specific, chemical identification is achieved via established databases, where the MIR spectra have been catalogued<sup>30,31</sup>. Thus MIR spectroscopy has become a workhorse for environmental monitoring and other chemical sensing applications<sup>32</sup>.

The number of absorbance lines in the MIR quickly increases with the size of the system. For proteins this is particularly true since the polypeptide chain has repeating peptide bonds as well as multiple instances of individual residues. As a result, MIR spectra of proteins have broad absorption bands that look nearly identical for different proteins. Protein MIR measurements suffer from the protein amide I band overlapping a liquid water strong vibrational line at  $\sim 1640 \text{ cm}^{-1}$ , similar to the background issues for THz measurements. The water overlap in the MIR is addressed by using  $\text{D}_2\text{O}$  as the solvent, with a strongly red shifted vibration to  $\sim 1210 \text{ cm}^{-1}$ . While MIR cannot be used for protein

fingerprinting, it is used to monitor changes in a protein's secondary structure during function or after a perturbation<sup>33–35</sup>.

Significant strides in the precision of IR determination of secondary structure content has been made using increasingly more sophisticated techniques such as 2D IR on deuterated samples<sup>36</sup>. As different proteins can have essentially identical secondary structure content, these measurements cannot be used as fingerprints of a specific protein. The vibrations in the far infrared (FIR), 2–400  $\text{cm}^{-1}$ , correspond to global motions which involve relative motions between domains and are dictated by the 3D structural arrangement of the alpha helices, beta sheets and unstructured regions. The vibrational frequencies and optical-dipole couplings (magnitude and direction) are dependent on the intramolecular interactions throughout the macromolecule, and thus are unique to each protein. While the long range vibrational spectra of proteins were first calculated many years ago<sup>37</sup>, they have been difficult to measure due in part to the overlap of the vibrational energies with bulk water absorption. In contrast to MIR, solvent replacement is not effective for FIR measurements as there is very little change in the strong broadband absorbance between  $\text{H}_2\text{O}$  and  $\text{D}_2\text{O}$ <sup>38</sup>.

While careful and precise isotropic absorbance measurements have detected changes in the broad absorption peaks with substrate/inhibitor binding for a variety of proteins<sup>39,40</sup>, the polarization dependent THz absorption methods used here remove the isotropic water absorption and increase detection selectivity, isolating resonant bands based on the direction of the transition dipoles. It is notable that the near field measurements reflect the net absorbance for a large population of biomolecules which are distributed over a rugged energy landscape. The fact that well-defined spectral resonances are observed for the ensemble suggests that there are preferential motions dictated by the overall architecture of the protein's energy landscape. These preferential motions can lead to biased sampling of specific regions of configurational space which enable efficient conformational transitions<sup>41</sup>. The changes in the PV-ATM spectra with inhibitor binding shown in Fig. 3 indicate a reorientation of the vibrational motions resulting in the backbone exploring a different region of configurational space. For example, Fig. 3a, b indicate that the motions with net displacements in a direction  $40^\circ$  from the [001] direction decrease with inhibitor binding, whereas motions at  $60^\circ$  from the [001] direction increase. Ideally specific motions could be identified through comparisons with calculations. While such spectral analysis may be possible for the ATM method, where the resonant bands for proteins have already been qualitatively reproduced using normal mode analysis, the PV-ATM simulations for sucrose indicate that it is essential to include the anisotropic refractive index to accurately calculate the spectrum. Current protein modeling cannot readily predict this level of dielectric response, making precise assignment of PV-ATM features out of reach at this time. Nevertheless, the simplicity of the technique along with high spectral definition makes the monitoring of biomacromolecule global structural dynamics easily accessible. The CEWL measurements for example illustrate the application of THz PV-ATM as a critical diagnostic to determine if complex formation is present in a biomolecular crystal before submission for high resolution crystallography.

We note that in addition to the spectral fingerprinting we show here, because PV-ATM is based on THz pulsed sources, there is an opportunity to time resolve changes in protein internal dynamics. For example PYP has been a test case for many time resolved dynamic characterization methods<sup>42–44</sup>, most notably time-resolved X-ray crystallography<sup>14,17–19,45–47</sup>.

Mutation studies<sup>48,49</sup>, NMR dynamical mappings<sup>2,50,51</sup>, and numerous computational studies<sup>52,53</sup> have indicated that distant structural regions in proteins are linked. Global vibrations of the

backbone are thought to provide a simple mechanism for that linkage, however accessible measurements of these motions have been lacking. PV-ATM is a simple fast measurement that can immediately determine if and how global vibrations impact biological function, and how these motions can be steered by mutation and binding within the dynamical network. The anisotropic absorbance overcomes the challenges from the large vibrational density of states and solvent background, enabling isolation of specific structural vibration bands. This is a considerable improvement over the featureless VDOS and isotropic absorbance spectra. While PV-ATM is far more accessible than ATM, it is still reliant on crystals to provide the uniform alignment of the macromolecules. Ideally one could realize molecular orientation control within solutions and perform these measurements in the far field. The rapid advancements in microfluidics may make this possible through either electrostatic or laminar flow alignment<sup>54</sup>.

## Methods

**Sample methods.** Sucrose powder is supplied by Sigma Aldrich. Sucrose was first dissolved in pure deionized water to saturation at  $23^\circ\text{C}$ . Then small sucrose seeds were grown by drying the saturated solution. Seeds were further hung in a beaker filled with clean saturated solution for several days at room temperature. The grown crystals were finally polished by hand along the *c* face to have a large area and are sufficiently thin so that the absorbance is within the dynamic range of the terahertz system<sup>22</sup>.

The sucrose, CEWL, CEWL-3NAG, and PYP samples measured by ATM and PV-ATM are mounted on a 200  $\mu\text{m}$  thick steel disk with a center drilled 200  $\mu\text{m}$  hole. The hDHFR and ecDHFR samples are mounted on a 100  $\mu\text{m}$  thick *c* face sapphire disk with a  $\sim 6 \mu\text{m}$  thick aluminum coating on the top surface and a 300  $\mu\text{m}$  aperture in the center of the metal coating. Samples are placed directly on top of the aperture. We have found good reproducibility between different crystals produced under similar growth conditions. We also have found some crystal aging effects, and believe this should be further investigated, as well as pH and salt concentration effects<sup>55–57</sup>.

Of particular importance is the maintenance of sample hydration for the protein crystals, which are typically greater than 0.5 g water/g protein. First the protein crystal samples are covered with a small drop of paraffin oil to maintain crystal hydration throughout the experiment, as is common practice in protein crystallography. The paraffin oil is transparent and isotropic at THz frequencies, as discussed in earlier papers. In addition, the sample sits in a hydration-controlled chamber for the measurement. The details of the construction of this chamber are in the ATM section below. Inspection of all samples before and after the measurements found no cracking associated with drying. The maintenance of hydration is critical for two main reasons. First, as the crystals are  $\sim 50\%$  water, removal of this water will decrease crystal order and protein alignment. Secondly, if the crystals are sufficiently dehydrated, the biological water that is necessary for the native 3D structure and dynamics will be impacted. In vivo structure requires a minimal hydration level, typically 0.3 g water/g protein. For all the biomacromolecules studied, the hydration level is higher than 0.5 g water/g protein. The CEWL and CEWL-3NAG crystals are aligned on the aperture such that the THz is incident on the (110) crystal face and the  $0^\circ$  THz polarization is aligned along the [001] axis of the protein crystal, as determined by the crystal facets. The PYP hexagonal crystal is aligned such that the  $0^\circ$  THz polarization is aligned along the *c* axis of the crystal. The ecDHFR crystal facets are not as easily identifiable and so the  $0^\circ$  THz polarization direction relative to the crystal alignment is unknown.

The free CEWL crystals are grown at room temperature with the sitting drop method<sup>4</sup>. The protein was purchased from Sigma-Aldrich. It was prepared at 60  $\text{mg ml}^{-1}$  in 0.1 M NaAc pH 5.2 buffer without further purification. The precipitant is 10% NaCl in the same buffer. The sitting drop contained 10  $\mu\text{l}$  of precipitant and 10  $\mu\text{l}$  of protein solution with 500  $\mu\text{l}$  of precipitant in the reservoir. CEWL crystals with 3NAG inhibitor are grown in the same manner, with the exception of an equal concentration of 3NAG as the protein added to the protein solution. The CEWL and CEWL-3NAG crystals have a tetragonal structure, space group  $P4_32_12$ .

*E. coli* DHFR (ecDHFR) is obtained from Sigma-Aldrich, and crystallized with a protein concentration of 9.4  $\text{mg ml}^{-1}$  and  $4^\circ\text{C}$  growth temperature using hanging drop method<sup>58</sup>. The recipe of the buffer includes 300 mM manganese (II) chloride, 14% polyethylene glycol 6000, and 20 mM imidazole. The buffer was adjusted to pH 7.0. The ecDHFR had space group  $P6_1$ . Characterization measurements after crystallization show the ecDHFR is MTX bound with NADPH not present.

PYP was directly purified from *E. halophila* strain BN9626, originally isolated by Meyer et al.<sup>59</sup>. First a microseed solution was made by grinding large, grossly twinned crystals in 3.2 M ammonium sulphate and 20 mM sodium phosphate, pH 7.0. Then the hanging drops of 10  $\mu\text{l}$  protein solution (15  $\text{mg ml}^{-1}$  PYP in 2.5 M ammonium sulphate and 20 mM sodium phosphate, pH 7.0) on silicized coverslips were injected with 0.5  $\mu\text{l}$  of microseed solution, over 500  $\mu\text{l}$  reservoirs of

2.5 M ammonium sulphate<sup>60</sup>. The resulting crystals are needle-like with a hexagonal structure, space group P6<sub>3</sub>. The long axis of the needle corresponds to the *c* axis and the THz light is incident on the (120) face.

RNA G-quadruplex crystallization conditions are using hanging drop method<sup>20</sup>. The recipe for the reservoir is 40 mM potassium cacodylate (pH 6.0), 35% 2-methyl-2,4-pentanediol, 5 mM spermine hydrochloride, 80 mM potassium chloride, 0.5 mM DHX36 fragment, and 20 mM barium chloride. The mixture of hanging drop contains 2 mM solution of the RNA, 10 mM HEPES potassium hydroxide (pH 7.5) and 100 mM potassium chloride, in a 1:1 ratio with the reservoir solution. The tetragonal crystals (space group P4<sub>2</sub>/2) were shaped as square plates, and the broad face of the crystals is perpendicular to the *c* axis. THz light is incident on the (001) face, with the reference direction along (100) or the equivalent (010).

**THz TDS.** TDS measurements were performed using THz time domain spectroscopy<sup>61,62</sup> using photoconductive gallium arsenide antenna generation<sup>63</sup> and zinc telluride electro optic detection. The 800 nm laser generated from a titanium doped sapphire laser (Tsunami, 100 femtosecond, Spectra-physics) is split into the pump and probe beams by a polarizing beamsplitter. The pump beam is focused on the antenna to generate THz emission, and then collected by a hyper-hemispherical silicon lens. The probe pulse propagates via a mechanical delay line for changing the delay time between pump and probe arms. Transmission measurements of single crystal sucrose are performed in 15° rotations from 0–360° where 0° denotes the THz polarization aligned along the *a* axis. A resolution of 0.68 cm<sup>-1</sup> was used for all measurements, THz TDS, ATM and PV-ATM. The resolution was chosen based on previous calculations of protein anisotropic bands<sup>10</sup> and the measured linewidths for single crystal molecular crystals such as sucrose, fructose and oxalic acid<sup>22,64</sup>. These crystals have strong lattice phonons in the THz range and typical linewidths of 2 cm<sup>-1</sup> or greater at room temperature.

**ATM.** Anisotropy Terahertz Microscopy (ATM)<sup>4</sup> is a near-field measurement, where the sample is in nearly direct contact with an electro-optic detection crystal. This enables measurements of samples, such as protein crystals, which are significantly smaller than the terahertz diffraction limited spot size. The sample is centered on a sample plate with a THz transparent aperture. The sample plate is placed directly on the electro optic detection crystal such that the aperture is over the near infrared (NIR) probe incident on the electro optic crystal from the bottom surface. The electro optic crystal also acts as the bottom of an environmental chamber for the sample. The hydration within the chamber is controlled by airflow from a Licor Dewpoint Generator. For the sucrose measurements, the hydration lid is in place, but no hydrated air is flowing. The THz light is incident on the sample from the top and transmission through the sample and aperture are probed with NIR which is incident from below the electro optic crystal. The NIR is reflected off of the top surface of the electro optic crystal, co propagates with the transmitted THz and is then guided to a balance detector, similar to a standard THz TDS setup. The sample is rotated by 15° increments using an automated rotation stage. For each rotation, the sample plate is lifted by automated actuators above the EO crystal, the sample plate is rotated and then lowered back onto the EO crystal. This allows for a controlled rotation measurement without breaking the dry nitrogen gas purge surrounding the THz system. This improves measurement times and resolution of previous ATM measurements<sup>4</sup>. At each orientation, a raster scan is performed to determine the position of the aperture center and three high resolution measurements are performed to obtain the transmitted THz electric field, at the center position and ± 30 μm off-center. The three waveforms are averaged for each orientation before performing a Fourier transform to determine the power spectrum. The Δabs is defined as difference in the sample absorption at a specific orientation relative to the reference orientation, as in the relation below

$$\Delta\text{abs}(\omega, \theta) = -2\ln \left[ \frac{|E_t(\omega, \theta)|}{|E_t(\omega, \theta_{\text{ref}})|} \right] = -2\ln \frac{F(\omega)|E_i(\omega)|e^{-\alpha(\omega, \theta)d/2}}{F(\omega)|E_i(\omega)|e^{-\alpha(\omega, \theta_{\text{ref}})d/2}} \quad (2)$$

$$= [\alpha(\omega, \theta) - \alpha(\omega, \theta_{\text{ref}})]d$$

Where Δabs is the relative absorption,  $E_t$  is the transmitted THz electric field,  $\omega$  is the frequency,  $\theta$  is the orientation,  $\theta_{\text{ref}}$  is the reference orientation 0°,  $d$  is the sample thickness and  $\alpha(\omega, \theta)$  is the absorption coefficient. This self-referencing removes any narrow band atmospheric water absorption from the sample hydration chamber.

**PV-ATM.** In order to do anisotropic measurements, one must rotate the sample relative to the THz polarization. This rotation of the sample is challenging due to its proximity to the electro optic detection crystal. Each time the sample is rotated relative to the electro optic crystal, the displacement of the sample center must be determined. In addition, one must account for the pixel-to-pixel variation in the electro optic crystal. As a first order method to speed up and simplify the measurement process, we rotate the terahertz polarization incident on the sample by rotating a wire grid polarizer mounted on an automated stage.

The polarization-varying ATM (PV-ATM) experimental setup is similar to the described ATM setup with the addition of a wire grid polarizer on an automated rotation stage which is located on the THz arm of the near field microscope. A

schematic of the PV-ATM system can be found in Fig. 1. The wire grid polarizer has an electric field extinction ratio,  $t_{\parallel}/t_{\perp}$ , of ~20:1 and a power extinction ratio of 400:1. Time domain THz transmission measurements to obtain the THz amplitudes and phases are taken in the near field taken at the center of the aperture every 20° of rotation for a sucrose measurement, and every 15° of rotation for a protein measurement. If the input THz light is linearly polarized, the transmitted THz changes by the factor  $\cos(\theta)$  where  $\theta$  is the angle of the polarizer relative to the initial polarization direction. The electro optic detection crystal is also polarization sensitive, so rotating the THz polarization also modulates the detection sensitivity<sup>65</sup>. To account for this, sample measurements at each THz polarization are referenced to an empty hole. A full PV-ATM measurement consists of first measuring an empty aperture for a full rotation of the polarizer followed by a full polarization rotation measurement of the sample. Sample measurements are referenced to the empty aperture and the relative absorption is determined using Eq. 1. All measurements are performed at room temperature. Hydration of the sample is maintained by a Licor Dewpoint Generator.

PV-ATM sucrose measurements are performed on the same sample as in the ATM measurements. The sample is not removed from the THz system between the two experiments, so the 0° orientation is the same for each experiment. Hydration control for the protein samples is achieved in the same way as in the ATM system, with both paraffin oil coating the sample and a hydration-controlled cell using hydrated air flow. For the sucrose measurements, the hydration lid is in place, but no hydrated air is flowing. We note the strong negative relative absorption feature at 80° THz polarization angle, present in the CEWL and CEWL-3NAG crystal results, but not present in the other spectra. This is due to a net rotation of the probe NIR polarization relative to the electro optic crystal axes, which was removed for subsequent measurements.

**Simulation and modeling methods.** The sucrose PV-ATM simulations are performed by modeling the THz dielectric response of a *c* cut sucrose crystal, and calculating the expected THz PV-ATM response. The dielectric response of the sample is modeled using the far-field THz absorption of the *c* cut sucrose crystal, shown in Fig. 2a. We fit the 0° and 90° absorption spectra with multiple Lorentzian peaks and a baseline function to obtain the peak amplitudes, frequencies, and peak widths ( $A_n$ ,  $\omega_n$ ,  $\gamma_n$ , respectively). The fit parameters are then entered into the following permittivity equation:

$$\epsilon(\omega) = \epsilon_{\text{DC}} + \sum_{n=0} \frac{1/2A_n}{\omega_n^2 - \omega^2 - i\omega\gamma_n}, \quad (3)$$

where  $\epsilon_{\text{DC}}$  is the DC dielectric constant,  $\omega$  is the frequency, and  $n$  is the peak number. We obtain two permittivity expressions  $\epsilon_c(\omega)$  and  $\epsilon_o(\omega)$  by using the 0° and 90° fit parameters, respectively. The DC dielectric constant is set for each permittivity equation such that the index of refraction,  $\text{Re}(\sqrt{\epsilon(\omega)})$ , at 0 cm<sup>-1</sup> is 1.8.

The expected THz signal is determined by first using Jones matrix calculus to obtain the transmitted THz magnitude and rotation angle through the polarizer and sucrose sample. The resulting transmitted electric field vector polarization relative to the EO crystal is then used to calculate the detected THz. As with the measurements, simulations are performed for both the sample and the reference. The initial THz polarization is set to the *x* direction of the lab frame. The electric field vector is then rotated into the reference frame of the polarizer, which lies at an angle  $\theta$  relative to the *x* direction, is transmitted through the polarizer, and then is rotated back to the lab frame. The transmitted electric field through the sample and polarizer is calculated similar to the reference with the addition of the sample frequency dependent anisotropic dielectric functions. For the anisotropic response of the *c* cut sucrose sample we label the *a* axis as the ordinary axis and the *b* axis as the extraordinary axis. The initial orientation of the sample is defined by the rotation of the *a* axis relative to the lab frame and is labeled  $\beta$ . The electric field vector transmitted through the polarizer is rotated into the sample frame, with rotation angle  $\beta$ , is transmitted through the sample, and then is rotated back to the lab frame. The transmission matrix for the sample is diagonal with  $t_{xx} = e^{i\sqrt{\epsilon_c(\omega)}2\pi\omega d}$  and  $t_{yy} = e^{i\sqrt{\epsilon_o(\omega)}2\pi\omega d}$ , where  $d$  is the sample thickness, and  $d = 0.0275$  cm for the sample measured. These procedures are used to obtain the transmitted THz magnitude and angle of rotation for both sample and reference.

Due to the polarization sensitivity of the EO response, the detected THz signal from the experiment modeled using the modified Planken expressions<sup>65</sup>:

$$\Delta I_{\text{ref}}(\omega, \theta) = |E_{\text{ref}}(\omega, \theta)| \left[ \cos(\phi_{\text{ref}}(\omega, \theta) + \phi_0) \sin(2\varphi) + 2 \sin(\phi_{\text{ref}}(\omega, \theta) + \phi_0) \cos(2\varphi) \right] \text{ and}$$

$$\Delta I_{\text{sample}}(\omega, \theta) = |E_{\text{sample}}(\omega, \theta)| \left[ \cos(\phi_{\text{sample}}(\omega, \theta) + \phi_0) \sin(2\varphi) + 2 \sin(\phi_{\text{sample}}(\omega, \theta) + \phi_0) \cos(2\varphi) \right], \quad (4)$$

where  $\phi$  and  $\varphi$  are the THz and NIR polarization angles relative to the [001] axis of a (110) cut EO crystal, and  $\phi_0$  is the initial THz polarization angle. The THz and NIR are incident on the [110] EO plane, perpendicular to the surface. The initial THz polarization angle,  $\phi_0$ , and the NIR angle,  $\varphi$ , are determined by fitting  $\Delta I_{\text{ref}}(33 \text{ cm}^{-1}, \theta)$  vs  $\theta$  and until there is good agreement between the data and simulation. In the case of the initial PV-ATM sucrose simulations  $\phi_0 = -5.12^\circ$  and  $\varphi = 48.45^\circ$  give the best agreement (see Supplementary Fig. 2).



The change in THz signal due to the polarizer is normalized by referencing each polarizer rotation angle through

$$\left| T_{\text{sample}}(\omega, \theta) \right| = \frac{\Delta I_{\text{sample}}(\omega, \theta)}{\Delta I_{\text{ref}}(\omega, \theta)}, \quad (5)$$

and the relative absorption is determined through

$$\Delta \text{abs}(\omega, \theta) = -2 \ln \left( \frac{\left| T_{\text{sample}}(\omega, \theta) \right|}{\left| T_{\text{sample}}(\omega, 0^\circ) \right|} \right). \quad (6)$$

This is calculated for  $\omega = 0, 0.1, \dots, 200$  and  $\theta = 0, 15, \dots, 360$  and the spectra is plotted as a surface plot as shown in the Fig. 3d. The orientation of the sucrose crystal relative to the THz polarization is unknown, so the sample orientation angle,  $\beta$ , is adjusted until the obtained spectra is in good agreement with the measurement. Here an orientation angle of  $\beta = -75^\circ$  is used as it gives the best agreement with the experimental data.

**Reporting summary.** Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

Data supporting the findings of this manuscript are available from the corresponding author upon reasonable request. A reporting summary for this Article is available as a Supplementary Information file.

Received: 23 July 2018 Accepted: 4 February 2019

Published online: 04 March 2019

### References

- Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. Dynamics of Ligand Binding to Myoglobin. *Biochemistry* **14**, 5355–5373 (1975).
- Bhabha, G. et al. Divergent evolution of protein conformational dynamics in dihydrofolate reductase. *Nat. Struct. Mol. Biol.* **20**, 1243–U1262 (2013).
- Toprak, E. et al. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat. Genet.* **44**, 101–U140 (2012).
- Acbas, G., Niessen, K. A., Snell, E. H. & Markelz, A. G. Optical measurements of long-range protein vibrations. *Nat. Commun.* **5**, 3076 (2014).
- Turton, D. A. et al. Terahertz underdamped vibrational motion governs protein-ligand binding in solution. *Nat Commun.* **5**, 3999 (2014).
- Liu, D. et al. Studies of phononlike low-energy excitations of protein molecules by inelastic X-ray scattering. *Phys. Rev. Lett.* **101**, 135501 (2008).
- Rheinstadter, M. C., Schmalz, K., Wood, K. & Strauch, D. Protein-protein interaction in purple membrane. *Phys. Rev. Lett.* **103**, 128104 (2009).
- Balog, E. et al. Direct determination of vibrational density of states change on ligand binding to a protein. *Phys. Rev. Lett.* **93**, 28103 (2004).
- Falconer, R. J. & Markelz, A. G. Terahertz spectroscopic analysis of peptides and proteins. *J. Infrared, Millim. THz Waves* **33**, 973–988 (2012).
- Niessen, K. A. et al. Moving in the right direction: protein vibrational steering function. *Biophys. J.* **112**, 933–942 (2017).
- Temiz, N. A. & Bahar, I. Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. *Proteins-Struct. Funct. Genet.* **49**, 61–70 (2002).
- Miller, D. W. & Agard, D. A. Enzyme specificity under dynamic control: a normal mode analysis of alpha-lytic protease. *J. Mol. Biol.* **286**, 267–278 (1999).
- Dhar, A. et al. Structure, function, and folding of phosphoglycerate kinase are strongly perturbed by macromolecular crowding. *Proc. Natl Acad. Sci. USA* **107**, 17586–17591 (2010).
- Tenboer, J. et al. Time-resolved serial crystallography captures high-resolution intermediates of photoactive yellow protein. *Science* **346**, 1242–1246 (2014).
- Miller, R. J. D. Femtosecond crystallography with ultrabright electrons and X-rays: capturing chemistry in action. *Science* **343**, 1108–1116 (2014).
- Schlichting, I. et al. Time-resolved X-Ray crystallographic study of the conformational change in Ha-Ras P21 protein on gtp hydrolysis. *Nature* **345**, 309–315 (1990).
- Jung, Y. O. et al. Volume-conserving trans-cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nat. Chem.* **5**, 212–220 (2013).
- Schotte, F. et al. Watching a signaling protein function in real time via 100-ps time-resolved Laue crystallography. *Proc. Natl Acad. Sci. USA* **109**, 19256–19261 (2012).
- Neutze, R. & Moffat, K. Time-resolved structural studies at synchrotrons and X-ray free electron lasers: opportunities and challenges. *Curr. Opin. Struct. Biol.* **22**, 651–659 (2012).
- Chen, M. C., Murat, P., Abecassis, K., Ferre-D'Amare, A. R. & Balasubramanian, S. Insights into the mechanism of a G-quadruplex-unwinding DEAH-box helicase. *Nucleic Acids Res.* **43**, 2223–2231 (2015).
- Kro"ll, J., Darmo, J. & Unterrainer, K. Terahertz optical activity of sucrose single-crystals. *Vib. Spectrosc.* **43**, 324–329 (2007).
- Singh, R., George, D. K., Benedict, J. B., Korter, T. M. & Markelz, A. G. Improved mode assignment for molecular crystals through anisotropic terahertz spectroscopy. *J. Phys. Chem. C* **116**, 10359–10364 (2012).
- Williams, M. R. C., Aschaffenburg, D. J., Ofori-Okai, B. K. & Schmuttenmaer, C. A. Intermolecular vibrations in hydrophobic amino acid crystals: experiments and calculations. *J. Phys. Chem. B* **117**, 10444–10461 (2013).
- Williams, M. R. C. et al. Terahertz spectroscopy of enantiopure and racemic polycrystalline valine. *Phys. Chem. Chem. Phys.* **13**, 11719–11730 (2011).
- Zhang, F., Wang, H. W., Tominaga, K. & Hayashi, M. Characteristics of low-frequency molecular phonon modes studied by THz spectroscopy and solid-state ab initio theory: polymorphs I and III of diflunisal. *J. Phys. Chem. B* **120**, 1698–1710 (2016).
- Fusco, D., Headd, J. J., De Simone, A., Wang, J. & Charbonneau, P. Characterizing protein crystal contacts and their role in crystallization: rubredoxin as a case study. *Soft Matter* **10**, 290–302 (2014).
- Achterhold, K. et al. Vibrational dynamics of myoglobin determined by the phonon-assisted Mo"ssbauer effect. *Phys. Rev. E* **65**, 051916 (2002).
- Meinhold, L., Merzel, F. & Smith, J. C. Lattice dynamics of a protein crystal. *Phys. Rev. Lett.* **99**, 138101 (2007).
- Whitmire, S. E. et al. Protein flexibility and conformational state: a comparison of collective vibrational modes of WT and D96N bacteriorhodopsin. *Biophys. J.* **85**, 1269–1277 (2003).
- Sharpe, S. W. et al. Gas-phase databases for quantitative infrared spectroscopy. *Appl. Spectrosc.* **58**, 1452–1461 (2004).
- Chu, P. M., Guenther, F. R., Rhoderick, G. C. & Lafferty, W. J. The NIST quantitative infrared database. *J. Res. Natl Inst. Stand. Technol.* **104**, 59–81 (1999).
- Haas, J. & Mizaikoff, B. Advances in mid-infrared spectroscopy for chemical analysis. *Annu. Rev. Anal. Chem., Vol. 9* **9**, 45–68 (2016).
- Von Dreele, R. Binding of N-acetylglucosamine to chicken egg lysozyme: a powder diffraction study. *Acta Crystallogr. Sect. D* **57**, 1836–1842 (2001).
- Bee, M. Localized and long-range diffusion in condensed matter: state of the art of QENS studies and future prospects. *Chem. Phys.* **292**, 121–141 (2003).
- Barth, A. & Zscherp, C. What vibrations tell us about proteins. *Q. Rev. Biophys.* **35**, 369–430 (2002).
- Baiz, C. R., Peng, C. S., Reppert, M. E., Jones, K. C. & Tokmakoff, A. Coherent two-dimensional infrared spectroscopy: quantitative analysis of protein secondary structure in solution. *Analyst* **137**, 1793–1799 (2012).
- Brooks, B. & Karplus, M. Normal modes for specific motions of macromolecules: application to the hinge bending mode of lysozyme. *Proc. Natl Acad. Sci. USA* **82**, 4995–4999 (1985).
- Doster, W., Cusack, S. & Petry, W. Dynamic instability of liquidlike motions in a globular protein observed by inelastic neutron scattering. *Phys. Rev. Lett.* **65**, 1080–1083 (1990).
- Cote, Y., Nomine, Y., Ramirez, J., Hellwig, P. & Stote, R. H. Peptide-protein binding investigated by far-ir spectroscopy and molecular dynamics simulations. *Biophys. J.* **112**, 2575–2588 (2017).
- Hielscher, R., Friedrich, T. & Hellwig, P. Far- and mid-infrared spectroscopic analysis of the substrate-induced structural dynamics of respiratory complex I. *Chemphyschem* **12**, 217–224 (2011).
- Schmidt, M. & Saldin, D. K. Enzyme transient state kinetics in crystal and solution from the perspective of a time-resolved crystallographer. *Struct. Dyn.* **1**, 14 (2014).
- Mataga, N. et al. Ultrafast photoinduced reaction dynamics of photoactive yellow protein (PYP): observation of coherent oscillations in the femtosecond fluorescence decay dynamics. *Chem. Phys. Lett.* **352**, 220–225 (2002).
- Mataga, N. et al. Ultrafast photoreactions in protein nanospaces as revealed by fs fluorescence dynamics measurements on photoactive yellow protein and related systems. *Phys. Chem. Chem. Phys.* **5**, 2454–2460 (2003).
- Maisuradze, G. G., Yu, X. & Leitner, D. M. Normal mode analysis and calculation of the cooling rates of the chromophore vibrations during isomerization of photoactive yellow protein. *J. Biol. Phys. Chem.* **7**, 25–29 (2007).
- Genick, U. K. et al. Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science* **275**, 1471–1475 (1997).
- Perman, B. et al. Energy transduction on the nanosecond time scale: early structural events in a xanthopsin photocycle. *Science* **279**, 1946–1950 (1998).
- Trincao, J., Hamilton, M. L., Christensen, J. & Pearson, A. R. Dynamic structural science: recent developments in time-resolved spectroscopy and X-ray crystallography. *Biochem. Soc. Trans.* **41**, 1260–1264 (2013).

48. Boehr, D. D. et al. A distal mutation perturbs dynamic amino acid networks in dihydrofolate reductase. *Biochemistry* **52**, 4605–4619 (2013).
49. Thompson, P. D. & Freisheim, J. H. Conversion of arginine to lysine at position 70 of human dihydrofolate reductase: generation of methotrexate-insensitive mutant enzyme. *Biochemistry* **30**, 8124–8130 (1991).
50. Tzeng, S. R. & Kalodimos, C. G. Dynamic activation of an allosteric regulatory protein. *Nature* **462**, 368–U139 (2009).
51. Tzeng, S. R. & Kalodimos, C. G. Protein dynamics and allostery: an NMR view. *Curr. Opin. Struct. Biol.* **21**, 62–67 (2011).
52. Wrabl, J. O. et al. The role of protein conformational fluctuations in allostery, function, and evolution. *Biophys. Chem.* **159**, 129–141 (2011).
53. Hilser, V. J., Wrabl, J. O. & Motlagh, H. N. Structural and energetic basis of allostery. *Annu. Rev. Biophys., Vol. 41* **41**, 585–609 (2012).
54. Trebbin, M. et al. Anisotropic particles align perpendicular to the flow direction in narrow microchannels. *Proc. Natl Acad. Sci. USA* **110**, 6706–6711 (2013).
55. Gaillard, T., Trivella, A., Stote, R. H. & Hellwig, P. Far infrared spectra of solid state L-serine, L-threonine, L-cysteine, and L-methionine in different protonation states. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **150**, 301–307 (2015).
56. Aoki, K., Shiraki, K. & Hattori, T. Salt effects on the picosecond dynamics of lysozyme hydration water investigated by terahertz time-domain spectroscopy and an insight into the Hofmeister series for protein stability and solubility. *Phys. Chem. Chem. Phys.* **18**, 15060–15069 (2016).
57. Vinh, N. Q. et al. High-precision gigahertz-to-terahertz spectroscopy of aqueous salt solutions as a probe of the femtosecond-to-picosecond dynamics of liquid water. *J. Chem. Phys.* **142**, 164502 (2015).
58. Carlson, J. C. T. et al. Designing protein dimerizers: The importance of ligand conformational equilibria. *J. Am. Chem. Soc.* **125**, 1501–1507 (2003).
59. Meyer, T. E. Isolation and characterization of soluble cytochromes, ferredoxins and other chromophoric proteins from the halophilic phototrophic bacterium *Ectothiorhodospira-Halophila*. *Biochim. Biophys. Acta* **806**, 175–183 (1985).
60. Borgstahl, G. E. O., Williams, D. R. & Getzoff, E. D. 1.4 angstrom structure of photoactive yellow protein, a cytosolic photoreceptor—unusual fold, active-site, and chromophore. *Biochemistry* **34**, 6278–6287 (1995).
61. Chen, J.-Y., Knab, J. R., Cerne, J. & Markelz, A. G. Large oxidation dependence observed in terahertz dielectric response for cytochrome C. *Phys. Rev. E. Rapid* **72**, 040901 (2005).
62. Chen, J.-Y., Knab, J. R., Ye, S., He, Y. & Markelz, A. G. Terahertz dielectric assay of solution phase protein binding. *Appl. Phys. Lett.* **90**, 243901 (2007).
63. Grischkowsky, D., Keiding, S., Exter, Mv. & Fattinger, C. Far-infrared time-domain spectroscopy with terahertz beams of dielectrics and semiconductors. *J. Opt. Soc. Am. B.* **7**, 2006–2015 (1990).
64. Walther, M., Fischer, B. M. & Jepsen, P. U. Noncovalent intermolecular forces in polycrystalline and amorphous saccharides in the far infrared. *Chem. Phys. Lett.* **288**, 261–268 (2003).
65. Planken, P. C. M., Nienhuys, H. K., Bakker, H. J. & Wenckebach, T. Measurement and calculation of the orientation dependence of terahertz pulse detection in ZnTe. *J. Opt. Soc. Am. B.* **18**, 313–317 (2001).

## Acknowledgements

All molecular dynamics calculations were performed using facilities provided by The Center for Computational Research, SUNY, Buffalo. We thank the National Science Foundation MRI<sup>^</sup>2 grant DBI2959989, IDBR grant DBI1556359, and MCB grant MCB1616529, and the Department of Energy BES grant DE-SC0016317 for support.

## Author Contributions

K.A.N. designed and developed apparatus, along with measurement and data analysis protocol, performed THz measurements on CEWL samples and analyzed results with A.G.M. providing overall supervision. M.X. developed data analysis protocol, performed THz measurements on ecDHFR, PYP and RNA samples and analyzed results with A.G.M. providing overall supervision. E.H.S. provided CEWL crystals and M. S. provided PYP crystals. J. P. synthesized, purified and crystallized ecDHFR with V.C. providing overall supervision. M.C.C. synthesized, purified and crystallized RNA quadruplex with A.R.F.-D. providing overall supervision. A.G.M. conceived of the overall study. A.G.M., K.A.N., D.K.G. and M. X. wrote and edited the manuscript. All authors discussed the results and contributed significantly to the writing of the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <https://doi.org/10.1038/s41467-019-08926-3>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Journal peer review information:** *Nature Communications* thanks Dongshan Wei, Petra Hellwig and the other anonymous reviewer for their contribution to the peer review of this work.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019