THE JOURNAL OF PHYSICAL CHEMISTRY B

pubs.acs.org/JPCB Article

Beyond pK_a: Experiments and Simulations of Nitrile Vibrational Probes in Staphylococcal Nuclease Show the Importance of Local Interactions

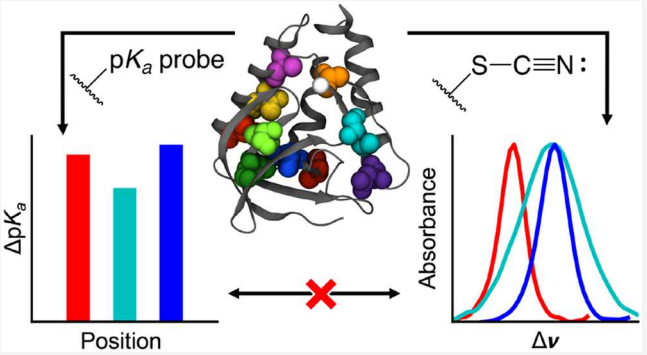
Published as part of The Journal of Physical Chemistry virtual special issue "Peter J. Rossky Festschrift".

Jeremy T. First,[‡] Elisa T. Novelli,[‡] and Lauren J. Webb*

Cite This: *J. Phys. Chem. B* 2020, 124, 3387−3399 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Electric fields are fundamentally important to biological phenomena, but are difficult to measure experimentally or predict computationally. Changes in pK_a of titratable residues have long been used to report on local electrostatic fields in proteins. Alternatively, nitrile vibrational probes are potentially less disruptive and more direct reporters of local electrostatic field, but quantitative interpretation is clouded by the ability of the nitrile to accept a hydrogen bond. To this end, we incorporated nitrile probes into 10 locations of staphylococcal nuclease (SNase) where pK_a shifts had already been determined. We characterized the local environment of each nitrile probe experimentally, through temperature-dependent spectroscopy, and computationally, through molecular dynamics simulations, and show that hydrogen bonding interactions dominate the spectral line shapes. We demonstrate that the information provided by the line shape of the nitrile spectra, compared to scalar values of pK_a shift or nitrile frequency shift, better describes local environments in proteins in a manner that will be useful for future computational efforts to predict electrostatics in complex biological systems.



INTRODUCTION

Electrostatic forces play a critical role in biological phenomena and are of long-standing interest to the biophysical community. In proteins, for example, biochemical processes such as folding, catalysis, and intermolecular interactions are regulated by electric fields resulting from the arrangement of partial charges on each of the many thousands of atoms within the protein structure.[1] Understanding the molecular detail of the structure, dynamics, and function of proteins therefore depends on accurate, quantitative experiments and calculations of highly heterogeneous electric fields inherent in protein structures.[2] However, because the Coulombic interactions between partially charged atoms decay over relatively long distances compared to other intermolecular forces, such as covalent bonds and van der Waals interactions,[3−8] electrostatic fields in proteins are difficult to measure experimentally or predict computationally.

Because of the sheer size and complexity of biological systems, calculations of electric field are approximations by necessity, and therefore the calculations require careful validation against experimental data. Historically, a convenient experimental window through which to assess the overall electrostatic environment in proteins relied on the change in pK_a (ΔpK_a) experienced by a titratable amino acid upon incorporation into the protein compared to aqueous solution. These measurements take advantage of the change in the free energy difference of the protonation reaction ($\Delta\Delta G_a$) associated with the placement of a titratable residue in a protein environment, as illustrated in eq 1:

$$\Delta \mathrm{p}K_a = \frac{\Delta\Delta G_a}{RT\ \ln(10)} \quad (1)$$

Equivalently, the $\Delta\Delta G_a$ term can be written as the difference between the charged and neutral states in the free energy of transferring the amino acid from solution to the protein, as shown in eq 2:

$$\Delta\Delta G_a = \Delta G_s^{\pm} - \Delta G_s^{0} \quad (2)$$

where $\Delta G_s^{\pm}$ and ΔG_s^{0} represent the difference in free energy of solvation in water and in the protein, for both the charged and

Received: January 27, 2020
Revised: March 24, 2020
Published: March 26, 2020

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

neutral states, respectively.[9−11] The overall preferential stabilization of one state versus the other, quantified by a ΔpK_a, has typically been interpreted as a product of the overall electrostatic environment of the protein interior, such as changes in local dielectric,[12] Coulombic interactions,[13] and local solvation,[14] although nonelectrostatic interactions can contribute to each ΔG_s. Despite the convolution of these other interactions, the accessibility of both the experimental and theoretical pK_a have made this a convenient measurement to directly compare calculations and experiments.

A concerted effort toward this comparison, known as the pK_a Cooperative, used experimental ΔpK_a values in a blind prediction test of various computational calculations of electrostatics.[15] In these studies, García-Moreno and co-workers systematically placed several ionizable residues in 25 buried locations of a highly stable engineered variant of staphylococcal nuclease, SNase Δ+PHS, hereafter referred to simply as "SNase".[16−19] The measured ΔpK_a values were then used as a target data set for different attempts to reproduce those values as a test of computational models of electrostatics. Models that primarily considered only electrostatic effects on both ΔG_s terms were typically unable to reproduce the experimental ΔpK_a values measured in SNase.[20,21] Models that explicitly include conformational flexibility, such as multi-conformation continuum electrostatics (MCCE)[22,23] or constant pH molecular dynamics (CPMD),[24−26] typically were able to reproduce experimental ΔpK_a values more accurately, highlighting the importance of non-Coulombic interactions on the ΔG_s terms in eq 2. Nevertheless, shifts in pK_a are still widely interpreted as primarily due to electrostatic effects.[15] However, because ΔpK_a values represent a difference in free energy between two charged states of the probe (and therefore two different states of the protein), it is difficult to interpret from a ΔpK_a any information about the electrostatic environment of a single, native state of the protein. An alternative benchmark that more directly measures electric field and reports on only one state of the protein would therefore be useful to further validate and test computational models. In particular, information about the electric fields in the native state of the protein can be used to understand electrostatically driven phenomena such as protein−protein interactions and substrate binding.

Our laboratory has been interested in building a data set with steady-state spectroscopic measurements that are directly correlated to electric field and could serve as an alternative benchmark for efforts to design, test, and validate computational models and strategies. Along with other laboratories, we have been pursuing vibrational Stark effect (VSE) spectroscopy of protein-bound nitrile probes to measure electric fields in biological systems.[27−38] Nitriles are convenient biomolecular probes because they are small, absorb in an uncluttered region of the protein vibrational spectrum, have a reasonable oscillator strength, and can be incorporated virtually anywhere within a biomolecular structure.[29,39−41] Vibrational spectra of nitriles have been collected in a variety of biologically relevant environments such as lipid membranes, protein interiors, protein−protein interfaces, and protein active sites.[42−45] Finally, shifts in the vibrational energy of nitrile groups can be related to changes in local electric field through the linear Stark equation:[46,47]

$$\Delta E = hc\Delta\nu = -\Delta\vec{\mu}\cdot\Delta\vec{F} \quad (3)$$

Here, ΔE is the change in absorption energy of the chromophore in response to a change in local electric field, $\Delta\vec{F}$, and $\Delta\vec{\mu}$ is the difference in the dipole moment between the ground and excited state of the vibrational transition. For nitriles in particular, $\Delta\vec{\mu}$ has been well-characterized,[36,48,49] allowing direct interpretation of differences in vibrational energy as differences in the magnitude and direction of the local electric field in the immediate vicinity of the nitrile. Previously, in superfolder green fluorescent protein (GFP), we demonstrated that three separate measurements of electrostatic environment (pK_a shifts of the titratable fluorophore, linear electronic Stark shifts of the fluorophore, and VSE shifts of a nearby nitrile probe) all responded similarly to changes in electric field resulting from amino acid mutations near the fluorophore.[50] This, along with previous investigations of nitriles,[36,48,49,51] led us to the conclusion that the vibrational spectra of nitrile probes could be used as a direct measurement of electric field.

One major caveat to interpreting nitrile spectra according to eq 3 is that nitriles can accept hydrogen bonds, which cause frequency shifts that are inherently quantum mechanical and thus not described by the Stark equation.[34] Moreover, the specific geometry of the hydrogen bonding interaction changes the magnitude and direction of the frequency shift. *Ab initio* calculations by Choi et al. showed that linear, σ-hydrogen bonds blue shift the nitrile frequency by withdrawing electron density from an antibonding orbital of the C≡N bond, and perpendicular, π-hydrogen bonds red shift the nitrile frequency by withdrawing electron density from a bonding orbital.[27] Recently, we demonstrated that these contributions from hydrogen bonding can dominate the observed frequency of nitrile probes in GFP, and therefore, interpretations of absolute values of nitrile absorption energy require careful control experiments, often coupled with molecular dynamics (MD) simulations, to reveal the extent of potential hydrogen bonding interactions.[52] In the same study, we attempted to resolve some of the challenges of the local effects on nitrile oscillators by demonstrating that the temperature dependence of the nitrile absorption energy, quantified by the frequency temperature line slope (FTLS), was a diagnostic tool for measuring hydrogen bonding, as had been postulated by Adhikary et al.[52,53] based on previous work by Huang et al.[54] We further demonstrated the necessity of such a tool; every single probe location experienced hydrogen bonding, suggesting it may be impossible to intentionally avoid hydrogen bonding to nitriles in a protein system. This study highlighted that using nitriles to measure electric fields according to eq 3 requires very careful additional control experiments and a detailed understanding of the local interactions.

Both ΔpK_a and VSE measurements of electric field require a site-specific molecular probe. pK_a probes are limited in that they must be titratable and the probe must change its charge state during the measurements. Large scale conformational changes of the protein around the charge state can make the interpretation of the ΔpK_a measurement difficult.[16,18,19,55−57] Furthermore, because this measurement is based on the difference between two states, the change in the number of hydrogen bond donors upon titration likely has a significant impact on the observed shift. While this can be useful for probing responses to pH driven phenomena, this is less useful for probing the steady-state electric field of a native protein. In contrast, VSE probes are sensitive to more subtle electrostatic perturbations and only rely on a vibrational transition. While

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

the dipole moment changes between the ground and excited vibrational states, this difference perturbs the protein structure significantly less, and because nitrile groups are small, the insertion of a nitrile is minimally disruptive to the protein structure.[58]

Ultimately, the goal of developing nitrile vibrational spectroscopy as a direct measurement of electric field in biological systems is to provide a robust and reliable data set to benchmark electric field calculations. It is our aim to provide a data set that can be used to validate electrostatic models by matching simulations to experiments on nitrile-containing proteins, which can then in turn be used to provide information about the electric fields in WT biological structures. Thus, additional work in our laboratory has focused on the calculation of electric fields from structures generated with MD simulations in order to interpret the nitrile frequencies. Our efforts have primarily followed two strategies: (1) direct computation of the field at the midpoint of the nitrile using MD force fields and (2) continuum electrostatics calculations of the potential using the Poisson−Boltzmann (PB) equation. For the most part, the results of these calculations have not been able to capture the measured nitrile absorption energies.[59−61] We hypothesize that this is due to two factors: (1) the contributions of hydrogen bonding effects and VSE shifts, which we demonstrated in GFP,[52] and (2) the fact that MD force fields are parameterized to sample structural ensembles accurately at room temperature but are not intended to quantitatively reproduce forces. More sophisticated models, such as the SolEFP model, have had more success in quantitatively capturing the vibrational shifts of nitriles and reproducing the line shapes of the vibrational spectra, and have demonstrated that the Coulombic term only plays a small role in the overall vibrational frequency.[62,63] Because such models are often computationally intensive and require specialized software and expertise to implement, it is still desirable to develop a simpler model using widely available tools to interpret vibrational spectra in electrostatic terms.

To investigate the utility of nitrile spectra as a direct probe of electric field in biological systems, here we report nitrile absorption measurements of cyanocysteine (CNC) nitrile probes in 10 locations in SNase (illustrated in Figure 1) at sites where ΔpK_a measurements have been previously reported.[18,19] Using both steady-state and temperature-dependent Fourier transform infrared spectroscopy (FTIR) experiments and MD simulations, we demonstrate that (1) the FTLS is a quantitative measurement of the hydrogen bonding status of nitrile probes,[52] (2) hydrogen bonding interactions dominate the observed vibrational frequency of nitriles, and (3) it may be impossible to position a nitrile probe in a protein system in a location that is intentionally occluded from hydrogen bonding interactions. Along with our previous report of using nitriles in GFP,[52] these observations in two different proteins strongly suggest that these results are generalizable to other proteins and must always be considered when attempting to understand the role of noncovalent forces in protein structure, function, and dynamics. Next, using these reported nitrile vibrational spectra as a benchmark electrostatic data set, we calculated the electric field at the midpoint of the nitrile from an ensemble of structures generated from MD simulations. We discovered that the distributions of these calculated fields were strongly correlated to the line shapes of the nitrile spectra. We conclude that while the average calculated fields are not correlated to the vibrational frequencies because of the contribution of hydrogen bonding effects and the shortcomings of nonpolarizable force fields, the distribution of fields qualitatively describes the line shape of the experimental spectra. Overall, this study demonstrates that the information contained in the line shapes of the nitrile spectra is invaluable for understanding the electrostatic environment around the probe. Further, additional experiments such as FTLS provide a diagnostic tool for quantifying the extent of hydrogen bonding to the nitrile, allowing the interpretation of the vibrational line shape according to the principles of eq 3. The additional information contained in the line width and specific line shape of nitrile vibrational spectra provide a robust, self-consistent experimental data set that enables quantitative interpretation of the electrostatic environment of any protein in the vicinity of a nitrile probe. Finally, we propose a simple model that can be used as a guide for interpreting nitrile spectra. This information can be used to benchmark electrostatic calculations, evaluate the need for additional complexity in force fields such as explicit polarizability, and improve the prediction and understanding of complex electric fields in biological systems.

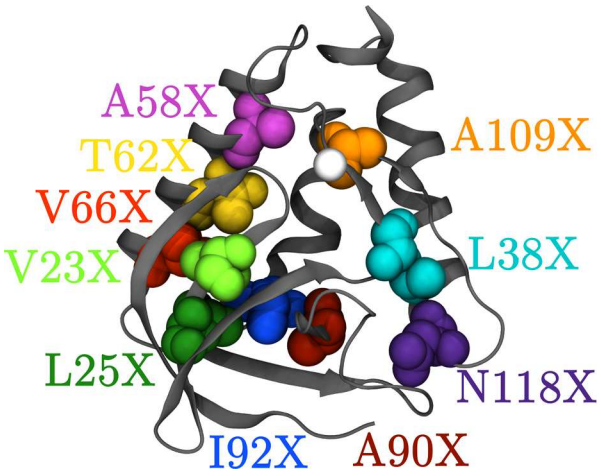


Figure 1. Locations of the nitrile probes in staphylococcal nuclease (SNase). Each location represented by the colored spheres was independently mutated to a cysteine, which was cyanylated through post-translational modification to cyanocysteine (CNC). The first letter indicates the one letter code of the WT residue that was replaced by CNC, denoted by "X". The dark gray ribbon shows the backbone of SNase, and the white sphere shows the location of the native Ca^{2+} ion.

MATERIALS AND METHODS

Mutagenesis, Expression, and Purification. Five pET42a plasmids containing the Δ+PHS variant of SNase and a mutation were generous gifts of Bertrand García-Moreno.[9] Additional mutations were made with the QuikChange Lightning Mutagenesis kit (Agilent). The mutated plasmids were transformed into BL21-DE3 cells (New England BioLabs) and expressed and purified from inclusion bodies based on a strategy described previously.[64] The CNC probe was incorporated into the unfolded protein as previously described.[39] The protein was then refolded by dialysis, diluted to 300 μM with deionized H_2O, and stored at −80 °C for further use. Prior to use, the protein samples were exchanged into 10 mM $CaCl_2$, pH 7.0. Protein refolding was confirmed by CD spectroscopy, and nitrile labeling was confirmed with mass spectrometry.

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

Fourier Transform Infrared Spectroscopy. All FTIR spectra were collected at pH 7.0 on a Bruker Vertex 70 instrument with an InSb detector as previously described.[43,50,52] Spectra were baseline corrected by fitting a fifth-order spline function to the baseline and were normalized to 1.

MD Simulations. A template SNase protein structure bound to Ca^{2+} was modeled by homology from PDB 3bdc using the MODELLER software package.[65,66] Mutations from the WT residue to cysteine were made with PyMol.[67] Using the Avogadro molecular editing package,[68] the mutated cysteine was modified to cyanocysteine (CNC) and independently minimized. Parameters for the CNC residue have been described previously.[35,69] All further minimizations and MD simulations were performed using the Gromacs 2016.3 molecular dynamics simulation package[70−76] and the Amber03 force field.[77,78] Production MD was prepared and run on each equilibrated system for 100 ns, using the same simulation parameters as described previously,[52] and a snapshot was saved every 4 ps.

RESULTS AND DISCUSSION

FTIR Spectra of Nitrile-Containing SNase Mutants. Measuring the pK_a shifts of either wild type (WT) or mutated ionizable residues has long been a strategy for estimating the effects of electrostatic forces in the interior of proteins.[9,10,15,79,80] More recently, vibrational spectroscopy of nitriles placed site-specifically in proteins has been an alternative method for quantifying electrostatic effects in such systems. The information provided by vibrational spectroscopy is fundamentally different than the information provided by a ΔpK_a measurement, since the nitrile measurement can be related to electric field through the Stark equation (eq 3), and the probe perturbs the protein structure significantly less than an ionizable residue. We speculated that nitrile spectroscopic data could be a more direct measurement of electric fields in proteins. In order to supplement the ΔpK_a data already available, we incorporated CNC nitrile probes into 10 separate locations in SNase (illustrated in Figure 1) for which there are already reported values of ΔpK_a.[18,19] These 10 positions were chosen to represent the full range of previously measured pK_a shifts, as well as to sample the spatial diversity of the interior of the protein, such as hydrophobic, hydrophilic, solvent-exposed, and solvent-excluded regions.

FTIR spectra were collected from each of the 10 nitrile-containing SNase variants, and are shown in Figure 2. The average and standard deviation of the peak frequencies for each nitrile location are given in Table 1, and the results, including a discussion of the unique spectral features of each mutant, are given in the Supporting Information. Additionally, we collected the FTIR spectra of methylthiocyanate (MeSCN) dissolved in H_2O and in DMSO, which served as standards for representing a nitrile probe in protic and aprotic environments, respectively. Standard deviations of the nitrile frequency measurements were all generally within 0.1 cm^{-1}. The full width at half maxima (fwhm) of the spectra, which qualitatively describe the environmental heterogeneity experienced by a probe, are also given in Table 1. Collectively, nitrile spectra from these 10 locations within the same protein displayed the diverse range of spectroscopic results that are possible from measuring electric fields in a biological sample.

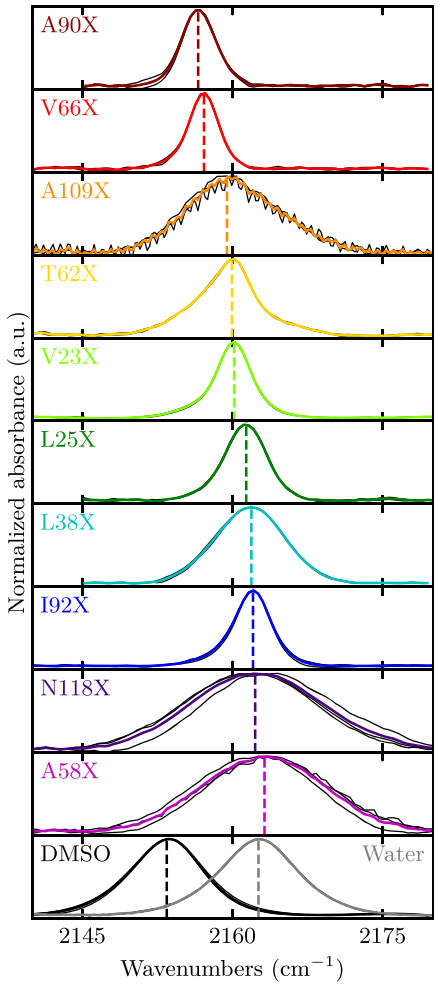


Figure 2. FTIR absorption spectra of CNC incorporated at each of the ten locations shown in Figure 1 at room temperature, arranged from lowest peak frequency to highest peak frequency. The spectra of MeSCN in both DMSO and water are included for comparison. The maximum absorbance of each spectrum was normalized to 1. Colored lines: averaged absorption spectra. Black traces: individual FTIR spectra (some traces lie directly underneath the averaged spectrum). Vertical dashed lines: peak frequencies reported in Table 1.

To our knowledge, this set of nitrile data is unique because there is additional pK_a information available at every nitrile location. While these results are derived from different physical phenomena (nitriles probe a single state, while ΔpK_a probes reflect the difference between two states), because pK_a shifts have often been used to benchmark computational methods, it is useful to compare the ΔpK_a of buried lysine and glutamate residues published by García-Moreno and co-workers against the measured nitrile frequencies at those same locations (Figure 1).[18,19] For convenience, the previously published pK_a values are reproduced in Table 1. Parts A and B of Figure S1 show that there is no agreement between the nitrile and pK_a probes at each nitrile location (r = 0.33 and −0.12 for lysine and glutamate, respectively). Moreover, when we compared the results of the two different pK_a probes, which differed in the sign of the charged state, there was poor agreement (Figure S1C, r = −0.69). This lack of agreement between the three molecular probes highlights that each one provides unique information about the environment at each location; this is discussed further in the Supporting Information. ΔpK_a values

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

Table 1. Nitrile Peak Vibrational Frequency, fwhm, and FTLS of each Nitrile Probe and Corresponding pK_a of Lysine and Glutamate Probes at 10 Locations in the Interior of SNase

	nitrile location			pK_a probe location	
amino acid position	$\tilde{\nu}$[a] (cm^{-1})	fwhm[a] (cm^{-1})	FTLS[b] (cm^{-1} °C^{-1})	lysine pK_a[c]	glutamate pK_a[d]
A90X	2156.58 ± 0.05	4.18 ± 0.23	−0.011 ± 0.000	8.6	6.4
V66X	2157.16 ± 0.01	3.38 ± 0.00	−0.006 ± 0.000	5.6	8.5
A109X	2159.46 ± 0.10	10.05 ± 0.60	−0.018 ± 0.002	9.2	7.9
T62X	2159.96 ± 0.03	5.63 ± 0.23	−0.014 ± 0.001	8.1	7.7
V23X	2160.18 ± 0.00	3.86 ± 0.00	−0.008 ± 0.001	7.3	7.1
L25X	2161.38 ± 0.04	4.98 ± 0.23	−0.008 ± 0.001	6.3	7.5
L38X	2161.89 ± 0.03	8.36 ± 0.23	−0.027 ± 0.002	10.4	6.8
I92X	2162.06 ± 0.03	4.34 ± 0.00	−0.016 ± 0.000	5.3	9.0
N118X	2162.27 ± 0.55	15.27 ± 0.23	−0.053 ± 0.006	10.4	4.5
A58X	2163.20 ± 0.40	14.71 ± 0.34	−0.027 ± 0.001	10.4	7.7
MeSCN in DMSO	2153.44 ± 0.15	8.36 ± 0.11	−0.008 ± 0.004	–	–
MeSCN in water	2162.60 ± 0.03	9.00 ± 0.11	−0.043 ± 0.002	–	–

[a]Errors are reported as the standard deviation of at least three measurements. [b]Errors are reported as the standard error of the fit. [c]Lysine pK_a values adapted from ref 19. [d]Glutamate pK_a values adapted from ref 18.

reflect larger, structural differences between the charged and neutral states of the amino acid side chain and possibly the surrounding protein, while nitriles report on the Coulombic forces present in the native state of the protein. These two data sets therefore complement each other and provide alternative information that can be used to benchmark computational models. For example, pK_a shifts could be used to interrogate the importance of local flexibility in protein electrostatic calculations, while nitrile spectra could be used to test the importance of including explicit polarizability in modern force fields. While significant effort has been invested in interpreting ΔpK_a shifts, a similar effort is required for a reliable interpretation of nitrile spectra according to eq 3.

Molecular Dynamics Simulations. In order to investigate the presence of hydrogen bonds to the 10 nitriles in SNase, we performed 100 ns MD simulations on each of the ten nitrile-containing protein systems and on the WT SNase. First, we monitored the backbone RMSD of the crystallized residues (residues 7−141) for each nitrile-containing mutant construct and compared them to the WT protein. These data are shown in Figure S2. The backbone RMSD values of the nitrile-containing systems were all similar to that of the WT protein, demonstrating that the nitrile perturbation to the local protein environment was small, at least on the time scale of the simulations, and that the simulations sufficiently sampled physically reasonable structures.

Second, we investigated the specific geometries of hydrogen bonding interactions in the MD simulations. The geometries were calculated using an in-house code described previously, based on criteria illustrated in Figure S3.[52] In order to be classified as a hydrogen bond, d_{NH} needed to be less than 2.45 Å, θ_1 needed to be greater than 99°, and θ_2 needed to be greater than 120°. Since θ_1 has been shown to have a large impact on the nitrile frequency,[27,52] we plotted θ_1 against d_{NH} of each interaction as a heat map, considering either only water molecules (Figure 3A) or protein O, N, and C_α atoms with a covalently bound hydrogen (Figure 3B) to be possible hydrogen bond donors. Figure 3A demonstrates that the nitrile locations experienced a wide range of solvent hydrogen bonding situations. The nitriles at positions N118X and A58X had a broad distribution of hydrogen bonding geometries that were donated from water molecules, consistent with partially solvent-exposed residues or interaction with unstructured

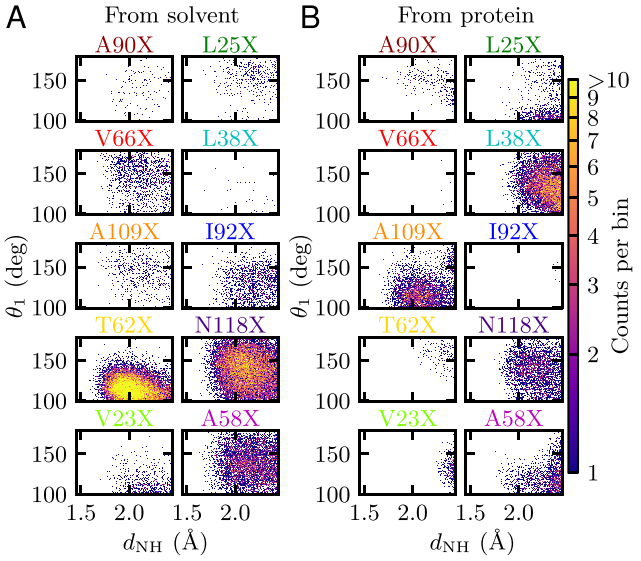


Figure 3. Heat maps of hydrogen bonding geometries to CNC at each probe location. (A) Hydrogen bonds that are donated from water. (B) Hydrogen bonds that are donated from the protein.

water. V66X and I92X experienced far fewer hydrogen bonds to solvent (less than 1500 occurrences each out of 25,000 frames analyzed) that were also broadly distributed around the measured geometries. This indicated a low level of interaction with unstructured water. A90X, A109X, V23X, L25X, and L38X experienced few hydrogen bonds to solvent (less than 600 occurrences each). Finally, T62X experienced a significant amount of hydrogen bonding to solvent (over 17,000 occurrences) with a narrow distribution of θ_1. Upon inspection, this population was found to result from a structural water that was stabilized by hydrogen bonding to the amide N atoms of both Ile18 and Asp19 and to the carboxyl O of Thr22. The water molecule in this interaction was found to be extremely stable and exchanged very infrequently. For example, one instance of a water molecule in this location had a residence time of over 40 ns. To visualize this, we plotted the number of hydrogen bonds that persisted for at least that length of time as a function of simulation time (Figure S4). The residence time of the water molecules in this

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

interaction with T62X was significantly longer than any other hydrogen bonding interaction to the nitrile observed in any of our simulations.

Likewise, the nitriles experienced a range of hydrogen bonding interactions from nearby protein donors (Figure 3B). The nitriles at positions A109X, L38X, and N118X experienced a significant amount of hydrogen bonding (over 3000 observations) from nearby protein donors. The nitrile at positions L25X and A58X experienced a smaller amount of hydrogen bonding from protein (less than 1700 occurrences). The nitriles at positions A90X, V66X, T62X, V23X, and I92X experienced almost no hydrogen bonding interactions from protein (less than 400 occurrences each). It is important to note that in all probe locations, the nitrile was observed to be involved in some amount of hydrogen bonding, either with water, protein, or both. This is likely true for pK_a probes in these same positions; crystal structures of SNase with pK_a probes at these locations show likely hydrogen bonding between the probe and nearby residues or crystallized water.[12] Taken together, these data demonstrate that with both pK_a and nitrile probes, hydrogen bonding interactions cannot be assumed to be absent or even consistent between probe locations. The local interactions at each location must be investigated and understood to interpret experimental results.

Despite the nitriles being buried in the hydrophobic core of SNase, our MD simulations revealed a range of hydrogen bonding interactions, consistent with our observations of a range of fwhm of the spectra in Figure 2. Beyond this, the two nitrile spectra with the largest fwhm, N118X and A58X, experienced broadly distributed hydrogen bonding geometries in our simulations, consistent with bulk-like solvent exposure. Conversely, the nitriles with the narrowest spectra, A90X, V66X, V23X, L25X, and I92X all experienced relatively few hydrogen bonds in our simulations. Finally, nitriles with intermediate spectral widths, A109X, T62X, and L38X, all experienced structural hydrogen bonds. The hydrogen bonds to A109X and L38X were donated by protein, and the vast majority of hydrogen bonding events to T62X were from a structurally rigid water molecule. The apparent correlation of observed hydrogen bonding in simulation and the experimentally measured fwhms suggests that hydrogen bond exchange was a significant source of heterogeneity for the nitrile probes in our system. Furthermore, because these locations were chosen intentionally to be buried in the hydrophobic core of a globular protein, it may be impossible to position a nitrile probe site-specifically in a protein location completely void of hydrogen bonding. This is consistent with our previous work in GFP, in which every nitrile positioned in a wide range of environments experienced hydrogen bonding.[52] Even in the buried hydrophobic core of a protein, such as in this case with SNase, one cannot assume a nitrile probe is free from hydrogen bonding, and in all systems, controls must be performed to account for these interactions. This is not unique to vibrational probes; the inclusion of any molecular probe, including a pK_a probe, requires careful consideration of local interactions that are difficult to predict *a priori*.

To investigate the effect of these hydrogen bonding interactions on our observed vibrational spectra, we calculated the most probable C≡N···H hydrogen bonding angle (θ_1 in Figure S3), in the manner described previously,[52] and we compared this angle to the observed vibrational frequencies. Because some of the nitrile locations in SNase experienced relatively few hydrogen bonding interactions, we weighted the regression based on the number of hydrogen bond observations to each nitrile. The most probable θ_1 was well-correlated to the observed vibrational frequency of the nitrile (Figure S5, r = 0.72). This is consistent with the *ab initio* calculations by Choi et al., where linear, σ-hydrogen bonds (θ_1 closer to 180°) were shown to blue shift the nitrile frequency and nonlinear, π-hydrogen bonds (θ_1 closer to 99°) were shown to red shift the nitrile frequency, and with our previous work in GFP.[27,52] Thus, we have demonstrated that (1) it is difficult if not impossible to position a nitrile in a location free from hydrogen bonding in a folded protein, and (2) any interpretation of nitrile frequencies must take the specific geometry of hydrogen bonding into consideration in order to interpret the observed results. The consistency of these two observations in two structurally diverse proteins (SNase and GFP) with two different molecular nitrile probes indicates that these observations are likely generalizable to other protein systems.

FTLS Is a Quantitative Measure of Hydrogen Bonding Interactions to Nitriles. In order to investigate the hydrogen bonding environments of the different probe locations experimentally, we measured the FTLS of the nitrile probe in each location. Adhikary et al. postulated an empirical relationship between the ability of a nitrile probe to accept a hydrogen bond and the dependence of the nitrile absorption frequency on changes in temperature.[53] This is due to an increase in hydrogen bond exchange and accompanying decrease in hydrogen bond lifetime with increasing temperature. In systems with fewer hydrogen bonds to the nitrile, less exchange is observed and the resulting temperature dependence of the frequency shift, quantified as the FTLS, is smaller. We measured the nitrile absorption frequency and fwhm of each probe at 10 °C intervals from 5 to 35 °C (Table S1). The spectra at each temperature are shown in Figure S6. The Δ+PHS construct of SNase was designed to be exceptionally thermally stable,[9,18,81] and representative circular dichroism (CD) spectra (Figure S7) demonstrate that the fold of the protein did not change upon heating to 35 °C. As expected, the change in vibrational frequency at each location over the temperature range was linear (Figure 4), and the slope of this linear dependence was taken to be the FTLS of each nitrile probe. These values are reported in Table 1. For reference, we also plotted the temperature dependent frequency shift of MeSCN in both H_2O and DMSO in Figure 4, which allowed us to compare the measured FTLS of the nitriles in protein to the amount of hydrogen bonding a nitrile experiences in protic and aprotic environments, respectively. The 10 nitrile-containing SNase constructs demonstrated a large range of FTLS. Of particular interest, the nitrile at position N118X (-0.053 cm^{-1} °C^{-1}) had FTLS that was larger than that of water (-0.043 cm^{-1} °C^{-1}), indicating a larger temperature response by the probe than MeSCN in water. At this position, the probe was located in a flexible portion of the protein, which we hypothesize can sample more environments as temperature increases. This increased flexibility results in a larger temperature dependence than a small molecule nitrile in water.

In order to investigate whether the FTLS measurements accurately served as a measure of hydrogen bonding according to the relationship postulated by Adhikary et al.,[53] we compared the FTLS to the amount of hydrogen bonding to the nitrile observed in our MD simulations. To this end, we first calculated the solvent accessible surface area (SASA) of

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

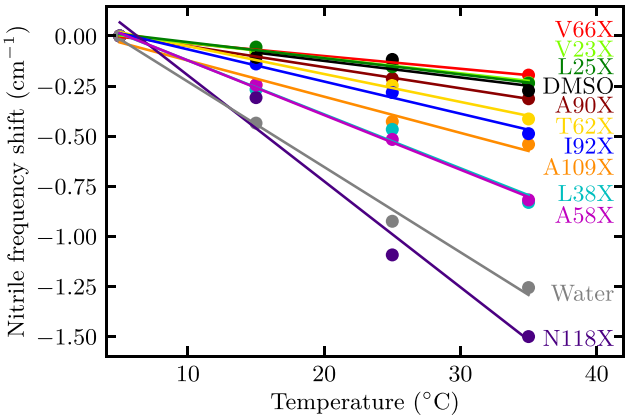


Figure 4. Average of peak nitrile frequencies of the 10 nitrile probes from 5 to 35 °C. The points for each mutant were fit using a linear regression. The slopes and the standard error of the fits are given in Table 1. Mutants are colored according to Figure 1. Peak nitrile frequencies of MeSCN in DMSO and water at the same temperatures are also shown in black and gray, respectively. Error bars were excluded for clarity but are listed in Table S1. No data were collected for MeSCN in DMSO at 5 °C because it is below the freezing point of DMSO.

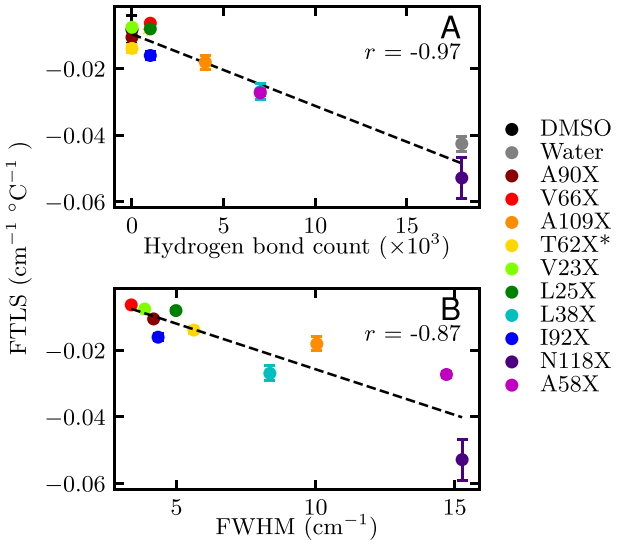


Figure 5. Frequency temperature line slope (FTLS) as an experimental measurement of hydrogen bonding. (A) Measured FTLS compared to the total number of hydrogen bonding observations from simulation. MeSCN in DMSO (black) was set at zero because it is aprotic. MeSCN in water (gray) was extrapolated from a 20 ns MD simulation of an acetyl-CNC-NH_2 peptide in water. The asterisk indicates that the hydrogen bonds donated from water for T62 were not included for this comparison, as discussed in the main text. (B) Measured FTLS compared to the experimental fwhm. MeSCN measured in water and DMSO are excluded because we do not expect a small molecule nitrile in solution to broaden similarly to a nitrile moiety in a protein environment.

the nitrile-containing residue (Figure S8), as this is often used as a first approximation of hydrogen bonding access. While residues such as N118X had larger SASA values, most nitriles had a small calculated SASA with relatively large fluctuations. These fluctuations led to standard deviations that were larger than the differences in the averages, despite the evidence that nitriles experienced a broad range of hydrogen bonding (see Figure 3). We therefore concluded that for the small CNC probe used here in SNase, SASA was not a useful metric to quantify hydrogen bonding environment. Instead, we compared the experimental FTLS to the number of observations of hydrogen bonding in the 100 ns MD simulations (Figure 5A). Because of the long-lived nature of the interaction between the nitrile at position T62X and water (see Figure S4), we reasoned that over the temperature range investigated in the FTLS experiments, these hydrogen bonds likely do not exchange more frequently with increased temperature and therefore do not contribute to the FTLS. We therefore did not include hydrogen bonds from water to residue T62 in the calculated number of hydrogen bonds. Excluding this interaction, we observed a strong linear relationship (r = −0.97) between the experimental FTLS and the number of hydrogen bonds observed in the simulations. This result is consistent with our previous result in GFP, where the calculated SASA of the larger *p*-cyanophenylalanine probe was well correlated to the measured FTLS.[52] Because the FTLS of two different types of nitrile probes are quantitatively related to measurements of hydrogen bonding environment in two structurally different protein systems, this suggests that the FTLS can generally be used to quantitatively measure the hydrogen bonding status of a nitrile probe *in situ*.

The FTIR vibrational frequencies, the fwhm of the nitrile spectra, and the FTLS results offer three orthogonal, experimental, steady-state measurements to assess the local environment of the nitriles. Indeed, when plotted against each other, the FTLS and fwhm are well-correlated (Figure 5B, r = −0.87) demonstrating that these two measurements are both responding to similar influences from chemical heterogeneity around the nitrile. The fwhm of vibrational spectra have been previously used to assess local environment in terms of polarity with a different molecular probe.[82] Our results suggest that while the peak frequency of the nitrile spectra reports on the specific geometry of nearby residues, the FTLS and fwhm of nitrile spectra generally report on a reaction field generated by local interactions that is much more specific than simple macroscopic polarity. Moreover, these techniques use routine FTIR instrumentation and methods that are straightforward to implement in essentially any laboratory. While ΔpK_a measurements give only a single scalar value and no straightforward method to diagnose hydrogen bonding, the nitrile spectra contain information from line shape and temperature dependence beyond that of a reported frequency. This strategy will be a useful generalizable tool for experimentally interrogating hydrogen bonding in proteins. Coupled with MD simulations, the peak frequency, fwhm, and FTLS measurements provide information-rich controls for understanding the contributions of hydrogen bonding to the nitrile spectra.

Electric Fields Contribute to Vibrational Line Shape. In order for the nitrile spectra to be a useful benchmark for electric field calculations, the nitrile spectra must contain direct information about the electric field in the vicinity of the probe. Xu et al. recently demonstrated that the distribution of Coulombic fields is a major contributor to spectral line shape and must be considered in conjunction with the effects of hydrogen bonding.[83] We therefore calculated the electric field at the midpoint of the nitrile in each MD trajectory. Due to the inherent limitations of calculations based on a fixed-charge force field and the impact of the geometry of local hydrogen bonding interactions,[50,59,60,69] we did not expect the simulations to be able to explain the absolute frequencies, but we hypothesized that the accurate modeling of the local

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

interactions could still yield the correct distributions of the fields that could be compared against the spectral line shapes obtained from the steady state FTIR experiments. In order to calculate the field distributions from the MD simulations, a dummy atom with a +1 charge was placed in the center of the nitrile bond; the electrostatic force on the dummy atom was calculated and projected along the nitrile bond vector. The partial charges on the carbon and nitrogen atoms of the nitrile group itself were neutralized because we have previously demonstrated that they result in a consistent and large field that overwhelms the contributions from the surrounding environment.[35,60] As expected, the averages of these fields, which we refer to as the total external fields, were not correlated to the experimental vibrational frequencies (Figure S9A, $r = -0.27$). Even in cases where there were fewer hydrogen bonds in the simulations (A90X, V66X, V23X, L25X, I92X), where we would expect Coulombic contributions to dominate the observed vibrational frequency, the calculated electric fields were still not correlated to the vibrational frequencies.

It has been suggested that continuum electrostatics models, such as the PB equation, may better capture electrostatic contributions from solvent than the direct calculation from the point-charge force field used for MD.[84,85] We repeated all the above calculations using a continuum solvent with the Adaptive PB Solver (APBS) software package.[84] In this approach, the total external field is split into two components: the solvent reaction field (SRF) that is treated with a continuum model and the protein Coulomb field (PCF) that is treated explicitly. The results of these two components were similar to those obtained using the MD force field and are summarized in Figure S10. Given that the PB based approach required significantly more computing time for similar results, we continued our investigation using the equivalent SRF and PCF component fields calculated directly from the MD force field. We reasoned that the peak frequencies of the nitrile spectra were not useful to interrogate the absolute electric fields within the protein systems because (1) it is likely impossible to avoid hydrogen bonds to nitriles even buried within a protein, (2) the frequencies of nitriles involved in hydrogen bonding are dominated by the angle of specific interactions, and (3) the total external fields from both methods were not correlated with the observed frequencies. However, a nitrile spectrum contains additional information beyond the peak frequency in the overall line shape of the spectrum. We therefore focus instead on the *distribution* of fields experienced by each probe for the remainder of this discussion.

To investigate the distribution of the fields experienced by the nitrile in each system, we binned the calculated SRF and PCF for each MD snapshot and fit a polynomial to the histograms of each component (Figure S11, parts B and D, respectively). We found that the distributions of the calculated fields qualitatively mirrored the line shapes observed in the experimental spectra. A brief comparison of each of these distributions to the corresponding experimental spectrum is given in the Supporting Information. Because of the apparent similarities between the histograms of the calculated SRF and PCF for each protein system and the experimentally observed line shape, we calculated the standard deviation of each calculated field over the length of the simulations to describe the distribution of these fields quantitatively. Alone, the standard deviation of either component, the SRF or the PCF, did not account for the experimentally observed fwhm (Figure S13, parts A and B, respectively). This indicated that while both the SRF and the PCF contributed, neither exclusively determined the observed line shape of the nitrile spectra. Further, the standard deviations of the total external field did not account for the experimentally observed fwhm (Figure S13C). We reasoned that this was the result of the cancellation of equal but opposite shifts in the component fields during the simulations for some probes (see, for example, A90X in Figure S11, parts A and C). This resulted in a narrow distribution of the total external field that did not reflect the variance in the environment sampled by the probe. To determine the overall effect of the distributions of both component fields, we combined their standard deviations according to eq 4:

$$\sigma_{\text{combined}} = \sqrt{{\sigma_{\text{SRF}}}^2 + {\sigma_{\text{PCF}}}^2} \tag{4}$$

Here, σ_{SRF} and σ_{PCF} are the individual standard deviations of the SRF and PCF distributions, respectively. The combined standard deviations were strongly correlated to the experimental fwhm from the nitrile spectra (Figure 6, $r = 0.95$), suggesting that the combined effect from both component fields determined the overall nitrile peak width at all 10 positions of SNase investigated here.

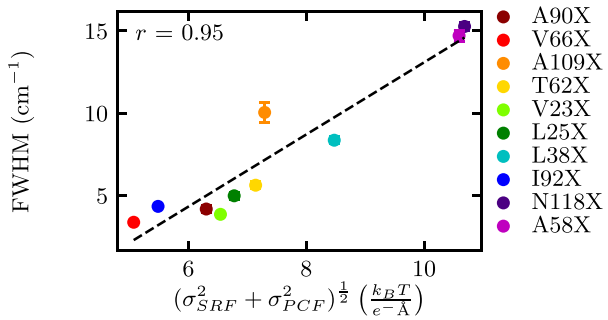


Figure 6. Combined standard deviations of the calculated solvent reaction field (SRF) and protein Coulomb field (PCF) against the experimental fwhm of the observed nitrile vibrational spectra. The combined standard deviation was calculated from the standard deviations of each component of the field through eq 4.

To demonstrate this further, we binned the total external calculated field (PCF+SRF) and show the resulting histograms in Figure 7 (gray), in which the units of field have been converted from k_BT/e^-Å to cm^{-1} and centered underneath the experimentally observed peak frequency. The distributions of the total external field were qualitatively similar to the overall line shapes of the experimental spectra (black dashed lines, reproduced from Figure 2). For example, the histogram of the total external field for T62X contained a minor red-shifted population. Correspondingly, the vibrational spectrum of T62X contained a red-shifted shoulder, although to a lesser degree than indicated by the histogram of the external field. While the range of the widths of the histograms and the range of the pooled standard deviations calculated using eq 4 were smaller than the range of fwhm of the experimental spectra, the histograms do not include broadening due to hydrogen bonding. Xu et al. demonstrated that the SolEFP model can accurately reproduce vibrational line shapes with a classicized linear response function, which accounts for broadening due to the vibrational transition lifetime.[83] Because nitrile vibrational lifetimes are affected by hydrogen bonding interactions,[86] we

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

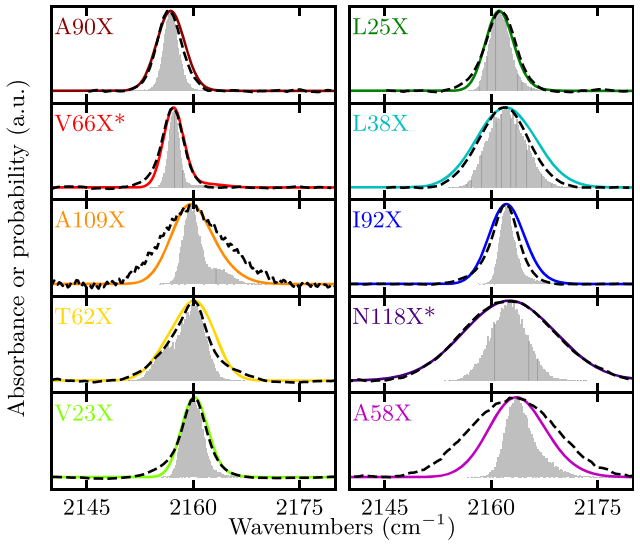


Figure 7. Broadened histograms of the total external field. Gray: calculated histogram of the total external field. Solid colored lines: sum of the broadened bins. Each bin was broadened by a Gaussian with a bandwidth proportional to the experimental FTLS. Asterisks for V66X and N118X indicate that these samples were used to gauge the proportion. Dashed black lines: experimental spectra from Figure 2 overlaid for reference.

reasoned that the additional broadening observed in the experimental spectra was due in part to hydrogen bonding. It is worth noting that, while a single hydrogen bond to a nitrile may actually narrow a vibrational spectrum due to effects such as motional narrowing,[83] a distribution of hydrogen bonding angles present during a steady-state experiment would result in broadening.

Because the FTLS of a nitrile provides quantitative information on the extent of hydrogen bonding, as demonstrated above, it can account for additional broadening not already captured by the field calculations. This is supported by the correlation between the fwhm of the spectral peaks and the experimentally measured FTLS (see Figure 5B). We therefore broadened each bin in the histograms of the total external field with a Gaussian with a bandwidth proportional to the experimentally measured FTLS, using the fwhm of V66X and N118X (the narrowest and widest spectra) to gauge the proportion. The sums of the broadened bins are shown as solid colored lines in Figure 7. For most of the spectra, our broadened distributions qualitatively matched the experimentally observed line shapes. For example, using this analysis we can now accurately account for the shape of the small shoulder in the T62X experimental spectrum, which was overestimated by the histogram of the external field alone. For A109X and A58X, however, the broadened field distribution seemed to underestimate the contribution of the more negative field (higher frequency). Nevertheless, the similarity between the broadened histograms and the experimentally measured vibrational spectra demonstrate that the details of the spectral line shape are important because they reveal the different microstates that the nitrile probe experiences. Even if classical fixed-charge force fields are unable to replicate overall shifts in vibrational frequencies, with the present MD simulations and orthogonal experimental measurements such as FTLS, we can qualitatively identify the populations of structures that give rise to these spectral features. The model proposed here outlines an approach that simulators can use to validate different electrostatic models. For example, by overlaying the simulated field distributions on top of the experimental spectra for T62X, we were able to identify which MD structures were capturing the contributions to the shoulder and were able to relate the specific hydrogen bonding interactions to the nitrile that resulted in the red-shifted frequencies. With refinement, this strategy could be used to isolate specific structures from MD that contribute to each observed spectral frequency. With the confidence that a particular model can reproduce the nitrile vibrational spectra observed for nitrile-containing proteins, the model can then be extended to proteins without the nitrile probe to provide information about the electric fields in WT, native biological structures.

The interpretation of nitrile spectra is complicated by the fact that the experimental nitrile stretching frequency is superimposed on top of a broad water absorption peak. Resolving the nitrile spectral line shape therefore requires a baseline correction that could systematically affect the results.[86] Though we observed agreement between broadened field calculations and experimentally reproducible line shapes, further quantitative investigation will require a more standardized approach to baseline correction and peak fitting. In spite of this, the vibrational line shape, fwhm, and FTLS were all self-consistent and could be quantified by electric field calculations enabled by the MD. In general, the suite of information offered by these experiments will be valuable for benchmarking methods for computing electrostatics moving forward.

In total, this work demonstrates that the overall vibrational frequencies of nitriles are difficult to relate to any one particular contribution. Specific hydrogen bonding interactions, nonspecific local interactions, and electric field all simultaneously contribute to the observed frequency. Indeed, eq 3 reflects this, because only *differences* in electric field may be related to *differences* in vibrational frequency. However, in order to interpret differences in vibrational frequency in this manner, the overall environment of the nitrile must remain constant. For measurements in proteins, this is likely only possible when the nitrile location is consistent and perturbations are made around the nitrile probe.[43,44,87] Here, we have shown that in the case of a single nitrile-containing construct, like any one of the SNase mutants in this study, it is difficult to interpret the relative vibrational frequency compared to the frequencies of nitrile probes in other locations. Nevertheless, the line shape of each nitrile spectrum can be used to identify individual populations for which the nitrile experiences a different projected field. By considering different populations for the same probe, it may be possible to interpret the distribution of frequencies inherent to a vibrational spectrum as differences in electric field according to eq 3. Further, in the event a nitrile experiences multiple populations with different hydrogen bonding interactions, the FTLS of individual components of the overall spectra can be monitored independently and may be used to quantify these different populations. An example of this is shown in Figure S14. The similarities between the broadened histograms of the calculated external field and the spectral line shape in this study (see Figure 7) demonstrate the feasibility of relating the electric field in different subpopulations to the line shape of a nitrile vibrational spectrum. Such close agreement between experiment and calculation demonstrates that only slightly more experimental information (i.e., fwhm and FTLS)

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

obtained with routine FTIR methodologies is required to interpret vibrational spectra quantitatively. This capability is greatly enhanced when experimental data can be compared to computational modeling as closely as possible.

CONCLUSION

Nitrile vibrational spectra are an alternative experimental measurement of electrostatic fields in biological molecules in that they (1) offer electric field information from the spectral line shape to provide information about subpopulations of the probe and (2) can be quantitatively accounted for with additional experiments, such as measuring FTLS. This demonstrates that libraries of nitrile spectra are a valuable resource for benchmarking electrostatic computational models and, with the proper control experiments, for reporting quantitative information about electric fields in proteins. This result, obtained with standard FTIR and MD simulations using widely accessible fixed-charged force fields, is generalizable to any protein of interest that is stable over a moderate temperature range. Further investigation of decoupling local structure from overall electronic contributions to nitrile vibrational spectra will allow for a greater understanding and increased predictability of the complex electric fields that dominate biological systems.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcb.0c00747.

Additional discussions of nitrile spectra, comparison of ΔpK_a and nitrile frequencies, APBS field calculations, comparison of nitrile spectra and field calculations, the FTLS of independent populations within nitrile spectra, additional details from Materials and Methods, a comparison of the three molecular probes of electrostatics at 10 locations in SNase, backbone RMSDs of the crystallized residues of each SNase construct, schematic of the hydrogen bonding geometric criteria, calculation of the observed hydrogen bond residence times, the most probable angle of hydrogen bonding from simulation plotted against the experimental peak vibrational frequency, FTIR absorption spectra of each SNase construct at each temperature, representative CD spectra, SASA calculations of the CNC residue, a comparison of the calculated fields and the experimental vibrational frequencies, a comparison of calculated fields using two different calculation methods, an analysis of the calculated components of the electric field, the dihedral angles of the CNC probe at each location, a comparison of the distribution of component fields against the experimental fwhm, an analysis of two Gaussians fit to spectra collected at various temperatures, and a table listing peak vibrational frequencies and fwhms of nitrile probes at each location at a range of temperatures (PDF)

AUTHOR INFORMATION

Corresponding Author

Lauren J. Webb − *Department of Chemistry, Texas Materials Institute, and Institute for Cell and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712-1224, United States;* orcid.org/0000-0001-9999-5500; Phone: 512-471-9361; Email: lwebb@cm.utexas.edu

Authors

Jeremy T. First − *Department of Chemistry, Texas Materials Institute, and Institute for Cell and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712-1224, United States;* orcid.org/0000-0001-8829-280X

Elisa T. Novelli − *Department of Chemistry, Texas Materials Institute, and Institute for Cell and Molecular Biology, The University of Texas at Austin, Austin, Texas 78712-1224, United States;* orcid.org/0000-0002-4046-1229

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpcb.0c00747

Author Contributions

‡J.T.F. and E.T.N. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the NSF (MCB-1714555) and The Welch Foundation (F-1722). We gratefully acknowledge the Texas Advanced Computing Center at the University of Texas at Austin for providing high performance computing resources that have contributed to the results reported within this paper. We acknowledge the Center for Biomedical Research Support Proteomics Core Facility and the Targeted Therapeutic Drug and Diagnostic Development Program, both at the University of Texas at Austin, for providing instrumentation. We would also like to thank Bertrand García-Moreno for providing the SNase DNA plasmids.

REFERENCES

(1) Honig, B.; Nicholls, A. Classical Electrostatics in Biology and Chemistry. *Science* **1995**, *268*, 1144.

(2) Warshel, A. Energetics of Enzyme Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75*, 5250.

(3) Warshel, A.; Russell, S. T. Calculations of Electrostatic Interactions in Biological Systems and in Solutions. *Q. Rev. Biophys.* **1984**, *17*, 283−422.

(4) Lee, F. S.; Warshel, A. A Local Reaction Field Method for Fast Evaluation of Long-range Electrostatic Interactions in Molecular Simulations. *J. Chem. Phys.* **1992**, *97*, 3100−3107.

(5) Norberg, J.; Nilsson, L. On the Truncation of Long-Range Electrostatic Interactions in DNA. *Biophys. J.* **2000**, *79*, 1537−1553.

(6) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Vattulainen, I. Lipid Bilayers Driven to a Wrong Lane in Molecular Dynamics Simulations by Subtle Changes in Long-Range Electrostatic Interactions. *J. Phys. Chem. B* **2004**, *108*, 4485−4494.

(7) Biava, H.; Schreiber, T.; Katz, S.; Völler, J.-S.; Stolarski, M.; Schulz, C.; Michael, N.; Budisa, N.; Kozuch, J.; Utesch, T.; et al. Long-Range Modulations of Electric Fields in Proteins. *J. Phys. Chem. B* **2018**, *122*, 8330−8342.

(8) Zhou, H.-X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2018**, *118*, 1691−1741.

(9) García-Moreno, B. E.; Dwyer, J. J.; Gittis, A. G.; Lattman, E. E.; Spencer, D. S.; Stites, W. E. Experimental Measurement of the Effective Dielectric in the Hydrophobic Core of a Protein. *Biophys. Chem.* **1997**, *64*, 211−224.

(10) Schutz, C. N.; Warshel, A. What are the Dielectric "Constants" of Proteins and How to Validate Electrostatic Models? *Proteins: Struct., Funct., Genet.* **2001**, *44*, 400−417.

(11) Sakipov, S. N.; Flores-Canales, J. C.; Kurnikova, M. G. A Hierarchical Approach to Predict Conformation-Dependent Histidine Protonation States in Stable and Flexible Proteins. *J. Phys. Chem. B* **2019**, *123*, 5024−5034.
(12) Robinson, A. C.; Schlessman, J. L.; García-Moreno E, B. Dielectric Properties of a Protein Probed by Reversal of a Buried Ion Pair. *J. Phys. Chem. B* **2018**, *122*, 2516−2524.
(13) Hwang, J.-K.; Warshel, A. Why Ion Pair Reversal by Protein Engineering is Unlikely to Succeed. *Nature* **1988**, *334*, 270−272.
(14) Tan, Y.-J.; Oliveberg, M.; Davis, B.; Fersht, A. R. Perturbed pKA-values in the Denatured States of Proteins. *J. Mol. Biol.* **1995**, *254*, 980−992.
(15) Nielsen, J. E.; Gunner, M. R.; García-Moreno E., B. The pKa Cooperative: A Collaborative Effort to Advance Structure-based Calculations of pKa Values and Electrostatic Effects in Proteins. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 3249−3259.
(16) Isom, D. G.; Cannon, B. R.; Castañeda, C. A.; Robinson, A.; García-Moreno E., B. High Tolerance for Ionizable Residues in the Hydrophobic Interior of Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17784−17788.
(17) Harms, M. J.; Castañeda, C. A.; Schlessman, J. L.; Sue, G. R.; Isom, D. G.; Cannon, B. R.; García-Moreno E., B. The pKa Values of Acidic and Basic Residues Buried at the Same Internal Location in a Protein Are Governed by Different Factors. *J. Mol. Biol.* **2009**, *389*, 34−47.
(18) Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; Velu, P. D.; García-Moreno E., B. Charges in the Hydrophobic Interior of Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16096−16100.
(19) Isom, D. G.; Castañeda, C. A.; Cannon, B. R.; García-Moreno E., B. Large Shifts in pKa Values of Lysine Residues Buried Inside a Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 5260−5265.
(20) Alexov, E.; Mehler, E. L.; Baker, N.; M. Baptista, A.; Huang, Y.; Milletti, F.; Nielsen, J. E.; Farrell, D.; Carstensen, T.; Olsson, M. H.; et al. Progress in the Prediction of pKa Values in Proteins. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 3260−3275.
(21) Arthur, E. J.; Yesselman, J. D.; Brooks, C. L. Predicting Extreme pKa Shifts in Staphylococcal Nuclease Mutants with Constant pH Molecular Dynamics. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 3276−3286.
(22) Georgescu, R. E.; Alexov, E. G.; Gunner, M. R. Combining Conformational Flexibility and Continuum Electrostatics for Calculating pKas in Proteins. *Biophys. J.* **2002**, *83*, 1731−1748.
(23) Gunner, M. R.; Baker, N. A. Continuum Electrostatics Approaches to Calculating pKas and Ems in Proteins. *Methods Enzymol.* **2016**, *578*, 1−20.
(24) Lee, M. S.; Salsbury, F. R., Jr; Brooks, C. L. Constant-pH Molecular Dynamics using Continuous Titration Coordinates. *Proteins: Struct., Funct., Genet.* **2004**, *56*, 738−752.
(25) Mongan, J.; Case, D. A.; McCammon, J. A. Constant pH Molecular Dynamics in Generalized Born Implicit Solvent. *J. Comput. Chem.* **2004**, *25*, 2038−2048.
(26) Swails, J. M.; York, D. M.; Roitberg, A. E. Constant pH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.* **2014**, *10*, 1341−1352.
(27) 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*, 134506.
(28) Fried, S. D.; Boxer, S. G. Measuring Electric Fields and Noncovalent Interactions Using the Vibrational Stark Effect. *Acc. Chem. Res.* **2015**, *48*, 998−1006.
(29) Getahun, Z.; Huang, C.-Y.; Wang, T.; De León, B.; DeGrado, W. F.; Gai, F. Using Nitrile-Derivatized Amino Acids as Infrared Probes of Local Environment. *J. Am. Chem. Soc.* **2003**, *125*, 405−411.
(30) Johnson, M. N. R.; Londergan, C. H.; Charkoudian, L. K. Probing the Phosphopantetheine Arm Conformations of Acyl Carrier Proteins Using Vibrational Spectroscopy. *J. Am. Chem. Soc.* **2014**, *136*, 11240−11243.
(31) McMahon, H. A.; Alfieri, K. N.; Clark, K. A. A.; Londergan, C. H. Cyanylated Cysteine: A Covalently Attached Vibrational Probe of Protein-Lipid Contacts. *J. Phys. Chem. Lett.* **2010**, *1*, 850−855.
(32) Oh, K.-I.; Choi, J.-H.; Lee, J.-H.; Han, J.-B.; Lee, H.; Cho, M. Nitrile and Thiocyanate IR probes: Molecular Dynamics Simulation Studies. *J. Chem. Phys.* **2008**, *128*, 154504.
(33) Park, E. S.; Thomas, M. R.; Boxer, S. G. Vibrational Stark Spectroscopy of NO Bound to Heme: Effects of Protein Electrostatic Fields on the NO Stretch Frequency. *J. Am. Chem. Soc.* **2000**, *122*, 12297−12303.
(34) Slocum, J. D.; Webb, L. J. Measuring Electric Fields in Biological Matter Using the Vibrational Stark Effect of Nitrile Probes. *Annu. Rev. Phys. Chem.* **2018**, *69*, 253−271.
(35) Stafford, A. J.; Ensign, D. L.; Webb, L. J. Vibrational Stark Effect Spectroscopy at the Interface of Ras and Rap1A Bound to the Ras Binding Domain of RalGDS Reveals an Electrostatic Mechanism for Protein-Protein Interaction. *J. Phys. Chem. B* **2010**, *114*, 15331−15344.
(36) Suydam, I. T.; Boxer, S. G. Vibrational Stark Effects Calibrate the Sensitivity of Vibrational Probes for Electric Fields in Proteins. *Biochemistry* **2003**, *42*, 12050−12055.
(37) Waegele, M. M.; Tucker, M. J.; Gai, F. 5-Cyanotryptophan as an Infrared Probe of Local Hydration Status of Proteins. *Chem. Phys. Lett.* **2009**, *478*, 249−253.
(38) Walker, D. M.; Hayes, E. C.; Webb, L. J. Vibrational Stark Effect Spectroscopy Reveals Complementary Electrostatic Fields Created by Protein-Protein Binding at the Interface of Ras and Ral. *Phys. Chem. Chem. Phys.* **2013**, *15*, 12241−12252.
(39) 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*, 13356−13357.
(40) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. Biosynthesis of Proteins Incorporating a Versatile Set of Phenylalanine Analogues. *ChemBioChem* **2002**, *3*, 235−237.
(41) Webb, L. J.; Boxer, S. G. Electrostatic Fields Near the Active Site of Human Aldose Reductase: 1. New Inhibitors and Vibrational Stark Effect Measurements. *Biochemistry* **2008**, *47*, 1588−1598.
(42) Shrestha, R.; Cardenas, A. E.; Elber, R.; Webb, L. J. Measurement of the Membrane Dipole Electric Field in DMPC Vesicles Using Vibrational Shifts of p-Cyanophenylalanine and Molecular Dynamics Simulations. *J. Phys. Chem. B* **2015**, *119*, 2869−2876.
(43) Slocum, J. D.; Webb, L. J. Nitrile Probes of Electric Field Agree with Independently Measured Fields in Green Fluorescent Protein Even in the Presence of Hydrogen Bonding. *J. Am. Chem. Soc.* **2016**, *138*, 6561−6570.
(44) Stafford, A. J.; Walker, D. M.; Webb, L. J. Electrostatic Effects of Mutations of Ras Glutamine 61 Measured Using Vibrational Spectroscopy of a Thiocyanate Probe. *Biochemistry* **2012**, *51*, 2757−2767.
(45) Walker, D. M.; Wang, R.; Webb, L. J. Conserved Electrostatic Fields at the Ras-Effector Interface Measured through Vibrational Stark Effect Spectroscopy Explain the Difference in Tilt Angle in the Ras Binding Domains of Raf and RalGDS. *Phys. Chem. Chem. Phys.* **2014**, *16*, 20047−20060.
(46) Fafarman, A. T.; Boxer, S. G. Nitrile Bonds as Infrared Probes of Electrostatics in Ribonuclease S. *J. Phys. Chem. B* **2010**, *114*, 13536−13544.
(47) Shields, J. M.; Pruitt, K.; McFall, A.; Shaub, A.; Der, C. J. Understanding Ras: 'It Ain't Over 'Til it's Over'. *Trends Cell Biol.* **2000**, *10*, 147−154.
(48) Andrews, S. S.; Boxer, S. G. Vibrational Stark Effects of Nitriles I. Methods and Experimental Results. *J. Phys. Chem. A* **2000**, *104*, 11853−11863.
(49) 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*, 469−477.

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

(50) Slocum, J. D.; First, J. T.; Webb, L. J. Orthogonal Electric Field Measurements near the Green Fluorescent Protein Fluorophore through Stark Effect Spectroscopy and pKa Shifts Provide a Unique Benchmark for Electrostatics Models. *J. Phys. Chem. B* **2017**, *121*, 6799−6812.
(51) 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*, 15250−15265.
(52) 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*, 6733−6743.
(53) Adhikary, R.; Zimmermann, J.; Dawson, P. E.; Romesberg, F. E. Temperature Dependence of CN and SCN IR Absorptions Facilitates Their Interpretation and Use as Probes of Proteins. *Anal. Chem.* **2015**, *87*, 11561−11567.
(54) Huang, C.-Y.; Wang, T.; Gai, F. Temperature Dependence of the CN Stretching Vibration of a Nitrile-derivatized Phenylalanine in Water. *Chem. Phys. Lett.* **2003**, *371*, 731−738.
(55) Harms, M. J.; Schlessman, J. L.; Sue, G. R.; García-Moreno E., B. Arginine Residues at Internal Positions in a Protein are Always Charged. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 18954−18959.
(56) Karp, D. A.; Stahley, M. R.; García-Moreno E., B. Conformational Consequences of Ionization of Lys, Asp, and Glu Buried at Position 66 in Staphylococcal Nuclease. *Biochemistry* **2010**, *49*, 4138−4146.
(57) Peck, M. T.; Ortega, G.; De Luca-Johnson, J. N.; Schlessman, J. L.; Robinson, A. C.; García-Moreno E, B. Local Backbone Flexibility as a Determinant of the Apparent pKa Values of Buried Ionizable Groups in Proteins. *Biochemistry* **2017**, *56*, 5338.
(58) Dippel, A. B.; Olenginski, G. M.; Maurici, N.; Liskov, M. T.; Brewer, S. H.; Phillips-Piro, C. M. Probing the Effectiveness of Spectroscopic Reporter Unnatural Amino Acids: a Structural Study. *Acta. Crystallog. D* **2016**, *72*, 121−130.
(59) Ritchie, A. W.; Webb, L. J. Optimizing Electrostatic Field Calculations with the Adaptive Poisson−Boltzmann Solver to Predict Electric Fields at Protein−Protein Interfaces. I. Sampling and Focusing. *J. Phys. Chem. B* **2013**, *117*, 11473−11489.
(60) Ritchie, A. W.; Webb, L. J. Optimizing Electrostatic Field Calculations with the Adaptive Poisson−Boltzmann Solver to Predict Electric Fields at Protein−Protein Interfaces II: Explicit Near-Probe and Hydrogen-Bonding Water Molecules. *J. Phys. Chem. B* **2014**, *118*, 7692−7702.
(61) Ritchie, A. W.; Webb, L. J. Understanding and Manipulating Electrostatic Fields at the Protein−Protein Interface Using Vibrational Spectroscopy and Continuum Electrostatics Calculations. *J. Phys. Chem. B* **2015**, *119*, 13945−13957.
(62) Błasiak, B.; Lee, H.; Cho, M. Vibrational Solvatochromism: Towards Systematic Approach to Modeling Solvation Phenomena. *J. Chem. Phys.* **2013**, *139*, 044111.
(63) 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*, 18094−18111.
(64) Shortle, D.; Meeker, A. K. Residual Structure in Large Fragments of Staphylococcal Nuclease: Effects of Amino Acid Substitutions. *Biochemistry* **1989**, *28*, 936−944.
(65) Castañeda, C. A.; Fitch, C. A.; Majumdar, A.; Khangulov, V.; Schlessman, J. L.; García-Moreno, B. E. Molecular Determinants of the pKa Values of Asp and Glu Residues in Staphylococcal Nuclease. *Proteins: Struct., Funct., Genet.* **2009**, *77*, 570−588.
(66) Martí-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sánchez, R.; Melo, F.; Šali, A. Comparative Protein Structure Modeling of Genes and Genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291−325.
(67) DeLano, W. L. *The PyMOL Molecular Graphics System*; www.pymol.org.
(68) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: an Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf.* **2012**, *4*, 17.
(69) Ensign, D. L.; Webb, L. J. Factors Determining Electrostatic Fields in Molecular Dynamics Simulations of the Ras/Effector Interface. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 3511−3524.
(70) 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.
(71) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-passing Parallel Molecular Dynamics Implementation. *Comput. Phys. Commun.* **1995**, *91*, 43−56.
(72) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(73) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: a Package for Molecular Simulation and Trajectory Analysis. *J. Mol. Model.* **2001**, *7*, 306−317.
(74) Páll, S.; Abraham, M. J.; Kutzner, C.; Hess, B.; Lindahl, E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. *Solving Software Challenges for Exascale* **2015**, *8759*, 3−27.
(75) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; et al. GROMACS 4.5: A High-throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845−854.
(76) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, Flexible, and Free. *J. Comput. Chem.* **2005**, *26*, 1701−1718.
(77) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A Point-charge Force Field for Molecular Mechanics Simulations of Proteins based on Condensed-phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(78) Sorin, E. J.; Pande, V. S. Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. *Biophys. J.* **2005**, *88*, 2472−2493.
(79) Warshel, A.; Dryga, A. Simulating Electrostatic Energies in Proteins: Perspectives and Some Recent sSudies of pKas, Redox, and Other Crucial Functional Properties. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 3469−3484.
(80) Warshel, A.; Papazyan, A. Electrostatic Effects in Macromolecules: Fundamental Concepts and Practical Modeling. *Curr. Opin. Struct. Biol.* **1998**, *8*, 211−217.
(81) Byrne, M. P.; Manuel, R. L.; Lowe, L. G.; Stites, W. E. Energetic Contribution of Side Chain Hydrogen Bonding to the Stability of Staphylococcal Nuclease. *Biochemistry* **1995**, *34*, 13949−13960.
(82) Manor, J.; Feldblum, E. S.; Zanni, M. T.; Arkin, I. T. Environment Polarity in Proteins Mapped Noninvasively by FTIR Spectroscopy. *J. Phys. Chem. Lett.* **2012**, *3*, 939−944.
(83) Xu, R. J.; Błasiak, B.; Cho, M.; Layfield, J. P.; Londergan, C. H. A Direct, Quantitative Connection between Molecular Dynamics Simulations and Vibrational Probe Line Shapes. *J. Phys. Chem. Lett.* **2018**, *9*, 2560−2567.
(84) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037.
(85) Wagoner, J.; Baker, N. A. Solvation Forces on Biomolecular Structures: A Comparison of Explicit Solvent and Poisson−Boltzmann Models. *J. Comput. Chem.* **2004**, *25*, 1623−1629.
(86) van Wilderen, L. J. G. W.; Kern-Michler, D.; Müller-Werkmeister, H. M.; Bredenbeck, J. Vibrational Dynamics and Solvatochromism of the Label SCN in Various Solvents and Hemoglobin by Time Dependent IR and 2D-IR Spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 19643−19653.

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399

(87) Novelli, E. T.; First, J. T.; Webb, L. J. Quantitative Measurement of Intrinsic GTP Hydrolysis for Carcinogenic Glutamine 61 Mutants in H-Ras. *Biochemistry* **2018**, *57*, 6356−6366.

https://dx.doi.org/10.1021/acs.jpcb.0c00747
J. Phys. Chem. B 2020, 124, 3387−3399