Published in final edited form as:

J Phys Chem B. 2023 June 29; 127(25): 5609–5619. doi:10.1021/acs.jpcb.3c02060.

The AMOEBA Force Field Predicts Accurate Hydrogen Bonds Counts of Nitriles in SNase by Revealing Water-Protein Interaction in Vibrational Absorption Frequencies

Yu-Chun Lin[a], **Pengyu Ren**[b], **Lauren J. Webb**[a,*]

[a]Department of Chemistry, Texas Materials Institute, and Interdisciplinary Life Sciences Program, The University of Texas at Austin, 105 E 24th St. STOP A5300, Austin, TX, 78712, USA

[b]Department of Biomedical Engineering, The University of Texas at Austin, Austin, TX, 78712, USA

Abstract

Precisely quantifying the magnitude and direction of electric fields in proteins has long been an outstanding challenge in understanding biological functions. Nitrile vibrational Stark effect probes have been shown to be minimally disruptive to the protein structure and can be better direct reporters of local electrostatic field in the native state of a protein than other measures such as pK_a shifts of titratable residues. However, interpretations of the connection between measured vibrational energy and electric field relies on the accurate molecular understanding of interactions of the nitrile group and its environment, particularly from hydrogen bonding. In this work, we compared the extent of hydrogen bonding calculated in two common force fields, the fixed charge force field Amber03 and polarizable force field AMOEBA, at 10 locations of cyanocysteine in staphylococcal nuclease (SNase) against the experimental nitrile absorption frequency in terms of full width half maximum (FWHM) and frequency temperature line slope (FTLS). We observed that the number of hydrogen bonds correlated well in AMOEBA trajectories with respect to both the FWHM (r = 0.88) and the FTLS (r = −0.85) whereas the correlation of Amber03 trajectories were less reliable because the Amber03 force field predicted more hydrogen bonds in some mutants. Moreover, we demonstrated that contributions from the interactions between cyanocysteine and nearby water molecules were significant in AMOEBA trajectories but were not predicted by Amber03. We conclude that although the nitrile absorption peak shape could be qualitatively predicted by the fixed charge Amber03 force field, the detailed electrostatic environment measured by the nitrile probe in terms of the extent of hydrogen bonding could only be accurately observed in the AMOEBA trajectories, where the permanent dipole, quadrupole, and dipole-induced-dipole polarizable interactions were all taken into account. The significance of this finding to the goal of accurately predicting electric fields in complex biomolecular environments is discussed.

*To whom correspondence should be addressed: lwebb@cm.utexas.edu.

Supporting Information:

FTIR spectra and FTLS of the 10 nitrile locations;[52] distances between the native Ca^{2+} and the nitrile in Amber03 and AMOEBA; RMSD values from Amber03[52] and AMOEBA trajectories; illustration of hydrogen bonding geometry and the χ_1 and χ_2 angles; Hydrogen bond counts with respect to time in Amber03 and AMOEBA trajectories; χ_1 and χ_2 angles in Amber03[52] and AMOEBA trajectories; parameters derivation for AMOEBA force field; considering water within 2.5 Å of the nitrile.

Introduction

Electric fields generated by the atomic charges from the amino acids, ions, and other molecules within a protein have long been recognized as a fundamental driving force for all biomolecular functions including folding, catalysis, ligand binding, protein-protein interactions, and other properties.[1–6] Quantifying either the magnitude or the direction of electric fields in proteins is a complex task because these fields are produced by a large number of partial charges interacting within and around a protein over a range of distances.[7–9] Previous work has demonstrated that the consideration of the electrostatic interactions from the residues that are sometimes far from the enzyme active site is necessary in order to replicate experimentally observed effects including reduction potential[10] and enzyme kinetic efficiency.[11–13] Taken together, both experimental and computational research have long demonstrated that an experimentally measurable indicator for protein electric field that could be supported by physical computational model is complicated but necessary. Taken together, both experimental and computational research have long demonstrated that an experimentally measurable indicator of integrated protein electric field that could be supported by physical computational model is complicated but necessary.

Although there are a number of experimental methods for measuring electric fields in proteins, including NMR spectroscopy[14–16] and vibrational Stark effect spectroscopy,[17,18] by far the most common method has been measuring pK_a shifts (ΔpK_a) of titratable residues and interpreting those shifts in terms of local electric field around the protons.[19–24] ΔpK_a values of targeted residues in a protein can be accurately measured from ^1H and ^{13}C chemical shifts by NMR or by absorption or emission energy changes of an internal spectroscopic probe.[14,15,17,19,25] These measurements reflect the tendency of a residue to stay in either its charged or neutral state determined by the electrostatic environment generated from the positions of partial charges of all atoms, ions, and water molecules in that protein.[23,24,26–28] Nevertheless, because the ΔpK_a value only represents the difference in free energy between two charged states of the target residue, it is difficult to interpret any information about the electrostatic environment of a single, native state of that protein from the measured ΔpK_a value. Therefore, the development of a more direct measurement of electric field for only one state of a protein would be useful to validate and improve computational models.

Our laboratory, along with other research groups, has explored the utility of using vibrational Stark effect (VSE) spectroscopy of nitrile probes to measure the local electric fields[29,30] in a variety of biologically relevant environments including lipid membranes, protein interiors, protein-protein interfaces, and protein active sites.[31–44] The nitrile vibrational frequency is a useful probe of different biological environments because of its small size, the ease with which it can be incorporated into biomolecules with minimal disruption of structure, and an absorption frequency located in a relatively empty region of the protein's vibrational background.[32,45–47] The relationship of the vibrational energy of the nitrile group to local electric field is given through the linear Stark equation (Equation 1):[48,49]

$$\Delta E = -\Delta\vec{\mu} \bullet \Delta\vec{F} = hc\Delta v \quad (1)$$

where ΔE is the change in absorption energy of the nitrile group with respect to the change in local electric field, $\Delta\vec{F}$, caused by differences in charge distributions near the probe upon making a chemical perturbation to the system, and $\Delta\vec{\mu}$ is the difference in the dipole moment between the ground and excited vibrational states. For the nitrile, $\Delta\vec{\mu}$ has been measured and verified in previous studies,[42,50,51] which has enabled changes in the vibrational frequency in Fourier transform infrared (FTIR) spectroscopy experiments, $\Delta\nu$, to be used in Equation 1 to determine $\Delta\vec{F}$. Our previous work on superfolder green fluorescent protein (GFP)[17] and staphylococcal nuclease (SNase)[52] that investigated nitriles at different locations throughout the protein have showed that VSE shifts respond to differences in the electric fields around the labeled nitrile. This work, along with other research that examined the nitriles in a wide range of biological environments,[38,48,53] all validate the measurement of local electric fields in a protein through VSE spectroscopy.

In previous work, we investigated the vibrational spectra of the nitrile at 10 different locations of a model protein SNase (Figure S1).[52] SNase has been used to understand the extreme pK_a values of the targeted residues by using constant pH molecular dynamics (MD) simulations.[54] To provide an alternative experimental benchmark as the verification of the computational models based on the SNase, our group previously investigated the nitrile vibrational frequency shifts in SNase and suggested that the pK_a shifts did not correlate with the nitrile vibrational frequency shifts, and the nitrile probe was a more direct measurement of electrostatic field because it can interrogate a single, native state of the protein, rather than measuring the free energy differences between two states (protonated and deprotonated).[52] This work, along with other experimental and computational investigations[29] has demonstrated that hydrogen bonding to the nitrile group can dominate the observed frequency of the nitrile probe in a way that is not captured in Equation 1. This is due to the intrinsic quantum mechanical factors of the nitrile's molecular orbitals: (1) a parallel σ-hydrogen bond with the nitrile lone pares that causes the nitrile frequency to shift to higher energy by withdrawing electrons from an antibonding orbital, and (2) a perpendicular π-hydrogen bond with the π cloud of the triple bond that leads to a shift of the nitrile frequency to lower energy by withdrawing electron density from a bonding orbital.[29] Therefore, to understand the measured vibrational frequency that could be contributed by the hydrogen bonds formed with the nitrile, the same work in our group[52] attempted to quantify the effect of hydrogen bonding on the nitrile absorption energy by measuring the frequency temperature line slope (FTLS, Figure S2),[52,55] which reflects the amount of hydrogen bonding that is perturbed with respect to temperature change.[52] In the same work, the MD simulations using a conventional, fixed charge, Amber03 force field[56–61] were conducted on all of the 10 positions of the nitrile to provide molecular details of the SNase structure, and quantified the extent and geometry of hydrogen bonds at each nitrile.[52] We hypothesized that some of the weaknesses of this approach that were observed

could be addressed with a force field that more accurately predicts electrostatic environment, not just structure.

In the work reported here, we generated the SNase structures of the 10 nitrile locations (Figure 1) in a polarizable force field, atomic multipole optimized energetics for biomolecular applications (AMOEBA), which includes many-body polarizations through atomic dipole induction that were not considered in our previous work.[52] The nitrile group was inserted into SNase as cyanocysteine (CNC), shown in Figure 2, at the 10 locations in Figure 1 to generate 10 constructs containing a single nitrile probe. The potential energy functions of the AMOEBA force field are described in Equation 2:[62]

$$U = U_{bond} + U_{angle} + U_{b\theta} + U_{oop} + U_{tor} + U_{vdw} + U_{ele}^{perm} + U_{ele}^{ind} \quad (2)$$

where the total potential energy of the system is a sum of potential energies from bond vibration (U_{bond}), angle bending (U_{angle}), vibrational and bending coupling ($U_{b\theta}$), out-of-plane bending (U_{oop}), dihedral torsions (U_{tor}), van der Waals interactions (U_{vdw}), permanent electrostatic interactions (U_{ele}^{perm}), and many-body induction or polarization(U_{ele}^{ind}).[63] Common classical force fields such as Amber03 use fixed charges on every atom of the system without considering the polarization effect (U_{ele}^{ind})[19,64,65] for computational speed. Although MD simulations using these fixed charge force fields are known to predict the secondary structures of peptides and proteins with a high degree of accuracy,[66,67] it has been found that they fail to generate structures that accurately predict the detailed electrostatic properties including pK_a values of buried residues,[19,24] and in fact can predict the opposite trends of protein-ligand binding free energies compared to the experimental results.[68]

In the work reported here, we show that the number of hydrogen bonds that were able to form with the nitrile at positions V23X, L25X, L38X, A58X, T62X, V66X, A90X, I92X, A109X, and N118X (where X stands for cyanocysteine, CNC) were less than the number predicted by Amber03 trajectories compared to the linear correlation against experimental FWHM and FTLS in most locations. We hypothesize this is because Amber03 over-stabilizes interactions of hydrogen bonds, and thus results in an unphysically large number of interactions with water.[69–71] Moreover, the AMOEBA trajectories better predicted the amount of hydrogen bonds for all ten mutants at a structurally diverse collection of nitrile locations in SNase. Both AMOEBA and Amber03 predicted similar results for the long range interactions to the native Ca^{2+} ion in SNase (Figure S3) and the overall nitrile absorption peak shape, but the correlations to number of hydrogen bonds and water molecules within 2.5 Å of the nitrile with respect to the experimental FTLS and FWHM were all higher when predicted by AMOEBA, even though both force fields started with the same initial configurations. We suggest that using the MD simulation with the fixed charge Amber03 force field is sufficient for the predictions in secondary structure of protein and the qualitative calculation of the nitrile absorption frequency shape; however, the polarizable AMOEBA force field is able to more accurately quantify total hydrogen bonding interactions with the nitrile group and the amount of water molecules around the nitrile

because of its more detailed consideration of electrostatic interactions including permanent dipole, quadrupole, and dipole-induced-dipole interactions.

Methods

Molecular Dynamics Simulations

The parameters for the AMOEBA polarizable force field for the artificial cyanocysteine (Figure 2) were derived from the AMOEBA parameter generator program Poltype2.[72,73] Psi4 was used for the QM calculations in Poltype2.[74] The initial SNase structures of all mutants were taken from our previous work, which describes how each structure was prepared.[52,75] We conducted MD simulations of V23X, L25X, L38X, A58X, T62X, V66X, A90X, I92X, A109X, and N118X, using the AMOEBA force field implemented in Tinker 9.[76] Packmol[77] was used to build a cubic 88 $Å^3$ simulation box containing 7 Cl^- to neutralize the charge of the SNase and water molecules.[63] After energy minimization of the system, MD simulations started at constant volume heating from 0 to 298 K for 6.6 ns with a time step of 2 fs, followed by constant pressure MD at 298 K for 2 ns at a time step of 2 fs. The production NVT MD simulations were done with an average constant volume for 60 ns in 2 fs steps with frames recorded every 4 ps. From the energy minimization process to every phase of the MD simulations, the permanent charge, dipole, quadrupole, and many-body polarization were all considered (Eq. 2). For all mutants, we calculated the root mean square deviation (RMSD) of protein backbone atoms between each MD structure and the initial energy minimized protein structure to determine when the AMOEBA structures became stable. These results are shown in Figure S4; we observed that the RMSD values throughout this production phase MD simulation were stable. We took the trajectories, including the water molecules, the native Ca^{2+} ion, and every atom in the protein, that were collected every 200 ps during the 60 ns MD simulations, to calculate the distance between the native Ca^{2+} ion and the nitrile group, the number of hydrogen bonds between the nitrile group and the nearby residues, and the number of water molecules within 2.5 Å of cyanocysteine.

The corresponding RMSD values from Amber03 trajectories for the protein backbone atoms for each captured structure and the initial energy minimized protein are shown in Figure S4.[52] Generation of the parameters for cyanocysteine was described in previous publications.[41,78] SNase structures were collected from MD simulations using the Amber03 force field[79,80] in the Gromacs 2016.3 software package and explicit solvation with TIP3P water.[56,61] We observed that the RMSD values remained low (<2.8 Å) throughout the last 60 ns. As previously observed, the AMOEBA polarizable force field gave rise to larger fluctuations (measured by RMSD) during MD simulations of proteins than fixed charge force fields.[24,81] We used trajectories from 40 ns to 100 ns, collected every 200 ps, that were generated from previous work[52] to further calculate the distance between the native Ca^{2+} ion and the nitrile group (Figure S3), the number of hydrogen bonds between the nitrile group and the nearby residues, and the number of water molecules within 2.5 Å of the nitrile.

Quantifying Hydrogen Bonding

For the last 60 ns of both AMOEBA and Amber03 trajectories, we used structures collected every 200 ps to calculate the number of hydrogen bonds between the nitrile group and

nearby molecules, including water and protein atoms. The number of hydrogen bonds were calculated based on the same criteria as the in-house code described in previous work.[17] Each hydrogen bond was composed of (1) a hydrogen bond acceptor, which was defined by the position of the nitrogen of the nitrile group on cyanocysteine, and (2) a hydrogen bond donor, which could be the hydrogen atom from the water molecule, or the hydrogen atom which was covalently bonded to the O, N, and C_α atoms on protein residues. Besides the criteria of the composition of atoms, there were three geometry criteria for the atoms to be counted as a hydrogen bond. First, the distance between the hydrogen bond acceptor and donor, d_{NH}, needed to be less than 2.45 Å. Second, the angle between the C≡N···H atoms, θ_1, was greater than 99°. Third, the angle between the N···H-R_{donor}, θ_2, was greater than 120°. These criteria are illustrated in Figure S5. These criteria were adapted from previous work[82] that reported each of these distances and angles of hydrogen bonding with nitrile groups obtained by *ab initio* molecular orbital calculations in 83 crystal structures that contains 95 nitrile hydrogen-bond interactions.[82] The cutoff distance and angles in this work were set to fall within 2 standard deviations of a typical hydrogen bond to a nitrile defined in the previous study.[82]

Calculating Experimental FTIR Absorption Peak Shape

For the last 60 ns of AMOEBA trajectories, we collected the MD snapshots every 200 ps to calculate the electric field at the center of the cyanide bond. To determine electric field, we inserted a dummy atom with +1 charge at the middle of the nitrile bond of the collected MD snapshots and calculated the electric field on that point charge. The electric field was first obtained by dividing the induced dipole (in Debye) of the dummy atom by its polarizability factor (0.0001 $Å^3$) then projecting the electric field vector of the dummy atom on to the C≡N bond vector. The unit of the electric field was then converted from D $Å^{-3}$ to MV cm^{-1} by multiplying 299.79 (rescaled speed of light in cgs units) as has been done previously,[83] then converted to cm^{-1} by using the $\Delta\vec{\mu}$ as 0.7 cm^{-1}/(MV/cm)[45] in equation 1, to obtain the unit of the infrared spectroscopy data. The most populated electric field distributions were placed at the maximum absorbance energy from the measured FTIR absorption spectra. The electric field distribution was then fit with Gaussian functions to compare with the experimental FTIR spectra.[52] The electric field of the dummy atom (+1 charge) at the center of C≡N bond along the C≡N bond vector in each of the collected snapshots of the last 60 ns of Amber03 trajectories were obtained in previous work[52] using the GROMACS force field calculation.[52,56,61]

Results and Discussion

Quantification of Hydrogen Bonding in Amber03 and AMOEBA

To quantify the sources of hydrogen bonds across the 10 nitrile groups in both Amber03 and AMOEBA force fields, we classified the hydrogen bonds from frames collected every 200 ps for the last 60 ns in both force fields trajectories. The hydrogen bond donor atom from water molecules and protein atoms are shown in Figure 3. The number of hydrogen bonds with respect to the simulation time for each mutant in the Amber03 and the AMOEBA force fields are shown in Figure S6. The total number of hydrogen bonds were less in the AMOEBA trajectories (Figure 3, red) than in the Amber03 trajectories (Figure 3, gray)

for most mutants (L25X, T62X, V66X, I92X, A109X, and N118X). This was expected because one of the most challenging obstacles for the development of AMBER force fields has been addressing overstabilization of secondary structures maintained by hydrogen bond interactions such as α-helices.[84,85] Because L25X is located within a α-helical component of SNase, it is not surprising that Amber03 predicted more hydrogen bonds than AMOEBA. Moreover, the fixed charge force fields like Amber03 also tend to erroneously over-stabilize protein-ion binding free energies by overestimating the effect of electrostatic interactions[68] and predicting lower solvation free energies of the charged state molecule.[24] Therefore, the reason that Amber03 predicted more hydrogen bonds than AMOEBA is likely because the fixed charged Amber03 structures over-stabilize the interactions of the partial negative and partial positive charge between the hydrogen bond acceptor and donor.[24,68]

The most significant exception to this pattern was A58X, where AMOEBA predicted almost twice as many hydrogen bonds per frame than Amber03 (0.56 ± 0.55 vs. 0.35 ± 0.53, respectively). This could be due to the differences in the χ_2 angles (shown in Figures 2 and S6.) predicted by the two MD strategies. The torsional factors in the AMOEBA force field trajectories, which were coupled with the more advanced electrostatic parameters,[69,86,87] could lead to the χ_2 angle in cyanocysteine to be different from the Amber03 trajectories. The smaller fluctuations in the χ_2 angle in Amber03 trajectories caused the nitrile group of A58X to remain close to the α-helical portion of SNase rather than rotate to be more exposed to solvent. The greater variability in the χ_2 angle in the AMOEBA structures therefore allowed the nitrile to form hydrogen bonds with more water molecules. L38X and A90X were also exceptions to this pattern, with nitriles at both sites exhibiting more hydrogen bonding in AMOEBA versus Amber03, although the differences are smaller, and in both cases both force fields predicted that most hydrogen bonds were to elements of the protein itself. Both locations are in portions of the SNase structure that are dominated by β-barrel features, and it is possible that AMOEBA's more accurate prediction of global energy minima were better able to stabilize this non α-helical structure, and increase overall nitrile-protein interactions.[88]

From Figure 3, hydrogen bond donors were more often from nearby water molecules in both sets of trajectories (V23X, A58X, T62X, V66X, I92X, A109X, and N118X) than from the nearby protein residues except for nitriles that were buried in the protein and surrounded by stable secondary structures (L25X and A90X). The observation that major sources of hydrogen bonds were from water molecules rather than the protein residues was not surprising because water molecules are small and have significantly more rotational degrees of freedom to optimize hydrogen bond geometry with the nitrile. However, there were some mutants (T62X, V66X, and I92X) where the sources of hydrogen bonds were almost exclusively water molecules and predicted many more of such interactions with Amber03. For example, in T62X, Amber03 predicted ~0.6 water hydrogen bonds more than with AMOEBA. Structural snapshots showed that a single water molecule resided around the nitrile probe more than 30 ns in the last 60 ns in the Amber03 trajectory,[52] but was not observed in the AMOEBA trajectory although both trajectories started with the same initial structures. It is likely that the over-stabilization of the interaction between water molecules and the nitrile group predicted the extremely long lifetime of the hydrogen bond in the Amber03 trajectory, as has been observed before between high charge density ions

and protein residues in calculations of binding energies between Ca^{2+} and Mg^{2+} among EF-hand proteins.[68] The same work demonstrated that the incorrect predictions of the Amber force field were due to the Coulomb interactions in the MD simulations over-stabilizing interactions with higher charge density in Mg^{2+} than in Ca^{2+}.[68] Moreover, the same work also demonstrated that the accurate predictions of greater binding by Ca^{2+} over Mg^{2+} could only be achieved by considering the dipole-induced-dipole polarization using AMOEBA trajectories.[68] Therefore, the long lifetime water molecule that caused a significantly larger amount of hydrogen bonding for T62X in the Amber03 trajectory could possibly be due to the over-stabilization of the hydrogen bonding interactions from the partial positive charge on water hydrogen and the partial negative charge on the nitrile group. The differences for V66X and I92X, where Amber03 predicted more hydrogen bonds from water molecules than AMOEBA, were likely from the fact that these two residues are located close to the surface of SNase, and so this over-stabilization of interactions with water would be pronounced.

Hydrogen Bonding Interactions Compared to Nitrile FTLS and FWHM

Although VSE spectroscopy with nitriles has many advantages, discussed above, the potential convolution of both electrostatic and hydrogen bonding factors on vibrational absorption energy shifts complicates its use. Because of this, there has been significant interest in finding orthogonal experiments to deconvolute these effects without the need for high quality structural information from MD simulations or experiment for every system. FTLS has been used in several systems to measure the amount of hydrogen bonding by changing the extent of molecular reorganization as a function of temperature, which should not effect steady-state local electric field around the nitrile.[55] Likewise, measuring FWHM and other features of the absorption energy peak shape has been used to understand conformational sampling of the nitrile within its local environment and identify populations of structures at steady-state that may be affected by hydrogen bonding to greater or lesser extent.[52] These data then provide a means of validating MD simulation strategies for the purpose of predicting both electric fields and hydrogen bonding in any system of interest. We compared measured SNase nitrile FTLS and FWHM results with the number of hydrogen bonds, with criteria defined by Le Questel et al,[82] from our MD simulations, discussed in Figure 3, with both Amber03 and AMOEBA force fields against the experimentally measured FTLS and FWHM. Figure 4 shows the correlation of the extent of hydrogen bonding with respect to FTLS in Amber03 and AMOEBA trajectories, which resulted in $r = -0.72$ and $r = -0.85$, respectively. As the number of hydrogen bonds for every mutant across two force fields all range from 0–2, the slightly higher correlation between the hydrogen bond counts in AMOEBA versus the FTLS indicates that AMOEBA trajectories are better able to predict the subtle but critical differences in the number of hydrogen bonds compared to the Amber03. The slightly higher correlation between the number of hydrogen bonds calculated in AMOEBA with FTLS than the correlation between the amount of hydrogen bonds in Amber03 with FTLS were mostly the result of differences in hydrogen bonding predicted in A58X, N118X, and especially T62X. As already discussed, the number of hydrogen bonds in T62X were noticeably greater in Amber03 compared to AMOEBA (0.68 ± 0.47 vs. 0.05 ± 0.22) even after considering the large standard deviation. The extremely long-lived hydrogen bond donated to T62X by a single water molecule in the

Amber03 trajectory (Figure 3) caused this residue to be a significant outlier when compared to experimental FTLS data (Figure 4a) that was not replicated in AMOEBA. For A58X, Amber03 trajectories predicted fewer hydrogen bonds (0.35 ± 0.53) than with AMOEBA (0.56 ± 0.55) which could be due to the observation that the nitrile χ_2 angle over the course of the Amber03 trajectory (Figure S7) consistently directed the nitrile group away from water. In the case of N118X, Amber03 predicted more hydrogen bonds than AMOEBA (0.79 ± 0.51 vs. 0.56 ± 0.54), possibly because χ_2 angles found in the Amber03 trajectories (Figure S7) allowed the nitrile to interact more often with nearby protein residues.

There were a number of nitrile positions that Amber03 and AMOEBA predicted similar hydrogen bonding environments, but Amber03 usually predicted significantly larger standard deviations of these values than AMOEBA, especially when the number of hydrogen bonds per frame was low (e.g. L25X: 0.06 ± 0.24 vs. 0.00 ± 0.06; A90X: 0.02 ± 0.28 vs. 0.10 ± 0.31; I92X: 0.05 ± 0.28 vs. 0.00 ± 0.06; A109X: 0.02 ± 0.15 vs. 0.01 ± 0.08 for Amber03 vs. AMOEBA, respectively). These mutants also all had FTLS values close to zero (L25X: −0.008 ± 0.001; A90X: −0.011 ± 0.000; I92X: −0.016 ± 0.000; A109X: −0.018 ± 0.002).[52] Because the error bar of the experimental measurement of FTLS were all less than 0.002 for these mutants, the larger standard deviation in the Amber03 trajectories, which in turn predicted the larger differences in hydrogen bonding interactions of these mutants, caused the fixed charge Amber03 model to fail to distinguish the subtle differences in the protein environment around each nitrile. On the other hand, for AMOEBA trajectories, the standard deviation values tended to be low, such as L25X: 0.00 ± 0.06, I92X: 0.00 ± 0.06, or A109X: 0.01 ± 0.08, which suggested that AMOEBA trajectories were able to predict subtle differences in a variety of protein environments, especially in cases where there were few stable hydrogen bonds. The capability of the AMOEBA force field for predicting low standard deviation of the mutants with few hydrogen bond counts indicates that both the accuracy and the precision of the prediction from AMOEBA were more reliable than the Amber03 model. The reason that the standard deviation values of Amber03 trajectories were high especially when a low total number of hydrogen bonds were predicted was because the stability of the local electrostatic conformation were mostly maintained by hydrogen bonds in the fixed charge force field. In other words, the fewer hydrogen bonds formed with the nitrile, the larger the degree of structural flexibility was predicted. On the other hand, when simulated with AMOEBA, the electrostatic interactions between every atom included charge-charge, dipole-dipole, and dipole-induced-dipole interactions which could stabilize the nitrile electrostatic environment with fewer or even no hydrogen bonds. Therefore, the complex multipole interactions in AMOEBA trajectories could predict more consistent local electrostatic environments even when hydrogen bonding was scarce, leading to low standard deviations.[88]

The comparison between the hydrogen bond counts and experimental FWHM revealed a similar but more pronounced trend (Figure 5); the correlation between simulation and experiment was significantly higher in AMOEBA (r = 0.88) than in Amber03 (r = 0.59). Two nitrile locations, T62X and A58X, were major outliers in this comparison in Amber03 (Figure 5a) but not in AMOEBA (Figure 5b), again because of the more realistic prediction of the hydrogen bonding environment around the nitrile at those two locations. In both Amber03 and the AMOEBA trajectories, A109X was a significant outlier because both

force fields predicted very few hydrogen bonds (0.02 ± 0.15 vs. 0.01 ± 0.08 for Amber03 and AMOEBA, respectively). This was possibly due to the fact that the χ_2 angles in both trajectories were pinned at exactly either ±180° for almost the entire simulation (Figure S7), therefore reducing the chance for A109X to form hydrogen bonds in both trajectories.

The reason that AMOEBA trajectories predicted higher correlations in both FTLS (Figure 4) and FWHM (Figure 5) experimental data in terms of the total number of hydrogen bonds than Amber03 trajectories could be because detailed multipole interactions were required to accurately predict the change in electrostatic environment in a range of different mutants, and this in turn allowed prediction of a more accurate picture of hydrogen bonding interactions. Previous research has compared the QM-calculated IR spectra of CN and SCN group in water with the CN and SCN group as for the systems with larger water clusters simulated by classical MD simulation. This has suggested that the hydrogen bonding interactions between solvent and the nitrile groups are critical to the prediction in IR peak shape.[29] Moreover, the same work further demonstrated that accurate calculations in IR peak width of the nitrile group were obtained by applying an appropriate QM model capable of correctly predicting that the lowest energy configuration would be obtained by a hydrogen bond angle of 180° between nitrile nitrogen and hydrogen atoms.[29] Adhikary et al. postulated that the FTLS, in which the nitrile absorption frequency changes with temperature, is the measurement of the ability of a nitrile probe to accept a hydrogen bond, which is then broken by increasing temperature.[55] In our previous work measuring the FWHM of the 10 nitrile probes and the FTLS of these probes from 5 to 35 °C, the correlation between the FTLS and FWHM was high ($r = -0.87$).[52] Therefore, more accurate prediction of the hydrogen bonding environment by AMOEBA from considering more detailed multipole interactions correlated more closely with the experimental FWHM and FTLS measurements. Finally, we performed the same analysis by instead counting the number of water molecules within 2.5 Å of the nitrile and observed the same trends. This is described in the Supplemental Information and Figures S13–S15.

Experimental FTIR peak shape calculation

The distribution of the experimental FTIR absorption energies of each nitrile position represents the range of electrostatic environments experienced by the probe over the course of a steady-state experiment. To investigate how the electric field of the nitrile probe itself contribute to the FTIR peak shape, we calculated electric field at the 10 nitrile locations by placing a point positive charge at the center of the nitrile bond and calculating the electric field on that dummy atom from the snapshots collected in last 60 ns of the trajectories. We then plotted the distribution of these calculated fields from both Amber03 (Figure 6a, taken from previous work)[52] and AMOEBA trajectories (Figure 6b) centered at the FTIR maximum absorbance wavenumber for each mutant. To better account for the FTIR peak broadening possibly due to the hydrogen bonding interactions which could not be fully considered in the dummy atom electric field calculation,[52] we broadened each histogram bin in the electric field distribution in a Gaussian function with a bandwidth proportional to the FTLS of each mutant, pinned by the FTLS of the narrowest (V66X) and the widest (N118X) spectra in the colored solid line in Figure 6. For most mutants, the calculated Gaussian distributions (colored solid line) were similar to the experimental peak shapes (dashed line)

in both Amber03 and AMOEBA trajectories, indicating that both force fields were able to adequately predict a wide range of electrostatic environments given a comprehensive set of snapshots. While it is not surprising that the nitrile probe locations for which both force fields predicted similar hydrogen bonding environment (V23X and A109X, Figure 3) would give similar predicted absorption spectra, as discussed above, most mutants had very different results when hydrogen bonds were quantified around the nitrile. The most extreme examples were A58X and T62X, which had substantially different average numbers of hydrogen bonding around the nitrile (A58X: 0.35 ± 0.53, 0.56 ± 0.55; T62X: 0.68 ± 0.47, 0.05 ± 0.22, for Amber03 and AMOEBA, respectively), yet the peak shapes in the two trajectories in Figure 6 were still very similar. The number of hydrogen bonds predicted by the two force fields trajectories alone therefore was not the major factor that contributed to the predicted peak shapes.

The clearest exception to this observation was AMOEBA's predicted absorption spectrum for N118X. In this case, the calculated peak (shown in gray in Figure 6b) was substantially shifted to lower energy than the experimentally observed spectrum. The FTLS of N118X was by far the steepest slope observed of any of the 10 nitrile positions examined here, steeper even than that of a cyanocysteine dissolved in water (Figure S2). Because hydrogen bonding information in the calculation of field is included through scaling by the magnitude of the FTLS, it is possible that this extreme hydrogen bonding environment exposes the limits of our strategy of using an empirical force field, not quantum mechanics, to describe the full electrostatic environment around the nitrile, as would be expected for cases that fall outside of the bounds established by a hydrophobic (DMSO) and hydrophilic (water) environment. It also reinforces the benefit of having an orthogonal experimental method for quantifying the extent of hydrogen bonding to the nitrile to aid interpretation of vibrational absorption energy shifts alone.

The similarity in the peak shapes predicted for most probe locations from both the fixed charge Amber03 and the polarizable AMOEBA trajectories indicates that the predicted line shapes over a long period of time in a steady-state experiment were dominated by Coulombic interactions. Xu et al. demonstrated that for nitrile buried inside a protein with very low exposure to water, the vibrational line shape was dominated by Coulombic interactions over several Ångstrom, particularly as simulation time was extended.[89] This was because for buried probes without much access to solvent, these long-range Coulombic interactions with charges on protein residues could persist for longer times than if the probe had significant exposure to solvent, thus dominating vibrational peak shapes for these locations.[89] Previous works investigated the peak positions of alanine[90] and alanine-phenylalanine based small peptide[91] for 1000–2000 cm^{-1} and 3300–3600 cm^{-1} region suggested that the AMOEBA MD trajectories accompanied with normal mode analysis could reproduce the accurate peak positions compared to the experiment.[90,91] In our case, for the 10 locations of the nitrile in SNase examined here, while most of the mutants were similarly buried within the protein and not interacting with water molecules, those that were more solvent exposed still exhibited this trend. This demonstrates that for the nitrile probe, qualitatively estimating the electric field by examining FTIR peak shape does not necessarily require the greater computation time and complexity of the polarizable force field.

Conclusion

In this work, we compared the experimental FWHM and FTLS with respect to the number of hydrogen bonds and amount of nearby water molecules in both a fixed charge force field Amber03 and a polarizable force field AMOEBA, and found that AMOEBA quantitatively predicted these experimental parameters better than Amber03, almost certainly because of the better snapshot of the local electrostatic environment from polarizability parameters that considered higher moment electrostatic interactions and the polarizabilities. Although the two trajectories started with the same initial configurations of each mutant in SNase, due to the considerations of dipole, quadrupole, and dipole-induced-dipole moment, AMOEBA avoided over-predicting hydrogen bonding interactions and could better distinguish and differentiate a variety of electrostatic environments at any given position of the nitrile in SNase. However, when comparing the calculated peak shapes of the nitrile with these two computational strategies, it appears that this qualitative measurement was in most cases accurately predicted by both models. Moreover, through this work, we demonstrated that the FTLS in particular is an ideal orthogonal experiment for identifying highly perturbed hydrogen bonding environments that might complicate simple interpretation of the nitrile absorption energy purely in terms of an electrochromic Stark effect. This is the first demonstration of the utility of using AMOEBA to understand electrostatic fields measured through vibrational spectroscopy within a protein. We suggest that the protocol established in this work, especially the computational approach of using the AMOEBA polarizable force field in MD simulations to quantify FTLS, FWHM, and other aspects of experimental spectra, could be used to investigate other vibrational modes. Specifically, it would be interesting to investigate the local secondary structure of a protein by comparing the computational spectra calculated from an AMOEBA MD simulation with an isotopically labelled ^{18}O amide I experimental vibrational spectra at the targeted part of a protein. Ongoing work in our laboratory is focused on applying the nitrile probe to investigate complex types of electrostatic interactions such as protein-protein interfaces and within the structure of the lipid bilayer membrane to advance our understanding of these crowded and dynamics electrostatic environments.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the NSF (Grant No. MCB-1714555) and The Welch Foundation (Grant No. F-1722) for support. PR is grateful for support by NIH (Grant No. GM114237).

We gratefully acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high performance computing resources that have contributed to the results reported within this paper.

References

(1). Hayes DM; Kollman PA Role of Electrostatics in a Possible Catalytic Mechanism for Carboxypeptidase A. J. Am. Chem. Soc 1976, 98, 7811–7816. [PubMed: 993499]

(2). Warshel A Energetics of Enzyme Catalysis. Proc. Natl. Acad. Sci. U. S. A 1978, 75, 5250–5254. [PubMed: 281676]

(3). Honig B; Nicholls A Classical Electrostatics in Biology and Chemistry. Science. 1995, 268, 1144–1149. [PubMed: 7761829]

(4). Gunner MR; Nicholls A; Honig B Electrostatic Potentials in Rhodopseudomonas Viridis Reaction Centers: Implications for the Driving Force and Directionality of Electron Transfer. J. Phys. Chem 1996, 100, 4277–4291.

(5). Warshel A; Papazyan A Electrostatic Effects in Macromolecules: Fundamental Concepts and Practical Modeling. Curr. Opin. Struct. Biol 1998, 8, 211–217. [PubMed: 9631295]

(6). Romei MG; Lin C; Mathews II; Boxer SG Electrostatic Control of Photoisomerization Pathways in Proteins. Science. 2020, 367, 76–79. [PubMed: 31896714]

(7). Fersht GSAR Rapid, Electrostatically Assisted Association of Proteins. J. Dairy Res 1979, 3, 427–431.

(8). Li L; Alper J; Alexov E Cytoplasmic Dynein Binding, Run Length, and Velocity Are Guided by Long-Range Electrostatic Interactions. Sci. Rep 2016, 6, 1–12. [PubMed: 28442746]

(9). Fadrná E; Hladečková K; Koča J Long-Range Electrostatic Interactions in Molecular Dynamics: An Endothelin-1 Case Study. J. Biomol. Struct. Dyn 2005, 23, 151–162. [PubMed: 16060689]

(10). Hosseinzadeh P; Marshall NM; Chacón KN; Yu Y; Nilges MJ; New SY; Tashkov SA; Blackburn NJ; Lu Y Design of a Single Protein That Spans the Entire 2-V Range of Physiological Redox Potentials. Proc. Natl. Acad. Sci. U. S. A 2016, 113, 262–267. [PubMed: 26631748]

(11). Savile CK; Janey JM; Mundorff EC; Moore JC; Tam S; Jarvis WR; Colbeck JC; Krebber A; Fleitz FJ; Brands J, et al. Biocatalytic Asymmetric Synthesis of Sitagliptin Manufacture. Science. 2010, 329, 305–310. [PubMed: 20558668]

(12). Jiménez-Osés G; Osuna S; Gao X; Sawaya MR; Gilson L; Collier SJ; Huisman GW; Yeates TO; Tang Y; Houk KN The Role of Distant Mutations and Allosteric Regulation on LovD Active Site Dynamics. Nat. Chem. Biol 2014, 10, 431–436. [PubMed: 24727900]

(13). Osuna S; Jiménez-Osés G; Noey EL; Houk KN Molecular Dynamics Explorations of Active Site Structure in Designed and Evolved Enzymes. Acc. Chem. Res 2015, 48, 1080–1089. [PubMed: 25738880]

(14). Markley JL Observation of Histidine Residues in Proteins by Means of Nuclear Magnetic Resonance Spectroscopy. Acc. Chem. Res 1975, 8, 70–80.

(15). Bradbury JH; Scheraga HA Structural Studies of Ribonuclease. XXIV. The Application of Nuclear Magnetic Resonance Spectroscopy to Distinguish between the Histidine Residues of Ribonuclease. J. Am. Chem. Soc 1966, 88, 4240–4246.

(16). Yu B; Pletka CC; Montgomery Pettitt B; Iwahara J De Novo Determination of Near-Surface Electrostatic Potentials by NMR. Proc. Natl. Acad. Sci. U. S. A 2021, 118, 1–7.

(17). First JT; Slocum JD; Webb LJ Quantifying the Effects of Hydrogen Bonding on Nitrile Frequencies in GFP: Beyond Solvent Exposure. J. Phys. Chem. B 2018, 122, 6733–6743. [PubMed: 29874077]

(18). Błasiak B; Ritchie AW; Webb LJ; Cho M Vibrational Solvatochromism of Nitrile Infrared Probes: Beyond the Vibrational Stark Dipole Approach. Phys. Chem. Chem. Phys 2016, 18, 18094–18111. [PubMed: 27326899]

(19). Slocum JD; First JT; Webb LJ Orthogonal Electric Field Measurements near the Green Fluorescent Protein Fluorophore through Stark Effect Spectroscopy and pK_a Shifts Provide a Unique Benchmark for Electrostatics Models. J. Phys. Chem. B 2017, 121, 6799–6812. [PubMed: 28650636]

(20). Isom DG; Castañed CA; Cannon BR; García-Moreno BE Large Shifts in pK_a Values of Lysine Residues Buried inside a Protein. Proc. Natl. Acad. Sci. U. S. A 2011, 108, 5260–5265. [PubMed: 21389271]

(21). Langsetmo K; Fuchs JA; Woodward C The Conserved, Buried Aspartic Acid in Oxidized Escherichia Coli Thioredoxin Has a PKa of 7.5. Its Titration Produces a Related Shift in Global Stability. Biochemistry 1991, 30, 7603–7609. [PubMed: 1854757]

(22). Antosiewicz J; Mccammon JA; Gilson MK The Determinants of pK_as in Proteins. Biochemistry 1996, 2960, 7819–7833.

(23). Nielsen JE; Gunner MR; García-Moreno E, The pKa Cooperative B: A Collaborative Effort to Advance Structure-Based Calculations of pK_a Values and Electrostatic Effects in Proteins. Proteins Struct. Funct. Bioinforma 2011, 79, 3249–3259.

(24). Lin YC; Ren P; Webb LJ AMOEBA Force Field Trajectories Improve Predictions of Accurate pK_aValues of the GFP Fluorophore: The Importance of Polarizability and Water Interactions. J. Phys. Chem. B 2022, 126, 7806–7817. [PubMed: 36194474]

(25). Antosiewicz J; Mccammon JA; Gilson MK The Determinants of pK_as in Proteins. Biochemistry 1996, 35, 7819–7833. [PubMed: 8672483]

(26). Forsyth WR; Antosiewicz JM; Robertson AD Empirical Relationships between Protein Structure and Carboxyl pK_a Values in Proteins. Proteins Struct. Funct. Genet 2002, 48, 388–403. [PubMed: 12112705]

(27). Gibas CJ; Subramaniam S Explicit Solvent Models in Protein pK_a Calculations. Biophys. J 1996, 71, 138–147. [PubMed: 8804597]

(28). Grimsley GR; Scholtz JM; Pace CN A Summary of the Measured pK_a Values of the Ionizable Groups in Folded Proteins. Protein Sci. 2009, 18, 247–251. [PubMed: 19177368]

(29). Choi JH; Oh KI; 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.

(30). Oh KI; Choi JH; Lee JH; Han JB; Lee H; Cho M Nitrile and Thiocyanate IR Probes: Molecular Dynamics Simulation Studies. J. Chem. Phys 2008, 128.

(31). Fried SD; Boxer SG Measuring Electric Fields and Noncovalent Interactions Using the Vibrational Stark Effect. Acc. Chem. Res 2015, 48, 998–1006. [PubMed: 25799082]

(32). Getahun Z; Huang CY; Wang T; De León B; DeGrado WF; Gai F Using Nitrile-Derivatized Amino Acids as Infrared Probes of Local Environment. J. Am. Chem. Soc 2003, 125, 405–411. [PubMed: 12517152]

(33). Shrestha R; Cardenas AE; Elber R; Webb LJ 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. [PubMed: 25602635]

(34). Slocum JD; Webb LJ 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. [PubMed: 27128688]

(35). Stafford AJ; Walker DM; Webb LJ Electrostatic Effects of Mutations of Ras Glutamine 61 Measured Using Vibrational Spectroscopy of a Thiocyanate Probe. Biochemistry 2012, 51, 2757–2767. [PubMed: 22385209]

(36). Walker DM; Wang R; Webb LJ 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. [PubMed: 25127074]

(37). Johnson MNR; Londergan CH; Charkoudian LK Probing the Phosphopantetheine Arm Conformations of Acyl Carrier Proteins Using Vibrational Spectroscopy. J. Am. Chem. Soc 2014, 136, 11240–11243. [PubMed: 25080832]

(38). McMahon HA; Alfieri KN; Clark KAA; Londergan CH Cyanylated Cysteine: A Covalently Attached Vibrational Probe of Protein-Lipid Contacts. J. Phys. Chem. Lett 2010, 1, 850–855. [PubMed: 20228945]

(39). Park ES; Thomas MR; Boxer SG 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.

(40). Slocum JD; Webb LJ Measuring Electric Fields in Biological Matter Using the Vibrational Stark Effect of Nitrile Probes. Annu. Rev. Phys. Chem 2018, 69, 253–271. [PubMed: 29677466]

(41). Stafford AJ; Ensign DL; Webb LJ 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. [PubMed: 20964430]

(42). Suydam IT; Boxer SG Vibrational Stark Effects Calibrate the Sensitivity of Vibrational Probes for Electric Fields in Proteins. Biochemistry 2003, 42, 12050–12055. [PubMed: 14556636]

(43). Waegele MM; Tucker MJ; Gai F 5-Cyanotryptophan as an Infrared Probe of Local Hydration Status of Proteins. Chem. Phys. Lett 2009, 478, 249–253. [PubMed: 20161057]

(44). Walker DM; Hayes EC; Webb LJ 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. [PubMed: 23771025]

(45). Fafarman AT; Webb LJ; Chuang JI; Boxer SG 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. [PubMed: 17031938]

(46). Kirshenbaum K; Carrico IS; Tirrell DA Biosynthesis of Proteins Incorporating a Versatile Set of Phenylalanine Analogues. ChemBioChem 2002, 3, 235–237. [PubMed: 11921403]

(47). Webb LJ; Boxer SG Electrostatic Fields near the Active Site of Human Aldose Reductase: 1. New Inhibitors and Vibrational Stark Effect Measurements. Biochemistry 2008, 47, 1588–1598. [PubMed: 18205401]

(48). Fafarman AT; Boxer SG Nitrile Bonds as Infrared Probes of Electrostatics in Ribonuclease S. J. Phys. Chem. B 2010, 114, 13536–13544. [PubMed: 20883003]

(49). Shields JM; Pruitt K; McFall A; Shaub A; Der CJ Understanding Ras: 'It Ain't over "Til It's Over." Trends Cell Biol. 2000, 10, 147–154. [PubMed: 10740269]

(50). Andrews SS; Boxer SG Vibrational Stark Effects of Nitriles. I. Methods and Experimental Results. J. Phys. Chem. A 2000, 104, 11853–11863.

(51). Andrews SS; Boxer SG Vibrational Stark Effects of Nitriles II. Physical Origins of Stark Effects from Experiment and Perturbation Models. J. Phys. Chem. A 2002, 106, 469–477.

(52). First JT; Novelli ET; Webb LJ Beyond p Ka: Experiments and Simulations of Nitrile Vibrational Probes in Staphylococcal Nuclease Show the Importance of Local Interactions. J. Phys. Chem. B 2020, 124, 3387–3399. [PubMed: 32212657]

(53). Slocum JD; First JT; Webb LJ Orthogonal Electric Field Measurements near the Green Fluorescent Protein Fluorophore through Stark Effect Spectroscopy and pK_a Shifts Provide a Unique Benchmark for Electrostatics Models. J. Phys. Chem. B 2017, 121, 6799–6812. [PubMed: 28650636]

(54). Arthur EJ; Yesselman JD; Brooks CL Predicting Extreme pK_a Shifts in Staphylococcal Nuclease Mutants with Constant pH Molecular Dynamics. Proteins Struct. Funct. Bioinforma 2011, 79, 3276–3286.

(55). Adhikary R; Zimmermann J; Dawson PE; Romesberg FE Temperature Dependence of CN and SCN IR Absorptions Facilitates Their Interpretation and Use as Probes of Proteins. Anal. Chem 2015, 87, 11561–11567. [PubMed: 26523838]

(56). Abraham MJ; Murtola T; Schulz R; Páll S; Smith JC; Hess B; Lindah E Gromacs: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. SoftwareX 2015, 1–2, 19–25.

(57). Berendsen HJC; van der Spoel D; van Drunen R GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. Comput. Phys. Commun 1995, 91, 43–56.

(58). Hess B; Kutzner C; Van Der Spoel D; Lindahl E GRGMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. J. Chem. Theory Comput 2008, 4, 435–447. [PubMed: 26620784]

(59). 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.

(60). Pronk S; Páll S; Schulz R; Larsson P; Bjelkmar P; Apostolov R; Shirts MR; Smith JC; Kasson PM; Van Der Spoel D; Hess B; Lindahl E GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. Bioinformatics 2013, 29, 845–854. [PubMed: 23407358]

(61). Van Der Spoel D; Lindahl E; Hess B; Groenhof G; Mark AE; Berendsen HJC GROMACS: Fast, Flexible, and Free. J. Comput. Chem 2005, 26, 1701–1718. [PubMed: 16211538]

(62). Ponder JW; Wu C; Ren P; Pande VS; Chodera JD; Schnieders MJ; Haque I; Mobley DL; Lambrecht DS; Distasio RA, et al. Current Status of the AMOEBA Polarizable Force Field. J. Phys. Chem. B 2010, 114, 2549–2564. [PubMed: 20136072]

(63). Ponder JW; Wu C; Ren P; Pande VS; Chodera JD; Schnieders MJ; Haque I; Mobley DL; Lambrecht DS; Distasio RA Current Status of the AMOEBA Polarizable Force Field. J. Phys. Chem. B 2010, 114, 2549–2564. [PubMed: 20136072]

(64). Reuter N; Lin H; Thiel W Green Fluorescent Proteins: Empirical Force Field for the Neutral and Deprotonated Forms of the Chromophore. Molecular Dynamics Simulations of the Wild Type and S65T Mutant. J. Phys. Chem. B 2002, 106, 6310–6321.

(65). Nifosì R; Tozzini V Molecular Dynamics Simulations of Enhanced Green Fluorescent Proteins: Effects of F64L, S65T and T203Y Mutations on the Ground-State Proton Equilibria. Proteins Struct. Funct. Genet 2003, 51, 378–389. [PubMed: 12696049]

(66). First JT; Webb LJ Agreement between Experimental and Simulated Circular Dichroic Spectra of a Positively Charged Peptide in Aqueous Solution and on Self-Assembled Monolayers. J. Phys. Chem. B 2019, 123, 4512–4526. [PubMed: 31038968]

(67). Collier G; Vellore NA; Yancey JA; Stuart SJ; Latour RA Comparison between Empirical Protein Force Fields for the Simulation of the Adsorption Behavior of Structured LK Peptides on Functionalized Surfaces. Biointerphases 2012, 7, 1–19. [PubMed: 22589044]

(68). Jing Z; Liu C; Qi R; Ren P Many-Body Effect Determines the Selectivity for Ca^{2+} and Mg^{2+} in Proteins. Proc. Natl. Acad. Sci. U. S. A 2018, 115, 7495–7501.

(69). Mao Y; Demerdash O; Head-Gordon M; Head-Gordon T Assessing Ion-Water Interactions in the AMOEBA Force Field Using Energy Decomposition Analysis of Electronic Structure Calculations. J. Chem. Theory Comput 2016, 12, 5422–5437. [PubMed: 27709939]

(70). Shi Y; Jiao D; Schnieders MJ; Ren P Trypsin-Ligand Binding Free Energy Calculation with AMOEBA. Proc. 31st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. Eng. Futur. Biomed. EMBC 2009 2009, 2328–2331.

(71). Jing Z; Rackers JA; Pratt LR; Liu C; Rempe SB; Ren P Thermodynamics of Ion Binding and Occupancy in Potassium Channels. Chem. Sci 2021, 12, 8920–8930. [PubMed: 34257893]

(72). Wu JC; Chattree G; Ren P Automation of AMOEBA Polarizable Force Field Parameterization for Small Molecules. Theor. Chem. Acc 2012, 131, 1–11.

(73). Walker B; Liu C; Wait E; Ren P Automation of AMOEBA Polarizable Force Field for Small Molecules: Poltype 2. J. Comput. Chem 2022, 43, 1530–1542. [PubMed: 35778723]

(74). Turney JM; Simmonett AC; Parrish RM; Hohenstein EG; Evangelista FA; Fermann JT; Mintz BJ; Burns LA; Wilke JJ; Abrams ML, et al. Psi4: An Open-Source Ab Initio Electronic Structure Program. Wiley Interdiscip. Rev. Comput. Mol. Sci 2012, 2, 556–565.

(75). Harms MJ; Castañeda CA; Schlessman JL; Sue GR; Isom DG; Cannon BR; García-Moreno E, The B pK_a 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. [PubMed: 19324049]

(76). Adjoua O; Lagardère L; Jolly LH; Durocher A; Very T; Dupays I; Wang Z; Inizan TJ; Célerse F; Ren P, et al. Tinker-HP: Accelerating Molecular Dynamics Simulations of Large Complex Systems with Advanced Point Dipole Polarizable Force Fields Using GPUs and Multi-GPU Systems. J. Chem. Theory Comput 2021, 17, 2034–2053. [PubMed: 33755446]

(77). Martínez L; Andrade R; Birgin EG, Packmol MJM: A Package for Building Initial Configurations for Molecular Dynamics Simulations. J. Comput. Chem 2009, 30, 2157–2164. [PubMed: 19229944]

(78). Ensign DL; Webb LJ Factors Determining Electrostatic Fields in Molecular Dynamics Simulations of the Ras/Effector Interface. Proteins Struct. Funct. Bioinforma 2011, 79, 3511–3524.

(79). Duan Y; Wu C; Chowdhury S; Lee MC; 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. [PubMed: 14531054]

(80). Sorin EJ; Pande VS Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. Biophys. J 2005, 88, 2472–2493. [PubMed: 15665128]

(81). Shi Y; Xia Z; Zhang J; Best R; Wu C; Ponder JW; Ren P Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. J. Chem. Theory Comput 2013, 9, 4046–4063. [PubMed: 24163642]

(82). Le Questel JY; Berthelot M; Laurence C Hydrogen-Bond Acceptor Properties of Nitriles: A Combined Crystallographic and Ab Initio Theoretical Investigation. J. Phys. Org. Chem 2000, 13, 347–358.

(83). Fried SD; Wang LP; Boxer SG; Ren P; Pande VS Calculations of the Electric Fields in Liquid Solutions. J. Phys. Chem. B 2013, 117, 16236–16248. [PubMed: 24304155]

(84). Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, S. C Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. PROTEINS Struct. Funct. Bioinformatic 2006, 65, 712–725.

(85). Maier JA; Martinez C; Kasavajhala K; Wickstrom L; Hauser KE; Simmerling C Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. J. Chem. Theory Comput 2015, 11, 3696–3713. [PubMed: 26574453]

(86). Das AK; Demerdash ON; Head-Gordon T Improvements to the AMOEBA Force Field by Introducing Anisotropic Atomic Polarizability of the Water Molecule. J. Chem. Theory Comput 2018, 14, 6722–6733. [PubMed: 30428257]

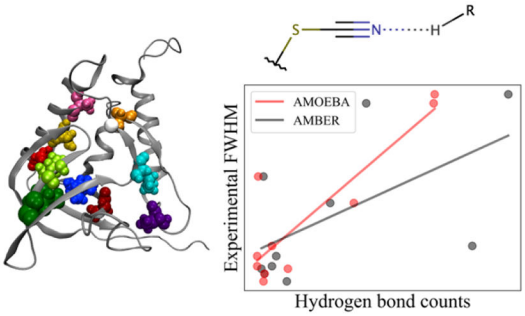
(87). Ren P; Wu C; Ponder JW Polarizable Atomic Multipole-Based Molecular Mechanics for Organic Molecules. J. Chem. Theory Comput 2011, 7, 3143–3161. [PubMed: 22022236]

(88). Ji C; Mei Y Some Practical Approaches to Treating Electrostatic Polarization of Proteins. Acc. Chem. Res 2014, 47, 2796–2803.

(89). Xu RJ; Blasiak B; Cho M; Layfield JP; Londergan CH A Direct, Quantitative Connection between Molecular Dynamics Simulations and Vibrational Probe Line Shapes. J. Phys. Chem. Lett 2018, 9, 2560–2567. [PubMed: 29697984]

(90). Bowles J; Jähnigen S; Vuilleumier R; Calvo F; Clavaguéra C; Agostini F Influence of the Environment on the Infrared Spectrum of Alanine: An Effective Mode Analysis. J. Chem. Phys 2023, 158.

(91). Thaunay F; Dognon JP; Ohanessian G; Clavaguéra C Vibrational Mode Assignment of Finite Temperature Infrared Spectra Using the AMOEBA Polarizable Force Field. Phys. Chem. Chem. Phys 2015, 17 (39), 25968–25977. [PubMed: 26214153]



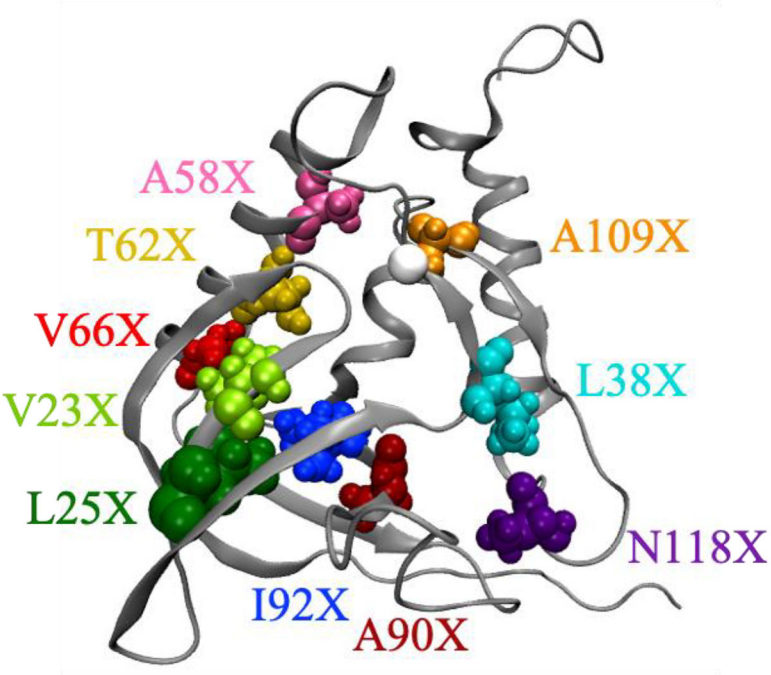


Figure 1.

The positions of the cyanocysteine in staphylococcal nuclease (SNase). Each position showed in multiple color spheres is an amino acid that was individually mutated to cyanocysteine (X). The secondary structure of SNase is shown in gray. The native Ca^{2+} is shown as a white sphere.

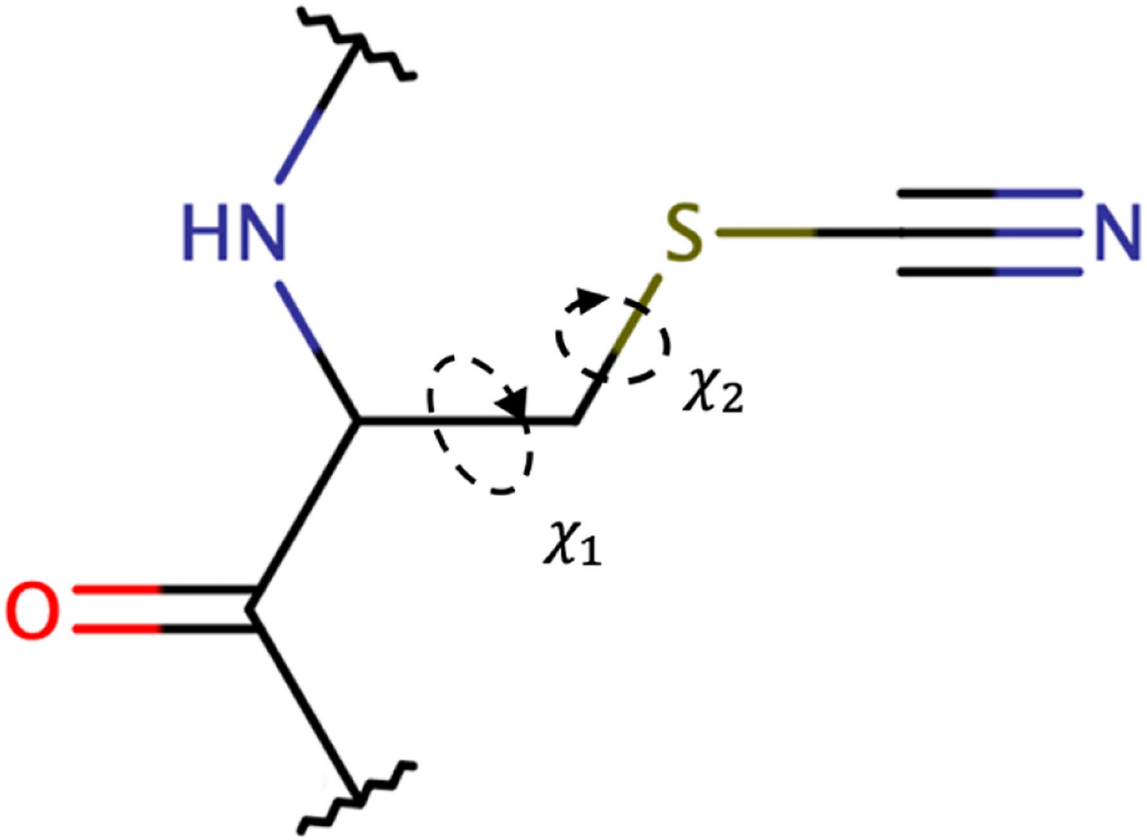


Figure 2.
The structure of cyanocysteine used to parameterize the molecule for AMOEBA simulation with the dihedral angles χ_1 and χ_2 indicated.

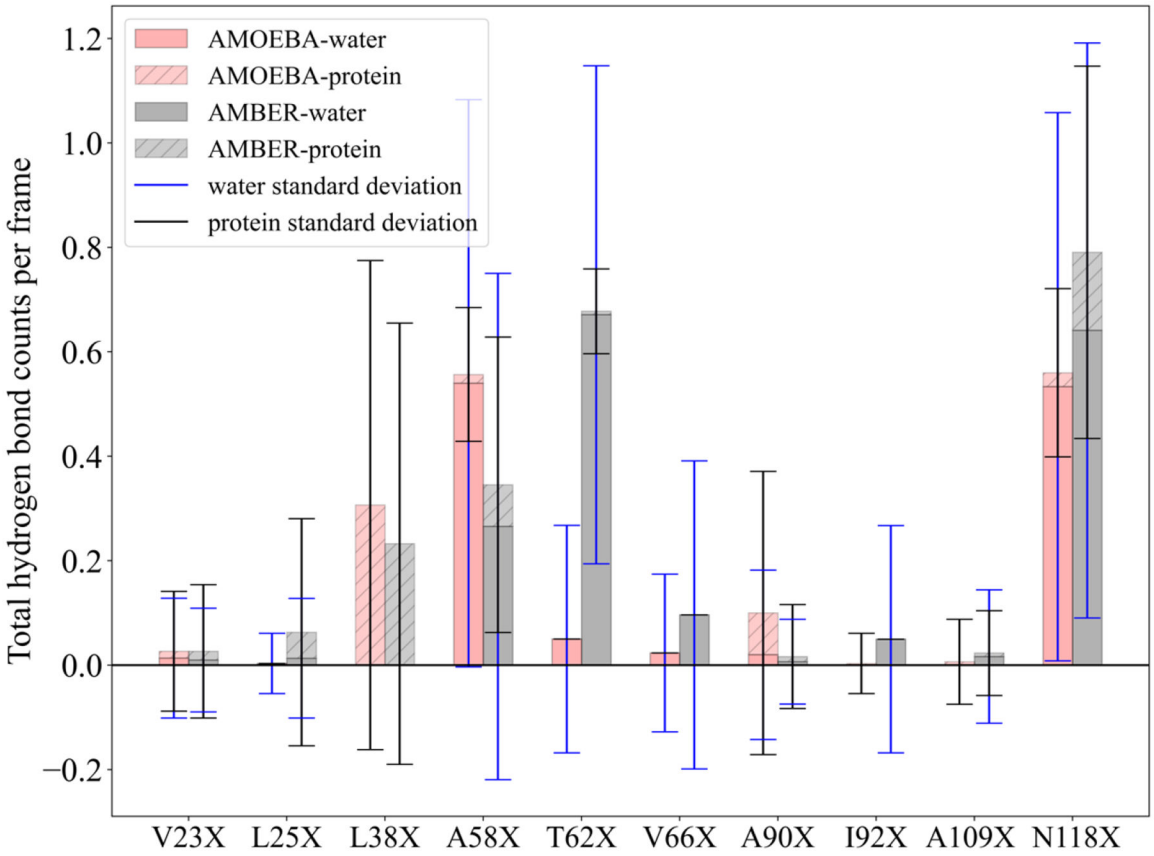


Figure 3.

Number of hydrogen bonds of cyanocysteine with nearby residues per simulation structure frame in terms of the hydrogen bonds donated by both water (solid bar, AMOEBA: red; Amber03: gray) and protein (hashed bar, AMOEBA: light red; Amber03: light gray). Standard deviations represent the last 60 ns of each trajectory.

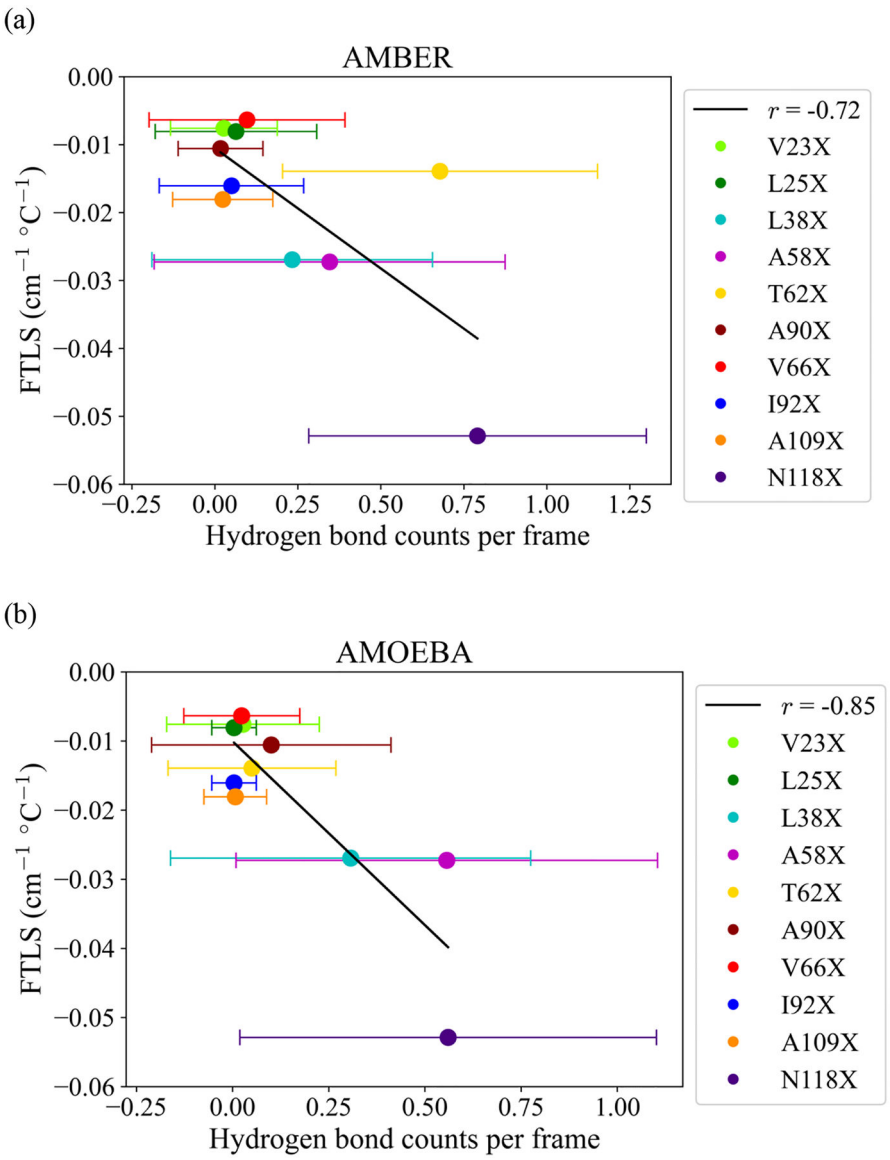


Figure 4.

Frequency temperature line slope (FTLS) compared to total hydrogen bond counts from the last 60 ns MD simulations in (a) Amber03 and (b) AMOEBA.

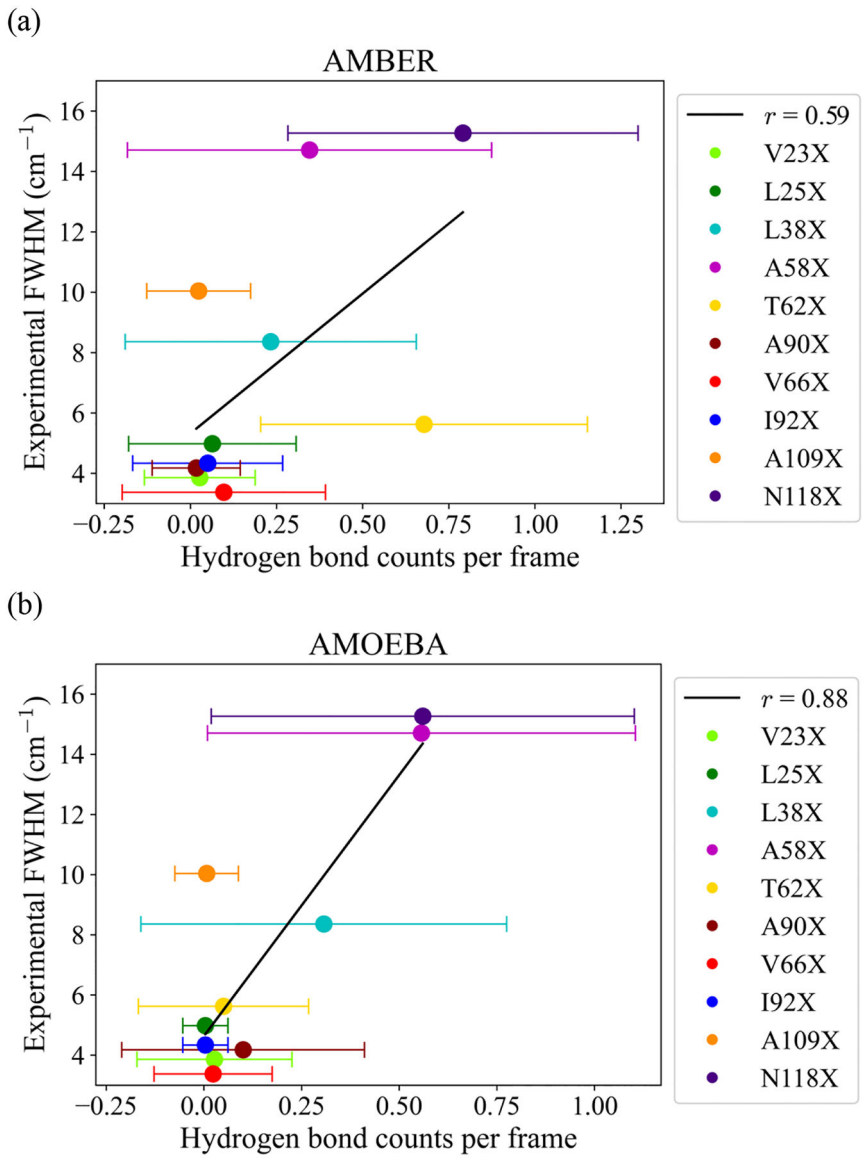


Figure 5.

Full width half maximum (FWHM) compared to total hydrogen bond counts from the last 60 ns MD simulations in (a) Amber03 and (b) AMOEBA.

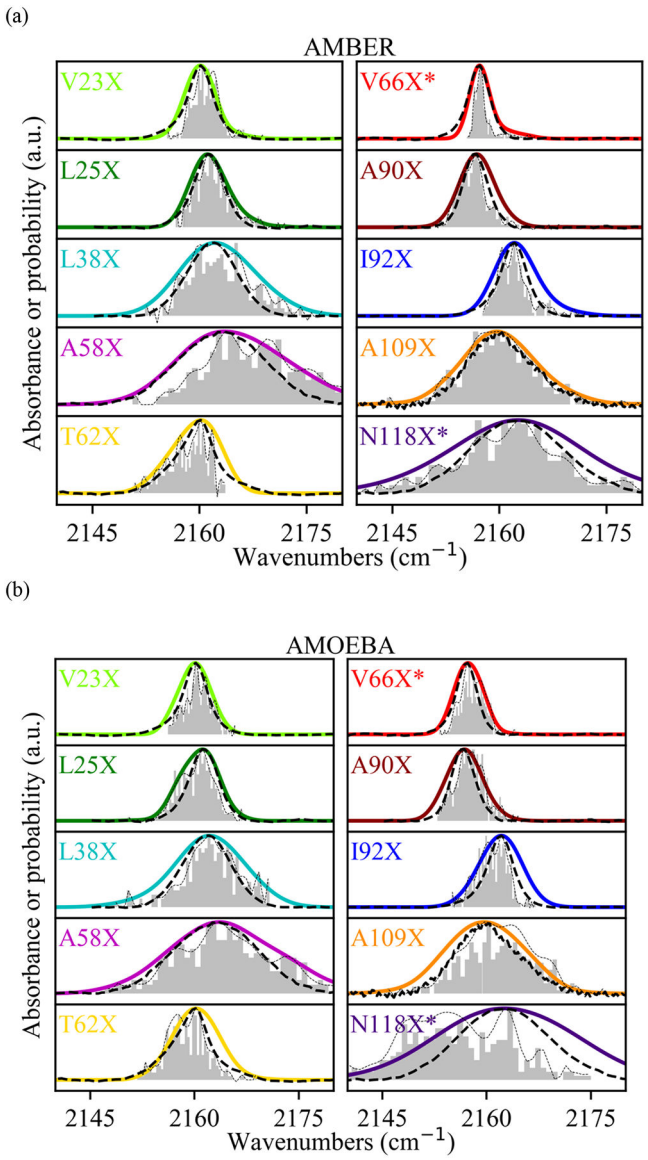


Figure 6.

Vibrational absorption energy (in cm^{-1}) calculated from the electric field on the +1 dummy atom at the center of the nitrile bond projected alone the nitrile bond axis. Gray: calculated histogram. Thin gray dashed line: Gaussian functions for the histogram bins. Solid colored line: Gaussian functions generated by the sum of the broadened histogram bins with the broadened bandwidth proportional to the experimental FTLS. The * sign at the V66X and N118X indicates that the FTLS of these mutants were used to pin the endpoints of the bandwidth that was used to broaden the histogram bin. Dashed black line: experimental spectrum from Figure S1.[52]