

# Nitrile Infrared Intensities Characterize Electric Fields and Hydrogen Bonding in Protic, Aprotic, and Protein Environments

Jared Bryce Weaver, Jacek Kozuch, Jacob M. Kirsh, and Steven G. Boxer\*



Cite This: *J. Am. Chem. Soc.* 2022, 144, 7562–7567



Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information

**ABSTRACT:** Nitriles are widely used vibrational probes; however, the interpretation of their IR frequencies is complicated by hydrogen bonding (H-bonding) in protic environments. We report a new vibrational Stark effect (VSE) that correlates the electric field projected on the  $-C\equiv N$  bond to the transition dipole moment and, by extension, the nitrile peak area or integrated intensity. This linear VSE applies to both H-bonding and non-H-bonding interactions. It can therefore be generally applied to determine electric fields in all environments. Additionally, it allows for semiempirical extraction of the H-bonding contribution to the blueshift of the nitrile frequency. Nitriles were incorporated at H-bonding and non-H-bonding protein sites using amber suppression, and each nitrile variant was structurally characterized at high resolution. We exploited the combined information available from variations in frequency and integrated intensity and demonstrate that nitriles are a generally useful probe for electric fields.

**N**oncovalent interactions play pivotal roles in chemistry, from microscopic solvation environments to electrode interfaces, homo- and heterogeneous catalysts, biological membranes, and enzyme active sites. Despite their importance, it can be difficult to experimentally quantify the many types of interactions present in the condensed phase. One quantitative metric for these interactions is the electric field, often determined through the vibrational Stark effect (VSE) on the frequency.<sup>1</sup>

The first vibrational Stark spectra in an external applied electric field were analyzed for simple nitriles ( $-C\equiv N$ ).<sup>2–4</sup> Nitriles are often found in drugs,<sup>5,6</sup> and several methods have been developed to introduce nitrile probes at a range of positions in proteins,<sup>7–10</sup> nucleic acids,<sup>11,12</sup> and biological membrane components,<sup>13</sup> as well as in nonbiological settings.<sup>14–16</sup> Nitrile IR transitions are relatively strong and, unlike carbonyls ( $-C=O$ ), occur in an uncluttered region of the IR spectrum. Unfortunately, the interpretation of nitrile frequency shifts using the language of the VSE is complicated by a well-known blueshift in H-bonding solvents,<sup>10,17–19</sup> undermining the utility of nitrile frequency shifts as direct quantitative probes of local electric fields, though changes in frequency can be useful at a qualitative level.<sup>20</sup>

The vibrational frequency of a well-isolated oscillator such as  $-C\equiv N$  or  $-C=O$  has been approximated to first order in the field by the linear (dipolar) VSE:<sup>1</sup>

$$\bar{\nu}(\vec{F}) = \bar{\nu}_0 - \Delta\vec{\mu} \cdot \vec{F} \quad (1)$$

where  $\Delta\vec{\mu}$  is the vibrational Stark tuning rate or difference dipole moment which determines the sensitivity of a probe frequency to the electric field,  $\vec{F}$ , projected by the environment onto the probe's difference dipole moment (typically parallel to the bond axis)<sup>2</sup> and  $\bar{\nu}_0$  is the zero-field (vacuum) frequency. This frequency–field dependence can be calibrated by measuring vibrational solvatochromism and calculating average fields using

molecular dynamics (MD) simulations, as shown for a carbonyl group in Figure 1A. This approach has been used with functionally relevant carbonyls to correlate electric fields at enzyme active sites with activation free energies.<sup>21–23</sup> While a similar linear frequency–field correlation also applies to nitriles for aprotic solvents, the aforementioned blueshift in protic solvents lies far from this correlation, shown for benzonitrile's  $-C\equiv N$  in Figure 1B. Computational approaches that account for H-bonding frequency shifts are available but are difficult to implement experimentally,<sup>17,24</sup> and experiments characterizing H-bonding<sup>25,26</sup> have not been used to correct H-bond frequency shifts in order to obtain electric fields.

While studying nitriles introduced as probes of electric fields in simple solvents and proteins, we observed a substantial variation of the nitrile intensity, an effect noted before but not analyzed quantitatively for nitriles or carbonyls.<sup>15,27–29</sup> Intriguingly, we found that the integrated intensity correlates with the solvent field. As shown in the following, this new VSE circumvents the complication of the blueshift in H-bonding environments. The sensitivity of the transition dipole moment (TDM),  $\vec{m}$ , to an electric field can be expressed as

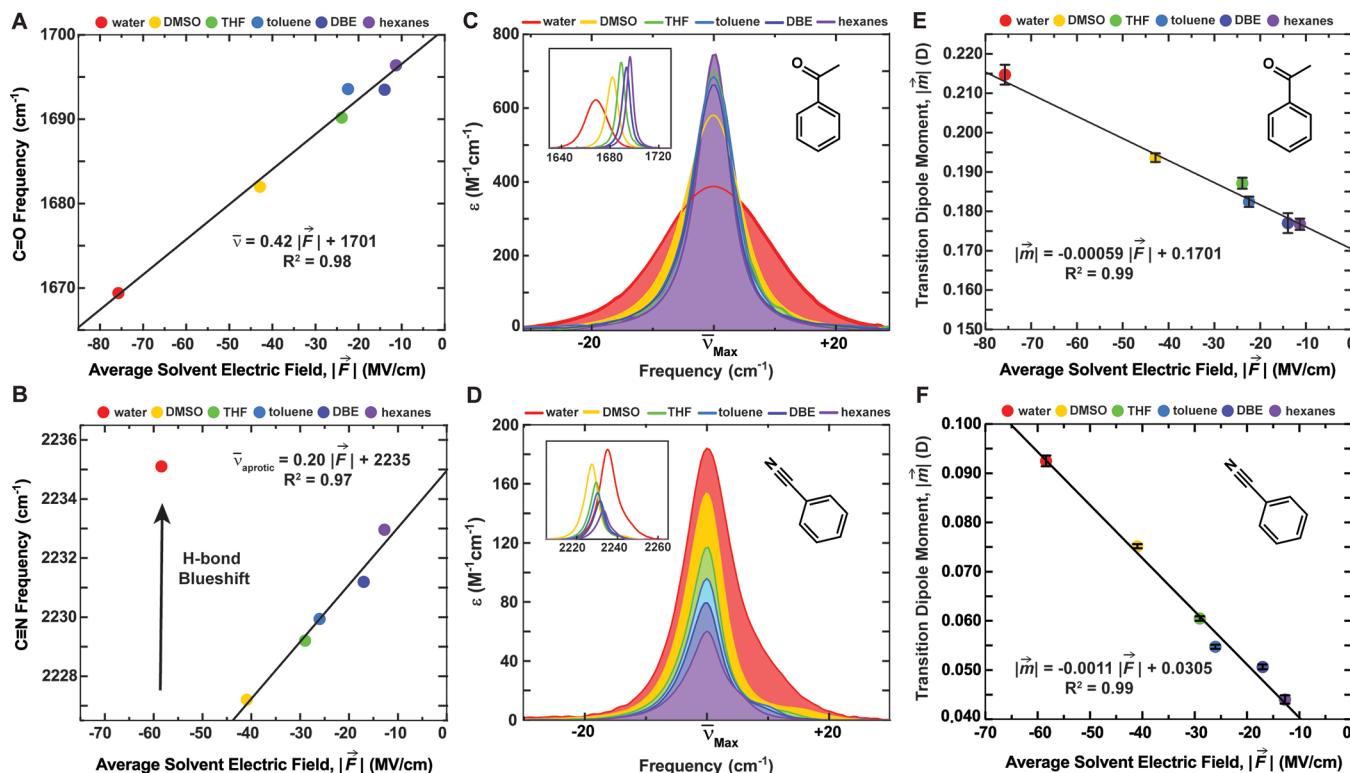
$$\vec{m}(\vec{F}) = \vec{m}_0 - \vec{F} \cdot \underline{\Delta} - \vec{F} \cdot \underline{\underline{B}} \cdot \vec{F} + \theta(F^3) + \dots \quad (2)$$

where  $\underline{\Delta}$  is the transition polarizability,  $\underline{\underline{B}}$  is the transition hyperpolarizability, and  $\vec{m}_0$  is the zero-field TDM.<sup>30</sup> Paralleling eq 1,  $\vec{m}$  depends on  $\underline{\Delta}$  in a manner analogous to  $\bar{\nu}$ 's dependence on  $\Delta\vec{\mu}$ . The  $|\vec{m}|$  value can be obtained directly from the concentration normalized peak area by

Received: January 18, 2022

Published: April 25, 2022





**Figure 1.** Peak frequency and peak area both change as a function of solvent electric field. (A, B) Plots of peak frequencies vs average solvent electric field calculated with polarizable MD for (A) ACP and (B) BZN. In (B) an arrow highlights the blueshift in  $-\text{C}\equiv\text{N}$  frequency due to H-bonding in water. (C, D) Overlays of IR spectra of (C) ACP and (D) BZN where the IR peak position ( $\bar{\nu}_{\text{Max}}$ ) has been centered in the plot and peaks are shaded to highlight peak area changes. The peak area increases with solvent polarity. Insets show noncentered spectra. (E, F) Plots of the TDM vs the calculated average solvent electric field for both ACP (E) and BZN (F).

$$|\vec{m}| = \sqrt{\frac{3\epsilon_0 hc \ln 10}{2\pi^2 N_A}} \int \frac{\epsilon(\bar{\nu})}{\bar{\nu}} d\bar{\nu} \quad (3)$$

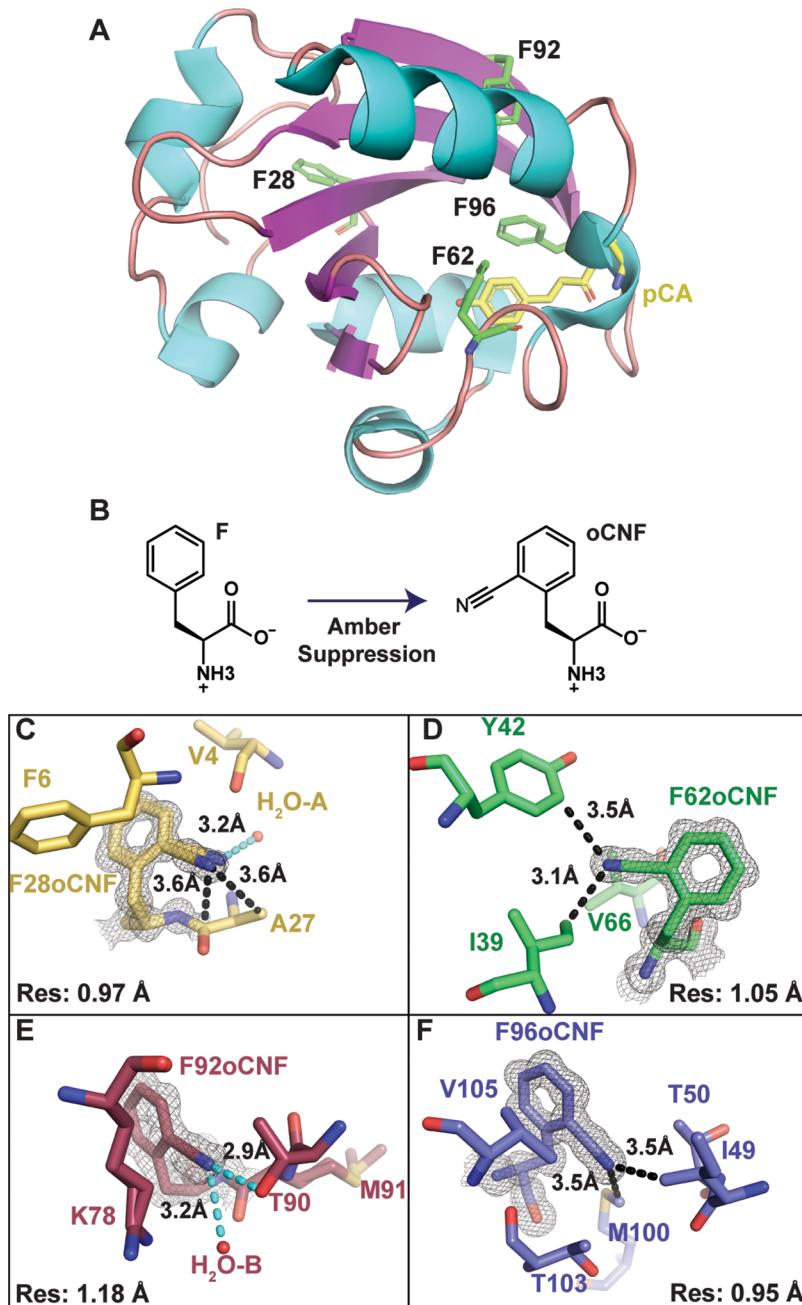
where  $\epsilon(\bar{\nu})$  is the extinction coefficient,  $\int \epsilon(\bar{\nu}) d\bar{\nu}$  is the integrated peak area, and  $\epsilon_0$ ,  $h$ ,  $c$ ,  $N_A$ , and  $\bar{\nu}$  are the vacuum permittivity, Planck's constant, the speed of light, Avogadro's number, and the wavenumber frequency, respectively.<sup>31</sup>

FTIR spectra were measured for acetophenone (ACP,  $-\text{C}=\text{O}$  band) and benzonitrile (BZN,  $-\text{C}\equiv\text{N}$  band) in a variety of aprotic and protic solvents of increasing polarity (Figure 1C–D). The relative increase in peak area from hexanes to water is only 48% for the ACP  $-\text{C}=\text{O}$  but is 342% for the BZN  $-\text{C}\equiv\text{N}$  (Figure S4). MD simulations using the polarizable AMOEBA09 force field<sup>32</sup> in Tinker 8.7<sup>33</sup> were used to calculate the average solvent electric field,  $|\vec{F}|$ , projected onto the  $-\text{C}=\text{O}$  and  $-\text{C}\equiv\text{N}$  probes in the same solvents, as described previously.<sup>34–37</sup> We demonstrate a linear trend when correlating  $|\vec{m}|$ , obtained from the peak areas using eq 3, with the average  $|\vec{F}|$  exerted on the ACP  $-\text{C}=\text{O}$  (Figure 1E). Importantly, for the BZN  $-\text{C}\equiv\text{N}$ , a linear relationship is also observed but with a steeper slope (Figure 1F), indicating a larger sensitivity of TDM to the field relative to the ACP  $-\text{C}=\text{O}$  (see Section S6 in the Supporting Information). In the context of the VSE on  $|\vec{m}|$  (eq 2), the linear correlation between  $|\vec{m}|$  and the average  $|\vec{F}|$  demonstrates that higher order terms beyond  $A$  do not meaningfully contribute to tuning  $|\vec{m}|$ , as expected.<sup>38</sup> Moreover, the linear relation between  $|\vec{m}|$  and the average  $|\vec{F}|$  is strictly monotonic for all solvents and is well captured by the VSE, in

contrast to the complicated picture for the  $-\text{C}\equiv\text{N}$  frequency dependence observed in Figure 1B.<sup>25,36</sup> By measuring the effect of solvent electric field on peak area, we are offered a direct, orthogonal calibration for measuring electric fields at nitriles that applies regardless of nitrile H-bonding status.

To demonstrate the utility of this new observable, we labeled photoactive yellow protein (PYP) with *o*-cyanophenylalanine (*o*CNF; Figure 2A,B) introduced site-specifically at multiple sites by amber suppression.<sup>39</sup> PYP is a 14 kDa protein with an endogenous chromophore that can be used to measure protein concentration. Four native phenylalanine sites were targeted for nitrile probe incorporation (Figure 2A and Figure S11) such that the nitrile is placed in electrostatically distinct non-H-bonding and H-bonding environments. High-resolution X-ray structures (<1.2 Å; Figure 2C–F) show that the nitrile probes are in a single orientation and occupy distinct environments: F28oCNF possesses one H-bond donor (3.2 Å heavy atom distance), F92oCNF has two donors (2.9 and 3.1 Å), and F62oCNF and F96oCNF are both in nonpolar environments (only carbons within 3.5 Å of the nitrile nitrogen).

The IR spectra of these PYP variants reveal significant variation of both the  $-\text{C}\equiv\text{N}$  peak area and frequency (Figure 3A). A calibration curve was made for *o*-tolunitrile (*o*TN) as a model compound for oCNF (Figure S7C). The peak areas for the nitrile in F96oCNF and F62oCNF are similar to oTN's peak area in hexanes, while the areas for F28oCNF and F92oCNF are considerably larger and indicate water-like environments, consistent with expectations from X-ray structures. Translating the peak areas to TDMs gives values for  $|\vec{m}|$  ranging from  $0.050 \pm 0.002$  D in F96oCNF to  $0.096 \pm 0.002$  D in F92oCNF (Table



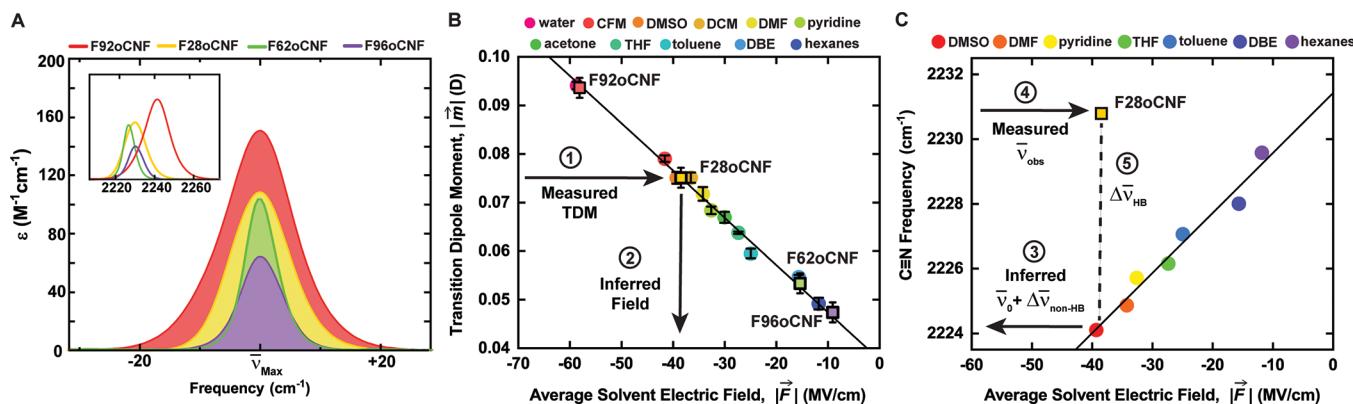
**Figure 2.** Overview of nitrile incorporation sites in PYP and their structural characterization. (A) The four phenylalanine sites targeted for nitrile probe incorporation are displayed in green, and the PYP chromophore, *p*-coumaric acid (pCA), is shown in yellow. (B) Native PYP phenylalanines (F) were replaced with genetically encoded *o*-cyanophenylalanine (oCNF) via amber stop codon suppression in four separate variants. Crystallographic structures of the nitrile site are displayed for F28oCNF (C), F62oCNF (D), F92oCNF (E), and F96oCNF (F) PYPs. Electron density maps for the oCNF in each structure are displayed ( $2m_{\text{Fo}} - DF_{\text{c}}$  contoured at  $1\sigma$ ) along with the resolution of the crystal structure (bottom corner) and the distance between the nitrile nitrogen and the nearest two atoms (H-bonds denoted in cyan). For F28oCNF, two highly similar oCNF conformers are present: the distance of H<sub>2</sub>O-A to the closer nitrile conformer is shown.

1). Using the oTN TDM calibration (Figure S7C), we extract average electric fields which range from  $-13 \pm 2$  to  $-61 \pm 2$  MV/cm. Note that the electric field projected on the nitrile of F96oCNF is  $\sim 4$  times smaller than that on the F28oCNF nitrile even though the peak frequencies are quite similar, illustrating that TDM tuning allows for identification of nonequivalent fields between species possessing nearly equivalent frequencies.

Various additive models have been developed to describe the behavior of the  $-\text{C}\equiv\text{N}$  stretching frequency in the presence of H-bonds (electrostatic multipole expansion;<sup>24</sup> separation into

forces of various origin;<sup>17</sup> empirical models,<sup>19</sup> see discussion in Section S17 of the Supporting Information). We adopt a simple semiempirical, experimentally useful model and partition the observed nitrile frequency ( $\bar{\nu}_{\text{obs}}$ ) relative to vacuum ( $\bar{\nu}_0$ ) into additive contributions from the shift due to the dipolar VSE ( $\Delta\bar{\nu}_{\text{non-HB}}$ ; obtained from solvatochromism in aprotic solvents) and from the frequency shift due to H-bonding not captured by the dipolar VSE ( $\Delta\bar{\nu}_{\text{HB}}$ ):

$$\bar{\nu}_{\text{obs}} = \Delta\bar{\nu}_{\text{HB}} + \Delta\bar{\nu}_{\text{non-HB}} + \bar{\nu}_0 \quad (4)$$



**Figure 3.** IR spectra of the  $-\text{C}\equiv\text{N}$  stretch from incorporated oCNF residues at four sites in PYP (cf. Figure 2) and extracted parameters. (A) A shaded IR overlay of the four labeled sites highlights the significant change in peak area from protein nitrile environments. The inset shows noncentered spectra. (B) Using the measured TDM (1), the electric field for proteins can be inferred from the oTN TDM vs field calibration curve (2) (Figure S7C), as demonstrated for F28oCNF. (C) The aprotic solvent field vs frequency calibration curve for oTN. Using the field inferred from the TDM (2), one can infer  $\bar{\nu}_0 + \Delta\bar{\nu}_{\text{non-HB}}$  (3), and by measuring  $\bar{\nu}_{\text{obs}}$  (4), the blueshift ( $\Delta\bar{\nu}_{\text{HB}}$ ) can be deduced (5) (Table 1).

**Table 1. Extracted IR Peak Parameters for the oCNF Nitrile Incorporated at Different PYP Sites**

environment	$\bar{\nu}_{\text{obs}}$ (cm⁻¹)	fwhm (cm⁻¹) <sup>a</sup>	$ \vec{m} $ (D)	$ \vec{F} $ (MV/cm)	$\Delta\bar{\nu}_{\text{non-HB}}$ (cm⁻¹)	$\Delta\bar{\nu}_{\text{HB}}$ (cm⁻¹)
hexanes	2229.6	5.1	$0.049 \pm 0.001$	$-12 \pm 1$	-2	0
water	2231.7	9.5	$0.094 \pm 0.001$	$-59 \pm 1$	-11	11
F28oCNF	2230.8	12.7	$0.075 \pm 0.002$	$-38 \pm 2$	-6	7
F62oCNF	2228.1	6.5	$0.053 \pm 0.002$	$-15 \pm 2$	-3	0
F92oCNF	2241.3	14.4	$0.094 \pm 0.002$	$-58 \pm 2$	-11	21
F96oCNF	2231.2	8.4	$0.047 \pm 0.002$	$-9 \pm 2$	-2	2

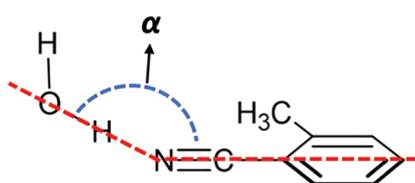
<sup>a</sup>Peak positions ( $\bar{\nu}_{\text{obs}}$ ) and fwhms have an error of  $<0.1$  cm⁻¹.

as illustrated in Figure 3C. Using the aprotic oCNF nitrile frequency-field calibration ( $|\Delta\bar{\nu}| = 0.19 \frac{\text{cm}^{-1}}{\text{MV/cm}}$ ,  $\bar{\nu}_0 = 2231.4$  cm⁻¹; Figure S8B) and fields inferred from the TDM-field calibration (Figure 3B),  $\Delta\bar{\nu}_{\text{non-HB}}$  was determined (Figure 3C). Per eq 4,  $\Delta\bar{\nu}_{\text{HB}}$  is the difference between  $\bar{\nu}_{\text{obs}}$  and  $\bar{\nu}_0 + \Delta\bar{\nu}_{\text{non-HB}}$  (Figure 3C). Using this approach, we observe a large range in  $\Delta\bar{\nu}_{\text{HB}}$ , with the most H-bonding variants, F28oCNF and F92oCNF, possessing  $\Delta\bar{\nu}_{\text{HB}}$  values of 7 and 21 cm⁻¹, respectively (Table 1). This is an intriguing observation, as  $\Delta\bar{\nu}_{\text{HB}}$  has been shown to be sensitive to H-bond geometry (i.e., H-bond angles and distances)<sup>24,40</sup> and the geometry determines H-bond strength.<sup>41</sup>

To further refine our analysis, we calculated the dependence of  $\Delta\bar{\nu}_{\text{HB}}$  on the heavy atom H-bond angle ( $\text{O}\cdots\text{N}\equiv\text{C}$ ;  $\alpha$  in Figure 4; used because proton positions are not resolved in the X-ray structures) for a comparison with the experimental results. A geometry optimization and vibrational frequency analysis (b3lyp/6-311++g(d,p)) were performed on oTN and a single water where  $\alpha$  was varied from  $70^\circ$  to  $175^\circ$  (see Section S16 in the Supporting Information). The H-bond distance was similar

(3.0–3.2 Å over  $90$ – $175^\circ$ ) following geometry optimization, and DFT-calculated frequencies demonstrate an angle dependence on the H-bond shift for oTN (Figure S19D) with the most significant blueshifts near  $180^\circ$ , similar to the findings of Choi et al.<sup>40</sup> Given the H-bond angles observed between F92oCNF and T90 ( $174^\circ$ ) and between F28oCNF and  $\text{H}_2\text{O}\cdots\text{A}$  ( $110^\circ$ ), these DFT calculations predict  $\Delta\bar{\nu}_{\text{HB}}$  values of  $\sim 4$  and  $\sim 22$  cm⁻¹ for F28oCNF and F92oCNF, respectively. While the H-bond distance and number of donors also contribute to  $\Delta\bar{\nu}_{\text{HB}}$ ,<sup>17</sup> we observe good agreement between the predicted  $\Delta\bar{\nu}_{\text{HB}}$  values and observed values (7 and 21 cm⁻¹). These results indicate the potential of  $\Delta\bar{\nu}_{\text{HB}}$  to report on H-bonding geometries, though further benchmarking is necessary.

In summary, a combined analysis of peak frequency and integrated intensity tuning with nitrile probes can be used to extract environmental electric fields and characterize H-bonding. A TDM analysis enables the quantification of significantly different fields for nitriles whose frequencies are nearly identical. We have utilized a simple phenomenological approach in this communication, motivating a more comprehensive theoretical description of TDM and frequency tuning. It may be useful to re-evaluate intensity changes seen in other studies.<sup>15,28</sup> Moreover, nitrile intensities can be used to determine electric fields and H-bonding in the extensive array of systems in which nitriles have been incorporated.<sup>7,9–11,13,14,16,42</sup>



**Figure 4.** Heavy-atom H-bond angle ( $\text{O}\cdots\text{N}\equiv\text{C}$ ;  $\alpha$ ) between a water and the nitrile of oTN, with the angle shown in blue.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.2c00675>.

Experimental details on sample preparation and IR spectroscopy, IR spectra and TDM calibration curves for oTN, fixed-charge TDM calibration curves for ACP and BZN, Stark spectroscopy comparison and discussion, UV-vis spectra and visible extinction coefficient determination of PYP variants, mass spectra of PYP variants, crystallography, DFT of H-bond shift calculations, and H-bond shift discussion ([PDF](#))

## ■ AUTHOR INFORMATION

### Corresponding Author

Steven G. Boxer – Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States; [orcid.org/0000-0001-9167-4286](https://orcid.org/0000-0001-9167-4286); Email: [sboxer@stanford.edu](mailto:sboxer@stanford.edu)

### Authors

Jared Bryce Weaver – Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States; [orcid.org/0000-0002-3823-422X](https://orcid.org/0000-0002-3823-422X)

Jacek Kozuch – Experimental Molecular Biophysics, Department of Physics, Freie Universität Berlin, 14195 Berlin, Germany; [orcid.org/0000-0002-2115-4899](https://orcid.org/0000-0002-2115-4899)

Jacob M. Kirsh – Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States; [orcid.org/0000-0002-1444-2913](https://orcid.org/0000-0002-1444-2913)

Complete contact information is available at:

<https://pubs.acs.org/10.1021/jacs.2c00675>

### Notes

The authors declare no competing financial interest.

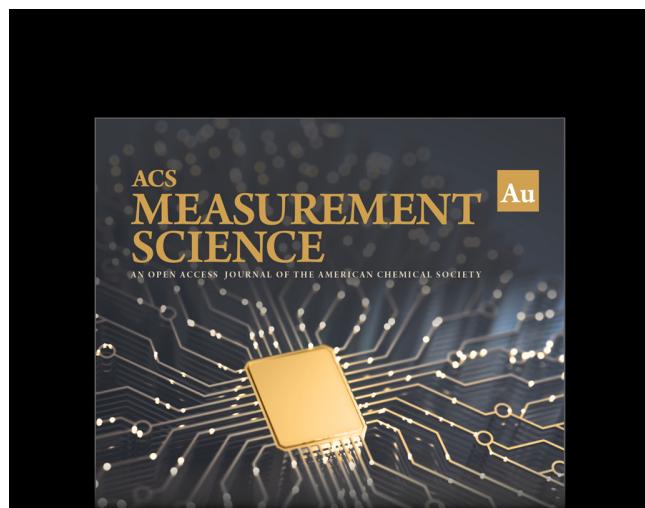
## ■ ACKNOWLEDGMENTS

J.B.W. was supported by a Stanford Center for Molecular Analysis and Design Fellowship. J.M.K. was supported by a Burt and Deedee McMurtry Stanford Graduate Fellowship. This work was supported by grants from the NSF (MCB-1915727) and NIH (R35GM118044) to S.G.B. and the DFG (Project ID 221545957-SFB 1078/B9) to J.K. We thank Dr. Chi-Yun Lin for helpful discussions regarding the extraction of transition polarizabilities from vibrational Stark spectral fits.

## ■ REFERENCES

- (1) Fried, S. D.; Boxer, S. G. Measuring Electric Fields and Noncovalent Interactions Using the Vibrational Stark Effect. *Acc. Chem. Res.* **2015**, *48*, 998–1006.
- (2) Chattopadhyay, A.; Boxer, S. G. Vibrational Stark Effect Spectroscopy. *J. Am. Chem. Soc.* **1995**, *117* (4), 1449–1450.
- (3) Andrews, S. S.; Boxer, S. G. Vibrational Stark Effects of Nitriles I. Methods and Experimental Results. *J. Phys. Chem. A* **2000**, *104* (51), 11853–11863.
- (4) Andrews, S. S.; Boxer, S. G. Vibrational Stark Effects of Nitriles II. Physical Origins of Stark Effects from Experiment and Perturbation Models. *J. Phys. Chem. A* **2002**, *106* (3), 469–477.
- (5) Suydam, I. T.; Snow, C. D.; Pande, V. S.; Boxer, S. G. Electric Fields at the Active Site of an Enzyme: Direct Comparison of Experiment with Theory. *Science* **2006**, *313*, 200–204.
- (6) Wang, X.; Wang, Y.; Li, X.; Yu, Z.; Song, C.; Du, Y. Nitrile-Containing Pharmaceuticals: Target, Mechanism of Action, and Their SAR Studies. *RSC Medicinal Chemistry* **2021**, *12* (10), 1650–1671.
- (7) Fafarman, A. T.; Webb, L. J.; Chuang, J. I.; Boxer, S. G. Site-Specific Conversion of Cysteine Thiols into Thiocyanate Creates an IR Probe for Electric Fields in Proteins. *J. Am. Chem. Soc.* **2006**, *128* (41), 13356–13357.
- (8) Fafarman, A. T.; Sigala, P. A.; Schwans, J. P.; Fenn, T. D.; Herschlag, D.; Boxer, S. G. Quantitative, Directional Measurement of Electric Field Heterogeneity in the Active Site of Ketosteroid Isomerase. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (6), E299–E308.
- (9) Adesina, A. S.; Swiderek, K.; Luk, L. Y. P.; Moliner, V.; Allemann, R. K. Electric Field Measurements Reveal the Pivotal Role of Cofactor–Substrate Interaction in Dihydrofolate Reductase Catalysis. *ACS Catal.* **2020**, *10* (14), 7907–7914.
- (10) Baiz, C. R.; Blasiak, B.; Bredenbeck, J.; Cho, M.; Choi, J.-H.; Corcelli, S. A.; Dijkstra, A. G.; Feng, C.-J.; Garrett-Roe, S.; Ge, N.-H.; Hanson-Heine, M. W. D.; Hirst, J. D.; Jansen, T. L. C.; Kwac, K.; Kubarych, K. J.; Lonergan, C. H.; Maekawa, H.; Reppert, M.; Saito, S.; Roy, S.; Skinner, J. L.; Stock, G.; Straub, J. E.; Thielges, M. C.; Tominaga, K.; Tokmakoff, A.; Torii, H.; Wang, L.; Webb, L. J.; Zanni, M. T. Vibrational Spectroscopic Map, Vibrational Spectroscopy, and Intermolecular Interaction. *Chem. Rev.* **2020**, *120* (15), 7152–7218.
- (11) Silverman, L. N.; Pitzer, M. E.; Ankomah, P. O.; Boxer, S. G.; Fenlon, E. E. Vibrational Stark Effect Probes for Nucleic Acids. *J. Phys. Chem. B* **2007**, *111* (40), 11611–11613.
- (12) Gai, X. S.; Fenlon, E. E.; Brewer, S. H. A Sensitive Multispectroscopic Probe for Nucleic Acids. *J. Phys. Chem. B* **2010**, *114* (23), 7958–7966.
- (13) Hu, W.; Webb, L. J. Direct Measurement of the Membrane Dipole Field in Bicelles Using Vibrational Stark Effect Spectroscopy. *J. Phys. Chem. Lett.* **2011**, *2* (15), 1925–1930.
- (14) Xiong, H.; Lee, J. K.; Zare, R. N.; Min, W. Strong Electric Field Observed at the Interface of Aqueous Microdroplets. *J. Phys. Chem. Lett.* **2020**, *11* (17), 7423–7428.
- (15) Staffa, J. K.; Lorenz, L.; Stolarski, M.; Murgida, D. H.; Zebger, I.; Uttesch, T.; Kozuch, J.; Hildebrandt, P. Determination of the Local Electric Field at Au/SAM Interfaces Using the Vibrational Stark Effect. *J. Phys. Chem. C* **2017**, *121* (40), 22274–22285.
- (16) Bhattacharyya, D.; Videla, P. E.; Cattaneo, M.; Batista, V. S.; Lian, T.; Kubiak, C. P. Vibrational Stark Shift Spectroscopy of Catalysts under the Influence of Electric Fields at Electrode–Solution Interfaces. *Chem. Sci.* **2021**, *12* (30), 10131–10149.
- (17) Blasiak, B.; Ritchie, A. W.; Webb, L. J.; Cho, M. Vibrational Solvatochromism of Nitrile Infrared Probes: Beyond the Vibrational Stark Dipole Approach. *Phys. Chem. Chem. Phys.* **2016**, *18* (27), 18094–18111.
- (18) Reimers, J. R.; Hall, L. E. The Solvation of Acetonitrile. *J. Am. Chem. Soc.* **1999**, *121* (15), 3730–3744.
- (19) Aschaffenburg, D. J.; Moog, R. S. Probing Hydrogen Bonding Environments: Solvatochromic Effects on the CN Vibration of Benzonitrile. *J. Phys. Chem. B* **2009**, *113* (38), 12736–12743.
- (20) Kraskov, A.; von Sass, J.; Nguyen, A. D.; Hoang, T. O.; Buhrk, D.; Katz, S.; Michael, N.; Kozuch, J.; Zebger, I.; Siebert, F.; Scheerer, P.; Mruginski, M. A.; Budisa, N.; Hildebrandt, P. Local Electric Field Changes during the Photoconversion of the Bathy Phytochrome Agp2. *Biochemistry* **2021**, *60* (40), 2967–2977.
- (21) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *Science* **2014**, *346* (6216), 1510–1514.
- (22) Schneider, S. H.; Kozuch, J.; Boxer, S. G. The Interplay of Electrostatics and Chemical Positioning in the Evolution of Antibiotic Resistance in TEM  $\beta$ -Lactamases. *ACS Central Science* **2021**, *7*, 1996–2008.
- (23) Zheng, C.; Mao, Y.; Kozuch, J.; Atsango, A.; Ji, Z.; Markland, T.; Boxer, S. Electric Field Orientations in Solution and Enzyme Active Site Revealed by a Two-Directional Vibrational Probe. *Nat. Chem.* **2021**, DOI: [10.33774/chemrxiv-2021-578vg](https://doi.org/10.33774/chemrxiv-2021-578vg).
- (24) Lee, H.; Choi, J.-H.; Cho, M. Vibrational Solvatochromism and Electrochromism. II. Multipole Analysis. *J. Chem. Phys.* **2012**, *137* (11), 114307.
- (25) First, J. T.; Slocum, J. D.; Webb, L. J. Quantifying the Effects of Hydrogen Bonding on Nitrile Frequencies in GFP: Beyond Solvent Exposure. *J. Phys. Chem. B* **2018**, *122* (26), 6733–6743.
- (26) Acharyya, A.; Mukherjee, D.; Gai, F. Assessing the Effect of Hofmeister Anions on the Hydrogen-Bonding Strength of Water via

- Nitrile Stretching Frequency Shift. *J. Phys. Chem. B* **2020**, *124* (52), 11783–11792.
- (27) Eaton, G.; Pena-Nuñez, A. S.; Symons, M. C. R. Solvation of Cyanoalkanes [CH<sub>3</sub>CN and (CH<sub>3</sub>)<sub>3</sub>NN]. An Infrared and Nuclear Magnetic Resonance Study. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84* (6), 2181–2193.
- (28) Delley, M. F.; Nichols, E. M.; Mayer, J. M. Interfacial Acid–Base Equilibria and Electric Fields Concurrently Probed by *In Situ* Surface-Enhanced Infrared Spectroscopy. *J. Am. Chem. Soc.* **2021**, *143* (28), 10778–10792.
- (29) Yan, J.; Wilson, R. W.; Buck, J. T.; Grills, D. C.; Reinheimer, E. W.; Mani, T. IR Linewidth and Intensity Amplifications of Nitrile Vibrations Report Nuclear-Electronic Couplings and Associated Structural Heterogeneity in Radical Anions. *Chem. Sci.* **2021**, *12* (36), 12107–12117.
- (30) Bublitz, G. U.; Boxer, S. G. Stark Spectroscopy: Applications in Chemistry, Biology, and Materials Science. *Annu. Rev. Phys. Chem.* **1997**, *48* (1), 213–242.
- (31) Park, E. S.; Boxer, S. G. Origins of the Sensitivity of Molecular Vibrations to Electric Fields: Carbonyl and Nitrosyl Stretches in Model Compounds and Proteins. *J. Phys. Chem. B* **2002**, *106* (22), 5800–5806.
- (32) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114* (8), 2549–2564.
- (33) Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. Tinker 8: Software Tools for Molecular Design. *J. Chem. Theory Comput.* **2018**, *14* (10), 5273–5289.
- (34) Fried, S. D.; Bagchi, S.; Boxer, S. G. Measuring Electrostatic Fields in Both Hydrogen-Bonding and Non-Hydrogen-Bonding Environments Using Carbonyl Vibrational Probes. *J. Am. Chem. Soc.* **2013**, *135* (30), 11181–11192.
- (35) Schneider, S. H.; Boxer, S. G. Vibrational Stark Effects of Carbonyl Probes Applied to Reinterpret IR and Raman Data for Enzyme Inhibitors in Terms of Electric Fields at the Active Site. *J. Phys. Chem. B* **2016**, *120* (36), 9672–9684.
- (36) Deb, P.; Haldar, T.; Kashid, S. M.; Banerjee, S.; Chakrabarty, S.; Bagchi, S. Correlating Nitrile IR Frequencies to Local Electrostatics Quantifies Noncovalent Interactions of Peptides and Proteins. *J. Phys. Chem. B* **2016**, *120*, 4034–4046.
- (37) Kozuch, J.; Schneider, S. H.; Zheng, C.; Ji, Z.; Bradshaw, R. T.; Boxer, S. G. Testing the Limitations of MD-Based Local Electric Fields Using the Vibrational Stark Effect in Solution: Penicillin G as a Test Case. *J. Phys. Chem. B* **2021**, *125*, 4415.
- (38) Reimers, J. R.; Zeng, J.; Hush, N. S. Vibrational Stark Spectroscopy. 2. Application to the CN Stretch in HCN and Acetonitrile. *J. Phys. Chem.* **1996**, *100* (5), 1498–1504.
- (39) Tharp, J. M.; Wang, Y. S.; Lee, Y. J.; Yang, Y.; Liu, W. R. Genetic Incorporation of Seven Ortho-Substituted Phenylalanine Derivatives. *ACS Chem. Biol.* **2014**, *9*, 884–890.
- (40) Choi, J.-H.; Oh, K.-I.; Lee, H.; Lee, C.; Cho, M. Nitrile and Thiocyanate IR Probes: Quantum Chemistry Calculation Studies and Multivariate Least-Square Fitting Analysis. *J. Chem. Phys.* **2008**, *128* (13), 134506.
- (41) Steiner, T. The Hydrogen Bond in the Solid State. *Angew. Chem., Int. Ed.* **2002**, *41* (1), 48–76.
- (42) Hopmann, K. H. Full Reaction Mechanism of Nitrile Hydratase: A Cyclic Intermediate and an Unexpected Disulfide Switch. *Inorg. Chem.* **2014**, *53* (6), 2760–2762.



**ACS MEASUREMENT SCIENCE**  
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

 Editor-in-Chief  
**Prof. Shelley D. Minteer**  
University of Utah, USA

**Open for Submissions** 

pubs.acs.org/measureau  ACS Publications  
Most Trusted. Most Cited. Most Read.

**Nitrile infrared (IR) intensities characterize electric fields and hydrogen bonding in protic, aprotic, and protein environments**

Jared Bryce Weaver,<sup>1</sup> Jacek Kozuch,<sup>2</sup> Jacob M. Kirsh,<sup>1</sup> and Steven G. Boxer<sup>1,\*</sup>

<sup>1</sup>*Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States*

<sup>2</sup>*Experimental Molecular Biophysics, Department of Physics, Freie Universität Berlin, 14195 Berlin, Germany*

Table of Contents

S1	Sample Preparation, IR Spectra Acquisition, and Spectra Processing .....	S2
S2	Incorporation of oCNF in PYP .....	S4
S3	IR Spectra of ACP, BZN, and oTN .....	S9
S4	ACP (-C=O), BZN (-C≡N), and oTN (-C≡N) Peak Areas Relative to Hexanes .....	S10
S5	Molecular Dynamics (MD) Solvent Electric Field Calculation.....	S12
S6	IR Transition Dipole Moment (TDM) as a Function of Solvent Field .....	S14
S7	IR Peak Frequency as a Function of Solvent Electric Field .....	S16
S8	IR Peak Frequencies, FWHM, TDMs, and Fields Calculated from MD .....	S17
S9	Vibrational Stark Spectroscopy of oTN and Comparison to Solvatochromism-based Frequency-Field and TDM-Field Calibrations.....	S19
S10	Targeted Sites for oCNF Incorporation in PYP.....	S22
S11	Crystallization of PYP, X-ray Data Collection, and Refinement of PYP Structures.....	S23
S13	PYP Extinction Coefficient Determination .....	S28
S14	PYP UV-Vis Spectra .....	S33
S15	Protein IR Spectra Compared with oTN spectra in Water and Hexanes.....	S34
S16	DFT Calculations of Frequency Dependence on H-bonding Angle .....	S35
S17	Commentary on the Additive Model of Main Text Equation 4 .....	S39
S18	References.....	S42

## S1 Sample Preparation, IR Spectra Acquisition, and Spectra Processing

Acetophenone (ACP, Sigma-Aldrich), benzonitrile (BZN, Sigma-Aldrich), and *o*-tolunitrile (*o*TN, TCI America) were all prepared to a concentration of 20 mM in each solvent, except for when the solutes were dissolved in water. All solvents were anhydrous and of 99.0% purity or higher (Sigma-Aldrich). Aqueous solutions of ACP, *o*TN, and BZN were prepared to a concentration of 4.2 mM, 3.9 mM, and 5.0 mM, respectively, to ensure solutes were entirely dissolved. Appropriate solute concentrations for IR were determined by verifying solutes were still miscible at concentrations that were at least double the concentration used in IR experiments. Sample concentrations were determined by mass, where solute masses had a relative error of <1% and solvent volume had a relative error of  $\pm 0.1\%$ . For ACP, BZN and *o*TN dissolved in water, solute mass error was estimated to be <2% of total mass. Mass error was estimated by reweighing all solute masses to estimate reproducibility. In the case of *o*TN, samples were prepared three times in each solvent and three concentration normalized IR spectra were acquired for each sample preparation. Solutions were prepared fresh the day spectra were acquired.

FTIR spectra were obtained on a Bruker Vertex 70 spectrometer with a liquid-nitrogen-cooled mercury cadmium telluride (MCT) detector and continuous purging of the sample chamber with dry air. Following previous protocol,<sup>1,2</sup> 20-30  $\mu$ L of sample was loaded into a demountable IR cell composed of two CaF<sub>2</sub> optical windows (19.05 mm diameter, 3.00 mm thickness, Lambda Research Optics, Inc.), separated by two nonoverlapping 56  $\mu$ m spacers yielding a pathlength of approximately 56  $\mu$ m. More exact pathlengths (distance known to  $\pm 0.2 \mu$ m, error on pathlength is for an unfilled sample cell) were determined by measuring the interference fringes present in the empty assembled IR cell prior to inserting sample through the injection port. Once solutions (or protein samples) were injected into the assembled sample cell, the entire sample cell was wrapped in parafilm at all possible junctions with outer air and multiple IR spectra were acquired in succession (~7 min intervals) to ensure that sample was not evaporating over time. For ACP and BZN, six spectra were acquired for solute dissolved in each solvent (for *o*TN, 9 spectra were acquired per solvent). The sample compartment was purged with 5-10 minutes of dry air prior to acquisition of spectra. An aperture of 3 to 6 mm was used during spectra acquisition and 256 to 512 scans were averaged to obtain each spectrum. For all protein IR measurements, proteins were exchanged into a pH 6.0 buffer containing 20 mM NaH<sub>2</sub>PO<sub>4</sub> and 10 mM NaCl. All IR spectra were acquired at room-temperature except for those in Figure S10 which were acquired at 77K. A UV-vis absorption spectrum of the sample in the IR cell was acquired prior to IR absorption measurement. The IR sample compartment was purged for 10 min prior to spectra acquisition and 9 spectra (each averaging 512 scans) were acquired.

IR spectra were minimally processed, with a shallow (<2% maximum peak intensity) polynomial (five-root or less) baseline applied. Each of the 6-9 spectra acquired per solvent/protein were processed independently. For small molecule solvents (water/organics), the solvent was used as the blank and for protein samples, WT PYP was used as a blank. Solvents in the main text and in Sections S3-S4 and S6-S8 are abbreviated as follows in parenthesis: dimethyl sulfoxide (DMSO), tetrahydrofuran (THF), dibutyl ether (DBE), chloroform (CFM), dichloromethane (DCM), and dimethyl formamide (DMF). Peaks were fit using a pseudo-Voigt lineshape in OPUS 6.5. Fits for *o*TN in organic solvents and water were improved by the inclusion of a second smaller peak (<20% the area of the major peak) to account for a small shoulder (Figure S3). That said, two-peak fitting didn't affect analysis of nitrile frequencies: positions

determined both using one-peak fits and OPUS software's peak picker (2<sup>nd</sup> Derivative mode) were always similar (within 0.3 cm<sup>-1</sup>) to positions determined for the major peak in two-peak fits.

Peak areas were determined in Excel; for each solvent, absorbance for the baselined spectrum (not the fits) was integrated from 2180 cm<sup>-1</sup> to 2280 cm<sup>-1</sup>. Peak areas determined in OriginPro using Peak Analyzer gave similar values (within 4%).

## S2 Incorporation of oCNF in PYP

### *Cell Grown and Protein Expression*

MAX Efficiency™ DH10B Competent Cells (ThermoFisher) were doubly transformed with pEVOL-pyIT-N346AC348A (Addgene) and either pBAD-PYP (for WT protein expression), pBAD-PYP-F28TAG, pBAD-PYP-F62TAG, pBAD-PYP-F92TAG, or pBAD-PYP-F96TAG. PYP plasmids were generated from the pBAD-PYP plasmid using an Agilent QuikChange Lightning Site-Directed Mutagenesis Kit (see gene sequences below). Doubly transformed cells were then expressed following literature procedure.<sup>3</sup> Briefly, for each liter of protein expression for a given transformed cell strain, a 14 mL round-bottom culture tube with 5 mL of Luria-Bertaini growth medium (LB, Millipore) was inoculated with transformed cells and grown for 12 to 16 hours (or until saturated). LB was supplemented with 100 µg/mL ampicillin and 34 µg/mL chloramphenicol for pBAD and pEVOL plasmid selection markers, respectively. Starter cultures were then poured into 1 L of auto-induction media where phenylalanine had been omitted,<sup>4</sup> also containing 100 µg/mL ampicillin (Sigma-Aldrich) and 34 µg/mL chloramphenicol (Sigma-Aldrich), which had been warmed to 37° C for 1 hour. As a note, auto-induction media contains arabinose (Sigma-Aldrich) which is added following sterilization of media via autoclave. Media was shaken for 30 minutes at 250 RPM, 37° C and then 10 mL of 100 mM oCNF was added. All oCNF was purchased from ChemImpex and 100 mM oCNF was prepared ~1-2 hrs prior to its addition to cell media. Protein expression was then auto-induced for ~19 hours at 250 RPM, 37° C. Cells were pelleted via centrifugation.

### *Apo-PYP Isolation*

Apo-PYP was purified following literature protocol.<sup>5</sup> Pelleted cells from centrifugation were resuspended in Buffer A (20 mM Tris HCl (Fisher Scientific), 10 mM NaCl (Fisher Scientific), pH 8.0) with 2 mL buffer added for every 1 mL of pelleted cells. Cells were lysed via homogenization in a C3 Emulsiflex, and insoluble cell debris was spun down by two successive rounds of 90 min. of centrifugation at 26,700 G. Cell lysate was collected and kept on ice while PYP chromophore activation was performed.

### *PYP Chromophore Activation*

To a dry 250 mL round-bottom flask, 1.54 mmol (253.3 mg) of *p*-coumaric acid (pCA, Sigma-Aldrich), 1.72 mmol (279.5 mg) carbonyl diimidazole (Sigma-Aldrich), and a dry stir bar were added. A rubber septum was attached, and the vessel was purged for 30 minutes with argon. Approximately 24 mL of anhydrous tetrahydrofuran (Sigma-Aldrich) was added while stirring rapidly and the reaction was allowed to proceed for 1 hour. The reaction mixture was then rotavapped under gentle heat (<45° C) until all THF had evaporated and a yellow oil remained.

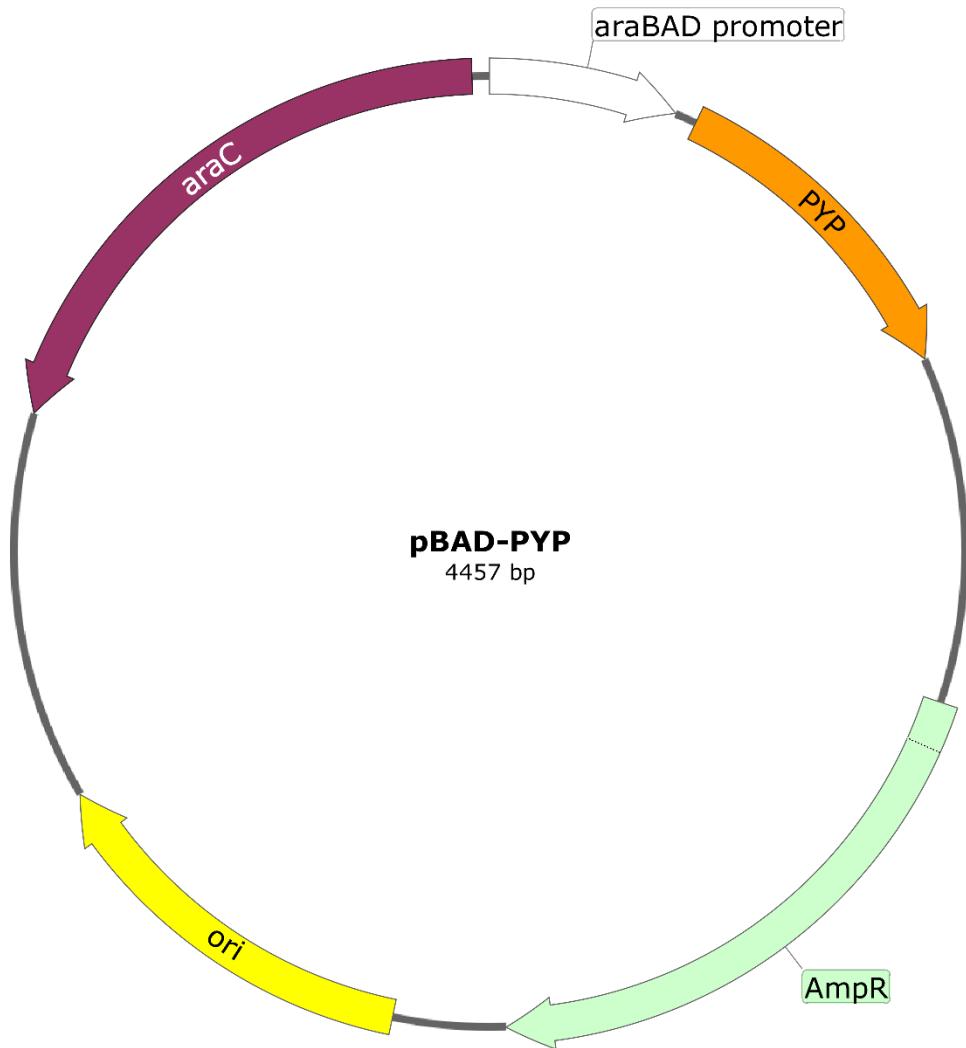
### *PYP Chromophore Incorporation and PYP Purification*

The crude apo-PYP cell lysate from one liter of expressed cells was added to the activated PYP chromophore immediately following THF removal. The activated chromophore was scraped off the sides of the round bottom flask and into the cell lysate. The mixture was allowed to stir gently (<200 RPM) overnight such that foam was not introduced by stirring. The following morning, insoluble reaction precipitate was spun down with 90 minutes of centrifugation at 26,700 G.

The supernatant (crude PYP) was then bound to a gravity Ni-nitrilotriacetic acid affinity column (Qiagen) and the column was washed with two column-volumes (CVs) of Buffer A. Two CVs of Wash Buffer (Buffer A + 20 mM Imidazole) were added taking care to collect PYP that had begun to elute towards the end of the 2<sup>nd</sup> CV of Wash Buffer added. PYP was eluted via addition of 2 CVs of Elution Buffer (Buffer A + 200 mM Imidazole) and eluted yellow fractions from the end of the PYP Wash buffer addition and the Elution Buffer addition were pooled.

Eluted PYP was purified via anion exchange chromatography where Buffer A was loaded onto an FPLC anion exchange column (3x5 mL Q HP columns connected in series) which had been pre-equilibrated in Buffer A. Protein was purified by running a gradient of 0–40% Buffer B (1 M NaCl, 20 mM Tris HCl, pH 8.0) at a flow rate of 5 mL/min over 180 mL. Protein which eluted between 25% and 32% Buffer B was pooled. Pooled fractions were placed in 15 mL 10 kDa Centrifugal Filter Units (Millipore) and NaCl was removed through buffer exchange with Buffer A. Protein was then concentrated to 8 mg/mL and the N-terminal hexahistidine affinity tag was removed via addition of 1 U enterokinase (New England Biolabs, Catalog #P8070S) for every 0.025 mg of protein and 2 µL of 1 M CaCl<sub>2</sub> for every 1 mL of protein solution. The enterokinase and PYP were softly shaken at 25 RPM in the dark at 25 °C for 16 hours. The mixture was then re-purified with anion exchange chromatography as performed following the Ni-NTA column. Cleaved PYP fractions eluted between 15 and 19% Buffer B. Fractions with maximal Abs<sub>446nm</sub>/Abs<sub>280nm</sub> ratios (~2.4 for WT) were pooled. As before, NaCl was removed through buffer exchange with Buffer A on a 10 kDa Amicon centrifugal filter unit. For sample intended for protein IR, PYP was concentrated further on a 0.5 mL 10 kDa Amicon centrifugal filter unit by concentrating PYP to ~0.02 mL and then spinning a further hour at 14,000 G at 4 °C. Concentrated PYP was collected, pipetted up and down to ensure sample was homogenous, and was transferred to a 0.6 mL microcentrifuge tube. Immediately prior to use for FTIR spectroscopy, sample was spun down for one hour at 21,000 G at 4°C.

*Plasmid Information*



**Figure S1.** Plasmid map of the WT pBAD-PYP plasmid, with the PYP gene, the high copy ColE1 origin (ori) of replication, Ampicillin resistance gene, and the araC gene and araBAD promoter displayed.

**pBAD-PYP (orange=PYP):**

```

AAGAAACCAATTGTCCATATTGCATCAGACATTGCCGTCACTGCGTCTTTACTGGCTCTCGCTAACCAAACCG
GTAACCCCGCTTATAAAAGCATTCTGTAACAAAGCGGGACAAAGCCATGACAAAAACGCGTAACAAAAGTGTG
TATAATCACGGCAGAAAAGTCCACATTGATTATTGCACGGCGTCACACTTGCTATGCCATAGCATTATCCAT
AAGATTAGCGGATCCTACCTGACGTTTATCGCAACTCTACTGTTCTCCATACCCGTTTTGGGCTAACAGG
AGGAATTAACCATGGGGGGTTCTCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGG
GTCGGGATCTGTACGACGATGACGATAAGGATCGAGACTCGAGGGATGACGATGACAAAATGGAACACGTAGCC
TTCGGTAGCGAGGACATCGAGAACACCCCTGCCAAGATGGACGACGGCCAGCTCGACGGCCTGGCCTGGCGC
CATCCAGCTCGACGGCGACGGCAACATCCTCAGTACAACGCCGGAGGGCGACATCACCGCCGACCCGAA
GCAGGTCATCGGCAAGAACTTCTCAAGGACGTGGCCCCGTGCACTGACAGCCCAGGAGTTCTACGGCAAGTTCAA
GGAGGGGGTGGCCTGGGCAACCTGAACACGATGTTGAGTACACCTCGATTACCAAATGACGCCACGAAGGT
GAAGGGTGCACATGAAGAAGGCCCTCTCCGGCGACAGCTACTGGTCTCGTAAGCGCGTCAAAAGCTGGCTG

```

TTTGGCGGATGAGAGAAGATTCAGCCTGATACAGATTAAATCAGAACGCAGAACGGCTGTATAAAACAGAA  
TTGCCTGGCGGAGTAGCGCGTGGTCCACCTGACCCCATGCCGAACTCAGAAAGTGAACGCCGTAGGCCGA  
TGGTAGTGTGGGTCTCCCCTGCGAGAGTAGGAACTGCCAGGCATCAAATAAACGAAAGGCTAGTCGAAA  
GACTGGGCCCTTCGTTTATCTGTTGTCGGTAACGCTCTCTGAGTAGGACAATCCGCCGGAGCGGATT  
TGAACGTTGCGAAGCAACGGCCGGAGGGTGGCGGGCAGGACGCCGCATAAAACTGCCAGGCATCAAATAAG  
CAGAAGGCCATCCTGACGGATGCCCTTTGCGTTACAAACTCTTGTTATTTCTAAATACATTCAAATAT  
GTATCCGCTCATGAGACAATAACCTGATAAAATGCTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACAT  
TCCGTGTCGCCCTTATCCCTTTGCGGCATTTGCCCTGCTTTGCTACCCAGAAACGCTGGTAAAGATCCTGAG  
AAAGATGCTGAAGATCAGTTGGTGACGAGTGGTTACATCGAACTGGATCTCAACAGCGTAAGATCCTGAG  
AGTTTCGCCCGAAGAACGTTCCAATGATGAGCACTTAAAGTTCTGCTATGTGGCGCGTATTATCCGTG  
TGACGCCGGCAAGAGCAACTCGTCGCCGCATAACTATTCTCAGAATGACTGGTGAGTACTCACAGTCACA  
GAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCG  
GCCAACTTACTCTGACAACGATCGGAGGACGAAGGAGCTAACGCTTTTGACAAACATGGGGATCATGTA  
ACTCGCCTGATCGTGGAACCGGAGCTGAATGAAGCCATACCAAAACGACGAGCGTGACACCACGATGCCGTG  
GCAATGGCAACAACGTTGCGAAACTATTAACTGGCGAACTACTTACTCTAGCTCCCGCAACAATTAAAGACT  
GGATGGAGGCGGATAAAGTTGAGGACCACTCTCGCCTGGCTGGCTGGTTATTGCTGATAAAT  
CTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCGACTGGGCCAGATGGTAAGCCCTCCGTATCGTAG  
TTATCTACACGACGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGTGAGATAGGTGCTCACTGA  
TTAAGCATTGTAACTGTCAGACCAAGTTACTCATATATACTTTAGATTGATTAAAACCTCATTAAATTAAAA  
GGATCTAGGTGAAGATCCTTTGATAATCTCATGACCAAATCCCTAACGTGAGTTCTGTTCCACTGAGCGTC  
GACCCGCTAGAAAAGATCAAAGGATCTTCTGAGATCCTTTCTCGCGTAATCTGCTGCTGCAAACAAAAAA  
ACCAACGCTACCAGCGGTGGTTGCTGCCGATCAAGAGCTACCAACTCTTCCGAAGGTAACTGGCTTCAGC  
AGAGCGCAGATAACAAACTGTCTCTAGTGTAGCGTAGTTAGGCCACCTCAAGAAACTCTGTAGCACCGC  
CTACATACCTCGCTGCTAATCTGTTACCGTAGGCTGCCAGTGGCGATAAGTCGTCTTACGGGTTGGA  
CTCAAGACGATAGTTACCGATAAGGCGCAGCGGTGGCTGAACGGGGGTCGTGACACAGCCAGCTGG  
AGCGAACGACCTACACCGAACTGAGACACTACAGCGTGAGCTATGAGAAAGGCCACGCTCCGAAGGGAGA  
AAGGCGGACAGGTATCCGTAAGCGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCAAGGGGAAAC  
GCCTGGTATTTAGTCCTGCGGTTGCCACCTCTGACTTGAGCGTCGATTTTGATGCTCGTCAGGGGG  
GCCGGAGCCTATGGAAAAAACGCCAGCAACCGCCCTTTACGGTCTGCCCTTGCTGGCTTGCTCACATG  
TTCTTCTCGTTATCCCTGATTCTGAGATAACCGTATTACCGCCTTGAGTGAGCTGATACCGCTGCCAG  
CCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCTGATGCCGTATTTCTCCTACGC  
ATCTGTGCGGTATTCACACCGCATATGGTCACTCTCAGTACAATCTGCTGATGCCGATAGTTAACCGAGTAT  
ACACTCCGCTATCGCTACGTGACTGGTCATGGCTGCCCGACACCCGCAACACCCGCTGACGCCCTGACG  
GGCTGTCTGCCGGCATCCGCTACAGACAAGCTGTGACCGCTCCGGAGCTGCATGTCAGAGGTTCA  
CCGTCATACCGAAACCGCGAGGCAGCAGATCAATTGCGCGGAAGGCCAGCGCATGATAATGTGCTG  
TCAAATGGACGAAGCAGGGATTCTGCAAACCCCTATGCTACTCCGTCAGCGTCATTGCTGATTGTTACCAATT  
ATGACAACCTGACGGCTACATCATTCACTTTCTTACAACCGGCACGGAACTCGCTGGGCTGGCCCGGTGCA  
TTTTAAATACCGCGAGAAATAGAGTTGATCGTAAACACATTGCGACCGACGGTGGCGATAGGCATCCG  
GGTGGTGCTAAAGCAGCTTCGCTGGCTGATACGTTGGCTCGCGCAGCTTAAGACGCTAATCCCTAAGTGC  
TGGCGAAAAGATGTGACAGACCGACGGGACAAGCAAACATGCTGTGCGACGCTGGGATATCAAATTGCT  
GTCTGCCAGGTGATCGTACTGACAAGCCTCGCTACCGATTATCCATGGTGGATGGAGCGACTCGTTA  
ATCGCTTCCATCGCCCGAGTAACAATTGCTCAAGCAGATTATCGCCAGCAGCTCCGAATAGGCCCTCCCTG  
CCCGCGTTAATGATTGCCAACAGGTCGCTGAAATGCCGGCTGGTGCCTCATCGGGCGAAAGAACCCGT  
ATTGGCAAATATTGACGGCCAGTTAACGCCATTGCGAGTAGGCCAGTGGCGCGGACGAAAGTAAACCAACTGGTGATA

CCATTCGCAGCCTCCGGATGACGACCGTAGTGATGAATCTCTGGCGGGAACAGCAAATATCACCCGGTCG  
GCAAACAAATTCTCGTCCCTGATTTTCAACCACCCCTGACCGCGAATGGTGGAGATTGAGAATATAACCTTCATT  
CCAGCGGTGGTCGATAAAAAAATCGAGATAACC GTGGCCTCAATGGCGTTAACCCGCCACCAGATGGGCAT  
TAAACGAGTATCCGGCAGCAGGGATCATTTGCCTCAGCCATACTTTCATACTCCGCCATTAGAG

*Mutated PYP Genes*

**F28TAG (mutation highlighted in blue):**

ATGGGGGGTTCTCATCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGTCGGATCTG  
TACGACGATGACGATAAGGATCCGAGACTCGAGGATGACGATGACAAAATGGAACACGTAGCCTCGGTAGCGA  
GGACATCGAGAACACCCCTGCCAAGATGGACGACGCCAGCTGACGCCCTGGCCTAGGGGCCATCCAGCTCG  
ACGGCGACGGCAACATCCTCAGTACAACGCCGCGGAGGGCGACATCACCGGCCGACCCGAAGCAGGTATCG  
GGCAAGAACTTCTCAAGGACGTGGCCCCGTGCACTGACAGCCGGAGTTCTACGGCAAGTTCAAGGAAGGGGT  
GGCCTCGGGCAACCTGAACACGATGTTGAGTACACCTCGATTACCAATGACGCCACGAAGGTGAAGGTGCA  
CATGAAGAAGGCCCTCTCCGGCAGCTACTGGGTCTCGTCAAGCGCGTCAA

**F62TAG (mutation highlighted in blue):**

ATGGGGGGTTCTCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGTCGGATCTG  
TACGACGATGACGATAAGGATCCGAGACTCGAGGATGACGATGACAAAATGGAACACGTAGCCTCGGTAGCGA  
GGACATCGAGAACACCCCTGCCAAGATGGACGACGCCAGCTGACGCCCTGGCCTCGGCCATCCAGCTCGA  
CGCGACGGCAACATCCTCAGTACAACGCCGCGGAGGGCGACATCACCGGCCGACCCGAAGCAGGTATCG  
GCAAGAACTTAGTTCAAGGACGTGGCCCCGTGCACTGACAGCCGGAGTTCTACGGCAAGTTCAAGGAAGGGGT  
GCCTCGGGCAACCTGAACACGATGTTGAGTACACCTCGATTACCAATGACGCCACGAAGGTGAAGGTGCA  
ATGAAGAAGGCCCTCTCCGGCAGCTACTGGGTCTCGTCAAGCGCGTCAA

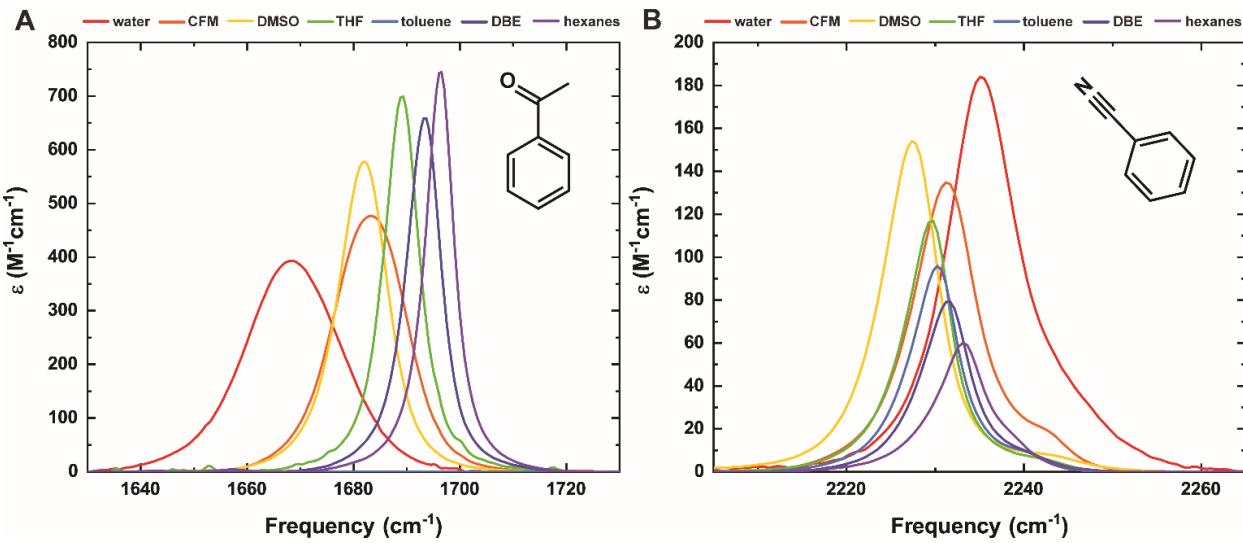
**F92TAG (mutation highlighted in blue):**

ATGGGGGGTTCTCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGTCGGATCTG  
TACGACGATGACGATAAGGATCCGAGACTCGAGGATGACGATGACAAAATGGAACACGTAGCCTCGGTAGCGA  
GGACATCGAGAACACCCCTGCCAAGATGGACGACGCCAGCTGACGCCCTGGCCTCGGCCATCCAGCTCGA  
CGCGACGGCAACATCCTCAGTACAACGCCGCGGAGGGCGACATCACCGGCCGACCCGAAGCAGGTATCG  
GCAAGAACTTCTCAAGGACGTGGCCCCGTGCACTGACAGCCGGAGTTCTACGGCAAGTTCAAGGAAGGGGT  
GCCTCGGGCAACCTGAACACGATGTTGAGTACACCTCGATTACCAATGACGCCACGAAGGTGAAGGTGCA  
ATGAAGAAGGCCCTCTCCGGCAGCTACTGGGTCTCGTCAAGCGCGTCAA

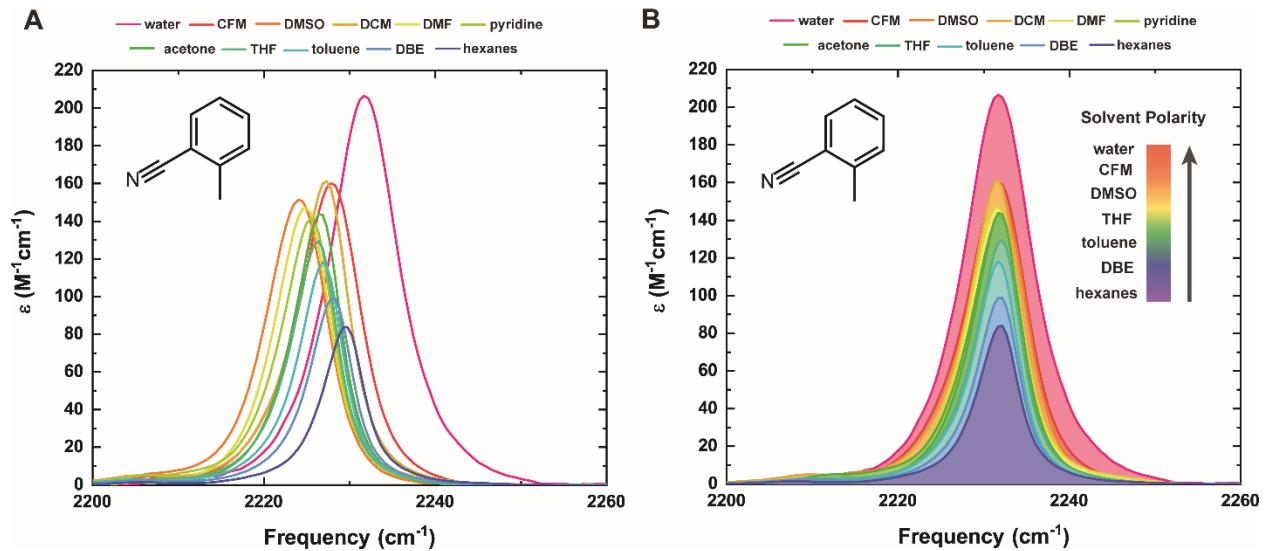
**F96TAG (mutation highlighted in blue):**

ATGGGGGGTTCTCATCATCATCATGGTATGGCTAGCATGACTGGTGGACAGCAAATGGTCGGATCTG  
TACGACGATGACGATAAGGATCCGAGACTCGAGGATGACGATGACAAAATGGAACACGTAGCCTCGGTAGCGA  
GGACATCGAGAACACCCCTGCCAAGATGGACGACGCCAGCTGACGCCCTGGCCTCGGCCATCCAGCTCGA  
CGCGACGGCAACATCCTCAGTACAACGCCGCGGAGGGCGACATCACCGGCCGACCCGAAGCAGGTATCG  
GCAAGAACTTCTCAAGGACGTGGCCCCGTGCACTGACAGCCGGAGTTCTACGGCAAGTTCAAGGAAGGGGT  
GCCTCGGGCAACCTGAACACGATGTTGAGTACACCTAGGATTACCAATGACGCCACGAAGGTGAAGGTGCA  
ATGAAGAAGGCCCTCTCCGGCAGCTACTGGGTCTCGTCAAGCGCGTCAA

S3 IR Spectra of ACP, BZN, and oTN

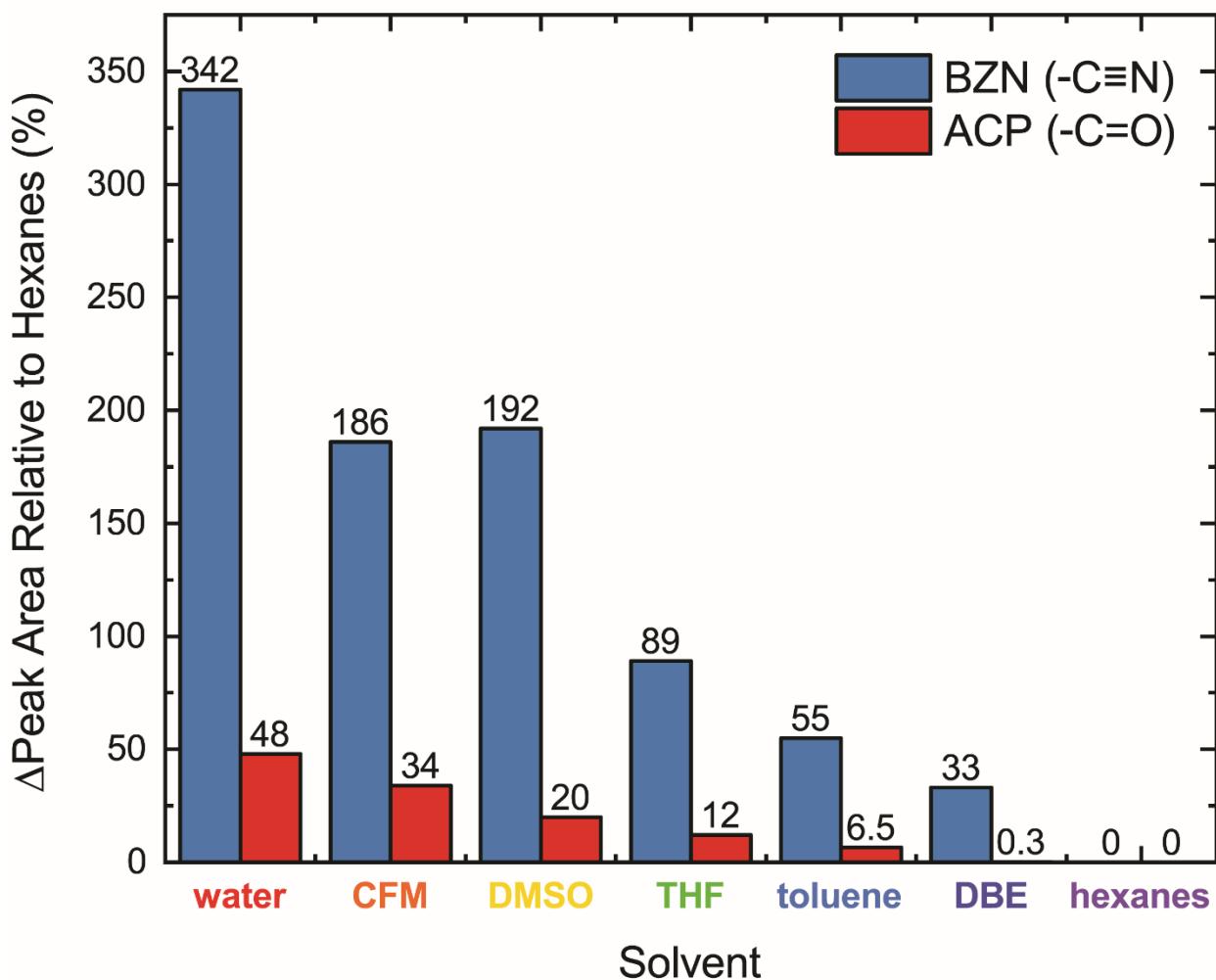


**Figure S2.** Representative room-temperature FTIR absorption spectra of (A) the ACP  $-\text{C=O}$  and (B) BZN  $-\text{C}\equiv\text{N}$  stretch solvated in a range of solvents where solvents are colored from violet to red in order of increasing polarity. Peak area changes are not obvious for ACP but are clear for BZN (see Figure 1 and Figure S3).

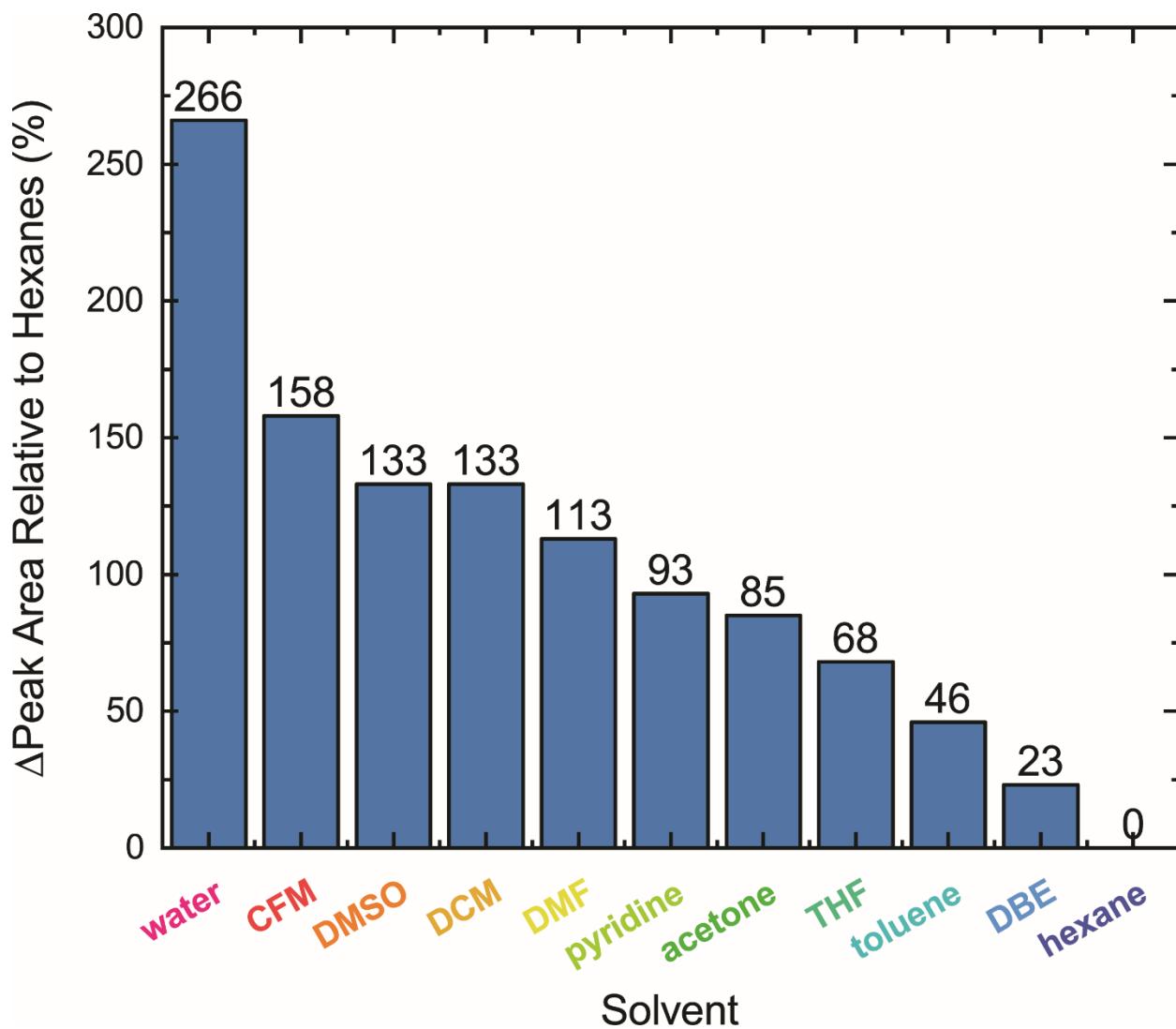


**Figure S3.** IR spectra of oTN. (A) Protic and aprotic solvents show significant tuning of peak area, seen most clearly in (B) where IR spectra are overlaid and peaks are shaded to highlight peak area.

S4 ACP (-C=O), BZN (-C≡N), and oTN (-C≡N) Peak Areas Relative to Hexanes



**Figure S4.** Percent change in peak area of BZN ( $-C\equiv N$ ) and ACP ( $-C=O$ ) in a variety of solvents relative to the peak area in hexanes. Solvent labels are colored from red to violet in order of decreasing solvent polarity.



**Figure S5.** Percent change in peak area of oTN ( $-C\equiv N$ ) in a variety of solvents relative to the peak area in hexanes. Solvent labels are colored from red to dark blue in order of decreasing solvent polarity.

## S5 Molecular Dynamics (MD) Solvent Electric Field Calculation

Solvent electric fields discussed and employed in TDM-field and frequency-field calibrations within the main text (Figure 1A-B, Figure 1E-F and Figure 3B-C) were derived from atomistic MD simulations performed with polarizable force fields (polarizable MD). MD simulations performed with fixed-charge (FC) force fields (FC MD) were also performed for comparison to prior electric field calibrations utilizing FC MD.<sup>2,6,7</sup> FC MD is utilized only in Figure S6, Figure S7A-B, Figure S8A, Figure S9A and S19B and related results are displayed in Table S3-S4.

### *Fixed-Charge Molecular Dynamics Simulation and Electric Field Calculation*

FC MD simulations were performed using the generalized Amber force field (GAFF)<sup>8,9</sup> in GROMACS 2018<sup>10</sup> as described previously.<sup>1,2,7</sup> Briefly, the solute (oTN) was geometry optimized in Gaussian 16<sup>11</sup> with a B3LYP density functional and a 6-311++G(d,p) basis set. The resulting structure was then parameterized using antechamber in AmberTools18 (AM1-BCC charges).<sup>12</sup> The solute was placed at the center of a cubic box with edges located at a distance of 2.0 nm from the solute; periodic boundary conditions were used. Organic solvent parameters were taken from virtualchemistry.org.<sup>13,14</sup> and water was modeled using TIP3P parameters.<sup>15</sup> The system was then energy minimized (steepest decent until < 1000 kJ mol<sup>-1</sup> nm<sup>-1</sup>) and equilibrated as an NPT ensemble for 100 ps at 300 K and 1 bar. An MD production run was carried out over 2 ns as an NPT ensemble (300 K, 1 bar) as performed previously.<sup>1,2</sup> A 2 fs step size was utilized along with a Bussi thermostat with  $\tau = 0.1$  ps, a Berendsen barostat, a PME method, a LINCS algorithm, and cutoff distances for coulomb summation and van der Waals interactions set at 1.0 nm.<sup>2</sup> Full coordinates and forces at the C and N atoms of the -C≡N bond of interest were extracted every 100 steps (0.2 ps). The same trajectory was rerun with fixed atomic coordinates and neutralized solvent partial charges to determine forces in the absence of intermolecular electrostatic interactions. The difference between the forces from both trajectories yields the electrostatic force due to the solvent. Once the electrostatic force is divided by the partial charge at the C and N, respectively, the solvent electric field at the C ( $\vec{F}_C$ ) and N ( $\vec{F}_N$ ) is known. The field projected on the -C≡N bond vector (C → N),  $|\vec{F}|$ , is then obtained by averaging the dot product of the unit -C≡N bond vector ( $\hat{u}_{CN}$ ) and  $\vec{F}_C$  with the dot product of  $\hat{u}_{CN}$  and  $\vec{F}_N$  such that  $|\vec{F}| = \frac{1}{2}(\hat{u}_{CN} \cdot \vec{F}_C + \hat{u}_{CN} \cdot \vec{F}_N)$ .  $|\vec{F}|$  is calculated at each snapshot and the average solvent  $|\vec{F}|$  for the entire MD trajectory is determined (Tables S3-S4).

### *Polarizable Molecular Dynamics Simulation and Electric Field Calculation*

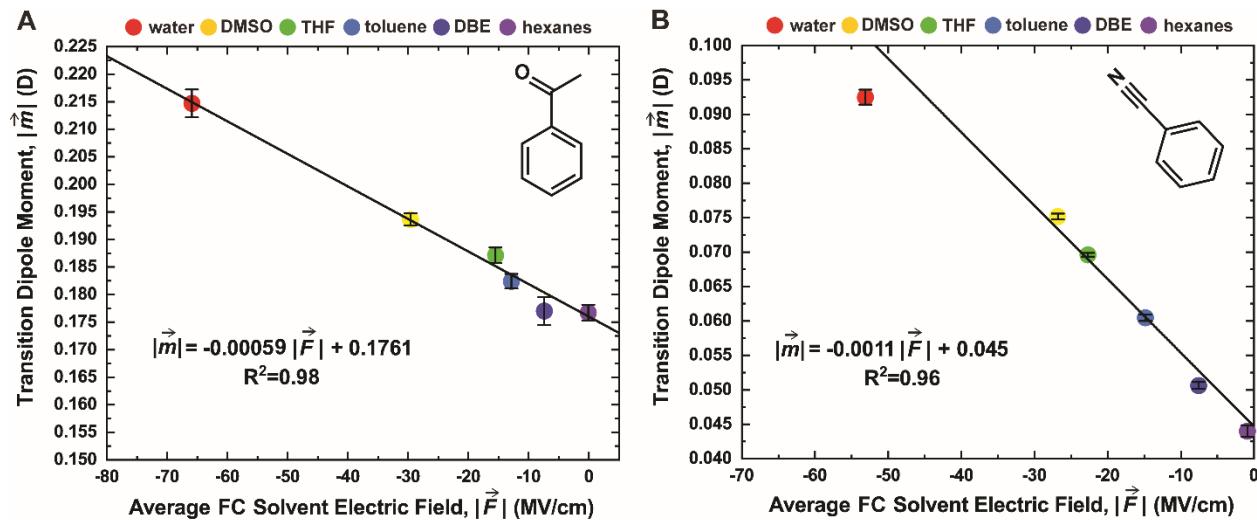
Polarizable molecular dynamics simulations were performed in Tinker 8.7<sup>16</sup> using the AMOEBA09 force field<sup>17</sup> as described previously.<sup>2,18</sup> Parameters for solvents and the ACP solute were taken as implemented in Tinker and from our previous work.<sup>2</sup> The oTN and BZN solutes were parameterized using Poltype with a multipole fitting to the electrostatic potentials from MP2/6-311g++(2d,2p) calculations in Gaussian 16<sup>11</sup>; solvent parameters were taken from previous work.<sup>2</sup> The solutes were placed into a solvent-filled cubic box with edges at a distance of 20 Å from the solute with periodic boundary conditions. The solute/solvent systems were energy minimized (steepest decent until < 0.5 kcal mol<sup>-1</sup> Å<sup>-1</sup>), and 50 ps-NVT (300 K) and 100 ps-NPT (300 K; 1 bar) equilibration steps were performed. 1 fs steps were utilized with a Beeman integrator using an induced dipole convergence threshold of 10<sup>-2</sup> D and a van der Waals cutoff of 9.0 Å. An electrostatics cutoff of 7.0 Å was used along with a particle mesh Ewald method, analytical van der Waals correction, a Velocity-Verlet integrator, an Andersen thermostat, and a Berendsen barostat with time constant of 1 ps during the NPT step. MD production runs were performed over 1 ns with an induced dipole moment threshold of 10<sup>-5</sup> D, a van

der Waals cutoff of 12.0 Å, a Monte-Carlo barostat with inverse friction of 1 ps, and “molecular” volume-scaling. Otherwise, similar conditions to the NPT equilibration step were used. Induced dipoles were extracted every 0.1 ps in the presence and absence of the solvent (at similar atomic coordinates). The difference yielded the induced dipoles due to the solvent, which were divided by the atomic polarizabilities to obtain the electric field vectors on C ( $\vec{F}_C$ ) and N ( $\vec{F}_N$ ) atoms. The average electric field projected along the -C≡N,  $|\vec{F}|$ , was determined as described for FC MD where  $|\vec{F}| = \frac{1}{2}(\hat{u}_{CN} \cdot \vec{F}_C + \hat{u}_{CN} \cdot \vec{F}_N)$ .

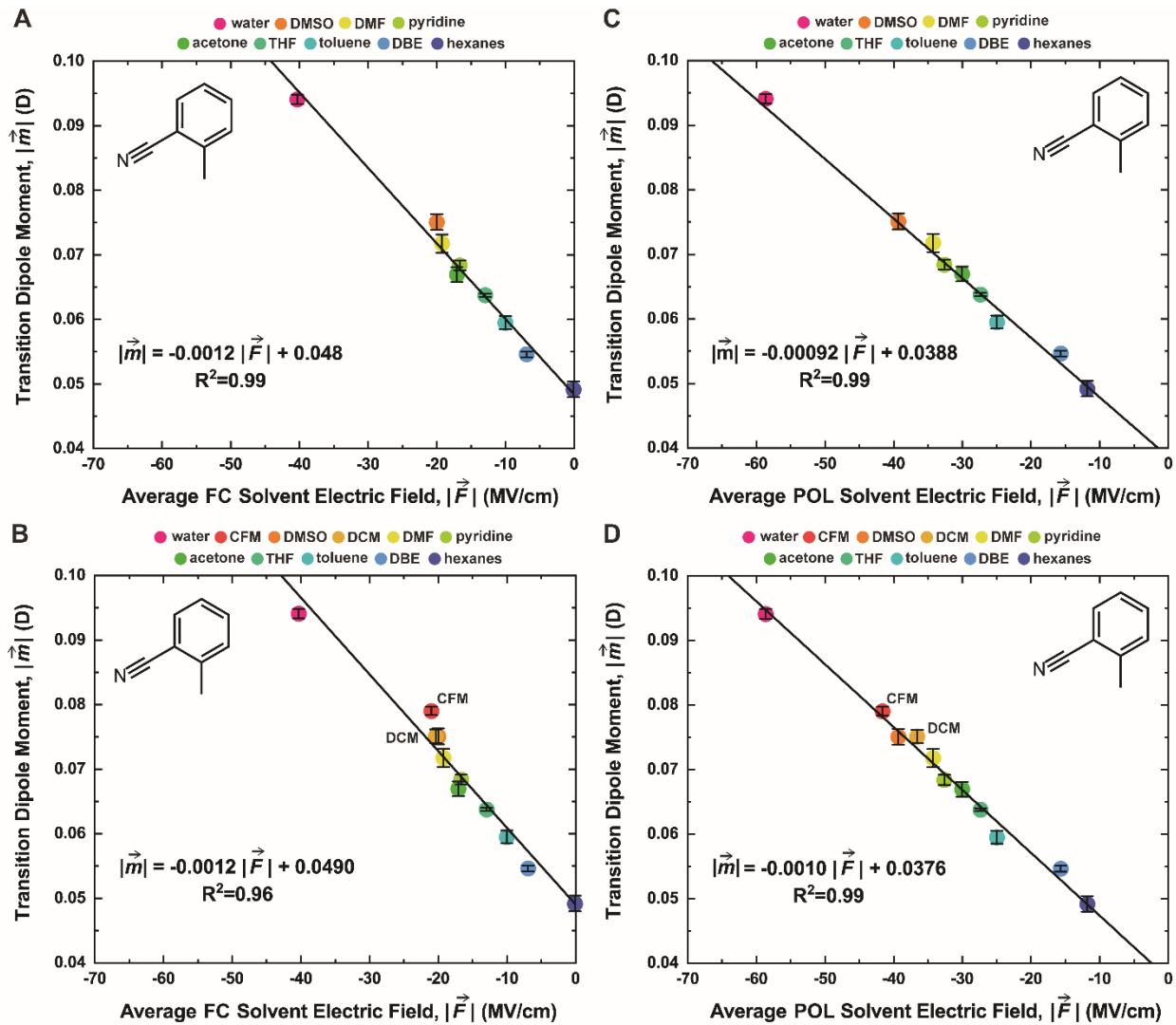
We note that while only fixed-charge and polarizable MD were performed here, it was observed that average calculated solvent electric fields on a similar rigid vibrational probe were very similar when grabbing frames from a fixed-charge MD simulation and optimizing with DFT, grabbing frames from a polarizable MD simulation and optimizing with DFT, or simply using the fields extracted from polarizable MD; as such, we did not perform additional quantum mechanical optimization to our polarizable MD runs.<sup>19</sup>

## S6 IR Transition Dipole Moment (TDM) as a Function of Solvent Field

In the TDM vs Field calibrations in the main text (Figures 1 and 3) and in this section (Figures S6 and S7), all  $|\vec{m}|$  which we calculate using main text Equation 3 technically contain an additional local field correction factor,  $f_{\text{opt}}$ , where  $|\vec{m}| = f_{\text{opt}} |\vec{m}_{\text{abs}}|$  and  $|\vec{m}_{\text{abs}}|$  is the absolute TDM (see SI Section S6 of Reference 20).<sup>20,21</sup> This  $f_{\text{opt}}$  is caused by the difference in the applied optical field from the light source relative to the optical field experienced by the probe due to environmental polarization in response to the light field.<sup>20,21</sup> The same is also true of the  $A_{zz}$  that we measure from the slope of TDM-field calibration and Stark spectroscopy (Section S8), where  $A_{zz} = f_{\text{opt}} A_{zz,\text{abs}}$ . This  $f_{\text{opt}}$  should not be confused with the local field correction factor,  $f$ , associated with the externally applied field in Stark experiments.  $f$  is expected to be larger than  $f_{\text{opt}}$  due to the much slower time scale (Hz) of the applied field relative to the optical field (THz), allowing for nuclear rearrangement in the solvent and resulting in greater environmental polarization. That said, because we calibrate  $|\vec{m}|$  with calculated *absolute* solvent electric fields,  $f_{\text{opt}}$  does not affect our ability to extract absolute fields. In the main text (Figure 1E-F) and here (Figures S6-S7), it is noted that the TDMs of the -C≡N in oTN and BZN are significantly more sensitive to the electric field than -C=O in ACP and therefore possess a higher transition polarizability ( $A_{zz}$ ). While the origin of the -C≡N's larger  $A_{zz}$  compared to the -C=O's is difficult to determine without further experiments and/or a more comprehensive theoretical model, it seems likely that the higher bond order in oTN and BZN (triple bond) than in ACP (double bond) gives rise to a larger polarizability and accompanying larger transition polarizability.

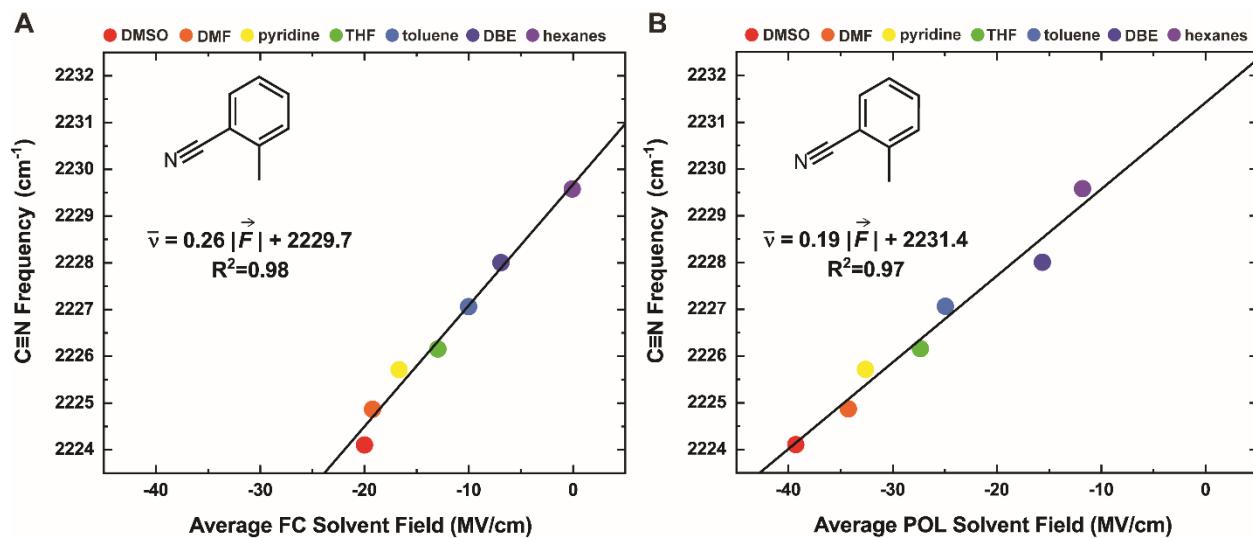


**Figure S6.** Plots of the TDM,  $|\vec{m}|$ , versus the average solvent electric field calculated for each of the seven solvents with fixed-charge force fields for (A) the -C=O stretch in ACP and (B) the -C≡N stretch in BZN. Note that a deviation from linearity in the TDM-field calibration above is possible due to hyperpolarizability contributions but is not observed over the field range from these solvents for oTN (Figure S7), BZN (Figure 1F, Figure S6B), or ACP (Figure 1E, Figure S6A).

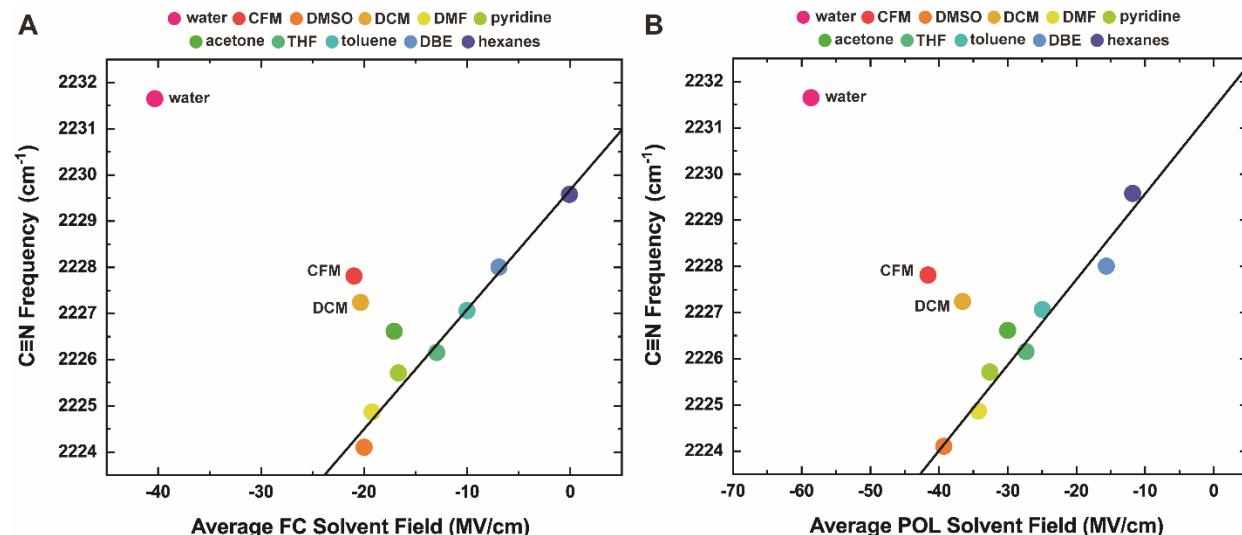


**Figure S7.** Plots of the oTN TDM,  $|\vec{m}|$ , versus the average solvent electric field calculated for each solvent with (A & B) fixed-charge and (C & D) polarizable force fields. TDM data in (B) and (D) are replicas of (A) and (C), respectively, with two additional solvents, DCM and CFM, included. These points were not included previously as polarizable and fixed-charge parametrization for halogenated solvents is currently difficult to model due to the  $\sigma$ -hole on the chlorine atom,<sup>22</sup> limiting our confidence in the average fields calculated for these points and motivating further investigation and benchmarking for halogenated solvents. We note that despite CFM's and DCM's deviations from the frequency-field linear calibration (Figure S9), CFM and DCM appear to follow the same linear trend as other solvents in TDM-field calibration. Note that a deviation from linearity in the TDM-field calibration above is possible due to hyperpolarizability contributions but is not observed over the field range from these solvents for oTN (Figure S7), BZN (Figure 1F, Figure S6B), or ACP (Figure 1E, Figure S6A).

S7 IR Peak Frequency as a Function of Solvent Electric Field



**Figure S8.** Aprotic solvent field-frequency calibration curves for oTN using (A) fixed-charge and (B) polarizable force fields. The slopes of these lines are the frequency Stark tuning rates.



**Figure S9.** Field-frequency calibration curves for oTN including both protic and non-protic solvents using (A) fixed-charge and (B) polarizable force fields. This highlights the blue-shifted deviations from the line drawn through the non-protic solvents for water, CFM and DCM (cf. Figure S7 where this deviation is not observed for the TDM-field calibration).

S8 IR Peak Frequencies, FWHM, TDMs, and Fields Calculated from MD

**Table S1.** Extracted IR peak parameters for oTN as a function of solvent<sup>a</sup>

Environment	Peak (cm <sup>-1</sup> )	FWHM (cm <sup>-1</sup> )	$\vec{m}$   (D)
water	2231.7 ± 0.04	9.5 ± 0.03	0.0941 ± 0.0007
CFM	2227.8 ± 0.03	7.9 ± 0.05	0.0790 ± 0.0007
DMSO	2224.1 ± 0.01	7.7 ± 0.01	0.0751 ± 0.0012
DCM	2227.2 ± 0.01	5.9 ± 0.01	0.0751 ± 0.0011
DMF	2224.9 ± 0.09	6.7 ± 0.1	0.0718 ± 0.0014
pyridine	2225.7 ± 0.01	7.5 ± 0.01	0.0684 ± 0.0008
acetone	2226.6 ± 0.06	5.4 ± 0.4	0.0670 ± 0.0011
THF	2226.2 ± 0.01	5.5 ± 0.5	0.0637 ± 0.0002
toluene	2227.1 ± 0.15	5.9 ± 1.1	0.0595 ± 0.0010
DBE	2228.0 ± 0.08	5.8 ± 0.3	0.0546 ± 0.0004
hexane	2229.6 ± 0.01	5.1 ± 0.01	0.0492 ± 0.0012

<sup>a</sup> For ACP and BZN frequencies and FWHM, refer to Fried *et al.*<sup>7</sup> and Bagchi *et al.*<sup>6</sup>

**Table S2.** Extracted IR | $\vec{m}$ | for ACP and BZN as a function of solvent

Environment	ACP   $\vec{m}$   (D)	BZN   $\vec{m}$   (D)
water	0.214 ± 0.002	0.0925 ± 0.0011
CFM	0.205 ± 0.001	0.0744 ± 0.0009
DMSO	0.194 ± 0.001	0.0752 ± 0.0004
THF	0.187 ± 0.001	0.0605 ± 0.0005
toluene	0.182 ± 0.001	0.0547 ± 0.0004
DBE	0.177 ± 0.003	0.0506 ± 0.0005
hexanes	0.177 ± 0.001	0.0440 ± 0.0004

**Table S3.** Average Solvent MD fields for oTN using polarizable and fixed-charge MD

Environment	POL $ \vec{F} $ (MV/cm)	FC $ \vec{F} $ (MV/cm)
water	-58.6	-40.3
CFM	-41.6	-21.0
DMSO	-39.3	-20.0
DCM	-36.6	-20.4
DMF	-34.3	-19.2
pyridine	-32.6	-16.7
acetone	-30.0	-17.1
THF	-27.4	-13.0
toluene	-25.0	-10.0
DBE	-15.7	-6.9
hexane	-11.8	-0.1

**Table S4.** Average Solvent MD Fields for ACP and BZN using polarizable and fixed-charge MD<sup>a</sup>

Environment	ACP POL $ \vec{F} $ (MV/cm)	ACP FC $ \vec{F} $ <sup>a</sup> (MV/cm)	BZN POL $ \vec{F} $ (MV/cm)	BZN FC $ \vec{F} $ <sup>b</sup> (MV/cm)
water	-75.7	-65.9	-58.4	-53.1
CFM	-59.0	-28.5	-43.0	-23.6
DMSO	-42.9	-29.6	-40.9	-26.8
THF	-23.9	-15.5	-29.0	-22.7
toluene	-22.4	-12.9	-26.0	-14.9
DBE	-13.9	-7.4	-17.0	-7.6
hexanes	-11.3	-0.1	-12.8	-0.9

<sup>a</sup> All FC MD average fields for ACP were obtained from Fried *et al.*<sup>7</sup> except for toluene; toluene was simulated and its average MD field was calculated in this work.

<sup>b</sup> All FC MD average fields for BZN were obtained from Fried *et al.*<sup>7</sup> and Bagchi *et al.*<sup>6</sup>

## S9 Vibrational Stark Spectroscopy of oTN and Comparison to Solvatochromism\*-based Frequency-Field and TDM-Field Calibrations

\*Solvatochromism definitionally refers to frequency shifts as a function of solvent; in this section we use the word “solvatochromism” more broadly to refer to the simultaneous tuning of frequencies and TDMs as a function of solvent.

Vibrational Stark spectroscopy (VSS), spectroscopy in a defined external electric field, allows for an independent, experimental determination of the electrooptic parameters  $|\Delta\vec{\mu}|$  and  $A_{zz}$ . The  $|\Delta\vec{\mu}|$  and  $A_{zz}$  determined from VSS complement the  $A_{zz}$  determined from a TDM-field calibration (Figures 1E-F and S6-S7) and the  $|\Delta\vec{\mu}|$  determined from an aprotic frequency-field calibration (Figures 1A-B and S8), where nitrile and carbonyl solvent average fields are obtained computationally. In vibrational Stark spectroscopy, a defined external electric field,  $F_{ext}$ , is applied to an isotropic, immobilized sample (often a frozen glass) and the change in absorption following field application, i.e. the Stark spectrum, is measured.  $|\Delta\vec{\mu}|$  and  $A_{zz}$  are then extracted from Stark spectral fitting using a classical sum-of-derivative analysis that has been described extensively in other work.<sup>1,7,23</sup> Specifically, the Stark spectrum is fit to a linear combination of the zeroth-, first-, and second-derivative of the absorption spectrum:

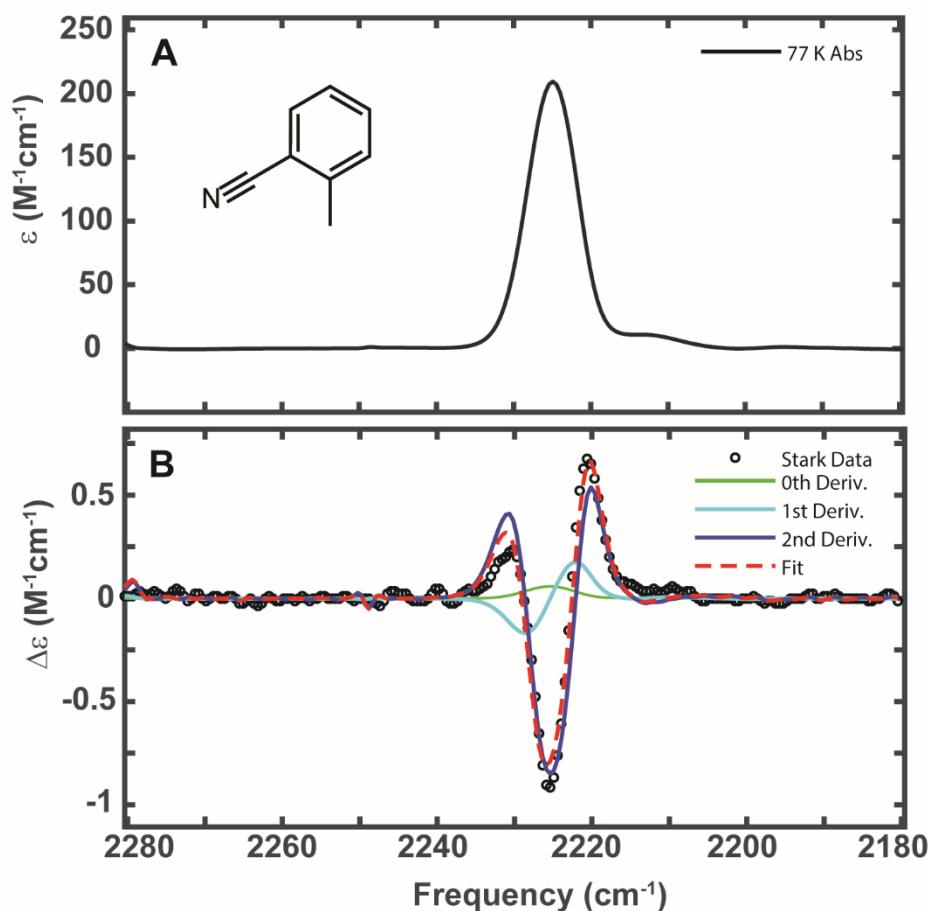
$$\Delta A(\bar{v}, F_{ext}) = (f F_{ext})^2 \left[ A_\chi A(\bar{v}) + \frac{B_\chi}{15hc} \bar{v} \frac{\partial}{\partial \bar{v}} \left( \frac{A(\bar{v})}{\bar{v}} \right) + \frac{C_\chi}{30h^2c^2} \bar{v} \frac{\partial^2}{\partial \bar{v}^2} \left( \frac{A(\bar{v})}{\bar{v}} \right) \right] + \dots \quad (\text{S1})$$

In Equation S1,  $F_{ext}$  is the magnitude of the externally applied electric field,  $f$  is the local field correction factor (discussed below), and all higher order field terms ( $\vartheta(F^4)$  and higher) are considered to contribute negligibly to  $\Delta A$ .  $A_\chi$ ,  $B_\chi$ , and  $C_\chi$  are the zeroth-, first-, and second-derivative contributions of  $A(\bar{v})$  to  $\Delta A(\bar{v}, F_{ext})$ , respectively. The Stark spectrum for oTN is displayed below and Stark characterization of ACP and BZN has been described previously.<sup>7,24</sup> Note that VSS spectra are obtained in frozen glasses to avoid molecular reorientation upon application of the external field. This contrasts with solvatochromism data which are obtained in fluid solution.

Unlike the  $|\Delta\vec{\mu}|$  and  $A_{zz}$  determined from solvatochromism-based frequency-field and TDM-field calibrations,  $|\Delta\vec{\mu}|$  and  $A_{zz}$  determined from VSS are both convolved with the local field correction factor,  $f$ , caused by the difference between the applied external field and the local field experienced by the vibrational probe (-C≡N or -C=O here).<sup>7,20,24</sup> Thus, VSS allows for the extraction of  $f|\Delta\vec{\mu}|$  and  $fA_{zz}$ , where the  $f$  associated with both parameters should be the same (noting that  $f$  is a scalar approximation to the true local field correction tensor). By comparing  $f|\Delta\vec{\mu}|$  to  $|\Delta\vec{\mu}|$  from solvatochromism-based frequency-field calibrations and  $fA_{zz}$  to  $A_{zz}$  from TDM-field calibrations,  $f$  can be evaluated. Note, since  $fA_{zz}$  is extracted from the zeroth-derivative contribution to the Stark fit which contributes least to the overall fit (Figure S10 and Table S5),  $fA_{zz}$  has a larger error than  $f|\Delta\vec{\mu}|$ . That said, the  $f$  for BZN and oTN determined from  $fA_{zz}$  are remarkably similar to the  $f$  determined from  $f|\Delta\vec{\mu}|$  (Table S6). While extraction of  $fA_{zz}$  via VSS is performed for oTN and BZN, it cannot be performed for ACP since ACP’s larger transition hyperpolarizability (B) is non-negligible (evidenced by negative zeroth-derivative in Stark fit) and impedes reliable extraction of  $fA_{zz}$  and comparison with  $A_{zz}$  from solvatochromism (Figure 1E).<sup>7,23</sup>

Stark spectra for oTN were acquired as described previously.<sup>1,2,7</sup> Briefly, ~20 μL of 100 mM oTN dissolved in 2-methyltetrahydrofuran (2-MeTHF) was injected into a sample cell composed of two offset CaF<sub>2</sub> optical windows (1 mm thick, 12.7 mm diameter) which were coated in 45 Å of Ni (deposited by Tom Carver of Stanford Nano Shared Facilities) and separated by two ~26 μm spacers. Upon loading, the sample was immediately plunged into a custom-built cryostat<sup>25</sup> filled with liquid nitrogen. Vibrational Stark spectra were recorded with electric fields of 0.5 – 1.25 MV/cm applied (applied

potential was 1 kV – 2.5 kV) using a Trek 10/10 high-voltage power amplifier on a Bruker Vertex 70 FTIR. Stark spectra had a  $1\text{ cm}^{-1}$  resolution and 64 scans were acquired for both field-on and field-off measurements at each applied field. The angle,  $\chi$ , between the experimentally applied field and the optical electrical field vector was set at  $90^\circ$ .



**Figure S10.** (A) Absorption and (B) vibrational Stark spectrum of oTN in 2-methyltetrahydrofuran (2-MeTHF) at 77 K. This spectrum was obtained in an external field of  $1.25\text{ MV/cm}$ , and  $\Delta\varepsilon$  was scaled to an applied field of  $1\text{ MV/cm}$  based on the field-squared dependence for an isotropic, immobilized sample (Eqn. S1). Fit parameters are given in Table S5.

**Table S5.** Fit parameters for oTN Stark spectra shown in Figure S10.

oTN Absorption Fit Parameters			oTN Stark Fit Parameters			
Peak ( $\text{cm}^{-1}$ )	FWHM ( $\text{cm}^{-1}$ )	Peak Area ( $\text{M}^{-1}\text{cm}^{-2}$ )	$f^2A_\chi^a$ $\left(\frac{1}{\text{MV}^2/\text{cm}^2}\right)$	$f^2B_\chi^a$ $\left(\frac{\text{cm}^{-1}}{\text{MV}^2/\text{cm}^2}\right)$	$f^2C_\chi^a$ $\left(\frac{\text{cm}^{-2}}{\text{MV}^2/\text{cm}^2}\right)$	$f \Delta\vec{\mu} ^b$ $\left(\frac{\text{cm}^{-1}}{\text{MV}/\text{cm}}\right)$
2225.1	7.4	1855	0.000306	-0.0608	1.39	$0.68 \pm 0.01$

<sup>a</sup> Where  $f^2A_\chi$ ,  $f^2B_\chi$ , and  $f^2C_\chi$  are the contributions of the zeroth-, first-, and second-derivatives of the absorption to the Stark fit (Equation S1).

<sup>b</sup> Since the angle,  $\zeta$ , between the TDM and  $\Delta\vec{\mu}$  is estimated to be  $\sim 0^\circ$ ,<sup>26</sup>  $f|\Delta\vec{\mu}|$  can be derived from  $f^2C_\chi$ , where  $f|\Delta\vec{\mu}| = \sqrt{\frac{f^2C_\chi}{3}}$ .

**Table S6.** Comparison of the transition polarizability ( $A_{zz}$ ) and Stark tuning rate ( $|\Delta\vec{\mu}|$ ) determined from vibrational Stark spectroscopy and from solvatochromic TDM-field and frequency-field calibration for oTN and BZN.

Solute	$f^2A_\chi^a$ $(\frac{1}{MV^2/cm^2})$	$fA_{zz}$ $(\frac{D}{MV/cm})$	POL $A_{zz}$ $(\frac{D}{MV/cm})$	FC $A_{zz}$ $(\frac{D}{MV/cm})$	$f^b$ (POL, $A_{zz}$ )	$f^c$ (FC, $A_{zz}$ )
oTN	0.000306	0.00327	0.000923	0.00117	3.5	2.8
BZN <sup>d</sup>	0.000260	0.00335	0.00108	0.00107	3.1	3.1
Solute	$f^2C_\chi^e$ $(\frac{cm^{-2}}{MV^2/cm^2})$	$f \Delta\vec{\mu} $ $(\frac{cm^{-1}}{MV/cm})$	POL $ \Delta\vec{\mu} ^f$ $(\frac{cm^{-1}}{MV/cm})$	FC $ \Delta\vec{\mu} ^f$ $(\frac{cm^{-1}}{MV/cm})$	$f^g$ (POL, $ \Delta\vec{\mu} $ )	$f^h$ (FC, $ \Delta\vec{\mu} $ )
oTN	1.39	0.68	0.19	0.26	3.6	2.6
BZN <sup>d</sup>	1.08	0.61	0.19	0.22	3.2	2.8

<sup>a</sup> Contribution of the zeroth derivative component of the 77K absorption spectra used to fit Stark spectra.  $A_{zz}$  (also called  $A_{||}$ )<sup>24</sup> is derived from  $f^2A_\chi$  of the Stark fit, where we approximate all components of the transition polarizability tensor  $\underline{A}$  which are not along the nitrile bond axis (perpendicular and off-diagonal components) to be zero, such that  $fA_{zz} = |\vec{m}| \sqrt{5f^2A_\chi}$ ; we approximate the transition hyperpolarizability ( $\underline{B}$ ) as negligible – see References 23 and 24 for detailed analysis.<sup>23,24</sup>

<sup>b</sup> The local field correction factor determined by comparing the Stark  $fA_{zz}$  with TDM-field  $A_{zz}$ , where solvent fields are calculated with *polarizable* MD.

<sup>c</sup> The local field correction factor determined by comparing the Stark  $fA_{zz}$  with TDM-field  $A_{zz}$ , where solvent fields are calculated with *fixed-charge* MD.

<sup>d</sup> Stark-derived values for BZN were obtained from Andrews *et al.*<sup>24</sup>

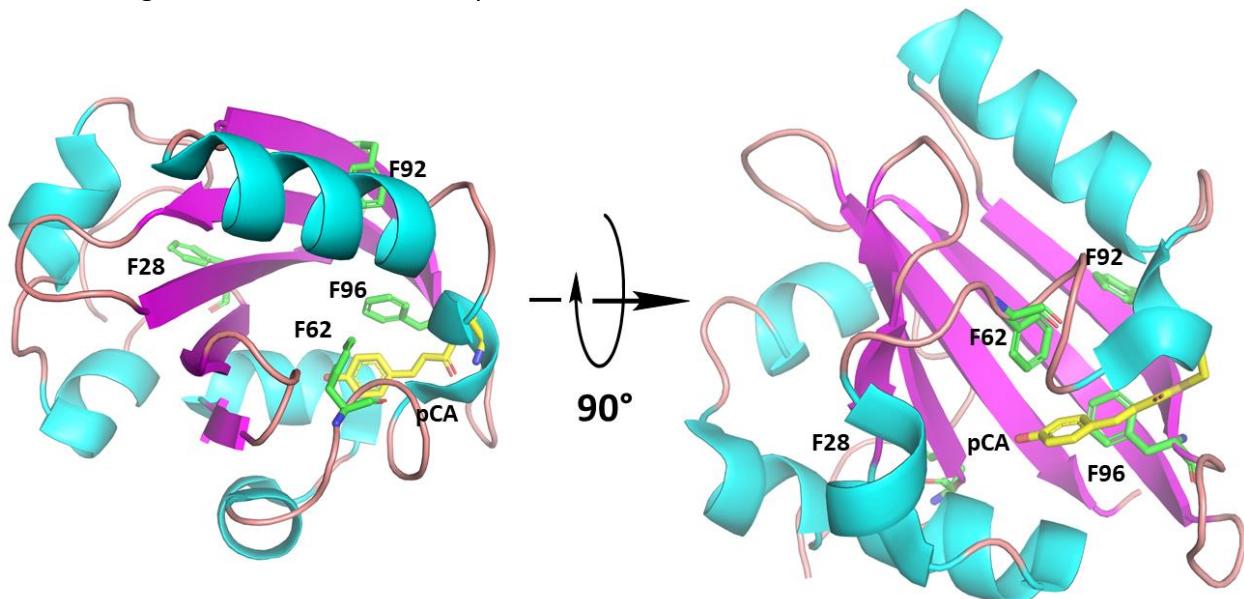
<sup>e</sup> Contribution of the second derivative component of the 77K absorption spectra used to fit Stark spectra. Note  $f|\Delta\vec{\mu}|$  can be derived from  $f^2C_\chi$ , where  $f|\Delta\vec{\mu}| = \sqrt{\frac{f^2C_\chi}{3}}$ .

<sup>f</sup> Aprotic frequency-field calibrations were used to determine  $|\Delta\vec{\mu}|$  (Figure S8).

<sup>g</sup> The local field correction factor determined by comparing the Stark  $f|\Delta\vec{\mu}|$  with the solvatochromism-based frequency-field  $|\Delta\vec{\mu}|$ , where solvent fields are calculated with *polarizable* MD.

<sup>h</sup> The local field correction factor determined by comparing the Stark  $f|\Delta\vec{\mu}|$  with the solvatochromism-based frequency-field  $|\Delta\vec{\mu}|$ , where solvent fields are calculated with *fixed-charge* MD.

S10 Targeted Sites for oCNF Incorporation in PYP



**Figure S11.** PYP phenylalanine targets for replacement with *o*-cyanophenylalanine (WT PYP, Protein Data Bank ID: 1NWZ) depicted from two perspectives. Phenylalanines targeted for replacement are colored green and the PYP chromophore, *p*-coumaric acid (pCA), is colored yellow.

## S11 Crystallization of PYP, X-ray Data Collection, and Refinement of PYP Structures

### *Protein Crystallization*

All PYPs were crystallized using the hanging-drop method as previously established.<sup>27</sup> Freshly purified PYP was buffer exchanged into an aqueous buffer at pH 6.0, 20 mM potassium phosphate (Sigma-Aldrich) and concentrated to 20 mg/mL. Hanging drops were set manually using 24-well VDX plates with sealant (Hampton Research) and 22 x 22 mm plastic coverslips (Ted Pella, Inc.). To each coverslip, 1 µL of protein was mixed with 1 µL of mother liquor and closed over wells containing 500 µL worth of mother liquor. Trays were kept at room temperature and covered in aluminum foil to minimize exposure to light.

PYPs were set in well solutions containing 1M NaCl and in ammonium sulfate (Sigma-Aldrich) ranging from 1.8 – 2.6 M in 0.1 M increments. In proteins that spontaneously crystallized, crystals were often observed within the first several days at higher ammonium sulfate concentrations. If no crystal formation was noted within one to two weeks, wells were seeded with WT PYP crystals which had been crushed with two 2-mm ceramic beads (Hampton Research). Drops were seeded using a clean cat whisker. Best results were found when seeds were diluted 10 – 1000-fold. Forced nucleation caused crystal formation within two days.

All crystals of sufficient quality were placed on 0.1 – 0.4 mm Mounted CryoLoops (Hampton Research), briefly (1-2 seconds) incubated in polyfluoroether cryo-oil (Hampton Research), rapidly frozen in liquid nitrogen, and placed into a Stanford Synchrotron Radiation Lightsource (SSRL) cassette for X-ray diffraction data collection.

### *X-ray Data Collection and Refinement Characterization for PYP Structures*

X-ray diffraction was performed at SSRL (Menlo Park, CA) Beam Lines 12-1 and 12-2 (see Table S7) at 100K. Information on data collection strategy can be found in Section 4.9 of Romei *et al.*<sup>27</sup> Data processing occurred using the X-ray Detector Software (XDS)<sup>28,29</sup> with the autoxds script.<sup>30</sup> Molecular replacement was performed in Phenix<sup>31</sup> with phenix.phaser using the WT PYP (PDB Entry: 1NWZ) structure as the search model for all PYP variants. The oCNF restraint file (.cif) was built by geometry optimizing oCNF in Gaussian 16<sup>11</sup> and inputting optimized coordinates into eLBOW in PHENIX. Numerous rounds of model building and refinement were carried out in Coot<sup>32</sup> and then with phenix.refine. In some cases, PDB-REDO was used to further refine structures.<sup>33</sup> Refinement continued until the R-free score could no longer be lowered. Resulting PDB codes and data collection/refinement statistics can be found in Table S7.

Table S7. X-ray data collection and refinement statistics for PYP variants.

PDB Entry	F28oCNF <b>7SPX</b>	F62oCNF <b>7SPW</b>	F92oCNF <b>7SPV</b>	F96oCNF <b>7SJ</b>
<b>Data collection statistics</b>				
Beamline	BL 12-2	BL 12-1	BL 12-1	BL 12-1
Wavelength (Å)	0.7749	0.9794	0.7126	0.7126
Detector distance (mm)	188	200	150	150
Resolution range (Å)	33.01 – 0.97 (1.01 – 0.97)	35.03 – 1.05 (1.09 – 1.05)	33.90 – 1.18 (1.22 – 1.18)	33.29 – 0.95 (0.99 – 0.95)
Space group	P 6 <sub>3</sub> (No. 173)	P 6 <sub>5</sub> (No. 170)	P 6 <sub>5</sub> (No. 170)	P 6 <sub>3</sub> (No. 173)
Unit cell dimensions a, b, c (Å) α, β, γ (°)	66.03, 66.03, 40.64 90.0, 90.0, 120.0	40.45, 40.45, 117.27 90.0, 90.0, 120.0	40.88, 40.88, 117.49 90.0, 90.0, 120.0	66.16, 66.16, 40.90 90.0, 90.0, 120.0
Total observations	782867 (68681)	1772253 (62954)	1528859 (141145)	2506853 (147373)
Unique reflections	59721 (5864)	49689 (4375)	36391 (3621)	61539 (4365)
Redundancy	13.1 (11.6)	35.7 (14.4)	42.0 (38.9)	40.7 (33.8)
Completeness (%)	99.9 (99.2)	98.5 (86.2)	99.9 (99.7)	98.4 (95.1)
Mean I/σI	21.6 (1.8)	19.3 (2.8)	27.3 (1.6)	22.7 (1.5)
Wilson B-factor (Å <sup>2</sup> )	7.68	10.8	23.3	8.36
R <sub>meas</sub>	0.070 (1.38)	0.139 (0.84)	0.075 (2.12)	0.090 (2.32)
CC <sub>1/2</sub>	1 (0.697)	0.998 (0.886)	1 (0.809)	1 (0.650)

	F28oCNF	F62oCNF	F92oCNF	F96oCNF
<b>Refinement statistics</b>				
Reflections used in refinement	59672	49686	36369	61729
Reflections used for R <sub>free</sub>	3046	2485	1809	3084
R <sub>work</sub>	0.1444	0.1455	0.1723	0.1415
R <sub>free</sub>	0.1658	0.1575	0.1926	0.1602
Number of non-H Atoms:				
Protein	1331	1201	1162	1337
Ligands	1185	1049	1027	1187
Solvent	37	24	24	24
Solvent	109	128	111	126
Protein residues	123	124	123	124
RMSD bond lengths (Å)	0.012	0.009	0.014	0.013
RMSD bond angles (°)	1.31	1.41	1.63	1.57
Ramachandran favored (%)	96.55	97.50	97.48	97.50
Ramachandran allowed (%)	3.45	2.50	2.52	2.50
Ramachandran outliers (%)	0.00	0.00	0.00	0.00
Rotamer outliers (%)	3.15	0.88	2.70	1.57
Clashscore	3.78	2.86	7.77	3.81
Average B-factor (Å <sup>2</sup> ):				
Protein	11.91	14.13	18.78	12.85
Ligand	11.13	13.26	17.84	12.29
Solvent	7.02	8.98	14.19	6.64
	22.08	22.27	28.44	19.27

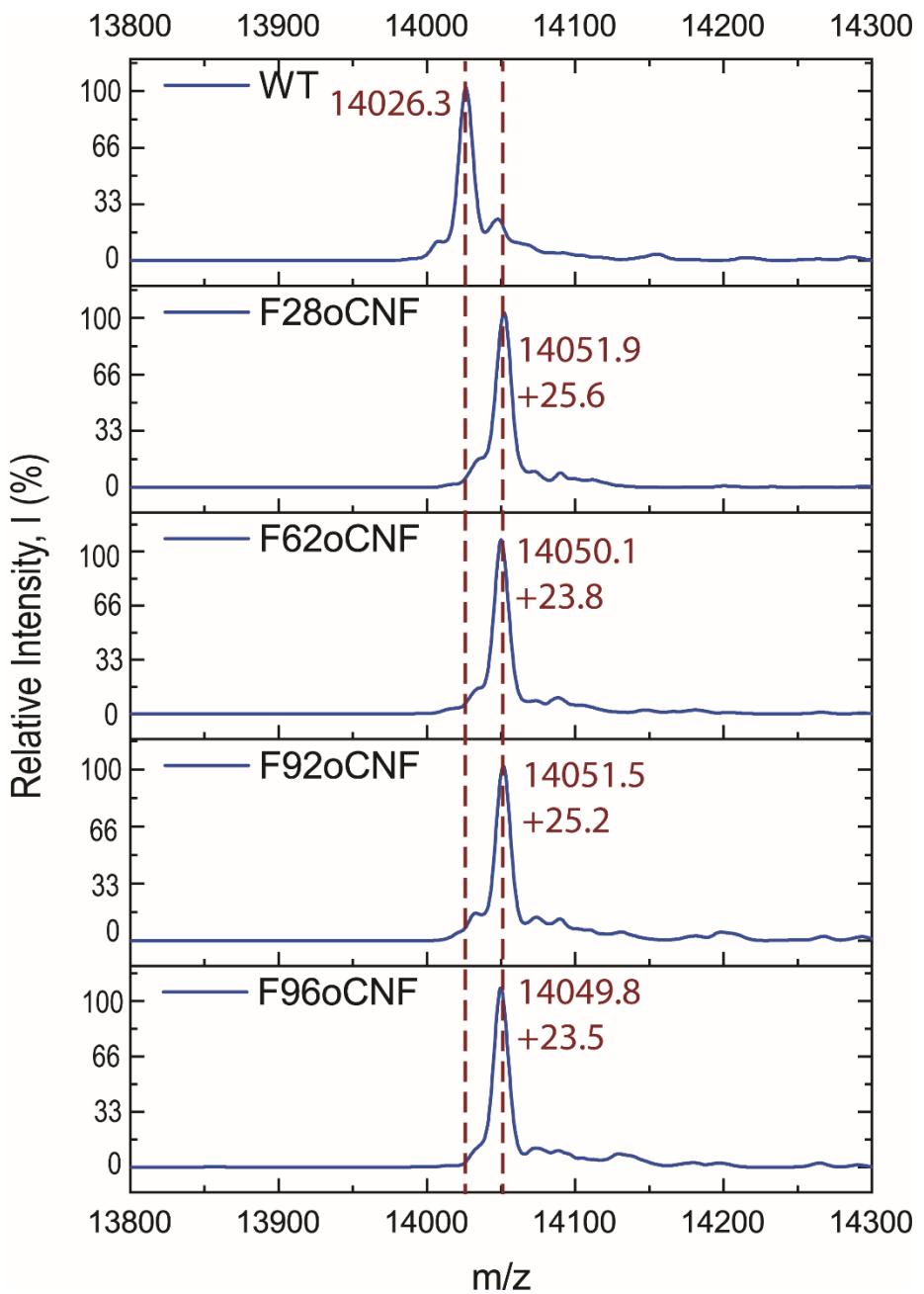
## S12 PYP MS Characterization

Protein liquid chromatography-electrospray ionization mass spectrometry (LC-ESI/MS) was performed on PYP to confirm replacement of Phe with oCNF. A Waters 2795 Acquity UPLC with a ZQ single quadrupole detector at the Stanford University Mass Spectrometry (SUMS) facility was utilized with a C8 column attached. Protein was eluted from the UPLC column at ~5.1 min by applying the following gradient:

**Table S8.** LC Gradient utilized in PYP LC-ESI/MS acquisition.

Time (min)	LC Buffer A (%)	LC Buffer B (%)
0.00	98.0	2.0
2.00	98.0	2.0
6.00	5.0	95.0
8.00	5.0	95.0
8.50	98.0	2.0
9.50	5.0	95.0
10.50	5.0	95.0

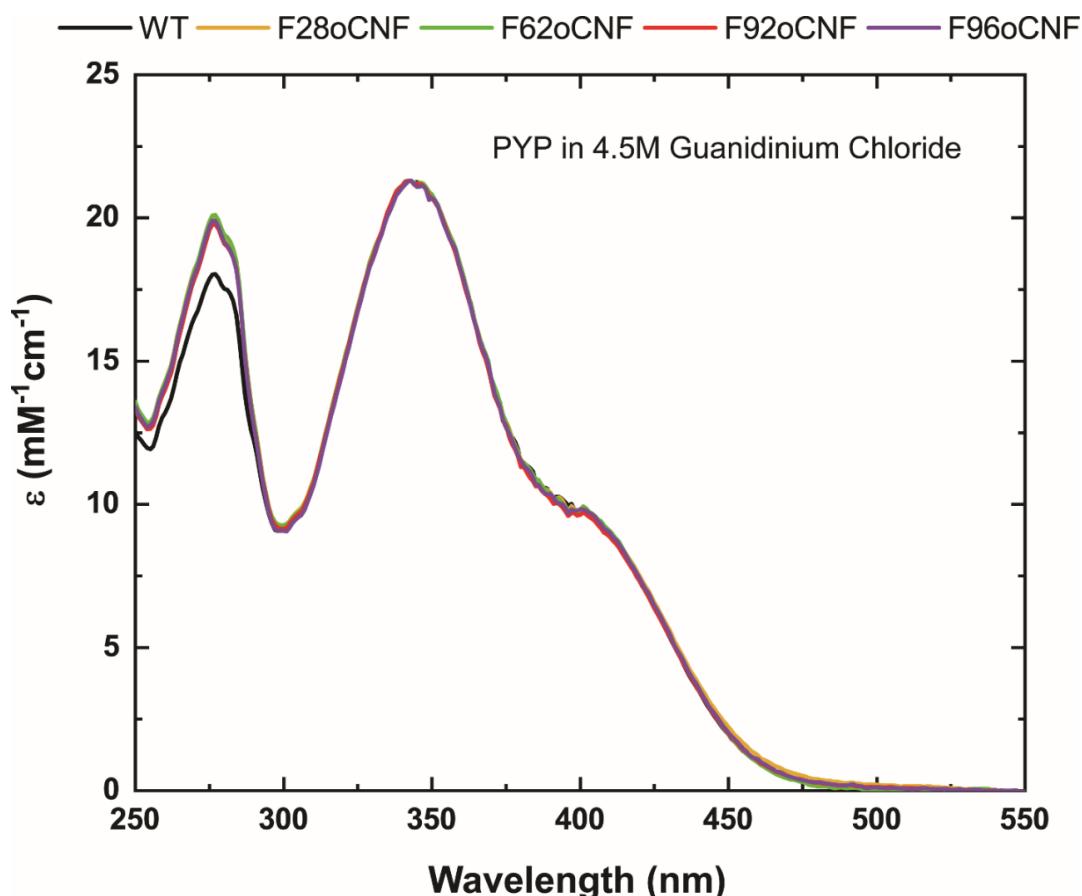
The LC Buffer A was 0.1% formic acid in water (Optima, LS118-4) and LC Buffer B was 0.1% formic acid in acetonitrile (Optima, LS120-4). A flowrate of 0.30 mL/min was utilized and the gradient was run at 40.0 °C.



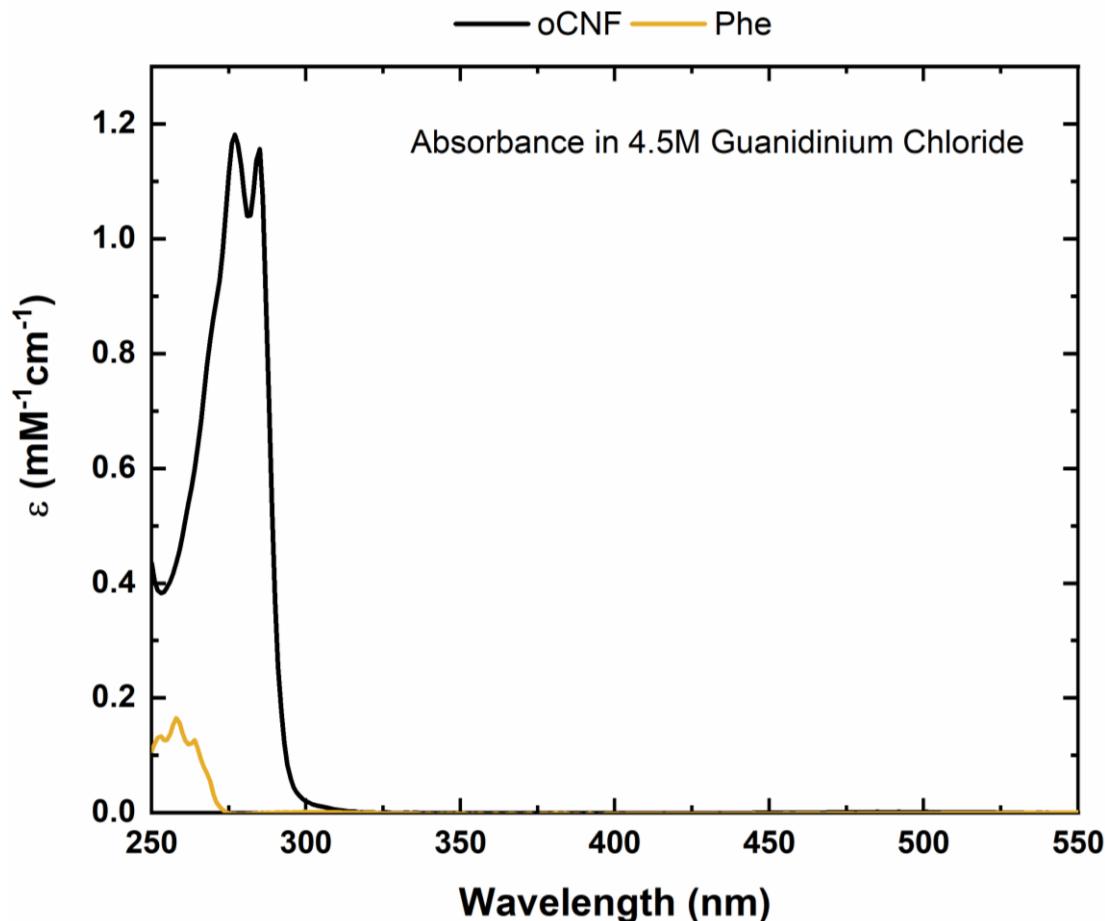
**Figure S12.** Mass spectra of PYP samples where Phe at a given site has been replaced with oCNF ( $\text{F} \rightarrow \text{oCNF}$ ). Dashed red lines indicate the WT Peak (14026 Da) and the PYP variant peak (14051 Da), with exact PYP variant masses and change in mass relative to WT indicated. All MS spectra indicate an  $\sim 25$  Da shift, matching phenylalanine replacement with oCNF and display high ( $\sim 100\%$ ) fidelity of oCNF incorporation. Note that high fidelity was observed for the incorporation of the PYP chromophore, pCA, in both WT and the oCNF PYP variants as no peak intensity is observed at 13880 Da (WT – pCA) or at 13905 Da (oCNF PYP – pCA).

### S13 PYP Extinction Coefficient Determination

Extinction coefficients for the pCA visible chromophore of PYP variants were necessary to determine PYP concentrations and, subsequently, nitrile concentrations for quantitative IR analysis. A guanidinium denaturation method was used to determine extinction coefficients in the UV-vis region for PYP variants in this study, similar to base denaturation used to determine extinction coefficients for various GFP constructs.<sup>20,34,35</sup> PYP fully denatures at guanidinium concentrations over 4 M<sup>36</sup> and absorption of the pCA chromophore in denatured PYP is unaffected by changes in the protein sequence like replacement of Phe with oCNF. This can be observed in the near identical spectral band shape for PYP variants (Figure S13), particularly in the 300 – 500 nm region where the denatured chromophore would be expected to absorb. The primary change between WT and PYP variants is a small increase in absorption at 280 nm caused by oCNF's more significant UV absorption (Figure S14) compared to the phenylalanine that oCNF replaces. Since WT PYP is known to have an extinction coefficient of 45.5 mM<sup>-1</sup>cm<sup>-1</sup>,<sup>37,38</sup> its concentration could be determined prior to denaturation. An  $\epsilon_{343\text{nm}} = 21.3 \text{ mM}^{-1}\text{cm}^{-1}$  was determined for denatured WT PYP and PYP variants. Absorption for each PYP variant was measured before and after denaturation and, utilizing the  $\epsilon_{343\text{nm}}$  of denatured PYPs, concentration normalized electronic (Figure S16) and vibrational spectra (Figures 3 and S17) were obtained for each PYP variant.



**Figure S13.** PYP absorption in the visible region following protein denaturation with 4.5 M guanidinium chloride.



**Figure S14.** Absorption spectra of oCNF and Phe in 4.5 M guanidinium chloride demonstrating oCNF's higher extinction coefficient at 280 nm relative to Phe.

Little variation in extinction coefficient spectra of PYP variants is observed ( $43.8\text{--}47.7\text{ mM}^{-1}\text{cm}^{-1}$  at the pCA absorption maxima, Table S9). This matches values from theoretical models<sup>20,39</sup> (Table S10) that predict little change in PYP extinction coefficients and meets naïve expectations given the wavelength of the pCA absorption maximum (446nm in WT) changes by less than 3nm ( $150\text{ cm}^{-1}$ ) in PYP variants.

**Table S9.** Extinction Coefficients ( $\epsilon$ ) Determined via Guanidinium Chloride (GuCl) Denaturation

Sample	Abs (pre-GuCl)	Abs (post-GuCl)	Conc. ( $\mu\text{M}$ )	$\epsilon_{\max}$ ( $\text{mM}^{-1}\text{cm}^{-1}$ )	$\epsilon_{\max}$ Avg. ( $\text{mM}^{-1}\text{cm}^{-1}$ )
WT	0.6060	0.2839	13.3	45.5 <sup>a</sup>	
F28oCNF	0.5930	0.2882	13.5	43.8	$44.1 \pm 0.2$
F62oCNF	0.6322	0.2965	13.9	45.5	$45.0 \pm 0.4$
F92oCNF	0.6771	0.3154	14.8	45.8	$44.6 \pm 1.0$
F96oCNF	0.5349	0.2420	11.4	47.1	$47.7 \pm 0.7$

<sup>a</sup> The extinction coefficient for WT PYP is taken from the literature.<sup>37,38</sup>

**Table S10.** Comparison between Observed and Predicted Extinction Coefficients

Sample	$\lambda_{\max}$ (nm)	$\epsilon_{\max}$ Observed (mM <sup>-1</sup> cm <sup>-1</sup> )	$\epsilon_{\max}$ Theoretical (mM <sup>-1</sup> cm <sup>-1</sup> )
WT	446	45.5 <sup>a</sup>	— <sup>b</sup>
F28oCNF	445	44.1	45.0
F62oCNF	449	45.0	46.8
F92oCNF	445	44.6	45.4
F96oCNF	449	47.7	47.4

<sup>a</sup> The extinction coefficient for WT PYP is taken from the literature.<sup>37,38</sup>

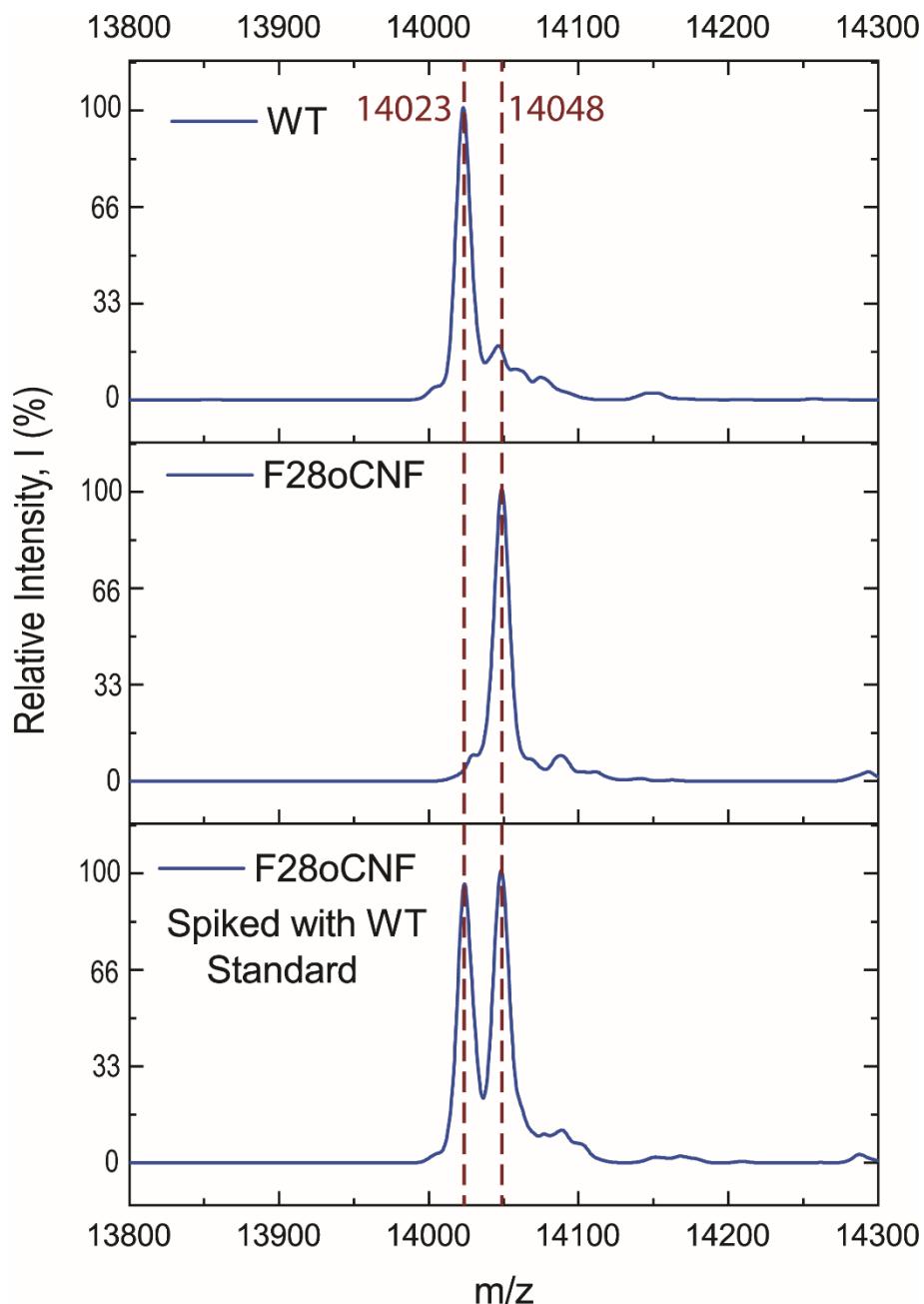
<sup>b</sup> Theoretical extinction coefficients are based on a model from the literature which predicts extinction coefficient changes starting from WT PYP's extinction coefficient.

#### *Alternate Methods of Determining Protein Extinction Coefficients*

While we chose to use UV-vis absorption following guanidinium denaturation to measure PYP concentration, multiple methods exist to quantify protein concentration. Often the extinction coefficient at 280 nm (or 205 nm)<sup>40</sup> can be a reliable means of protein concentration determination; in fact, all PYP variants had similar extinction coefficients at 280 nm (Figure S16). Additionally, protein concentration could have been determined by multiple other well-known methods, such as the Kjeldahl method,<sup>41</sup> amino acid analysis,<sup>42</sup> the Bradford assay,<sup>43–45</sup> commercially available fluorometric assays,<sup>46–48</sup> and mass spectrometry.<sup>49,50</sup>

We applied mass spectrometry as an alternative method of determining PYP concentrations. PYP chromophore extinction coefficients in the visible region can be estimated via mass spectrometry of samples with a known amount of WT PYP added (via UV-vis before and after spike addition). MS peak intensity at the WT peak (14023 Da) in the sample was compared to MS peak intensity of the PYP variant peak (Figure S15) to get a ratio of relative protein concentration. This was done three times for each PYP variant. Concentrations determined via mass spectrometry (Table S11) were slightly different from and had higher error than those observed for guanidinium denaturation (Table S9). This greater error could perhaps have been decreased with more extensive calibration of the WT standard, HPLC column preparation through more extensive column cleaning or passivation *prior* to loading any samples, and/or additional cleaning of the HPLC column *between* the acquisition of each spectrum. That said, mass spectrometry stands as an alternative, accessible method for protein quantification that does not require an endogenous chromophore as in PYP. Additionally, while we ultimately use concentration and integrated peak area to calculate TDMs, TDMs can also be directly determined through comparison of the Polarization-Selective Pump-Probe amplitude ( $\chi^{(3)}$ ) with the linear FTIR amplitude ( $\chi^{(1)}$ ).<sup>51</sup> We also note that while concentrations are needed to determine changes in absolute intensities, much IR spectroscopic work is performed following changes in conditions, e.g., exposure to light,<sup>52</sup> after which the difference spectrum in reference to t = 0 is observed. In such an experiment, the concentration is fixed and changes in intensity can be directly observed along with shifts of frequency. In this sense,

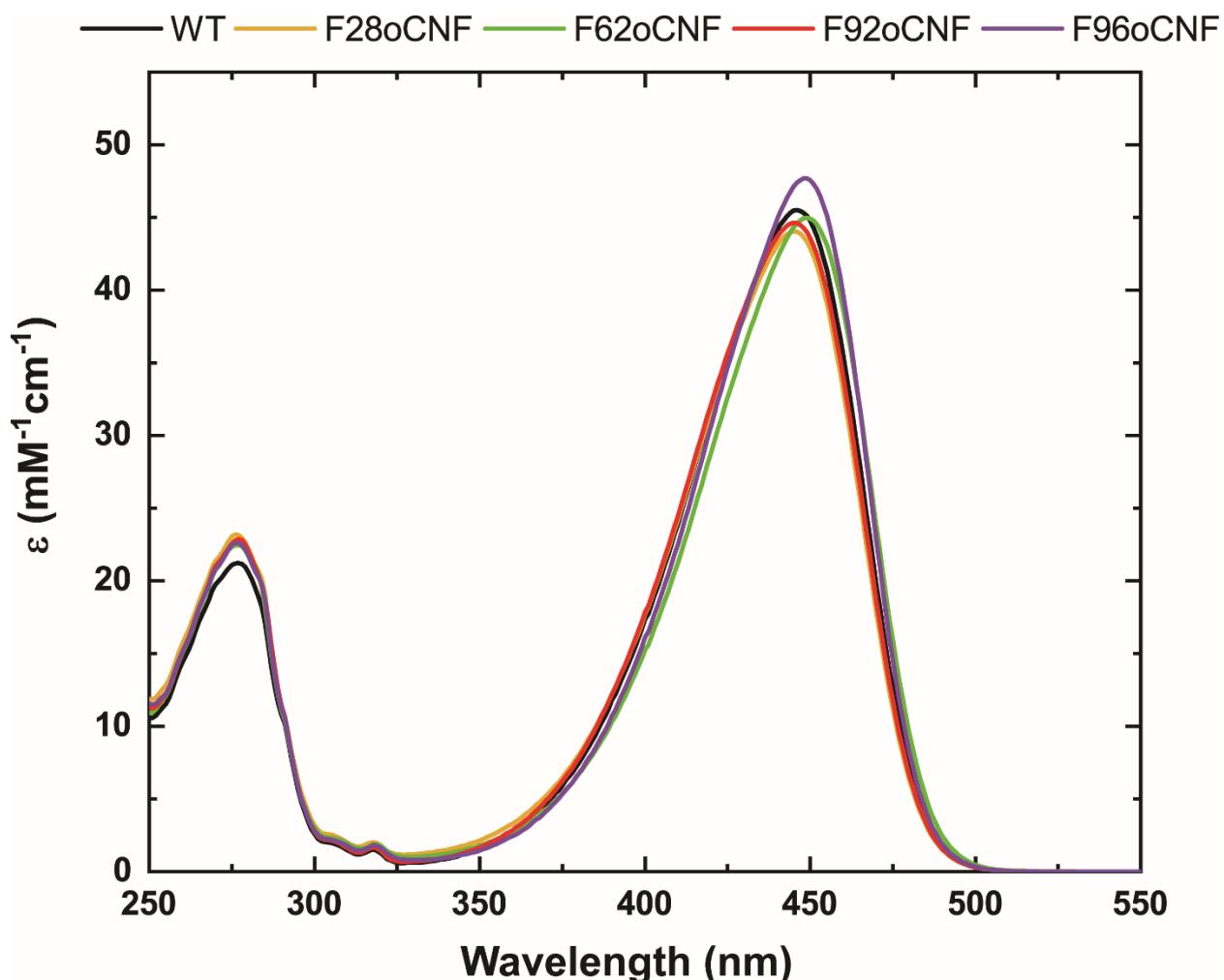
using the intensity as a direct observable can easily illuminate situations where similar frequencies (as seen in the case of F28oCNF and F96oCNF in Table 1) might lead one to believe the probe environment remains constant.



**Figure S15.** Mass spectrometry of samples with a known amount of WT PYP added (via UV-vis before and after spike addition) as an alternative means of determining PYP extinction coefficients in the UV-vis region. MS peak intensity at the WT peak (14023 Da) in the sample was compared to MS peak intensity of the PYP variant peak (F28oCNF in the example given above). Mass spectra were acquired with the same LC column, gradient, and instrument as described in Section S12.

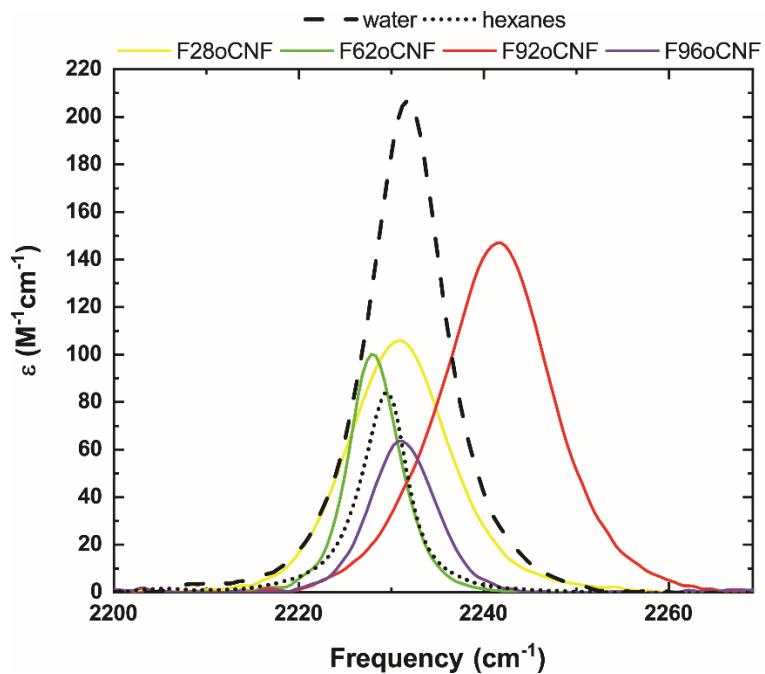
**Table S11.** MS Based Extinction Coefficient ( $\varepsilon$ ) Determination

Sample	WT Spiked ( $\mu\text{M}$ )	MS Sample I/WT I	Conc. Sample ( $\mu\text{M}$ )	Abs (Pre-spike)	$\varepsilon$ ( $\text{mM}^{-1}\text{cm}^{-1}$ )
F28oCNF	9.45	$1.11 \pm 0.06$	$10.5 \pm 0.6$	0.478	$45 \pm 2$
F62oCNF	9.85	$0.89 \pm 0.03$	$8.7 \pm 0.3$	0.391	$45 \pm 1$
F92oCNF	9.71	$0.87 \pm 0.04$	$8.4 \pm 0.4$	0.401	$48 \pm 2$
F96oCNF	9.65	$0.89 \pm 0.05$	$8.6 \pm 0.6$	0.435	$51 \pm 3$



**Figure S16.** Concentration normalized UV-vis spectra of WT and nitrile-incorporated PYP variants, where concentration was determined with guanidinium denaturation (Figure S13, Table S9). Nitrile incorporation did not significantly alter the PYP chromophore band's peak position or the extinction coefficient for any of the variants. Note while these spectra are similar, they do not necessarily indicate that the PYP photocycle kinetics or mechanism is unaffected.

S15 Protein IR Spectra Compared with oTN spectra in Water and Hexanes



**Figure S17.** IR absorption spectra of PYP variants compared with spectra for oTN dissolved in hexanes and water. Although the band in water has a peak intensity about 1.5 times higher than F92oCNF's band, they possess roughly the same integrated band area.

All protein nitrile IR spectra appear symmetric and single-peak fits were sufficient for fitting the bandshape. While not encountered for these protein variants, application of our oCNF TDM vs field calibration could become more difficult in scenarios where multiple populations (i.e. sample heterogeneity) are evidenced in the IR spectra. Such heterogeneous protein nitrile bandshapes can be evaluated by use of protein MD and IR band deconvolution.

## S16 DFT Calculations of Frequency Dependence on H-bonding Angle

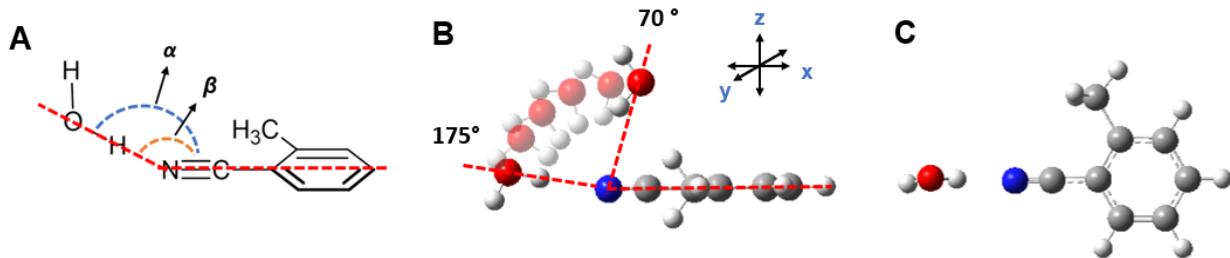
We performed a simple DFT-based estimation of the dependence of the H-bond frequency on the heavy atom H-bond angle ( $O\cdots N\equiv C$ ,  $\alpha$  in Figure 4 and Figure S18A) to allow for comparison with experimental results. Note that even with these high-resolution X-ray structures, proton positions are not resolved (Figure 2C-F). In these calculations, a water was placed such that the water O was 3 Å from the nitrile N with H-bond angles restrained to angles from  $70^\circ - 175^\circ$  in  $5^\circ$  increments. Geometry optimization and normal mode analysis were performed (Figure S18B) using b3lyp/6-311++g(d,p) in Gaussian 16.<sup>11</sup> Harmonic frequencies were scaled by 0.9679.<sup>2,53</sup> Following geometry optimization, the heavy atom H-bond distance ( $O\cdots N$ ) remained in a range of 3.0 – 3.2 Å between  $90^\circ - 175^\circ$  and increased to 3.2 – 3.5 Å between  $70^\circ - 90^\circ$ .

In a plot of calculated vibrational frequency vs H-bond angle (Figure S19A), a  $\sim 25 \text{ cm}^{-1}$  frequency span is observed over the  $70^\circ - 175^\circ$  range, and the dependence of the H-bond shift on the H-bond angle is similar to that reported previously.<sup>54</sup> We refer to these calculated vibrational frequencies as  $\bar{\nu}_{DFT}$  and, similar to the experimentally observed  $\bar{\nu}_{obs}$  in main text Equation 4, we take  $\bar{\nu}_{DFT}$  to be:

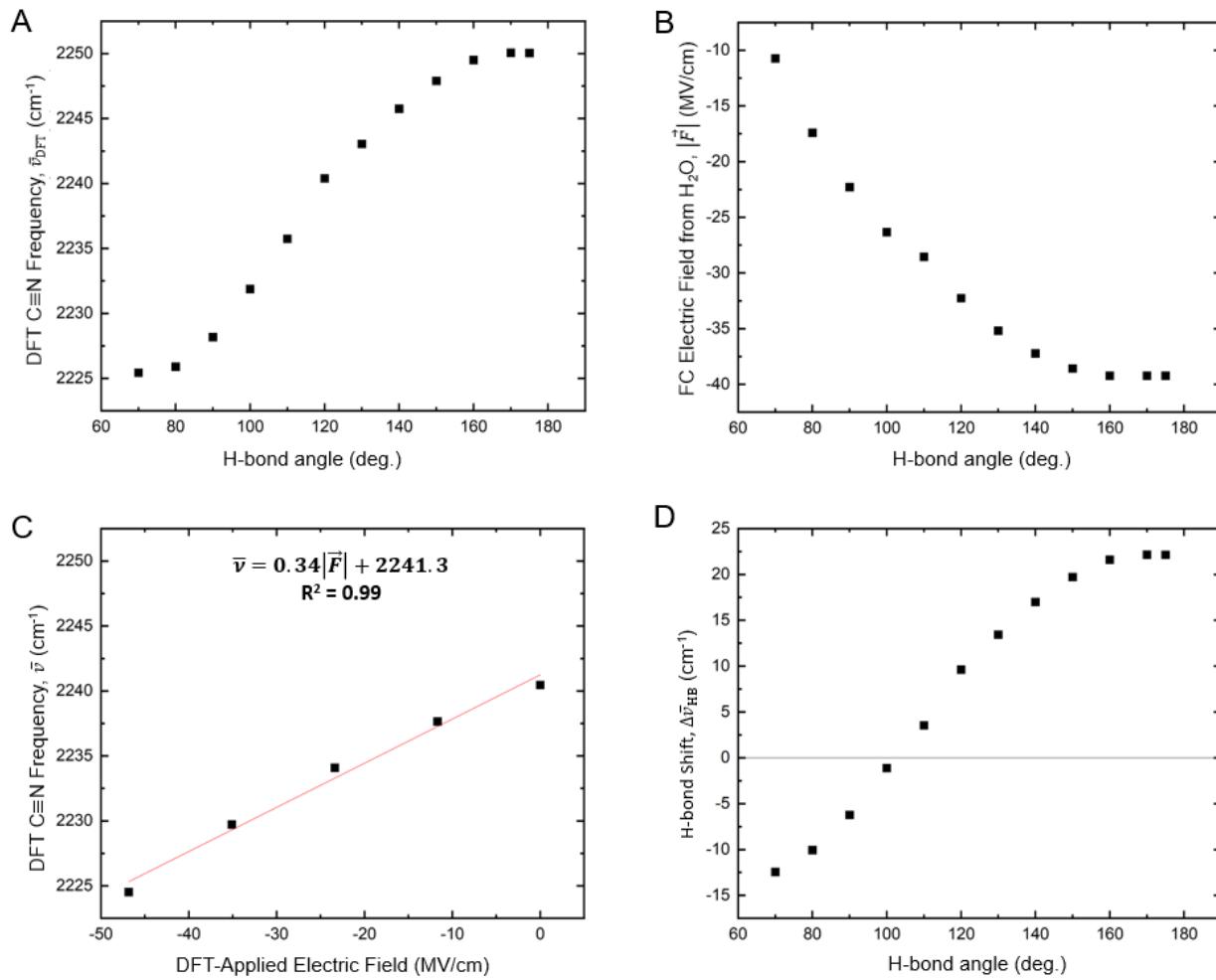
$$\bar{\nu}_{DFT} = \Delta\bar{\nu}_{HB} + \Delta\bar{\nu}_{non-HB} + \bar{\nu}_0 \quad (\text{S2})$$

Here, as in Equation 4 in the main text,  $\Delta\bar{\nu}_{non-HB}$  is the contribution to the frequency from the dipolar VSE,  $\Delta\bar{\nu}_{non-HB} = -\Delta\vec{\mu} \cdot \vec{F}$ ,  $\bar{\nu}_0$  is the vacuum (zero-field) frequency, and  $\Delta\bar{\nu}_{HB}$  is the H-bond contribution to frequency not captured by main text Equation 1. Determining an *in silico*  $\Delta\bar{\nu}_{non-HB}$  requires both the electric field along the nitrile bond at each H-bond angle and a field-frequency calibration to determine an *in silico* Stark tuning rate, or  $|\Delta\vec{\mu}|$ .<sup>19</sup> To extract the electric field projected onto oTN's  $-C\equiv N$  at each H-bond angle in Figure S19A, we utilized GAFF and tip3p parameters for oTN and water, respectively, at each given geometry (Figure S19B). To determine a DFT-based field-frequency calibration, we applied electric fields along the  $-C\equiv N$  bond in DFT calculations and extracted the resulting vibrational frequencies (Figure S19C), where  $\bar{\nu}_0$  was the vacuum frequency obtained at zero-applied field. Using the electric fields from each H-bond angle (Figure S19B) and the DFT-based field-frequency calibration in Figure S19C, we determined  $\Delta\bar{\nu}_{non-HB}$ . From Equation S2 and using the calculated  $\bar{\nu}_{DFT}$  and  $\Delta\bar{\nu}_{non-HB} + \bar{\nu}_0$ ,  $\Delta\bar{\nu}_{HB}$  was determined.

After removing  $\Delta\bar{\nu}_{non-HB} + \bar{\nu}_0$  from  $\bar{\nu}_{DFT}$ , we determine that the H-bond shift varies between  $22 \text{ cm}^{-1}$  and  $-12 \text{ cm}^{-1}$  at H-bond angles of  $175^\circ$  and  $70^\circ$ , respectively; at an H-bond angle of  $\sim 100^\circ$  the H-bond contribution to the frequency becomes negligible. The angle dependence of  $\Delta\bar{\nu}_{HB}$  is depicted in Figure S19D. We observe a similar dependence of  $\Delta\bar{\nu}_{HB}$  on H-bond angle as that observed for  $\Delta\bar{\nu}_{HB}$  in Choi *et al.*<sup>54</sup> and for  $\bar{\nu}_{DFT}$  in Figure S19A.



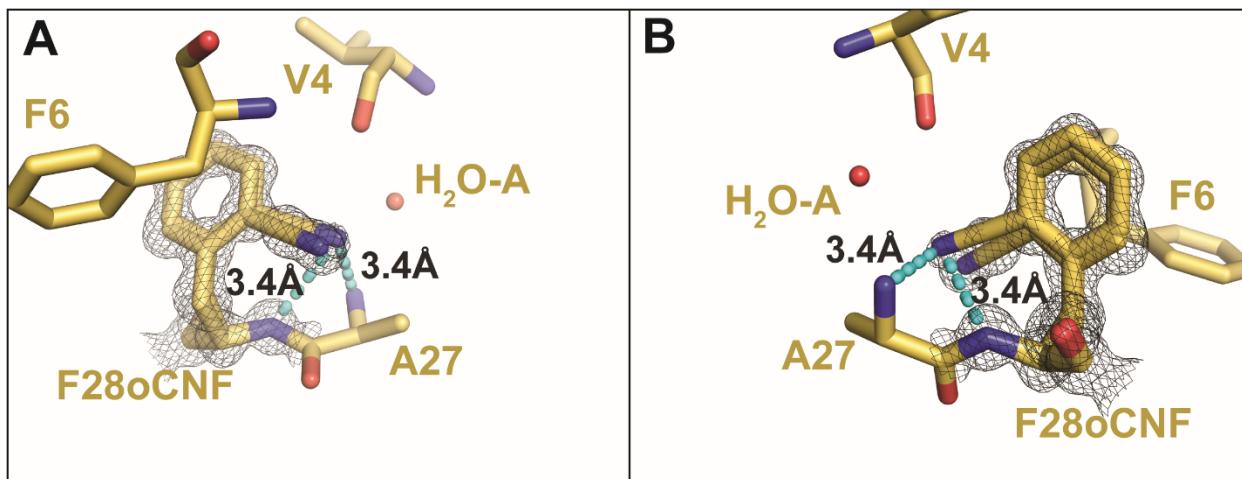
**Figure S18.** The H-bond angle  $\alpha$ , ( $\text{O}\cdots\text{N}\equiv\text{C}$ , shown in blue) between water and oTN's  $-\text{C}\equiv\text{N}$  that was scanned in DFT simulations (Figure S19) to analyze the angle-dependent response of the vibrational frequency. (A) The dependence of the nitrile frequency on the heavy atom H-bond angle ( $\text{O}\cdots\text{N}\equiv\text{C}$ ,  $\alpha$ ) was examined and plotted on Figure S19A-B and S19D. The  $\text{H}\cdots\text{C}\equiv\text{N}$  angle ( $\beta$ ) following geometry optimization was consistently  $0 - 5^\circ$  larger than the  $\text{O}\cdots\text{N}\equiv\text{C}$  angle,  $\alpha$ . This small deviation was due to the water O-H bond pointing towards the nitrile's lone pair rather than the center of the nitrogen atom. (B) Representative snapshots of the H-bond angle following geometry optimization in the gas phase where  $\alpha$  was scanned from  $175^\circ$  to  $70^\circ$ ; the snapshot overlay is depicted such that the water O-H and the nitrile  $-\text{C}\equiv\text{N}$  are in the XZ plane and the ring is in the XY plane. (C) Alternative perspective at the  $\alpha=150^\circ$  snapshot, where the ring of oTN is now in the XZ plane.



**Figure S19.** DFT-based estimation of the angle dependence of the *o*-tolunitrile (oTN) H-bond shift. (A) DFT-based -C≡N stretch frequencies were obtained from oTN H-bonded with one water molecule at restrained H-bond angles (O···N≡C;  $\alpha$  in Figure S18A) between 70° – 175° showing a similar trend as observed previously.<sup>54</sup> (B) The corresponding electric fields projected by the water molecule onto the -C≡N bond were determined at the 70° – 175° geometries using GAFF and tip3p parameters for oTN and water, respectively. (C) A DFT-based VSE calibration of oTN’s -C≡N frequency obtained by exposing oTN to an external electric field oriented along the -C≡N axis. Plotting the -C≡N frequency vs the electric field at the center of the -C≡N group (in analogy to the formalism used for MD in this work; SI Section S5) yields a calibration of  $\bar{\nu} = 0.34 |\vec{F}| + 2241.3$  cm<sup>-1</sup> ( $R^2 = 0.99$ ), where 2241.3 cm<sup>-1</sup> is the DFT-calculated zero-field frequency<sup>a</sup> and  $|\Delta \vec{\mu}| = 0.34 \frac{\text{cm}^{-1}}{\text{MV/cm}}$ . (D) Using the calibration in (C) and the electric fields in (B), the  $\Delta \bar{\nu}_{\text{non-HB}} + \bar{\nu}_0$  frequency contribution in (A) was removed resulting in an angle-dependent H-bond shift ( $\Delta \bar{\nu}_{\text{HB}}$ ) that varies between ~22 cm<sup>-1</sup> and ~ -12 cm<sup>-1</sup> in the range of 175° to 70° and vanishes at an H-bond angle of ~100°.

<sup>a</sup> The experimentally determined gas-phase frequency for oTN is 2226 cm<sup>-1</sup> (Table 2.6 in Reference 55)<sup>55</sup> while the scaled value predicted by DFT is 2241 cm<sup>-1</sup>; the relative difference (0.7%) observed between the experimental and calculated value is likely due to the scaling factor used here (0.9679). This scaling factor was determined via a best fit to a variety of compounds where a 0-2% differences were observed for 93.4% of the data.<sup>53</sup>

In the main text, the X-ray structure indicates the nitrile in the F28oCNF PYP variant is in an H-bonded environment (Figure 2C), and we observe an H-bond shift of  $7\text{ cm}^{-1}$  (Table 1). The F28oCNF PYP also has two additional potential H-bond donors, the backbone N atom of A27 and the backbone N of F28oCNF, both about  $3.4\text{ \AA}$  away from the nitrile. Both backbone nitrogens are farther from the nitrile than the  $\text{H}_2\text{O-A O}$  (Figure 2C) and have heavy atom H-bond angles ( $\text{N}\cdots\text{N}\equiv\text{C}$ ) of  $85^\circ$  and  $62^\circ$  for A27 N and F28 N, respectively, making them very poor potential H-bond donors. As such, these potential H-bonds were not discussed in Figure 2 but are shown in Figure S20 for completeness. It should be noted that the  $-\text{C}\equiv\text{N}$  of F28oCNF PYP predominantly points outside of PYP towards the bulk solvent, where multiple H-bonds with a broader distribution of angles are expected ( $130^\circ$ – $140^\circ$ );<sup>56</sup> estimating the consequence on vibrational frequency of this broader angle distribution would require a more elaborate analysis than the one presented in Figure S19. Nonetheless, based on the crystallographic heavy atom H-bond angle between F28oCNF and  $\text{H}_2\text{O-A}$  ( $110^\circ$ ), the DFT calculation predicts a  $\Delta\bar{v}_{\text{HB}}$  of  $\sim 4\text{ cm}^{-1}$  which is similar to the  $7\text{ cm}^{-1}$  H-bond shift observed experimentally (Table 1).



**Figure S20.** Alternate perspectives (A,B) of the F28oCNF nitrile and its interaction with the backbone N from A27 and the backbone N of F28oCNF. Given the larger distances (A27:  $3.4\text{ \AA}$ ; F28oCNF:  $3.4\text{ \AA}$ ) and angles (A27:  $85^\circ$ ; F28oCNF:  $62^\circ$ ) these donors adopt in comparison to  $\text{H}_2\text{O-A}$ , they were not considered as H-bond donors in main text Figure 2.

## S17 Commentary on the Additive Model of Main Text Equation 4

As shown in Figure 1B, nitriles demonstrate divergent frequency shift trends in H-bonding and non-H-bonding solvents. Here, we briefly review prior attempts in the literature to model the nitrile frequency shift in both protic and aprotic solvents as grounds for the model invoked in the main text. This commentary is by no means comprehensive but is instead meant to provide context for the model used in the main text which is attractive because it is directly accessible from experiment and physically reasonable. In the following,  $\bar{\nu}_{\text{obs}}$  is the observed nitrile frequency and  $\bar{\nu}_0$  is the vacuum frequency of the relevant nitrile-containing molecule.

### *Empirical Models*

Early work demonstrating frequency shifts (and changes of nitrile intensities) in aprotic and protic solvents can be found in Eaton *et al.*<sup>57</sup> Eaton *et al.* correlate the experimental acetonitrile -C≡N frequency against the empirical Gutmann acceptor number (AN), a measure of solvent (Lewis) acidity; they observe a linear correlation in aprotic solvents but no correlation in protic solvents (Figure 2a in Reference 57).<sup>57</sup> Other empirical models have been produced, such as an additional correlation between acetonitrile's experimental -C≡N frequency and the AN<sup>58</sup> and a correlation of benzonitrile's experimental -C≡N frequency against the Kamlet-Taft – and other – empirical parameters.<sup>59</sup> Some of the first attempts at quantitative, additive models for nitrile frequencies likewise utilized empirical parameters, as in Fawcett *et al.*<sup>60</sup>:

$$\bar{\nu}_{\text{obs}} - \bar{\nu}_{\text{ACN}} = \Delta\bar{\nu}_{\text{AN}} + \Delta\bar{\nu}_{\text{DN}} + \Delta\bar{\nu}_{\text{dielectric}} + \Delta\bar{\nu}_{\text{polarizability}} \quad (\text{S3})$$

where  $\bar{\nu}_{\text{ACN}}$  is the -C≡N frequency of neat, liquid solution acetonitrile, and  $\Delta\bar{\nu}_{\text{AN}}$ ,  $\Delta\bar{\nu}_{\text{DN}}$ ,  $\Delta\bar{\nu}_{\text{dielectric}}$ , and  $\Delta\bar{\nu}_{\text{polarizability}}$  refer to frequency shifts associated with Gutmann AN, Gutmann donor number (DN), solvent dielectric strength, and solvent polarizability, respectively. This model described the experimental acetonitrile frequencies well except for -C≡N frequencies in H-bond donating fluorinated alcohols; the model's development is discussed extensively in Section 4 of Reference 61.<sup>61</sup> While these models often worked reasonably well in non-H-bonding solvents, challenges in H-bonding solvents led others to use force-based variables instead of empirical parameters.

### *A Quantitative, Coarse-Grained Model*

Ben-Amotz *et al.*<sup>62</sup> produced one of the earliest attempts to understand nitrile frequency shifts with a force-based model (see Section 3 of Reimers and Hall<sup>61</sup> for brief review). Their model simplifies the solvent into a hard-sphere fluid and describes the solvent frequency shift of acetonitrile's experimental -C≡N frequency using the additive model:

$$\bar{\nu}_{\text{obs}} = \Delta\bar{\nu}_{\text{A,El}} + \Delta\bar{\nu}_{\text{A,Disp}} + \Delta\bar{\nu}_{\text{R}} + \Delta\bar{\nu}_{\text{C}} + \bar{\nu}_0 \quad (\text{S4})$$

where  $\Delta\bar{\nu}_{\text{A,El}}$ ,  $\Delta\bar{\nu}_{\text{A,Disp}}$ ,  $\Delta\bar{\nu}_{\text{R}}$ , and  $\Delta\bar{\nu}_{\text{C}}$  refer to frequency shifts from attractive electrostatic, attractive dispersive, short-range repulsive, and centrifugal forces (only relevant in the gas phase), respectively. This model performed reasonably well in the description of solvent- and pressure-dependence of nitrile

frequencies and demonstrated the importance of the interplay of electrostatic, dispersive, and repulsive interactions. However, the parameters used are based on mean solvent-solute forces, not specific solvent-solute interactions known to be important to modeling solvation phenomena and protic interactions more specifically.

#### *Quantitative Models based on Specific Solvent-Solute Interactions*

In Eaton *et al.*,<sup>57</sup> the different frequency shifts observed in non-H-bonding (redshifts) and H-bonding (blueshifts) environments were recognized and explained in terms of solvent interactions with either the nitrile's  $\pi$  bond or with the nitrile nitrogen's lone pair, respectively; this was the first suggestion that differences in solvent-nitrile interaction could lead to divergent frequency shifts. In subsequent work, Reimers and Hall performed DFT on a variety of acetonitrile-water clusters and found that linear H-bonds result in  $-C\equiv N$  blueshifts while distorted H-bonds lead to  $-C\equiv N$  redshifts;<sup>61</sup> similar DFT calculations on benzonitrile-methanol interactions recapitulate  $-C\equiv N$  frequency shift trends in analogous scenarios.<sup>59</sup>

Cho's group built on these studies and performed additional work focusing on the angle between the nitrile and the H-bond donor.<sup>54,63–65</sup> In Choi *et al.*,<sup>54</sup> H-bond angle dependent QM simulations for acetonitrile/methyl thiocyanate with a single water indicate blueshifts for H-bonds at  $175^\circ$  and redshifts at  $90^\circ$ , with a frequency shift sign change occurring at  $\sim 120^\circ$  (Figure 2 in Cho *et al.*).<sup>54</sup> With this result, the authors begin to build a model for nitrile frequency shifts, concluding that an electrostatic potential model can recapitulate calculated nitrile frequency shifts in some H-bonding scenarios. Similarly, a related DFT-based report found that the solvent-generated electrostatic potential and electric field along the bond axis were necessary to accurately model calculated  $-C\equiv N$  frequencies of H-bonded nitriles.<sup>66</sup>

In Lee *et al.*,<sup>63</sup> the *ab initio* frequency shifts of nitriles, such as acetonitrile (e.g. Figure 2 in Lee *et al.*),<sup>63</sup> were modeled with part of the Coulombic (electrostatic) multipole expansion, i.e. the additive contributions of the water's electrostatic interactions with the nitrile group's dipolar, quadrupolar, and octupolar components:

$$\bar{\nu}_{\text{obs}} = \Delta\bar{\nu}_{\text{dipole}} + \Delta\bar{\nu}_{\text{quadrupole}} + \Delta\bar{\nu}_{\text{octopole}} + \bar{\nu}_0 \quad (\text{S5})$$

They indicate the dipolar and quadrupolar interactions sum to provide the dominant contribution to  $-C\equiv N$  frequency tuning, indicating that the nitrile's interaction(s) with the solvent electric field(s) and electric field gradient(s) are most significant.<sup>63</sup>

In Cho's most recent model, Blasiak *et al.*,<sup>65</sup> solvation effects are described via an additive model of frequency contributions from Coulombic, exchange-repulsive, inductive, dispersive, and charge-transfer terms:

$$\bar{\nu}_{\text{obs}} = \Delta\bar{\nu}_{\text{Coul}} + \Delta\bar{\nu}_{\text{Ex-Rep}} + \Delta\bar{\nu}_{\text{Ind}} + \Delta\bar{\nu}_{\text{Disp}} + \Delta\bar{\nu}_{\text{CT}} + \bar{\nu}_0 \quad (\text{S6})$$

When analyzed without the charge-transfer term, this model indicated that the exchange-repulsion interactions predominately accounted for the blueshift in H-bonding environments (Figure 3 in Blasiak *et al.*).<sup>64</sup>

#### *Model Utilized in Main Text Equation 4*

In our current work, we take inspiration from the abovementioned models to semi-empirically model our experimental observations:

$$\bar{\nu}_{\text{obs}} = \Delta\bar{\nu}_{\text{non-HB}} + \Delta\bar{\nu}_{\text{HB}} + \bar{\nu}_0 \quad (\text{S7})$$

We partition our model in this way because our experimental frequency-field relation calibrates only for the electric field along the -C≡N bond axis, or the dipolar component of the multipole expansion ( $\Delta\bar{\nu}_{\text{non-HB}}$ ). We do not have experimental means of calibrating other multipole moments in Equation S5 or any of the other forces invoked in Equation S6; as such, we combine all other frequency contributions to nitriles in H-bonding solvents into the second term of Equation S7 ( $\Delta\bar{\nu}_{\text{HB}}$ ). Though our model does not dissect all the individual interactions as in Blasiak *et al.*,<sup>65</sup> it is clear that such an additive approach is in line with prior modeling and useful for our purposes as it is experimentally accessible. Since  $\Delta\bar{\nu}_{\text{HB}}$  appears to be significant only in protic solvents, and because it is sensitive to the geometry of nitrile protic interaction, we use it as an empirical means to estimate structural information (H-bond angles) regarding the nitrile H-bonding interaction.

## S18 References

- (1) Schneider, S. H.; Boxer, S. G. Vibrational Stark Effects of Carbonyl Probes Applied to Reinterpret IR and Raman Data for Enzyme Inhibitors in Terms of Electric Fields at the Active Site. *J. Phys. Chem. B* **2016**, *120* (36), 9672–9684.
- (2) Kozuch, J.; Schneider, S. H.; Zheng, C.; Ji, Z.; Bradshaw, R. T.; Boxer, S. G. Testing the Limitations of MD-Based Local Electric Fields Using the Vibrational Stark Effect in Solution: Penicillin G as a Test Case. *J. Phys. Chem. B* **2021**, *13*.
- (3) Tharp, J. M.; Wang, Y. S.; Lee, Y. J.; Yang, Y.; Liu, W. R. Genetic Incorporation of Seven Ortho-Substituted Phenylalanine Derivatives. *ACS Chem. Biol.*, **2014**, *9*, 884–890.
- (4) Hammill, J. T.; Miyake-Stoner, S.; Hazen, J. L.; Jackson, J. C.; Mehl, R. A. Preparation of Site-Specifically Labeled Fluorinated Proteins for 19F-NMR Structural Characterization. *Nat Protoc.*, **2007**, *2*, 2601–2607.
- (5) Thomson, B.; Both, J.; Wu, Y.; Parrish, R. M.; Martínez, T. J.; Boxer, S. G. Perturbation of Short Hydrogen Bonds in Photoactive Yellow Protein via Noncanonical Amino Acid Incorporation. *J. Phys. Chem. B* **2019**, *123* (23), 4844–4849.
- (6) Deb, P.; Haldar, T.; Kashid, S. M.; Banerjee, S.; Chakrabarty, S.; Bagchi, S. Correlating Nitrile IR Frequencies to Local Electrostatics Quantifies Noncovalent Interactions of Peptides and Proteins. *J. Phys. Chem. B* **2016**, *120*, 4034–4046.
- (7) Fried, S. D.; Bagchi, S.; Boxer, S. G. Measuring Electrostatic Fields in Both Hydrogen-Bonding and Non-Hydrogen-Bonding Environments Using Carbonyl Vibrational Probes. *J. Am. Chem. Soc.* **2013**, *135* (30), 11181–11192.
- (8) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry* **2004**, *25* (9), 1157–1174.
- (9) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *Journal of Molecular Graphics and Modelling* **2006**, *25* (2), 247–260.
- (10) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25.
- (11) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Rob, M. A.; Cheeseman, J. R.; Scalmani, G.; et al. *Gaussian 16*; Gaussian, Inc.: Wallingford, CT, **2016**.
- (12) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *Journal of Computational Chemistry* **2002**, *23* (16), 1623–1641.
- (13) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.* **2012**, *8* (1), 61–74.
- (14) van der Spoel, D.; van Maaren, P. J.; Caleman, C. GROMACS Molecule & Liquid Database. *Bioinformatics* **2012**, *28* (5), 752–753.
- (15) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.
- (16) Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. Tinker 8: Software Tools for Molecular Design. *J. Chem. Theory Comput.* **2018**, *14* (10), 5273–5289.

- (17) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114* (8), 2549–2564.
- (18) Fried, S. D.; Wang, L.-P.; Boxer, S. G.; Ren, P.; Pande, V. S. Calculations of the Electric Fields in Liquid Solutions. *J. Phys. Chem. B* **2013**, *117* (50), 16236–16248.
- (19) Zheng, C.; Mao, Y.; Kozuch, J.; Atsango, A.; Ji, Z.; Markland, T.; Boxer, S. Electric Field Orientations in Solution and Enzyme Active Site Revealed by a Two-Directional Vibrational Probe. *Nature Chemistry, in press* **2021**, <https://doi.org/10.26434/chemrxiv-2021-578vg>, Submitted 2021-09-02, Accessed 2021-11-09.
- (20) Lin, C.-Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S. G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *J. Am. Chem. Soc.* **2019**, *141* (38), 15250–15265.
- (21) Parson, W. W. Electronic Absorption. In *Modern Optical Spectroscopy: With Exercises and Examples from Biophysics and Biochemistry*; Parson, W. W., Ed.; Springer: Berlin, Heidelberg, **2015**; pp 138–139.
- (22) Mu, X.; Wang, Q.; Wang, L.-P.; Fried, S. D.; Piquemal, J.-P.; Dalby, K. N.; Ren, P. Modeling Organochlorine Compounds and the  $\sigma$ -Hole Effect Using a Polarizable Multipole Force Field. *J. Phys. Chem. B* **2014**, *118* (24), 6456–6465.
- (23) Bublitz, G. U.; Boxer, S. G. Stark Spectroscopy: Applications in Chemistry, Biology, and Materials Science. *Annu. Rev. Phys. Chem.* **1997**, *48* (1), 213–242.
- (24) Andrews, S. S.; Boxer, S. G. Vibrational Stark Effects of Nitriles I. Methods and Experimental Results. *J. Phys. Chem. A* **2000**, *104* (51), 11853–11863.
- (25) Andrews, S. S.; Boxer, S. G. A Liquid Nitrogen Immersion Cryostat for Optical Measurements. *Review of Scientific Instruments* **2000**, *71* (9), 3567–3569.
- (26) Chattopadhyay, A.; Boxer, S. G. Vibrational Stark Effect Spectroscopy. *J. Am. Chem. Soc.* **1995**, *117* (4), 1449–1450.
- (27) Romei, M. G.; Lin, C.-Y.; Boxer, S. G. Structural and Spectroscopic Characterization of Photoactive Yellow Protein and Photoswitchable Fluorescent Protein Constructs Containing Heavy Atoms. *Journal of Photochemistry and Photobiology A: Chemistry* **2020**, *401*, 112738.
- (28) Kabsch, W. XDS. *Acta Cryst D* **2010**, *66* (2), 125–132.
- (29) Kabsch, W. Integration, Scaling, Space-Group Assignment and Post-Refinement. *Acta Cryst D* **2010**, *66* (2), 133–144.
- (30) A Quick XDS Tutorial for SSRL  
[https://smb.slac.stanford.edu/facilities/software/xds/#autoxds\\_script](https://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script) (accessed 2021-11-07).
- (31) Adams, P. D.; Afonine, P. V.; Bunkóczki, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Cryst D* **2010**, *66* (2), 213–221.
- (32) Emsley, P.; Cowtan, K. Coot : Model-Building Tools for Molecular Graphics. *Acta Cryst.* **2004**, *60* (12), 2126–2132.
- (33) Joosten, R. P.; Joosten, K.; Cohen, S. X.; Vriend, G.; Perrakis, A. Automatic Rebuilding and Optimization of Crystallographic Structures in the Protein Data Bank. *Bioinformatics* **2011**, *27* (24), 3392–3398.
- (34) Lin, C.-Y.; Both, J.; Do, K.; Boxer, S. G. Mechanism and Bottlenecks in Strand Photodissociation of Split Green Fluorescent Proteins (GFPs). *PNAS* **2017**, *114* (11), E2146–E2155.
- (35) Ward, W. Properties of the Coelenterate Green-Fluorescent Proteins. In *Bioluminescence and Chemiluminescence*; 1981; pp 235–242.

- (36) Takeshita, K.; Imamoto, Y.; Kataoka, M.; Tokunaga, F.; Terazima, M. Thermodynamic and Transport Properties of Intermediate States of the Photocyclic Reaction of Photoactive Yellow Protein. *Biochemistry* **2002**, *41* (9), 3037–3048.
- (37) Meyer, T. E. Isolation and Characterization of Soluble Cytochromes, Ferredoxins and Other Chromophoric Proteins from the Halophilic Phototrophic Bacterium *Ectothiorhodospira halophila*. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1985**, *806* (1), 175–183.
- (38) Meyer, T. E.; Tollin, G.; Hazzard, J. H.; Cusanovich, M. A. Photoactive Yellow Protein from the Purple Phototrophic Bacterium, *Ectothiorhodospira halophila*. Quantum Yield of Photobleaching and Effects of Temperature, Alcohols, Glycerol, and Sucrose on Kinetics of Photobleaching and Recovery. *Biophysical Journal* **1989**, *56* (3), 559–564.
- (39) Lin, C.-Y.; Boxer, S. G. Unusual Spectroscopic and Electric Field Sensitivity of Chromophores with Short Hydrogen Bonds: GFP and PYP as Model Systems. *J. Phys. Chem. B* **2020**, *124* (43), 9513–9525.
- (40) Anthis, N. J.; Clore, G. M. Sequence-Specific Determination of Protein and Peptide Concentrations by Absorbance at 205 nm. *Protein Science* **2013**, *22* (6), 851–858.
- (41) Ballentine, R. Determination of Total Nitrogen and Ammonia. In *Methods in Enzymology*; Elsevier, **1957**; Vol. 3, pp 984–995.
- (42) Ozols, J. Amino Acid Analysis. In *Methods in Enzymology*; Deutscher, M. P., Ed.; Guide to Protein Purification; Academic Press, 1990; Vol. 182, pp 587–601.
- (43) Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry* **1976**, *72* (1), 248–254.
- (44) Kielkopf, C. L.; Bauer, W.; Urbatsch, I. L. Bradford Assay for Determining Protein Concentration. *Cold Spring Harb Protoc* **2020**, *2020* (4), pdb.prot102269.
- (45) Cui, W.; Xue, H.; Cheng, H.; Zhang, H.; Jin, J.; Wang, Q. Increasing the Amount of Phosphoric Acid Enhances the Suitability of Bradford Assay for Proteomic Research. *ELECTROPHORESIS* **2019**, *40* (7), 1107–1112.
- (46) Noble, J. E.; Bailey, M. J. A. Chapter 8 Quantitation of Protein. In *Methods in Enzymology*; Burgess, R. R., Deutscher, M. P., Eds.; Guide to Protein Purification, 2nd Edition; Academic Press, 2009; Vol. 463, pp 73–95.
- (47) Jones, L. J.; Haugland, R. P.; Singer, V. L. Development and Characterization of the NanoOrange® Protein Quantitation Assay: A Fluorescence-Based Assay of Proteins in Solution. *BioTechniques* **2003**, *34* (4), 850–861.
- (48) Fluorescent Protein Assays - US //www.thermofisher.com/us/en/home/life-science/protein-biology/protein-assays-analysis/protein-assays/fluorescent-protein-assays.html (accessed 2021-11-07).
- (49) Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; Lantz, C.; Jung, W.; Johnson, K. R.; Koller, A.; Wolff, J. J.; Campuzano, I. D. G.; Auclair, J. R.; Ivanov, A. R.; Whitelegge, J. P.; Paša-Tolić, L.; Chamot-Rooke, J.; Danis, P. O.; Smith, L. M.; Tsybin, Y. O.; Loo, J. A.; Ge, Y.; Kelleher, N. L.; Agar, J. N. Best Practices and Benchmarks for Intact Protein Analysis for Top-down Mass Spectrometry. *Nat Methods* **2019**, *16* (7), 587–594.
- (50) Pan, S.; Aebersold, R.; Chen, R.; Rush, J.; Goodlett, D. R.; McIntosh, M. W.; Zhang, J.; Brentnall, T. A. Mass Spectrometry Based Targeted Protein Quantification: Methods and Applications. *J. Proteome Res.* **2009**, *8* (2), 787–797.
- (51) Fica-Contreras, S. M.; Daniels, R.; Yassin, O.; Hoffman, D. J.; Pan, J.; Sotzing, G.; Fayer, M. D. Long Vibrational Lifetime R-Selenocyanate Probes for Ultrafast Infrared Spectroscopy: Properties and Synthesis. *J. Phys. Chem. B* **2021**, *125* (31), 8907–8918.

- (52) Kraskov, A.; von Sass, J.; Nguyen, A. D.; Hoang, T. O.; Buhrke, D.; Katz, S.; Michael, N.; Kozuch, J.; Zebger, I.; Siebert, F.; Scheerer, P.; Mroginski, M. A.; Budisa, N.; Hildebrandt, P. Local Electric Field Changes during the Photoconversion of the Bathy Phytochrome Agp2. *Biochemistry* **2021**, *60* (40), 2967–2977.
- (53) Andersson, M. P.; Uvdal, P. New Scale Factors for Harmonic Vibrational Frequencies Using the B3LYP Density Functional Method with the Triple- $\zeta$  Basis Set 6-311+G(d,p). *J. Phys. Chem. A* **2005**, *109* (12), 2937–2941.
- (54) Choi, J.-H.; Oh, K.-I.; Lee, H.; Lee, C.; Cho, M. Nitrile and Thiocyanate IR Probes: Quantum Chemistry Calculation Studies and Multivariate Least-Square Fitting Analysis. *J. Chem. Phys.* **2008**, *128* (13), 134506.
- (55) Chapter 2 - Nitriles, Isonitriles, and Dialkyl Cyanamides. In *Interpreting Infrared, Raman, and Nuclear Magnetic Resonance Spectra*; Nyquist, R. A., Ed.; Academic Press: San Diego, 2001; pp 27–43.
- (56) First, J. T.; Slocum, J. D.; Webb, L. J. Quantifying the Effects of Hydrogen Bonding on Nitrile Frequencies in GFP: Beyond Solvent Exposure. *J. Phys. Chem. B* **2018**, *122* (26), 6733–6743.
- (57) Eaton, G.; Pena-Nuñez, A. S.; Symons, M. C. R. Solvation of Cyanoalkanes [CH<sub>3</sub>CN and (CH<sub>3</sub>)<sub>3</sub>NN]. An Infrared and Nuclear Magnetic Resonance Study. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84* (6), 2181–2193.
- (58) Nyquist, R. A. Solvent-Induced Nitrile Frequency Shifts: Acetonitrile and Benzonitrile. *Appl Spectrosc* **1990**, *44* (8), 1405–1407.
- (59) Aschaffenburg, D. J.; Moog, R. S. Probing Hydrogen Bonding Environments: Solvatochromic Effects on the CN Vibration of Benzonitrile. *J. Phys. Chem. B* **2009**, *113* (38), 12736–12743.
- (60) Fawcett, W. R.; Liu, G.; Kessler, T. E. Solvent-Induced Frequency Shifts in the Infrared Spectrum of Acetonitrile in Organic Solvents. *J. Phys. Chem.* **1993**, *97* (37), 9293–9298.
- (61) Reimers, J. R.; Hall, L. E. The Solvation of Acetonitrile. *J. Am. Chem. Soc.* **1999**, *121* (15), 3730–3744.
- (62) Ben-Amotz, D.; Lee, M.; Cho, S. Y.; List, D. J. Solvent and Pressure-induced Perturbations of the Vibrational Potential Surface of Acetonitrile. *J. Chem. Phys.* **1992**, *96* (12), 8781–8792.
- (63) Lee, H.; Choi, J.-H.; Cho, M. Vibrational Solvatochromism and Electrochromism. II. Multipole Analysis. *The Journal of Chemical Physics* **2012**, *137* (11), 114307.
- (64) Błasiak, B.; Ritchie, A. W.; Webb, L. J.; Cho, M. Vibrational Solvatochromism of Nitrile Infrared Probes: Beyond the Vibrational Stark Dipole Approach. *Phys. Chem. Chem. Phys.* **2016**, *18* (27), 18094–18111.
- (65) Błasiak, B.; Londergan, C. H.; Webb, L. J.; Cho, M. Vibrational Probes: From Small Molecule Solvatochromism Theory and Experiments to Applications in Complex Systems. *Acc. Chem. Res.* **2017**, *50* (4), 968–976.
- (66) Torii, H. Unified Electrostatic Understanding on the Solvation-Induced Changes in the CN Stretching Frequency and the NMR Chemical Shifts of a Nitrile. *J. Phys. Chem. A* **2016**, *120* (36), 7137–7144.