JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

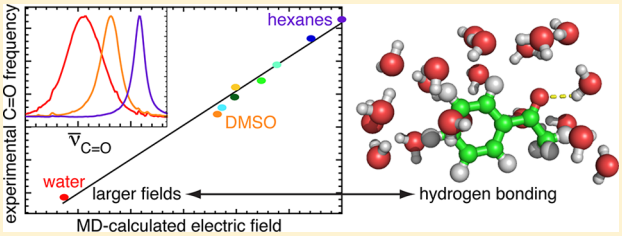
Measuring Electrostatic Fields in Both Hydrogen-Bonding and Non-Hydrogen-Bonding Environments Using Carbonyl Vibrational Probes

Stephen D. Fried, Sayan Bagchi,[†] and Steven G. Boxer*

Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States

Supporting Information

ABSTRACT: Vibrational probes can provide a direct readout of the local electrostatic field in complex molecular environments, such as protein binding sites and enzyme active sites. This information provides an experimental method to explore the underlying physical causes of important biomolecular processes such as binding and catalysis. However, specific chemical interactions such as hydrogen bonds can have complicated effects on vibrational probes and confound simple electrostatic interpretations of their frequency shifts. We employ vibrational Stark spectroscopy along with infrared spectroscopy of carbonyl probes in different solvent environments and in ribonuclease S to understand the sensitivity of carbonyl frequencies to electrostatic fields, including those due to hydrogen bonds. Additionally, we carried out molecular dynamics simulations to calculate ensemble-averaged electric fields in solvents and in ribonuclease S and found excellent correlation between calculated fields and vibrational frequencies. These data enabled the construction of a robust field−frequency calibration curve for the C═O vibration. The present results suggest that carbonyl probes are capable of quantitatively assessing the electrostatics of hydrogen bonding, making them promising for future study of protein function.



1. INTRODUCTION

Vibrational (IR) probes are sensitive to their local electrostatic environment and can be as small as two atoms. Therefore, these probes are ideally suited to study the intricate architecture of proteins, whose electrostatic characteristics vary steeply over very small length scales.[1−3] Furthermore, the sensitivity of a vibrational frequency to an electrostatic field can be experimentally calibrated in an external well-defined electric field using vibrational Stark spectroscopy.[4,5] An ideal vibrational probe of electrostatics would allow one to read an electric field, $\vec{F}$, experienced by a vibration due to its surrounding environment directly from the infrared (IR) spectrum, according to

$$\vec{F}\cdot\hat{u}_{\text{probe}} = \frac{hc(\bar{\nu}_{\text{obs}} - \bar{\nu}_0)}{|\Delta\vec{\mu}_{\text{probe}}|} \quad (1)$$

where $\bar{\nu}_{\text{obs}}$ is the observed vibrational frequency, $\bar{\nu}_0$ is a reference frequency calibrated to zero electric field, $|\Delta\vec{\mu}_{\text{probe}}|$ is the magnitude of the probe's difference dipole, which is determined by measuring the vibrational Stark effect and defines the sensitivity of vibrational shifts to electric field, and $\vec{F}\cdot\hat{u}_{\text{probe}}$ is the electric field experienced by the vibration projected onto the difference dipole vector, $\Delta\vec{\mu}_{\text{probe}}$. For the case of a decoupled vibrational mode that is confined to two atoms, $\hat{u}_{\text{probe}}$ is necessarily colinear with the bond axis of the vibration and aligned with the permanent bond dipole moment.[6] This criterion implies that when a field is oriented in a manner that stabilizes the bond's (difference) dipole, the vibration shifts to lower frequency.

The nitrile (C≡N) vibration has gained wide attention as an electrostatic probe in biophysical studies because it is intense, local, and absorbs in an uncluttered region of the IR spectrum and a number of approaches have been developed to introduce it into biological systems.[1−3,7−11] The goal to utilize eq 1 as a general tool to translate observed vibrational frequency shifts into electric fields is complicated by the possibility that the variation in a probe's absorption frequency may not be entirely due to electrostatics. For example, C≡N frequencies exhibit a characteristic blue shift upon accepting a hydrogen bond (H-bond) in a manner that is not described by the linear vibrational Stark effect.[12,13] It would be preferable to apply eq 1 uniformly across both H-bonding and non-H-bonding environments because many cases where protein electrostatics are functionally relevant (ligand binding, enzyme catalysis, protein−protein recognition) involve H-bonding or a transition between non-H-bonding and H-bonding states.

The carbonyl (C═O) stretching mode is also very local and generally more intense than nitriles.[14,15] Importantly, computational results have predicted that carbonyl frequencies vary linearly with electrostatic field in H-bonding environments.[16] A limitation to the C═O probe is that its characteristic frequency (1700 cm^{-1}) overlaps with the densely populated amide I region, making it more challenging to detect in proteins. However, we have overcome this limitation by carefully selecting a reference sample that is almost identical to the vibrational-probe-bearing sample, but alters the vibrational

Received: April 19, 2013
Published: June 28, 2013

   dx.doi.org/10.1021/ja403917z | J. Am. Chem. Soc. 2013, 135, 11181−11192

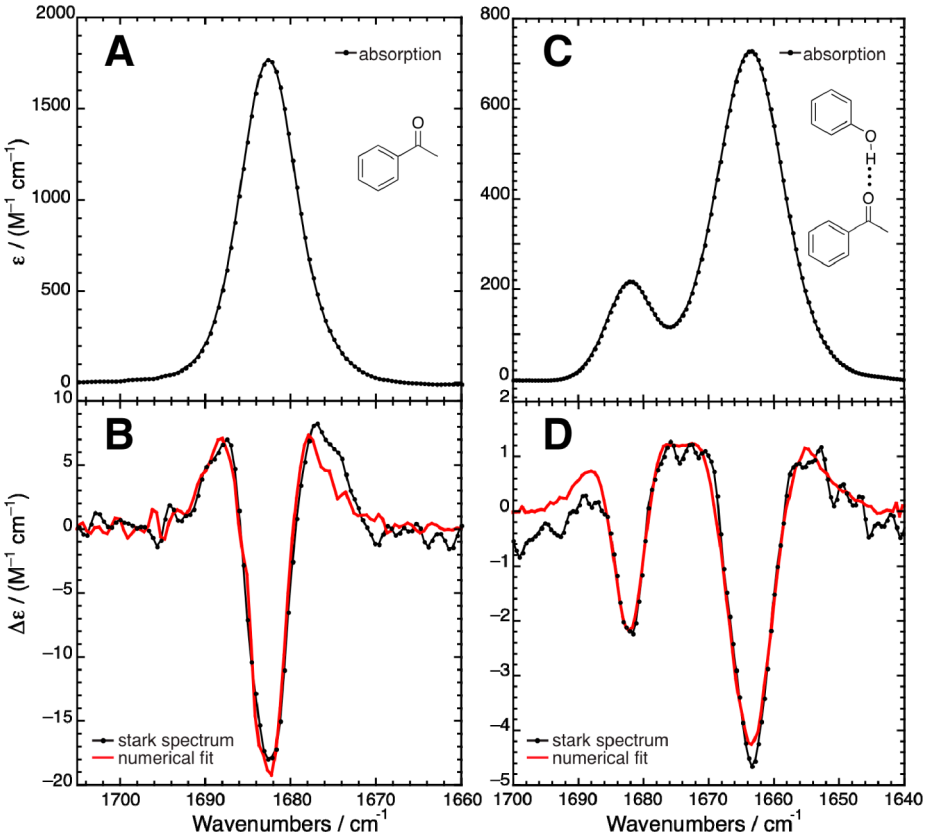


Figure 1. Vibrational Stark spectroscopy of the C═O stretch: absorption (A, C) and Stark (B, D) spectra of acetophenone (50 mM) at 77 K in 2-methyltetrahydrofuran (A, B) and in toluene with 1.25 equiv of phenol (C, D). Stark spectra are recorded at several applied fields from 0.5 to 1.4 MV/cm but are scaled to 1 MV/cm here for comparison (note that the Stark signal scales with the square of the external field[53]). In the Stark spectra, the black dots and trace are the experimental data and the red trace is a numerical fit (consisting of contributions of the zeroth, first, and second derivatives of the absorption) from which Stark parameters are extracted (see Table 1).

Table 1. Vibrational Stark Effect Data of Acetophenone[a]

	absorption			Stark			
system	peak position, cm^{-1}	fwhm, cm^{-1}	peak area, M^{-1} cm^{-2}	A	B, cm^{-1}	C, cm^{-2}	$\lvert\Delta\vec{\mu}\rvert_p$[b] $cm^{-1}/(MV/cm)$
2-MeTHF	1682.6	7.9	15600	$(-4.3 \pm 2.4) \times 10^{-4}$	$(1.53 \pm 1.03) \times 10^{-3}$	$(1.097 \pm 0.032) \times 10^{-1}$	1.05 ± 0.08
toluene + 1.25 equiv of phenol	1682.0	6.3	1900	$(-1.2 \pm 0.1) \times 10^{-3}$	$(-2.1 \pm 7.5) \times 10^{-4}$	$(1.136 \pm 0.031) \times 10^{-1}$	1.07 ± 0.08
	1663.7	10.9	9800				

[a]Data and fitting parameters correspond to absorption and Stark spectra in Figure 1 (see refs 4 and 6 for details on interpretation of fitting parameters). [b]Derived from the second derivative component, C. Assumes the angle between the vibration's difference dipole and transition dipole is 0°.

probe, and then calculating the difference spectrum.[17] This technique further requires that transmission spectra of sample and reference be closely matched and that the protein concentration be low enough that some light can transmit through the amide I frequency range. The carbonyl probe can be installed into proteins via the unnatural amino acid, *p*-acetyl-L-phenylalanine (p-Ac-Phe), through reassembly of a split protein, protein semisynthesis, or nonsense suppression.[18] Carbonyl groups are also often present on substrates, inhibitors, and drugs. In this study, we deployed the C═O probe into a protein by incorporating p-Ac-Phe as a residue into a polypeptide chain. To study a diverse range of solvents, we chose acetophenone as a solute because it maintains the same local structure around the C═O chromophore as p-Ac-Phe.

In this work, we test the viability of carbonyl vibrations as local electrostatic probes, with a focus on examining their response to strong interactions such as H-bonding. To do this, we performed vibrational Stark spectroscopy on C═O vibrations to determine the intrinsic sensitivity of the C═O vibrational frequency to an electrostatic field. We also obtained IR spectra of carbonyls in a variety of settings (such as in solvents of different polarities and H-bond donor strengths and in the model protein ribonuclease S (RNase S)) to see how their frequencies report on the local electrostatic fields of those environments. These observations were compared against molecular dynamics (MD) calculations of the electric fields in the respective environments, with which they displayed a remarkably high linear correlation. With this correlation, we developed a robust field−frequency calibration for carbonyl

dx.doi.org/10.1021/ja403917z | J. Am. Chem. Soc. 2013, 135, 11181−11192

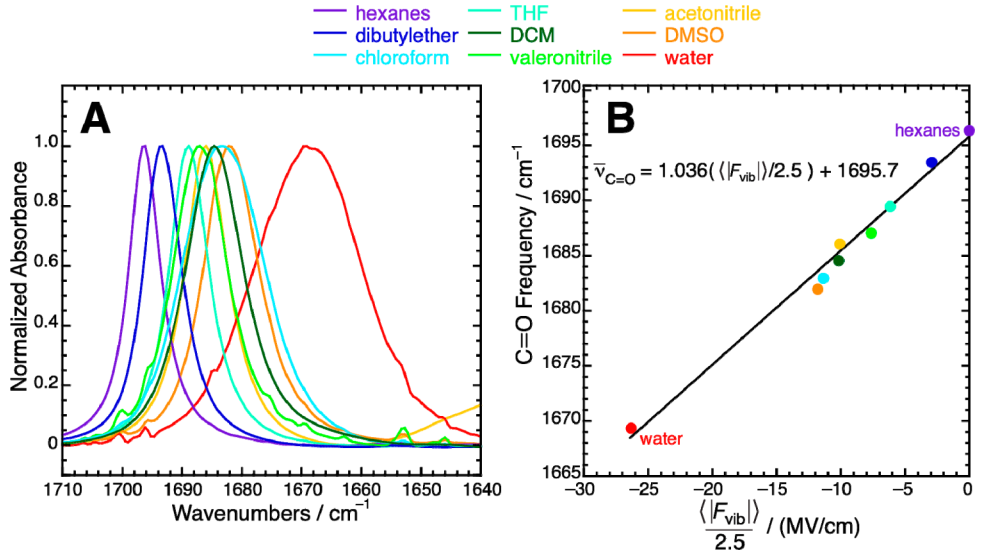


Figure 2. C═O vibrational frequencies shift due to solvent electric field. (A) Representative FTIR spectra of the C═O stretch band of acetophenone (10 mM) dissolved in a number of organic solvents and water. The spectra are colored from violet to red in order of increasing static dielectric constant. (B) Plot of acetophenone's peak C═O frequency in nine solvents compared against the average electric field C═O experiences in each of those nine solvents, as calculated by MD simulation. Note that electric fields are rescaled by a factor of 2.5 (see main text). The best-fit linear model is $\bar{\nu}_{C=O} = (1.036\langle|F_{vib}|\rangle/2.5) + 1695.7$ with $R^2 = 0.99$.

vibrations. Our results collectively indicate that carbonyl frequency shifts conform to the formalism in eq 1 and report on the relatively large electric fields that are created by H-bonding by way of a linear Stark effect. In particular, the latter feature makes them an attractive probe for future research in protein structure and function, and these applications will be described in future publications.

2. RESULTS AND DISCUSSION FOR SOLVENTS

2.1. Vibrational Stark Spectroscopy of Acetophenone. In vibrational Stark spectroscopy, the sensitivity of a vibration to an electric field is determined by applying a uniform field onto a vitrified sample with a high-voltage power source and then observing the response of the IR spectrum.[4] The absorption and Stark spectra of acetophenone, a model compound that contains a local C═O stretch, are displayed in Figure 1, and summaries of the spectral data are given in Table 1. Stark spectra are conventionally shown as the difference spectrum between the field-on and the field-off absorbances and are interpreted by numerically fitting the difference spectrum to contributions from the zeroth, first, and second derivatives of the absorption spectrum. Vibrational Stark spectra have been previously reported for the heme-bound CO in CO-myoglobin (wild-type and several mutants),[19] for the C═O in acetone,[20] and for 6-propionyl-2-(dimethylamino)naphthalene,[5] but not for a simple aromatic ketone like acetophenone.

The C═O stretch of acetophenone in 2-MeTHF has an intense peak molar extinction of 1800 M^{-1} cm^{-1} (Figure 1A, Table 1) and possesses a strong Stark signal (Figure 1B) whose line shape is almost identical to the second derivative of the absorption spectrum. As shown in Table 1, the fitting coefficients for the zeroth and first derivative contributions are essentially zero within experimental error. The dominance of the second derivative contribution implies that the sensitivity of the vibration to electric field is determined almost exclusively by the C═O vibration's difference dipole, $|\Delta\vec{\mu}|$.[4,6] The difference dipole, also known as the Stark tuning rate, specifies the extent to which a vibrational frequency shifts linearly with respect to electrostatic field. From the second derivative fitting parameter (Table 1), we calculate that $|\Delta\vec{\mu}|f = 1.05 \pm 0.08$ cm^{-1}/(MV/cm) where f is the local field correction factor, described further in section 2.5 below and experimental methods section 5.4.

To examine the effect of accepting an H-bond on the Stark tuning rate of C═O, acetophenone was dissolved in toluene along with 1.25 equiv of phenol, an H-bond donor. The low-temperature absorbance (Figure 1C) consists of two well-resolved peaks, and the minor band's peak position matches that of acetophenone when no phenol is present. The second peak is broader, has a peak frequency (1663.7 cm^{-1}) closer to that of acetophenone dissolved in water (1669.4 cm^{-1}), and increases in intensity upon increasing the concentration of phenol (data not shown). These observations confer strong evidence that acetophenone forms an H-bonding complex with phenol and that the second peak corresponds to the H-bonded population. As seen in Figure 1D, a satisfactory fit to the Stark features of both the H-bonded and non-H-bonded populations was obtained with a single set of parameters (Table 1). Moreover, the Stark tuning rate determined from this fit (1.07 cm^{-1}/(MV/cm)) was nearly identical to that found in 2-MeTHF (1.05 cm^{-1}/(MV/cm)). These results strongly suggest that the carbonyl frequency's intrinsic sensitivity to an electric field is the same in both non-H-bonding and H-bonding environments.[21]

In summary, the carbonyl vibration is more sensitive to electric fields than the nitrile vibration, whose Stark tuning rate is typically observed to be 0.6−0.7 cm^{-1}/(MV/cm).[4−6] Acetophenone's Stark tuning rate is more than that of acetone (0.8 cm^{-1}/(MV/cm)[20]) but less than that of 6-propionyl-2-(dimethylamino)naphthalene (1.8 cm^{-1}/(MV/cm)[5]), suggesting that for carbonyls, increasing the conjugation leads to larger sensitivity to electric field.

2.2. Vibrational Solvatochromism of Acetophenone. To explore the response of the C═O vibration to the effect of diverse condensed phase environments, we carried out

dx.doi.org/10.1021/ja403917z | J. Am. Chem. Soc. 2013, 135, 11181−11192

vibrational solvatochromism studies on acetophenone. Acetophenone is miscible in a wide range of solvents, and it can serve as a model compound for p-Ac-Phe, the unnatural amino acid that we employed to incorporate the carbonyl probe into proteins.

As shown in Figure 2A, the C═O stretching band in acetophenone progressively shifts to the red as it is dissolved in solvents of progressively greater polarity, with the peak frequency moving 14.4 cm^{-1} from hexanes (1696.4 cm^{-1}) to dimethyl sulfoxide (DMSO, 1682.0 cm^{-1}). Full spectral data are given in Table S1. The peak frequency of acetophenone dissolved in chloroform (1683.3 cm^{-1}) possesses a value near that of DMSO, which is striking given the relatively small dielectric constant of 4.7 for chloroform compared to DMSO's dielectric constant of 47. As discussed further in the simulation section, we speculated that this peculiar shift could be attributed to H-bonding because chloroform (as well as dichloromethane) possesses a polar H-atom unlike the other organic solvents examined, whose polar components only involve heavy atoms. The solvatochromic dispersion of acetophenone's C═O band is greater than that of benzonitrile's C≡N band, which, while possessing the same qualitative trend of shifting to the red in solvents of greater polarity, only shifts 5.8 cm^{-1} from hexanes to DMSO.[22] When dissolved in water (water refers to D_2O), acetophenone's C═O band shifts substantially to the red (to 1669.4 cm^{-1}) and broadens. According to eq 1, this behavior would be consistent with water exerting much larger electric fields compared to the other solvents, which can be rationalized by considering that an H-bond donor positions a large dipole very close to the reporter vibration because of the small van der Waals radius of hydrogen. This observation is suggestive that the C═O frequency reports on the electric fields in H-bonding environments. In contrast, the C≡N stretch of benzonitrile in water is strongly blue-shifted and has a frequency higher than benzonitrile dissolved in hexanes. Figure 2A also indicates that there is a strong correlation between the spectral line widths (related to the spread in electric fields) and peak frequency (related to the overall electric field magnitude), as examined further in Figure S1.

2.3. Solvation Simulation Overview. Motivated by the importance to describe both H-bonding and non-H-bonding environments in a consistent fashion, we used an atomistic approach to calculate solvation fields by MD simulation. Pioneering work by Jorgensen and co-workers demonstrated the capability of classical force fields to recapitulate bulk properties of organic liquids like heat of vaporization, density, and radial distribution functions.[23,24] In the following, we show that these models can also predict solvent-induced frequency shifts. The simulations were carried out with the general AMBER force field (GAFF),[25,26] following a recently published study[27] that benchmarked GAFF's and OPLS/AA's reliability for modeling a wide array of organic liquids including acetophenone. We chose to model solvent molecules with GAFF parameters to enable comparison to protein simulations that use AMBER force fields. Water was modeled using the TIP3P model.[28] In total, electric fields were calculated for nine solvents (acetonitrile, chloroform, dibutyl ether, dichloromethane, dimethyl sulfoxide, *n*-hexane, tetrahydrofuran, valeronitrile, and water). All simulations were set up and carried out in GROMACS, version 4.5.3.[29] For each snapshot, the electric field exerted onto the C═O vibration of acetophenone by the solvation environment was calculated by projecting onto the C═O bond vector and then averaging between the C-atom and at the O-atom. This quantity is denoted $|F_{vib}|$ and referred to as the field experienced by the C═O vibration. Similarly, the electric field drop over the bond length (denoted $|\Delta F_{vib}|$) was calculated by taking the difference between those two quantities instead of averaging. In both cases, an ensemble average was determined by calculating the mean across all snapshots in the trajectory, denoted $\langle|F_{vib}|\rangle$ and $\langle|\Delta F_{vib}|\rangle$. Simulation methodology and the method for calculating electric fields are described in detail in the computational methods sections 4.1 to 4.3.

2.4. MD Calculation of Solvation Fields. The ensemble electrostatic data for the C═O probe of acetophenone in various solvents are compiled in Table 2. These calculated fields display a strong correlation with solvent polarity: the more polar solvents exert an electric field of greater magnitude onto acetophenone's C═O bond. The negative sign associated with all calculated fields implies that the environment interacts favorably with the carbonyl moiety of the solute. It is important to point out how the MD simulations handled the cases where the solvent can form H-bonds with the C═O probe. Interestingly, MD predicted that chloroform exercises an average electric field (−28.5 MV/cm) of nearly equal magnitude to DMSO (−29.6 MV/cm), even though their dielectric constants are very different. This result is consistent with the FTIR measurements (Figure 2A), where it was found that acetophenone's C═O band has a very similar peak frequency when it is dissolved in DMSO and chloroform (1682.0 and 1683.3 cm^{-1}, respectively). Taken together, these observations suggest that chloroform's (and dichloromethane's) deviation from the correlation with solvent dielectric can be explained as a linear Stark effect when one accounts for the fact that chloroform's H-bonding to the carbonyl group results in larger electric fields than would be predicted from a continuum descriptor like dielectric constant. This concept can be expanded further to the case of water. The C═O stretch of acetophenone is shifted 12.6 cm^{-1} to the red in water relative to DMSO (Figure 2A). This large shift is recapitulated by the MD

Table 2. MD Electrostatic Data for C═O Bonds in Solvents and Proteins

	$\lvert F_{vib}\rvert$, MV/cm [a]		$\lvert\Delta F_{vib}\rvert$, MV/cm [b]	
	mean	std dev	mean	std dev
hexanes	−0.115	0.77	−0.029	0.76
dibutylether	−7.40	6.42	−0.28	3.83
chloroform	−28.5	13.6	−13.7	12.7
tetrahydrofuran	−15.5	9.36	−0.19	6.09
dichloromethane	−25.5	15.9	−8.66	11.6
valeronenitrile	−19.2	12.0	−0.67	7.46
acetonitrile	−25.3	13.6	−3.73	9.58
dimethylsulfoxide	−29.6	11.5	−1.53	10.1
water (TIP3P)	−65.9	23.1	−40.7	25.7
[p-Ac-Phe]S-peptide	−63.6	39.6	−38.1	33.1
	[−61.9][c]	[39.0][c]	[−38.6][c]	[33.3][c]
[p-Ac-Phe]RNase S	−13.4	7.05	−3.44	4.90
	[−11.5][c]	[7.00][c]	[−4.00][c]	[5.04][c]

[a]The electric field experienced by the C═O vibration, as defined by eq 4. [b]The electric field drop across the C═O vibration, as defined by eq 5. [c]For the bottom two entries, the set of values in brackets reflect calculations that *do* count the probe-bearing residue's backbone atoms as part of the environment (see text).

simulations, which indicate that the average electric field in water (−65.9 MV/cm) is more than 2 times the next largest field (−29.6 MV/cm, DMSO). Indeed, the stronger H-bonds furnished by water would be expected to exert larger electric fields onto the C═O bond. To further develop the concept that dichloromethane, chloroform, and water create increasingly large fields via H-bonds of increasing strength, we examined the trajectories and identified structural/dynamic properties that correlate with H-bonding capacity and strength. On average, the solvent atom closest to the O-atom of the C═O probe was able to get somewhat closer and stay the closest atom for longer in the H-bonding solvents (see Table S3). These features are both characteristic of H-bonding interactions.

An additional comparison can be made between the observed line widths of the solvatochromic spectra and the standard deviations of the calculated field distributions. These two variables correlate modestly well (with an R^2 of 0.72; see Figure S2 and accompanying analysis) but not perfectly. This is not surprising given that the line width is a dynamical property that depends on field correlations as well as field statistics.[30] Nonetheless, a qualitative connection between the range of the electric field distribution and IR line width is apparent.

We noticed there was a connection between H-bonding and the field drop, $|\Delta F_{vib}|$ (Table 2). For most of the solvents, the ensemble-averaged field drop was rather small, with absolute values clustered around and less than 1 MV/cm. Small field drops imply that the ensemble-averaged field is smooth and shallow, exactly as expected for fields that arise from nonspecific dielectric polarization. On the other hand, a rather steep field drop arises for the weak H-bonding hydrochlorocarbons and an enormous field drop (41 MV/cm) is manifested in water. These results are explained by the intuition that when an H-bond forms between the probe and a donor, it is only the H-bond accepting atom that enters into close contact with the H-bond donor's dipole. In other words, the large field associated with the H-bond is strongly weighted by the contribution from the H-bond accepting atom, resulting in a precipitous field drop with respect to the more distal atom of the vibrational probe.

An important observation about the MD fields is that they are consistently large relative to what one would expect from the solvent-induced frequency shifts and the Stark tuning rate. For example, the calculated field dispersion from *n*-hexane to DMSO is ∼30 MV/cm; however, the observed frequency dispersion over the same span in environments is only 14.4 cm^{-1}, and we expect a frequency shift of 1.05 cm^{-1} for every 1 MV/cm change in electric field from the measured Stark tuning rate. This finding suggests either that the MD simulations systematically overestimate the magnitude of the environment's electric fields or that the Stark tuning rate is overestimated. Possible origins for this discrepancy is discussed further in the following.

2.5. Discussion for Solvents. Plotting the observed C═O peak frequencies in the nine different solvents (Figure 2A) against the ensemble-average electric fields in those corresponding solvents (Table 2), we obtain a remarkable linear trend whose R^2 is 0.99 (Figure 2B). This excellent correlation provides strong evidence in support of the claim that solvation-induced frequency shifts are explained by the average electric fields created by the various solvents.[31] This trend extends equally to cases in which H-bonding is present: both weak H-bonds (chloroform and dichloromethane) and moderately strong H-bonds (water), with their increasingly larger electric fields, are found to produce frequency shifts with a consistently linear pattern. This analysis suggests that C═O's solvent-induced frequency variation can be essentially explained in terms of a model with a single electrostatic parameter (the average electric field, $\langle |F_{vib}| \rangle$), including cases with specific chemical interactions. The same cannot be said for nitriles, for which the field−frequency calibration curve only extends to non-H-bonding cases and for which additional analysis is necessary to first determine that the nitrile is not H-bonded before one can translate a vibrational frequency to a field.[22]

Despite this excellent linear correlation between observed frequency and calculated field, the slope of the regression line is 0.414 cm^{-1}/(MV/cm), which is significantly different from C═O's observed Stark tuning rate of 1.05 cm^{-1}/(MV/cm) (Figure 1 and Table 1). In Figure 2B, we have rescaled the calculated electric fields uniformly by introducing a correction factor of 2.5. Consequently, the MD-calculated fields plotted in Figure 2B ($\langle |F_{vib}| \rangle$/2.5) are 2.5 times smaller than the raw MD-calculated fields ($\langle |F_{vib}| \rangle$) in Table 2. The resultant regression line is $\bar{\nu}_{C=O}$ = 1.036($\langle |F_{vib}| \rangle$/2.5) + 1695.7; importantly, the intercept and the correlation coefficient (0.99) are independent of the rescaling factor.

We have two hypotheses for the origin of the 2.5-fold discrepancy between the simulated fields and the observed Stark tuning rate. Fixed-charge force fields (such as GAFF) represent nuclear degrees of freedom explicitly but do not represent electronic polarizability. In practice, the missing electronic degrees of freedom are compensated for by "prepolarizing" the atomic charges, which are noticeably larger than atomic charges in high-level quantum population analyses.[32] These larger charges would in turn lead to inflated electric fields. On the other hand, polarizable force fields use a more physical description of intermolecular interactions in the condensed phase,[32] so they might predict the scale of solvent fields differently (and presumably, more correctly). Use of polarizable models to calculate electric fields is the subject of ongoing work.

Another source of difference between experiment and theory could come from the local field correction to the Stark tuning rate, which refers to the uncertainty in the magnitude of the electric field experienced by a vibration in a Stark experiment owing to the difference between the externally applied field and the local field at the position of a chromophore.[4,33] The local field at a particular point (where a vibrational probe sits) will in general be somewhat larger (by a factor denoted *f*) than the external field, derived from the distance and applied voltage between the two electrodes, because of the extra contribution arising from the polarization of the surrounding environment.[33,34] The local field effect implies that the experimentally determined Stark tuning rate ($|\Delta\vec{\mu}|f$) will be larger than the microscopic Stark tuning rate ($|\Delta\vec{\mu}|$), which the slope of the field−frequency curve ought to reproduce. Classical electrostatic theory estimates the value of *f* of a frozen glass to be ∼1.3 and largely independent of the glass-former's polarity,[33] although its value could be larger. In summary, the factor of 2.5 discrepancy can be attributed both to overestimation of calculated solvent fields and to the local field correction factor.

A number of studies have now shown that a ∼2-fold correction is needed to harmonize experimental vibrational Stark effect shifts with those calculated with electrostatic models,[1,35,36] so the effect described above appears to be general. This factor was previously found to lead to better agreement between calculated and observed vibrational

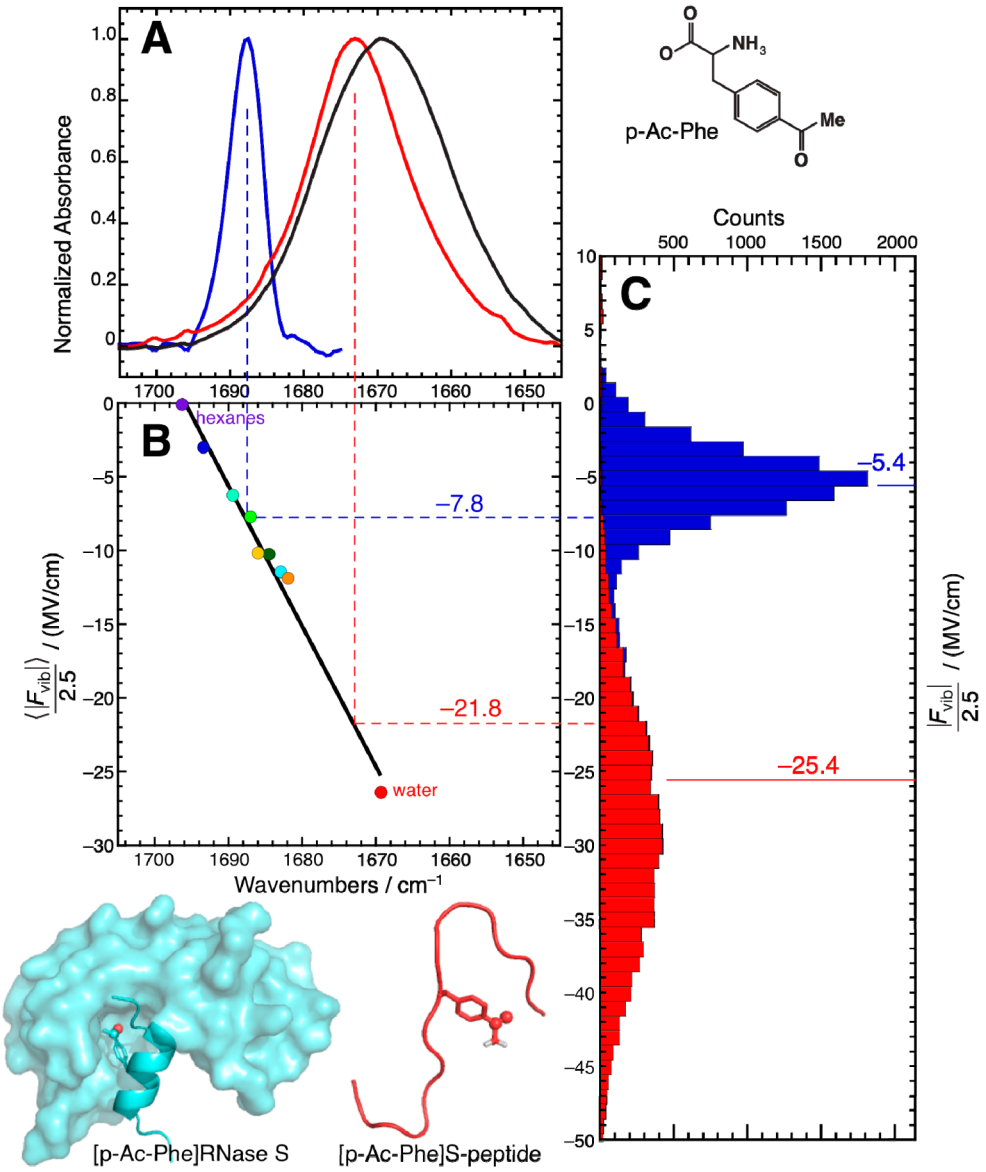


Figure 3. A solvatochromic model calibrates C═O vibrational frequency to electrostatic field. (A) FTIR spectra of p-Ac-Phe (black), [p-Ac-Phe]S-peptide (red), and [p-Ac-Phe]RNase S (blue). (B) The solvatochromic field−frequency correlation plot from Figure 2B (rotated by 90°) with rescaled fields used to convert the peak frequencies of the two spectra in (A) to predictions for ensemble-averaged electric fields (illustrated with dotted lines). This process yields values of −21.8 MV/cm for [p-Ac-Phe]S-peptide and −7.8 MV/cm for [p-Ac-Phe]RNase S. (C) Histograms of $|F_{vib}|/2.5$ from the MD simulations of [p-Ac-Phe]S-peptide and [p-Ac-Phe]RNse S. The mean values (marked with solid lines) are −5.4 and −25.4 MV/cm, which differ from their predicted values by 31% and 16%, respectively.

frequency shifts accompanying mutation and pH change,[1,35] as well as ultrafast vibrational frequency shifts in response to photoinduced charge transfer in a nearby dye.[36] The key result illustrated in Figure 2B is that observed IR frequencies correlate linearly with the computed average value of the electric field, across solvents of very different dielectric and H-bonding properties. Some uncertainty remains regarding the overall scale of solvent fields, the resolution of which awaits further study; nevertheless, meaningful comparisons between solvents and proteins are still possible, as described in the following.

3. RESULTS AND DISCUSSION FOR RIBONUCLEASE S

3.1. Vibrational Probes in RNase S. The FTIR spectrum of p-Ac-Phe in D_2O, 20 mM HEPES, pD = 8.0 is shown in black in Figure 3A. The peak position matched perfectly with that of acetophenone in D_2O (full spectral data given in Table S2). This experiment demonstrates, as expected, that the C═O vibration in acetophenone is an appropriate model for the C═O vibration in p-Ac-Phe and suggests that the two compounds possess identical vibrational frequencies when exposed to identical environments. It further demonstrates that the C═O stretch is indeed very local, and so its frequency is relatively independent of the distal parts of molecule in which it resides. In turn, this condition implies that acetophenone's field−frequency calibration should be transferable onto p-Ac-Phe, as discussed below. With this in mind, we set out to install p-Ac-Phe as a residue into a peptide and a protein to examine biomolecular electrostatic environments.

RNase S is a split protein consisting of a peptide fragment (the S^{15}-peptide, residues 1−15) and a protein fragment (the S-protein, residues 21−124).[37] The S^{15}-peptide is strongly bound to the S-protein by noncovalent interactions[38] such that the

dx.doi.org/10.1021/ja403917z | J. Am. Chem. Soc. 2013, 135, 11181−11192

reassembled two-chain protein retains essentially the same structure and activity as the single-chain protein, RNase A.[39] Residues 16−20 are not as important for binding, structure, or function. As reported previously,[35] RNase S is an ideal system for facile insertion of vibrational probes because the S^{15}-peptide can be exploited to deliver a probe through an unnatural amino acid. In the present case, p-Ac-Phe was incorporated into a synthetic S^{15}-peptide by replacing the natural Phe at position 8 with it. We refer to this construct as [p-Ac-Phe]S-peptide (Figure 3). Isolated S^{15}-peptide is preponderantly a random coil in aqueous solution,[40] so the C═O probe from p-Ac-Phe in the S^{15}-peptide is expected to be highly exposed to water. In contrast, when [p-Ac-Phe]S-peptide is combined with S-protein, the RNase S complex is formed (referred to as [p-Ac-Phe]RNase S, Figure 3) wherein the C═O is placed in a highly hydrophobic pocket. The procedure for preparing and purifying [p-Ac-Phe]RNase S is described in the experimental methods section 5.2. Kinetics and thermodynamic parameters for RNase S catalysis and reassembly are provided in Table 3.

Table 3. Kinetic and Thermodynamic Parameters for [p-Ac-Phe]RNase S Compared to Those of Wild-Type RNase

	wild-type RNase[a]	[p-Ac-Phe]RNase S
	Catalytic Activity[b]	
k_{cat}, s^{-1}	2.8 ± 0.1	0.8 ± 0.03
K_M, mM	1.25 ± 0.1	1.37 ± 0.2
R^2	0.994	0.991
	S-Peptide Binding Thermodynamics[c]	
K_D	110 ± 20 nM[d]	12.7 ± 1 μM
$\Delta G°$, kcal/mol	−9.5 ± 0.1 [d]	−6.7 ± 0.1
$\Delta H°$, kcal/mol	−39.3 ± 0.6 [d]	−16.7 ± 0.1
$\Delta S°$, cal/(mol K)	−100 [d]	−33.6

[a]Wild-type RNase refers to single-chain RNase A for activity measurements and the split-protein RNase S for binding measurements. [b]Michaelis−Menten kinetics for the hydrolysis of cyclic cytidine monophosphate. See Figure S3 for full data and analysis. [c]Thermodynamics for the binding of S-peptide to S-protein to form RNase S at 25 °C as determined by isothermal titration calorimetry. See Figure S6 for full data and analysis. [d]Data from ref 38.

Additional spectroscopic characterization and full data for the enzymological and calorimetric measurements are described in the Supporting Information (Figures S3−S6 and accompanying analysis).

The FTIR spectra for the two probe-bearing constructs ([p-Ac-Phe]S-peptide, red; [p-Ac-Phe]RNase S, blue) are displayed in Figure 3A. The C═O band position in the peptide is at 1673.1 cm^{-1} (close to the value of acetophenone in water) and very broad (14.7 cm^{-1}). Both of these features signify H-bonding to water molecules. On the other hand, there is a large shift to the blue (to 1687.6 cm^{-1}) and band narrowing (to 5.6 cm^{-1}) accompanying the replacement of water around p-Ac-Phe with a protein environment. These features are consistent with the nonpolar and structurally defined milieu characteristic of a protein interior. Under the same conditions except employing a nitrile as a probe (i.e., p-CN-Phe at position 8),[22] we also observed a band narrowing upon formation of [p-CN-Phe]RNase S. However, unlike the C═O probe which underwent a large blue shift, the C≡N probe shifted 4.0 cm^{-1} to the red. This shift was interpreted as a superposition of two competing factors: a blue shift due to lower electric fields and a larger red shift due to removing an H-bond.[22,41] In contrast, it appears that the C═O shift in RNase S can be rationalized purely with electrostatic arguments and without invoking a competing effect arising from H-bonding, highlighting the advantage of C═O probes when large changes in environment occur.

3.2. MD Calculation of Fields in RNase S. MD simulations were carried out on [p-Ac-Phe]S-peptide and [p-Ac-Phe]RNase S, and ensemble-average electric fields were calculated as described in computational methods section 4.3. In both cases, the calculated electrostatic fields are readily understood from qualitative arguments about the structures of the two systems and the IR frequencies observed for them. Just as [p-Ac-Phe]S-peptide places the carbonyl probe in a largely aqueous environment, the average field calculated for it (−63.6 MV/cm; −25.4 MV/cm rescaled by 2.5) is quite close to the field acetophenone experiences in water. Furthermore, the calculations validate the prediction that the hydrophobic environment associated with a protein core would lead to a significant attenuation in the electric field. The calculated field experienced by the probe in [p-Ac-Phe]RNase S is −13.4 MV/cm (−5.4 MV/cm rescaled), similar to that of THF. These two extra points fall squarely along the field−frequency curve spanned by the solvatochromic series, and their inclusion only lowers R^2 from 0.99 to 0.97. This point supports the idea that proteins do not introduce any new idiosyncratic contributions to the vibrational frequency that cannot simply be attributed to electrostatics. Another interesting observation is that the pronounced narrowing of p-Ac-Phe's C═O band (Figure 3A) upon formation of the RNase S complex is matched with a narrowing in the calculated electric field distribution (Figure 3C), supporting the link between field heterogeneity and line width established with the solvent series.

3.3. Discussion for RNase S. Our interest in solvation fields and solvent-induced frequency shifts stems from the possibility that the information can be used to construct a field−frequency calibration curve to semiempirically convert C═O frequencies measured in proteins into the protein's local electrostatic fields projected onto the probe's bond.[22] This strategy effectively applies eq 1 in the form of the regression line in Figure 3B. By using the peak frequencies found for the C═O probe in [p-Ac-Phe]S-peptide and in [p-Ac-Phe]RNase S and applying the solvatochromic model described above, the average electric field for the C═O probe in these two environments is predicted to be −21.8 and −7.8 MV/cm, respectively. The field-to-frequency conversion is illustrated in Figure 3. These predictions compare favorably with the average electric fields calculated for those two systems from direct MD simulation (−25.4 and −5.4 MV/cm, respectively), differing on average by a factor of 1.2 (note that this comparison is independent of the value chosen for the correction factor). It is difficult to say whether the discrepancy between fields calculated by direct simulation versus by semiempirical field−frequency mapping is due to force field accuracy, sampling error, or limitations in the solvatochromic calibration scheme. In any event, the correspondence is relatively robust and invites the use of solvatochromic models to predict electric fields in cases where they would not otherwise be possible to calculate because of computational limitations arising from system size, sampling requirements, or force field accuracy.

Although C═O probes in proteins are not technically new, following early work on myoglobin[15,19,42] they have been largely discarded and viewed as intractable owing to their overlap with proteins' amide I bands. Indeed, as local

vibrational probes have seen greater use in protein science, possessing an absorption band in an uncluttered region has become accepted as sine qua non;[3,5,7−11] the present work employing C═O in RNase S represents a departure from that axiom. We emphasize that C═O vibrations *are* observable amidst the protein background if one carefully selects a reference sample that is nearly identical to the vibrational-probe-bearing construct and leverages difference spectroscopy to bring buried bands into view. From a practical perspective, the experimental approach was *not* greatly limited by the transmission level in the amide I range but mostly by how precisely the sample's and the reference's spectra were matched. We therefore expect this approach to generalize to larger and more complicated proteins, especially if one employs more sophisticated forms of difference spectroscopy that employ isotopic replacement (used to locate the buried N≡O band in NO-myoglobin[43]) or photoactivation (used to locate buried vibrational bands in the reaction center[44] and bacteriorhodopsin[45]). With this in mind, we believe C═O vibrational probes constitute a powerful tool to interrogate the whole range of biomolecular processes such as binding, catalysis, and self-assembly because these are almost always mediated by H-bonds.

4. COMPUTATIONAL METHODS

4.1. Solvation Simulation Methods. Acetophenone was placed at the center of a cubic box filled with solvent molecules. The size of the box was such that the distance between acetophenone and the box edge was 1.0 nm; between 370 molecules (for *n*-hexane) and 3500 molecules (for water) of solvent were required to fill the volume. The solvated system was energy-minimized by 1000 steps of steepest descent, then equilibrated for 100 ps (2 fs time step) in an NPT ensemble with the temperature set to 300 K and the reference pressure set to 1 bar. The equilibration procedure employed the Bussi thermostat,[46] the Berendsen barostat, and the particle mesh Ewald (PME) method[47] for calculating long-range electrostatic interactions. The cutoff distances for Coulomb summation and for van der Waals interactions were both set to 1.0 nm. We constrained all bond vibrations with the LINCS algorithm.[48] Production dynamics were carried out for 2 ns (2 fs time step) in the NPT ensemble continuing from the final coordinates and velocities of the equilibration run. The temperature was set to 300 K, and the reference pressure was set to 1 bar utilizing the Parrinello−Rahman barostat.[49] These simulations employed a stochastic dynamics (SD) integrator, as implemented in GROMACS.[50] Stochastic dynamics were chosen as a temperature-coupling mechanism because it allows equilibrium properties of the thermal ensemble to converge more quickly. A relatively short trajectory of 2 ns was deemed sufficient to capture the average electrostatic field of a bulk solvent because the dominant relaxation components operative in solvation are generally very fast (on the order of the Debye time scale, ∼10 ps).[51] All other settings were the same as the equilibration run.

4.2. Calculation of Solvation Fields. A new method was developed to calculate the electric field exerted onto the vibrational probe from each snapshot. Snapshots were recorded every 100 steps (200 fs) and consisted of full-precision coordinates and forces. First, the total force on the C-atom and on the O-atom of the carbonyl in acetophenone was extracted for each snapshot. Then a charge-neutralized topology file was generated for the system in which the partial charges for all of the solvent atoms are set to zero (but all of the solute's atoms retain the same charges as used during dynamics) and all noncharge parameters are kept identical. The trajectory from the production dynamics was postprocessed with the charge-neutralized topology using GROMACS' rerun utility. In the resulting trajectory, different forces are present on each atom, due to the absence of any intermolecular electrostatic interaction. Likewise, the total force on the C and O atoms was extracted from each snapshot. With this information, the total electric field experienced by the vibration due to the environment in each snapshot was calculated with eqs 2−4.

$$\vec{f}^{\,i}_{\text{electro}} = \vec{f}^{\,i}_{\text{tot}} - \vec{f}^{\,i}_{\text{nonelectro}} \tag{2}$$

$$\vec{F}^{\,i} = \vec{f}^{\,i}_{\text{electro}} / q_i \tag{3}$$

$$|F_{\text{vib}}| = \frac{1}{2}(\vec{F}^{\text{C}}\cdot\hat{u}_{\text{CO}} + \vec{F}^{\text{O}}\cdot\hat{u}_{\text{CO}}) \tag{4}$$

In eqs 2−4, i is indexed over the C-atom and the O-atom of the carbonyl probe, $\vec{f}$ denotes force, and $\vec{F}$ denotes electric field. The subtraction of all nonelectrostatic forces (as determined by rerunning the trajectory with the charge-neutralized topology) from the total force results in a force exerted on a particular atom due only to electrostatic interactions (eq 2). The electrostatic force can be converted into an electrostatic field (eq 3) simply by dividing by the partial charge of the atom in question (either C or O of acetophenone or p-Ac-Phe). Finally, the electric field "experienced" by the vibration, $|F_{\text{vib}}|$, is calculated by projecting the field at either C or O onto the unit vector defining the vibration's bond axis and then averaging the two field projections between the two atoms (eq 4). $|F_{\text{vib}}|$, as defined in this way, is operationally equivalent to $\vec{F}\cdot\hat{u}_{\text{probe}}$ in eq 1. The electric field experienced by the vibration is averaged over 10 001 snapshots taken over 2 ns to obtain the ensemble-averaged electric field, $\langle|F_{\text{vib}}|\rangle$. Other statistical measures of the field distribution, such as standard deviation ($\sigma_{|F_{\text{vib}}|}$), are calculated as well. Additionally, we calculated the electric field drop over the carbonyl bond, $|\Delta F_{\text{vib}}|$, for each snapshot. This calculation involves the same first two steps as used to calculate $|F_{\text{vib}}|$, but the step in eq 4 is replaced with eq 5:

$$|\Delta F_{\text{vib}}| = \vec{F}^{\text{O}}\cdot\hat{u}_{\text{CO}} - \vec{F}^{\text{C}}\cdot\hat{u}_{\text{CO}} \tag{5}$$

Similar to $|F_{\text{vib}}|$, the ensemble average and standard deviation of the field drop distribution were calculated. Python scripts used to streamline these calculations are provided in the Supporting Information (section S3.2).

4.3. Ribonuclease Simulation Methods. We calculated ensemble-averaged fields and field drops for the C═O probe in the S-peptide system and the RNase S system, using methods similar to those described for the simple solvents. All simulations employed the AMBER99SB-ILDN force field[52] to describe the protein, and the TIP3P water model to describe the solvent. The p-Ac-Phe residue was parametrized and added to the force field using a simple procedure because of its high similarity to Phe. The details of this procedure are provided in the Supporting Information (section S3.1).

To simulate [p-Ac-Phe]RNase S, the 1.5 Å resolution crystal structure of the analogue [p-CN-Phe]RNase S (PDB code 3OQY[35]) was used to generate starting coordinates. To obtain starting coordinates for [p-Ac-Phe]S-peptide, the coordinates of the first chain from the corresponding RNase S structure

were taken. In both cases, the PDB2GMX utility with its default options was employed to protonate the starting structures, assign disulfide linkages, and assign protonation states to ionizable moieties except that all histidines were manually entered as protonated at the ϵ-nitrogen, as expected at pH 8.0 (the condition for all experiments).[53] The protein was solvated in a periodically replicated dodecahedral box filled with water, ensuring that the edges of the box were at least 1.0 nm separated from the ends of the protein. The system was neutralized and brought to a net ion concentration of ~20 mM (the experimental condition) by addition of sodium and chloride ions. Conditions for minimization and equilibration were similar to those stated for the solvation systems; vide supra. For the RNase S systems, equilibration was conducted first for 100 ps with 1000 kJ mol^{-1} nm^{-1} positional restraints applied isotropically to all protein atoms and then for 20 ns without positional restraints. The long equilibration time allowed the local structure of the S-protein to relax around the acetyl group appended onto Phe8 of the S^{15}-peptide. For the S-peptide system, equilibration runs were all 100 ps long. The second equilibration without positional restraints was conducted at 400 K and then at 300 K. The high-temperature equilibration step was meant to give the S-peptide an opportunity to relax nonintrinsic helical structure induced by its association with the S-protein.

Production dynamics were again performed in an NPT ensemble, with a stochastic integrator used for temperature-coupling. The only major difference was that the simulations were 20 ns long (rather than 2 ns), and full-precision snapshots with coordinates and forces were recorded every 1000 steps (rather than every 100 steps). Electric fields on the C and O atoms of the vibrational probe were determined in each snapshot and used to calculate the electric field experienced by the vibration (eq 4) and the electric field drop over the bond (eq 5) as for the solvent simulations. To create the charge-neutralized topology file for these system, the charges on all the water molecules, the ions, and the residues other than the probe-containing residue were set to zero.

4.4. Additional Considerations of the Current MD Approach to Calculating Fields. The methodology we developed here differs somewhat from previous strategies to extract electric fields from MD trajectories, which have relied on calculating the Coulombic force on a virtual test-charge particle inserted into the MD trajectory.[35,54,55] A problem with this method is that the calculation of the Coulombic interaction omits interactions arising from outside the main simulation box in a simulation with periodic boundary conditions (personal communication, J. Chodera). Calculation of forces in eq 2 using GROMACS' built-in functions ensures that direct Coulomb and particle mesh Ewald contributions are calculated and combined (as they would be during normal dynamics) for the purposes of calculating electric fields.

Another important consideration that arises in these calculations is the choice of a zero-field reference state against which one chooses to define the electric field. We have chosen to define the electric field as the total field impinging onto a target atom from its environment. The environment refers to all atoms that are *not* part of the same molecule as the target atom. What is intentionally excluded from this definition is the self-field an atom experiences due to a molecule's own nuclei and electronic density. By application of this definition, the electric field on any atom of any molecule in the gas phase is zero. The purpose for this exclusion is that we are ultimately interested in utilizing the electric field as a descriptor for *inter*molecular interactions. This description facilitates a mapping onto intermolecular energetics so long as the interactions in question are electrostatic in nature. On this point, we note that the H-bonding interactions furnished by water and chlorohydrocarbons appear to be well explained purely in terms of electrostatics, as has been noted for weak H-bond complexes as well.[56] Implicit in our definition of the total field is a strict demarcation between a molecule and its environment, which is not enforceable in a rigorous quantum mechanical formalism but which is readily adapted to MD force fields that systematically separate bonding (within the same molecule) from nonbonding (between molecules) interactions.

In practice, the valence terms in MD force fields (bonds, angles, and torsions) are only meant to capture local contributions to the bonding potential energy curve, and electrostatic interactions are allowed between two atoms of the same molecule if they are separated by a certain number of bonds. This treatment of intramolecular electrostatics is meant to approximately recapture long-range quantum mechanical interactions and is an essential component of MD force fields. However, these contributions lie outside our chosen definition of the electric field as being limited to those due to the environment. A simple and general way to remove the intramolecular electrostatics from our electric field calculations is to retain the solute's atomic charges in the charge-neutralized topology (we only zero the solvent's charges) so that the intramolecular electrostatics cancels out in the difference of eq 2. This demarcation becomes subtler for calculations on proteins, where we must arbitrate where the probe molecule or fragment ends and the protein environment (even though it may be on the same molecule) begins. We chose to discount the self-field arising from the probe-bearing amino acid residue and considered the other residues as well as solvent as the environment. Another possibility would be to additionally count the field arising from the probe-bearing residue's backbone (see values in brackets in Table 2), which also gave reasonable (and not very different) results. The definition we chose is purely operational and highlights a difficulty with unnatural amino acid based probes (the same difficulty would not be encountered when a vibrational probe is delivered on a ligand). One reason why the current electric field calculation method presented here is preferable is that it provides a reliable way to control what interactions are to be included in the field by choosing which charges to zero out in the charge-neutralized topology.

4.5. Control Simulations for Solvation Electrostatics. To calculate solvation fields, the methodology described above used stochastic dynamics and the RESP-fitting method to determine solvent atomic charges, as carried out by Caleman et al.[27] We followed up those simulations with three additional control scenarios to test how sensitive the resultant ensemble-average electric fields were to certain choices. In the first control, we used deterministic dynamics in place of stochastic dynamics to test whether the random force terms in the Langevin equation appreciably change the ensemble-average electrostatic quantities (Table S4). In the second control, we used the simplified AM1-BCC method[26] (as implemented by Antechamber in AmberTools 12) to assign the charges to the solvent molecules' atoms rather than the RESP fitting procedure used by Caleman et al. (Table S5).[27] This control was designed to examine the accuracy of the AM1-BCC method by testing how well it can recapitulate electrostatic

quantities derived from RESP charges. In a final third control, we replaced all the (RESP) solvent charges with their values multiplied by 0.9 (Table S6). This control was designed to test the sensitivity of the calculated electric fields on the charge parametrization. Given that there is some error in any force field's charge parameters, this simulation sheds light on what degree that inaccuracy endangers the dependability of the calculated electric fields. All of these modifications resulted in minor changes to ensemble-average solvation fields (see Tables S4−S6 and accompanying discussion). An important conclusion from these control simulations is that the overall magnitude of the MD-calculated fields (which required the 2.5 rescaling factor for the slope of the field−frequency curve to agree with the measured Stark tuning rate) was not a consequence of a particular choice of how to parametrize the charges or run the dynamics.

5. EXPERIMENTAL METHODS

5.1. Materials. Liquid acetophenone, 99% (Sigma Aldrich), was purchased and used without further purification. Anhydrous solvents including deuterium oxide (99.5% or higher) were purchased from Acros Organics and were used without further purification. Ribonuclease A from bovine pancreas and subtilisin A from *Bacillus licheniformis* (Sigma) were purchased and used without further purification.

5.2. Synthesis of [*p*-Ac-Phe]RNase S and Precursors. *N*-Fmoc-L-(*p*-acetyl-Phe) was purchased from Peptech and used without further purification. To convert the N-protected amino acid into the free amino acid, we perform a solution-phase base-catalyzed Fmoc removal reaction with piperidine according to previously established procedure.[22] Material derived from this reaction was recovered in 75% yield and was verified by ESI mass spectrometry and ^{1}H NMR in D_2O. ^{1}H NMR (D_2O, 300 MHz): δ 7.83 (d, J = 8 Hz, 2H, Hε), 7.29 (d, J = 8 Hz, 2H, Hδ), 3.87 (dd, J = 6 Hz, 5 Hz, 1H, Hα), 3.0−3.2 (m, 2H, Hβ), 2.51 (s, 3H, Hθ). LCMS: mass calcd for [M] $C_{11}H_{13}NO_3$ is 207.2. Found: (+) 209.18 [M + 2] and (−) 206.16 [M − 1].

To prepare [p-Ac-Phe]S-peptide (KETAAAK[p-Ac-Phe]-ERQHMDS), *N*-Fmoc-L-(*p*-acetyl-Phe) was transferred to Elim BioPharm (Hayward, CA) which employed standard Fmoc-based solid-phase peptide synthesis to generate the S-peptide. The material was verified by HPLC and by ESI mass spectrometry. HPLC: single peak detected by A_{220}, eluting at 8 min. LCMS: mass calcd for [M] $C_{75}H_{119}N_{23}O_{26}S$ is 1790.9. Found: (+) 1791.9 [M +1].

The S-protein fragment was isolated by combining 100 mg of RNase A (20 mg/mL) with 40 μL of subtilisin (10 mg/mL) in 100 mM Tris (pH 8.0) and letting the digest proceed overnight on ice. Afterward, the solution was acidified to pH 2 via dropwise addition of concentrated HCl and purified by HPLC as described previously.[35] To prepare [p-Ac-Phe]RNase S, [p-Ac-Phe]S-peptide (1.5 equiv) and the lyophilized S-protein fragment were dissolved in 20 mM HEPES (pH 8.0) and combined. The resultant solution was purified by cation exchange chromatography with a Hi-Trap SP XL column (GE Healthcare) on an FPLC setup (buffer A = 20 mM HEPES; buffer B = 20 mM HEPES, 1.0 M NaCl), running a gradient from 0−100% B over 15 column volumes. The S-peptide did not bind to the column, and the purified RNase S complex eluted between 15% and 28% B. The material was verified by ESI mass spectrometry. LCMS: two peaks eluted from the LC column at 4.64 min (S-peptide) and 6.90 min (S-protein). The RNase S complex is not stable under the acidic conditions of the LC column, but the existence of the S-peptide peak indicates that the RNase S complex was intact to begin with because S-peptide needed to be bound to S-protein to coelute during cation exchange. Mass calcd for S-peptide 1791; for S-protein 11542. Found at 4.64 min: (+) 1790. Found at 6.90 min: (+) 11 537. (Masses reflect the most likely mass as determined by maximum entropy deconvolution.) Additional biophysical characterization of the [p-Ac-Phe]RNase S construct is provided in the Supporting Information: UV−vis spectroscopy (Figure S4), circular dichroism spectroscopy (Figure S5), and isothermal titration calorimetric measurements of the S-peptide/S-protein binding thermodynamics (Figure S6).

5.3. FTIR Spectroscopy. All spectra were recorded on a Bruker Vertex 70 FTIR spectrometer in a fashion very similar to that described previously.[17] In general, samples were loaded into a demountable liquid cell (Bruker) with two windows (CaF_2, 0.750 in. thick, Red Optronics). The windows were separated using two offset semicircular mylar spacers (75 and 100 μm for solvatochromism; 50 and 75 μm for RNase S and precursors). For the solvatochromism experiments, a 5 min delay was applied to purge gaseous CO_2, and then 64 scans were acquired and averaged to obtain each transmission interferogram. For experiments on RNase S and its precursors, all conditions were kept constant except the purging time was increased to 10 min and 256 scans were taken. Band positions and fwhm values were calculated using the OPUS software's peak picker (Bruker) (which uses a second-derivative-based method) as well as fitting the curve with a Levenberg−Maquardt algorithm.

A background transmission spectrum was subtracted from the sample's transmission to obtain absorption spectra. For the solvatochromism experiments, the background was simply the pure solvent itself without acetophenone. For the experiments on RNase S and its precursors, backgrounds were by necessity chosen more judiciously: (1) for p-Ac-Phe, the background was phenylalanine at the same concentration (5 mM); (2) for [p-Ac-Phe]S-peptide, the background was [p-CN-Phe]S-peptide at the same concentration (5 mM); (3) for [p-Ac-Phe]RNase S, the background was [p-CN-Phe]RNase S at the same concentration (2.5 mM). All of these experiments were conducted in the same buffer system, namely, 20 mM HEPES in D_2O, pD = 8.0.

5.4. Vibrational Stark Spectroscopy. Vibrational Stark spectra were recorded as previously described.[4] Briefly, acetophenone was dissolved in the organic solvent to a concentration of 50 mM, loaded into a custom-made sample cell consisting of CaF_2 windows coated with a 4 nm thick layer of nickel metal, and frozen in liquid nitrogen using a custom-designed cryostat. The sample serves as a dielectric between the two conducting plates, which are connected to a high-voltage power source; in other words, the sample cell acts overall as a parallel-plate capacitor. Fields were applied in the range 0.2−1.0 MV/cm. Stark spectra were determined by calculating the difference between the field-on and field-off transmission spectra (64 scans apiece) and were repeated at three different field magnitudes. The linear Stark tuning rates were determined from a numerical fit of the Stark spectra with derivatives of the experimental absorption spectra.

Because of the residual dielectric response of a frozen solvent, the effective field exerted onto the solute is larger than the applied field by a small factor f, the local field correction factor.[33] For a particular frozen glass, f is expected to be a constant, and its value is estimated to be between 1.1 and 1.4. Because f is not independently measurable, we report the observed Stark tuning rate as a product between the true Stark tuning rate and the local field correction factor, $|\Delta\vec{\mu}|f$.

5.5. Enzymology. Kinetics traces were obtained by measuring the UV absorption at 286 nm every second for 5 min on a Lamda 25 UV−vis spectrometer (Perkin-Elmer) at room temperature. To obtain a trace, 50 μL of RNase (6 μM) in 20 mM Bis-Tris, 1.0 mM EDTA, pH 6.0, was pipetted into an Eppendorf tube containing 250 μL of cyclic cytidine monophosphate (cCMP, Sigma-Aldrich) in 20 mM Bis-Tris, 1.0 mM EDTA, pH 6.0, at concentrations ranging between 0.12 and 4.8 mM. The mixture was quickly aspirated, transferred to a 1.0 mm quartz cuvette, and placed into the spectrometer cuvette holder. The final mixture comprised RNase at 1 μM and substrate at 0.1−4.0 mM. The initial rate was determined by calculating the slope of the least-squares regression line of the absorbance vs time data during the second minute and converting from OD/min to mM/s using $\Delta\varepsilon_{286}$(cCMP−CMP) = 1.22 cm^{-1} mM^{-1}, as determined by letting an RNase-catalyzed reaction go to completion running overnight. The concentration of protein was determined using ε_{280}(RNase) = 9.63 cm^{-1} mM^{-1} (ExPASy protein parameter server). The nonlinear fit of the data to the Michaelis−Menten equation was carried out in Kaleidagraph.

5.6. Calorimetry. Isothermal titration calorimetry (ITC) was used to measure the binding thermodynamics of [p-Ac-Phe]S-peptide to S-protein. The titration was carried out in a buffer of 50 mM sodium acetate, 100 mM sodium chloride, pH 6.0 at 25.0 °C, following previous conditions.[38] S-protein solution was desalted and equilibrated into buffer using a PD-10 column (GE Healthcare), then diluted to 0.25 mM. An amount of 2 mL of this solution was used to fill the cell of a MicroCal VP-ITC system (GE Healthcare). [p-Ac-Phe]S-peptide was dissolved in 2 M triethylammonium acetate and lyophilized (to remove excess protons and prevent pH drops) and redissolved into the same buffer at a concentration of 4.5 mM, 300 μL of which was used to load the ITC microsyringe. The titration proceeded with 60 injections of 5 μL apiece, allowing 270 s between injections. The data were processed using MicroCal LLC's Origin package.

6. CONCLUDING REMARKS

We have assembled data from FTIR spectroscopy, vibrational Stark spectroscopy, and MD simulation to validate C═O vibrations for use as electrostatic probes in molecular biophysics. Our results show that the C═O vibrational frequency varies linearly with electrostatic field in a wide range of conditions: from nonpolar solvents and protein hydrophobic cores to the highly polar environment of an H-bond. These observations suggest that C═O frequencies can be converted into ensemble-average electrostatic fields using a linear model (eq 1). The present findings relied on the use of atomistic methods to calculate electric fields. We found that these methods capably represent the electrostatic consequences that accompany specific chemical interactions, a significant advantage over continuum descriptions of electrostatics (see Figure S7 and accompanying analysis). Computational models and vibrational Stark effect measurements are in good agreement up to an overall scaling factor of 2.5. Moreover, the need to invoke this rescaling parameter was independent of several choices in the simulation methodology (see computational methods section 4.5). The disagreement almost certainly reflects the role of polarization effects, which have not been accounted for in the solvent field calculations (which used a fixed-charge model) or in the Stark tuning rate (which is influenced by the local field correction[57]). We surmise that the burden of error is borne by both of these factors, which points to the necessity of including explicit polarization to reproduce electrostatics at the level of detail accessible to vibrational spectroscopic measurements. Therefore, we anticipate that vibrational probes will continue to play an important role in benchmarking electrostatic calculations. Finally, while RNase S served as a useful testing ground to assess the ability of a solvatochromic calibration scheme to measure electric fields in proteins with vibrational probes, the concepts developed here can be more generally applied to measure fields more intimately involved in function, which will be described shortly.

■ ASSOCIATED CONTENT

Supporting Information

IR data summaries, MD control simulations, IR line width analysis, biophysical characterization of RNase S including circular dichroism, isothermal titration calorimetry and analysis, parametrization of the unnatural amino acid, and scripts for implementing electric field calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

sboxer@stanford.edu

Present Address

[†]S.B.: Physical and Materials Chemistry Division, National Chemical Laboratory (CSIR), Pune 411008, India.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

S.D.F. thanks the NSF for a predoctoral fellowship and the Stanford Bio-X interdisciplinary graduate fellowship program for support. This work was supported in part by a grant from the NIH (Grant GM27738). MD simulations were carried out on Stanford's Certainty cluster, which is financially supported by the NSF (Award 0960306), under the American Recovery and Reinvestment Act of 2009 (Public Law 111-5). We thank Prof. V. S. Pande, Prof. J. Chodera, and Dr. L.-P. Wang for advice and feedback on the MD simulations.

■ REFERENCES

(1) Fafarman, A. T.; Sigala, P. A.; Schwans, J. P.; Fenn, T. D.; Herschlag, D.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, E299−E308.
(2) Kim, H.; Cho, M. *Chem. Rev.* [Online early access]. DOI: 10.1021/cr3005185. Published Online: May 16, **2013**.
(3) Lindquist, B. A.; Furse, K. E.; Corcelli, S. A. *Phys. Chem. Chem. Phys.* **2009**, *11*, 8119−8132.
(4) Andrews, S. S.; Boxer, S. G. *J. Phys. Chem. A* **2000**, *104*, 11853−11863.
(5) Suydam, I. T.; Boxer, S. G. *Biochemistry* **2003**, *42*, 12050−12055.
(6) Andrews, S. S.; Boxer, S. G. *J. Phys. Chem. A* **2002**, *106*, 469−477.
(7) Waegele, M. M.; Culik, R. M.; Gai, F. *J. Phys. Chem. Lett.* **2011**, *2*, 2598−2609.
(8) Bazewicz, C. G.; Lipkin, J. S.; Smith, E. E.; Liskov, M. T.; Brewer, S. H. *J. Phys. Chem. B* **2012**, *116*, 10824−10831.
(9) Zimmerman, J.; Thielges, M. C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2011**, *50*, 8333−8337.
(10) Waegele, M. M.; Tucker, M. J.; Gai, F. *Chem. Phys. Lett.* **2009**, *478*, 249−253.
(11) Urbanek, D. C.; Vorobyev, D. Y.; Serrano, A. L.; Gai, F.; Hochstrasser, R. M. *J. Phys. Chem. Lett.* **2010**, *1*, 3311−3315.
(12) Choi, J.-H.; Oh, K.-I.; Lee, H.; Lee, C.; Cho, M. *J. Chem. Phys.* **2008**, *128*, 134506.
(13) Aschaffenburg, D.; Moog, R. *J. Phys. Chem. B* **2009**, *113*, 12736−12743.
(14) Bagchi, S.; Falvo, C.; Mukamel, S.; Hochstrasser, R. *J. Phys. Chem. B* **2009**, *113*, 11260−11273.
(15) Merchant, K. A.; Noid, W. G.; Akiyama, R.; Finkelstein, I. J.; Goun, A.; McClain, B. L.; Loring, R. F.; Fayer, M. D. *J. Am. Chem. Soc.* **2003**, *125*, 13804−13818.
(16) Choi, J.-H.; Cho, M. *J. Chem. Phys.* **2011**, *134*, 154513.
(17) Fried, S. D.; Boxer, S. G. *J. Phys. Chem. B* **2012**, *116*, 690−697.
(18) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 56−61.
(19) Park, E. S.; Andrews, S. S.; Hu, R. B.; Boxer, S. G. *J. Phys. Chem. B* **1999**, *103*, 9813−9817.
(20) Park, E. S.; Boxer, S. G. *J. Phys. Chem. B* **2002**, *106*, 5800−5806.
(21) We found that the Stark tuning rate for H-bonded and non-H-bonded nitriles is also nearly identical (Levinson, Bagchi, and Boxer; results to be published). The important difference emphasized below is that carbonyl frequencies respond linearly to electric fields in both H-bonding and non-H-bonding environments; this is *not* the case for nitrile frequencies (see refs 12, 16, 22, 41).
(22) Bagchi, S.; Fried, S. D.; Boxer, S. G. *J. Am. Chem. Soc.* **2012**, *134*, 10373−10376.
(23) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(24) Kaminski, G.; Jorgensen, W. L. *J. Phys. Chem.* **1996**, *100*, 18010−18013.

(25) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(26) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623−1641.
(27) Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D. *J. Chem. Theory Comput.* **2012**, *8*, 61−74.
(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(29) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(30) Anderson, P. W. *J. Phys. Soc. Jpn.* **1954**, *9*, 316−339.
(31) Levinson, N. M.; Fried, S. D.; Boxer, S. G. *J. Phys. Chem. B* **2012**, *116*, 10470−10476.
(32) Ren, P.; Wu, C.; Ponder, J. W. *J. Chem. Theory Comput.* **2011**, *7*, 3143−3161.
(33) Bublitz, G. U.; Boxer, S. G. *Annu. Rev. Phys. Chem.* **1997**, *48*, 213.
(34) Kohler, B. E.; Woehl, J. C. *J. Chem. Phys.* **1995**, *102*, 7773−7781.
(35) Fafarman, A. T.; Boxer, S. G. *J. Phys. Chem. B* **2010**, *114*, 13536−13544.
(36) Jha, S. K.; Ji, M.; Gaffney, K. J.; Boxer, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16612−16617.
(37) Richards, F. M.; Vithayathil, P. J. *J. Biol. Chem.* **1959**, *234*, 1459.
(38) Connelly, P. R.; Varadarajan, R.; Sturtevant, J. M.; Richards, F. M. *Biochemistry* **1990**, *29*, 6108−6114.
(39) Kim, E. E.; Varadarajan, R.; Wyckoff, H. W.; Richards, F. M. *Biochemistry* **1992**, *31*, 12304−12314.
(40) Klee, W. A. *Biochemistry* **1968**, *8*, 2731−2736.
(41) Fafarman, A.; Sigala, P.; Herschlag, D.; Boxer, S. *J. Am. Chem. Soc.* **2010**, *132*, 12811−12813.
(42) Phillips, G. N.; Teodoro, M. L.; Li, T.; Smith, B.; Olson, J. S. *J. Phys. Chem. B* **1999**, *103*, 8817−8829.
(43) Park, E. S.; Thomas, M. R.; Boxer, S. G. *J. Phys. Chem. B* **2000**, *122*, 12297−12303.
(44) Gerwert, K.; Souvignier, G.; Hess, B. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9774−9778.
(45) Hienerwadel, R.; Grzybek, S.; Fogel, C.; Kreutz, W.; Okamura, M. Y.; Paddock, M. L.; Breton, J.; Nabedryk, E.; Mäntele, W. *Biochemistry* **1995**, *34*, 2832−2843.
(46) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
(47) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
(48) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116−122.
(49) Parrinello, M.; Rahman, A. *Phys. Rev. Lett.* **1980**, *45*, 1196−1199.
(50) Van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173−185.
(51) Stratt, R.; Maroncelli, M. *J. Phys. Chem.* **1996**, *100*, 12981−12996.
(52) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins: Struct., Funct., Bioinf.* **2010**, 1950−1958.
(53) Quirk, D. J.; Raines, R. T. *Biophys. J.* **1999**, *76*, 1571−1579.
(54) Suydam, I. T.; Snow, C. D.; Pande, V. S.; Boxer, S. G. *Science* **2006**, *313*, 200−204.
(55) Ensign, D. L.; Webb, L. J. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 3511−3524.
(56) Saggu, M.; Levinson, N. M.; Boxer, S. G. *J. Am. Chem. Soc.* **2011**, *133*, 17414−17419.
(57) Brewer, S. H.; Franzen, S. *J. Chem. Phys.* **2003**, *119*, 851−858. The fact that Stark tuning rates calculated in the gas phase by DFT tend to be significantly lower than corresponding experimental values indirectly supports this notion.

Supplementary Information for

Measuring Electrostatic Fields in Both Hydrogen Bonding and non-Hydrogen Bonding Environments using Carbonyl Vibrational Probes

Stephen D. Fried, Sayan Bagchi†, Steven G. Boxer*
Department of Chemistry, Stanford University; Stanford, California 94305–5012

† Current address: Physical and Materials Chemistry Division, National Chemical Laboratory (CSIR), Pune 411008, India

* Corresponding author. E-mail: sboxer@stanford.edu

Contents

1. Supplementary Tables

Table S1. Spectral data for vibrational solvatochromism of acetophenone [a,b]

		Peak Picking [d]		Curve Fitting [d]	
Solvent	**Dielectric Constant**	**Peak Position / cm^{-1}**	**FWHM / cm^{-1}**	**Peak Position / cm^{-1}**	**FWHM / cm^{-1}**
hexanes	1.89	1696.4	6.4	1696.3	6.4
cyclohexane	2.02	1695.5	5.7	1695.4	5.7
dibutylether	3.08	1693.5	7.6	1693.3	7.7
diethylether	4.34	1692.9	8.0	1692.8	8.1
chloroform	4.71	1683.3 [c]	16.1	1683.0 [c]	16.1
2-methyl-tetrahydrofuran	6.97	1690.2	8.4	1690.0	8.5
tetrahydrofuran	7.43	1689.5	7.9	1688.8	8.1
dichloromethane	8.93	1684.6 [c]	11.4	1684.5 [c]	11.7
valeronitrile	20.04	1687.1	10.8	1686.9	10.9
butyronitrile	24.83	1686.9	10.3	1686.7	10.5
acetonitrile	37.5	1686.1	8.6	1686.0	8.9
dimethyl sulfoxide	46.84	1682.0	10.4	1681.8	11.0
D_2O	78.54	1669.4 [c]	20.6	1668.8 [c]	21.4

Table S1. [a] See Figure 2A in main text for spectra of selected solvents. [b] All samples were at 10 mM concentration at room temperature. [c] Highlighted in red are the solvents that can H-bond to acetophenone, and whose frequencies are therefore not in sequence with dielectric constant. [d] Center frequencies and full-width at half maxima (FWHM) were determined using peak picking and curve fitting (see main text, experimental methods section V.3). For all analyses described in the main text, the data obtained with peak picking were used; however, their similarity to the values obtained with curve fitting attests to the high quality of the spectral data.

Table S2. IR data for [p-Ac-Phe]RNase S and its precursors [a,b]

	Peak Picking [c]		Curve Fitting [c]	
	Peak Position / cm^{-1}	**FWHM / cm^{-1}**	**Peak Position / cm^{-1}**	**FWHM / cm^{-1}**
p-Ac-Phe	1669.2 ± 0.2	20.3 ± 0.9	1668.4 ± 0.6	20.5 ± 1.0
[p-Ac-Phe]S-peptide	1673.1 ± 0.1	14.7 ± 0.8	1672.8 ± 0.2	15.9 ± 0.9
[p-Ac-Phe]RNase S	1687.6 ± 0.1	5.6 ± 0.4	1687.7 ± 0.3	5.5 ± 0.3

Table S2. [a] See Figure 3A in main text for spectra. [b] All samples were in 20 mM HEPES in D_2O, pD = 8.0, at room temperature. The concentrations of p-Ac-Phe and the S-peptide were 5 mM, whereas for RNase S, the concentration was 2.5 mM. For each sample, spectra were measured in triplicate. The values given for center frequency and full-width at half maximum (FWHM) are the mean from three measurements, and the error is the standard deviation. [c] Center frequencies and full-width at half maxima (FWHM) were determined using peak picking and curve fitting (see main text, experimental methods section V.3). For all analyses described in the main text, the data obtained with peak picking were used.

Table S3. H-bond analysis from MD trajectories

	Closest atom distance / Å [a]		Retention time / fs [b]	
	mean	std. dev.	mean	std. dev.
hexanes	2.77	0.24	63	133
dibutylether	2.69	0.21	69	148
chloroform	2.47	0.31	529	1190
tetrahydrofuran	2.65	0.20	49	121
dichloromethane	2.54	0.25	153	273
valeronenitrile	2.68	0.20	72	156
acetonitrile	2.69	0.22	55	126
dimethylsulfoxide	2.59	0.20	74	166
water (TIP3P)	1.95	0.22	311	514

[a] The mean and standard deviation for the distance between the closest solvent atom to the O-atom of the C=O probe. For all solvents, this atom was hydrogen ~99% of the time. [b] The mean and standard deviation for the time interval during which the identity of the closest atom remained the same.

For the six non-H-bonding solvents, the closest atom approaches the O-atom of acetophenone at a distance of 2.6–2.7 Å. In chloroform and dichloromethane, it is slightly closer (2.5 Å), and significantly closer for water (2.0 Å). Short O…H distances are characteristic of H-bonding, and the order of the average distance (water < $CHCl_3$ $\lesssim$ DCM) is consistent with the order of electric field strengths (water > $CHCl_3$ $\gtrsim$ DCM).

Because H-bond breaking requires surmounting a barrier, another feature of an H-bond is a persistence time longer than those of non-specific solvation interactions. One way of measuring this timescale is to keep track of how many timesteps it takes for the identity of the closest atom to change, on average. For the non-H-bonding solvents, this so-called retention time is very short, and is close to 60 fs. For the H-bonding solvents, it is considerably longer. Interestingly, chloroform has an exceptionally long retention time (530 fs), which can be explained by considering the fact that each chloroform molecule only has one candidate H-bond donor. Water and dichloromethane toggle more quickly because the density of H-bond donor candidates is larger.

Tables S4–S6. Controls simulations for electric field calculation. [a]

Table S4. Deterministic dynamics

	$\lvert F_{vib} \rvert$ / (MV/cm) [b]		$\lvert \Delta F_{vib} \rvert$ / (MV/cm) [c]	
	mean	**std. dev.**	**mean**	**std. dev.**
hexanes	–0.108	0.77	–0.014	0.77
dibutylether	–7.74	6.12	–0.32	3.91
chloroform	–24.5	16.6	–11.2	13.4
tetrahydrofuran	–15.7	9.58	–0.12	5.92
dichloromethane	–26.7	13.4	–8.94	11.1
valeronenitrile	–19.9	11.4	–1.18	7.48
acetonitrile	–24.7	14.2	–3.70	9.45
dimethylsulfoxide	–29.0	13.2	–0.95	9.84
water (TIP3P)	–64.0	28.5	–39.7	27.4

Table S5. AM1-BCC charges

	$\lvert F_{vib} \rvert$ / (MV/cm) [b]		$\lvert \Delta F_{vib} \rvert$ / (MV/cm) [c]	
	mean	**std. dev.**	**mean**	**std. dev.**
hexanes	–0.048	0.79	0.0075	0.77
dibutylether	–7.99	6.47	–0.011	3.99
chloroform	–23.1	11.9	–8.98	10.5
tetrahydrofuran	–15.6	9.93	–0.093	6.13
dichloromethane	–21.1	12.8	–5.98	9.16
valeronenitrile	–17.3	9.29	–1.00	6.71
acetonitrile	–22.8	11.5	–3.15	8.54
dimethylsulfoxide	–27.5	11.1	–0.92	9.04

Table S6. 0.9-fold charge attenuation

	$\lvert F_{vib} \rvert$ / (MV/cm) [b]		$\lvert \Delta F_{vib} \rvert$ / (MV/cm) [c]	
	mean	**std. dev.**	**mean**	**std. dev.**
hexanes	–0.071	0.70	–0.038	0.70
dibutylether	–6.86	5.01	–0.39	3.52
chloroform	–22.3	15.6	–9.86	12.3
tetrahydrofuran	–14.0	8.85	–0.058	5.69
dichloromethane	–24.5	11.8	–7.94	10.1
valeronenitrile	–18.5	9.14	–1.44	6.99
acetonitrile	–24.4	10.6	–3.49	8.75
dimethylsulfoxide	–27.6	10.8	–1.42	8.94

[a] See computational methods section IV.5 in the main text for an introduction to these simulations. [b] The electric field experienced by the C=O vibration, as defined by Eq. 4. [c] The electric field gradient across the C=O vibration, as defined by Eq. 5.

Deterministic dynamics (using a leap-frog integrator), the AM1-BCC charge scheme, and uniformly attenuated charges *all* produced minor changes to ensemble-average solvation fields. Compared to the method used most extensively in the main text (stochastic dynamics with RESP charges), the average unsigned change in solvation fields was 1.1 MV/cm (4.6%) for deterministic dynamics, and 2.1 MV/cm (16.2%) for AM1-BCC charges. The electric fields due to AM1-BCC charges were generally smaller in magnitude, whereas deterministic dynamics caused the solvation field both to increase and decrease. Consistent with expectations, applying a 10% decrease to the solvent's charges caused the solvation fields to be smaller, by 1.6 MV/cm (11.8%) on average. In principle, a perturbation to the charge magnitudes *might* have had an exaggerated effect on the ensemble-average electric fields due to the influence of sampling – i.e., configurations with smaller stabilizing fields would become sampled more frequently, causing the ensemble-average field to become even smaller. That this effect was not observed – in fact, the percent decrease in the fields was very close to the percent decrease in the charge magnitudes – suggests that small perturbations to electrostatics do not significantly bias the sampling of the trajectory.

2. Supplementary Figures with Analysis

Figure S1. Linewidth vs. peak position correlation plot

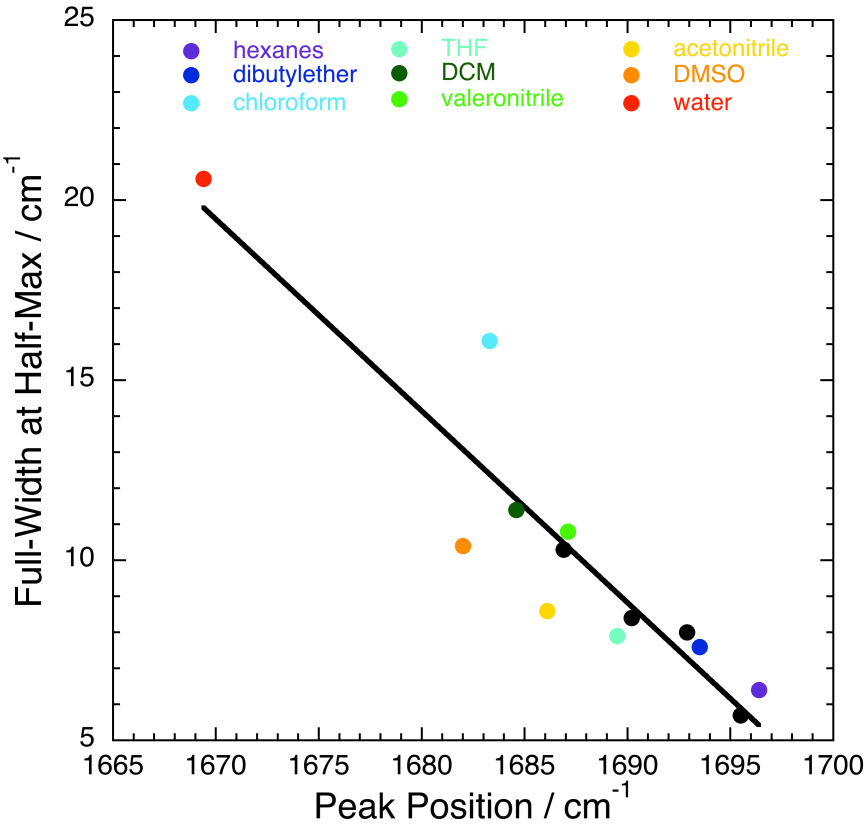


Figure S1. C=O band linewidth of acetophenone plotted against its peak position. A fairly consistent trend between lower peak frequency and wider peak is present, with an R^2 value of 0.85. This correlation suggests that if the probe vibration experiences larger electric fields overall (related to the peak position), it will also experience a larger range and broader distribution of electric fields (related to the linewidth, assuming inhomogeneous broadening dominates). The regression line is $FWHM = -0.532\bar{\nu} + 907.9$.

Figure S2. Linewidth vs. electric field distribution plot

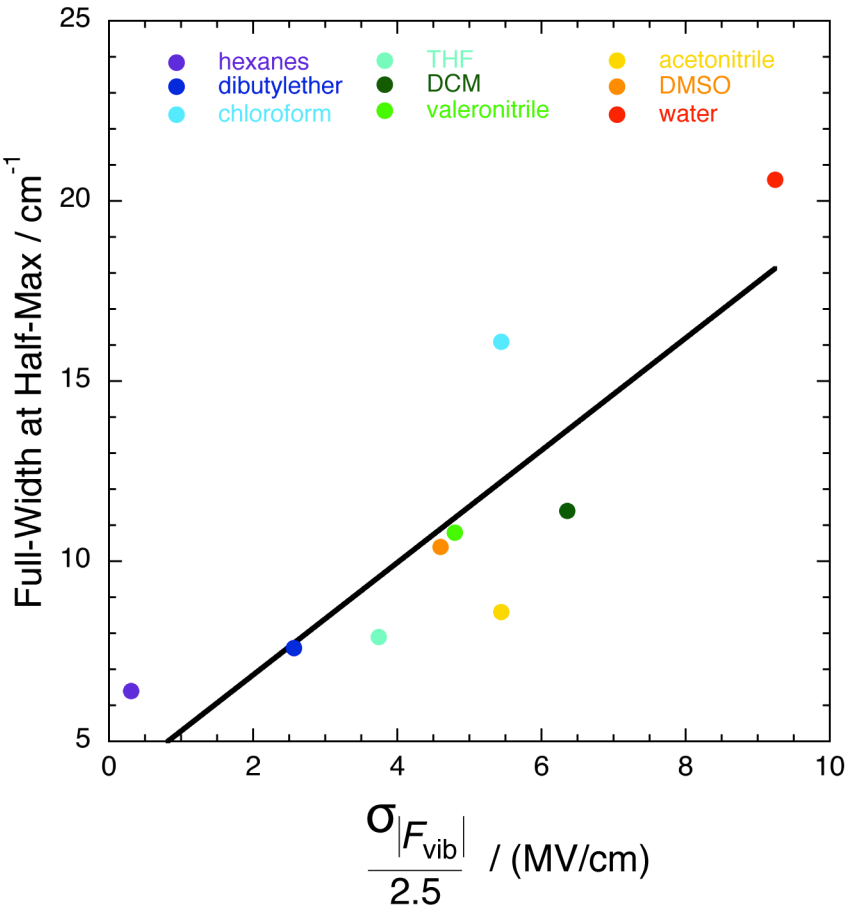


Figure S2. C=O band linewidth of acetophenone plotted against the rescaled standard deviation of the field distribution for each solvent. In each solvent, the C=O vibration samples a range of different electrostatic fields from its environment (described by $\sigma_{|F|}$), and therefore a range of frequencies via the linear Stark effect (Eq. 1). This effect contributes to the linewidth by inhomogeneous broadening. The regression line is $FWHM = 1.56\left(\sigma_{|F|}/2.5\right) + 3.73$.

The linewidth of an IR band is a dynamic quantity that depends on correlations in the frequency fluctuations rather than just the distribution of fields sampled. The strength of this linear correlation is only moderate ($R^2 = 0.72$), indicating that $\sigma_{|F|}$ is *not* sufficient to explain all the variation in FWHM, and suggesting that the field correlations are not uniform across the solvent series.

The intercept (3.73 cm^{-1}) is a measure of the intrinsic linewidth of the C=O band independent of environment heterogeneity (i.e., inhomogeneous broadening), which is qualitatively similar to the calculated lifetime-limited linewidth of C=O of ca. 5 cm^{-1}, assuming an average excited state lifetime of ca. 1 ps.[58] The slope (1.56 cm^{-1}/(MV/cm)) is a measure of how sensitive the linewidth is to field distribution.

Figure S3. Michaelis-Menton Plots of RNase A and [p-Ac-Phe]RNase S

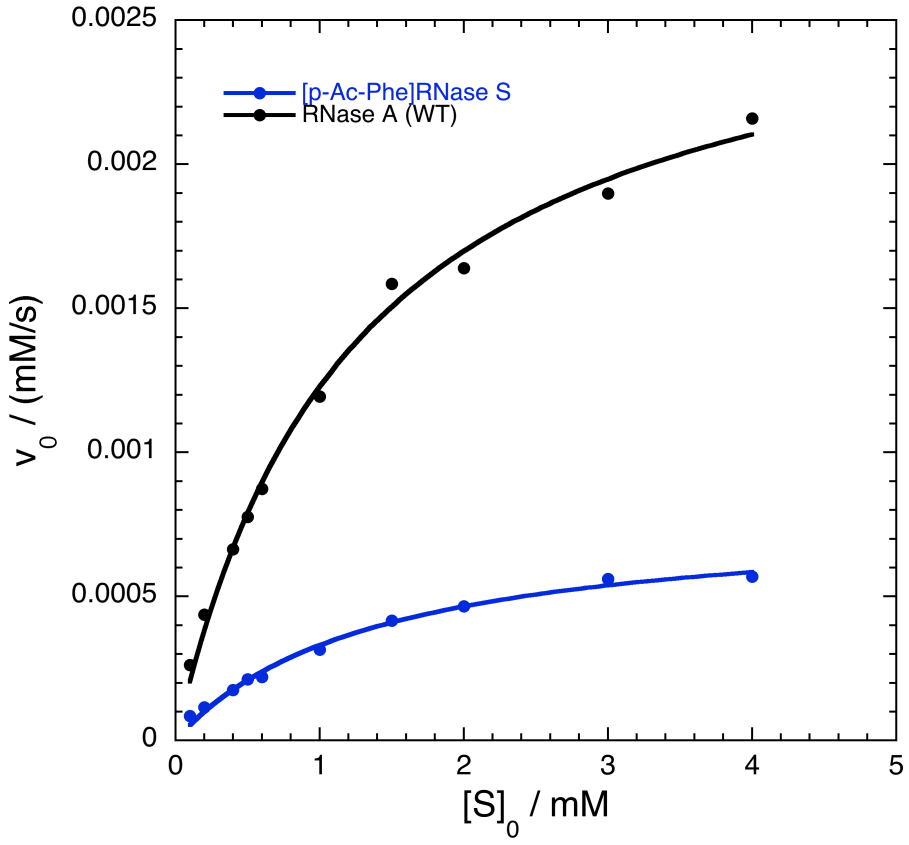


Figure S3. Initial rate of cyclic cytidine monophosphate (cCMP) hydrolysis catalyzed by ribonuclease at various initial concentrations of cCMP, all in 20 mM Bis-Tris pH 6.0. The assay was performed on the vibrational probe-bearing [p-Ac-Phe]RNase S construct (blue trace), as well as on wild-type RNase A (from bovine pancreas, black trace) as a control. The enzyme's affinity to monomeric substrate is unchanged within error, and the catalytic efficiency is decreased 3.5-fold (see Table 3).

Extrinsic vibrational probes, because of their small size, can be innocuously introduced into proteins at functionally-sensitive positions, and have small-to-negligible effect on the native function of the protein. This key asset has been noted and exploited for nitrile vibrational probes introduced into RNase S[35] and ketosteroid isomerase,[1] and interpreted to mean that the probe causes a minimal perturbation to the underlying protein structure. In the case of RNase S, the Phe at position 8 is located near the active cleft of the enzyme, ~5 Å away from His 12 and 119, the catalytic general acid-base pair. We employed a standard enzymatic assay of RNase as a control. The initial rate's dependence on initial substrate concentration is described very well by the Michaelis-Menton model ($R^2 > 0.99$), and the kinetic parameters extracted from fitting the data to the Michaelis-Menton equation are shown in Table 3. The values for k_{cat} and

K_M of RNase A were almost identical to those seen previously.[35] The carbonyl-probe modified RNase variant experienced a 3.5-fold rate deceleration relative to wild-type RNase A, although it had a nearly unchanged value for K_M. This rate-detriment can be attributed to a structural perturbation imparted by the vibrational probe near the active site of the enzyme. The rate-detriment of replacing Phe at position 8 with p-Ac-Phe (3.5-fold) was somewhat larger than that of p-CN-Phe (2-fold),[35] which is consistent with the fact that the acetyl probe is more sterically demanding than the nitrile probe (involving three new heavy-atoms rather than two), and a C=O bond is generally more polar than a C≡N bond. These observations are suggestive that acetyl probes may impart larger perturbations to host proteins than nitriles, which may be problematic in other systems.

Figure S4. UV-Vis spectroscopy of [p-Ac-Phe]RNase S

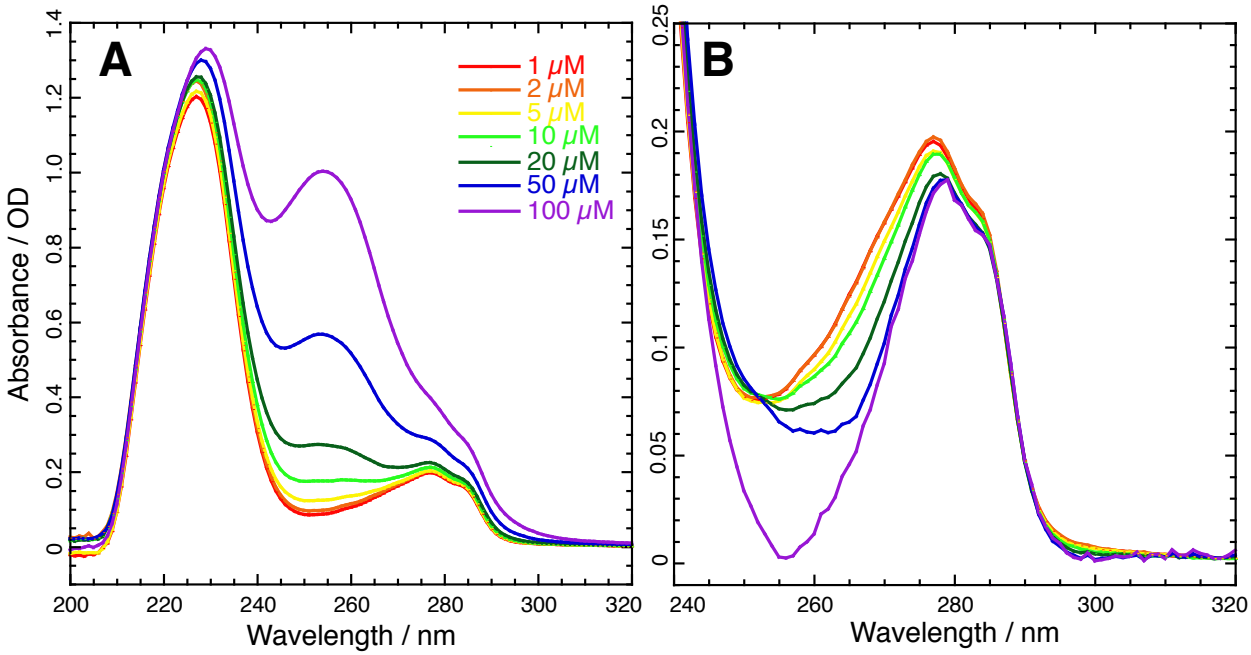


Figure S4. UV-Vis titration of the ribonuclease S-protein equilibrated with [p-Ac-Phe]S-peptide, all in 50 mM sodium acetate, 100 mM sodium chloride, pH 6.0. [S-protein] was kept constant at 20 μM, and [S-peptide] varied between 1 μM and 100 μM, according to legend. The spectra were baselined by setting the absorbance at 400 nm to zero.

(A) Averaged spectra over duplicate measurements. The spectral change is dominated by the absorption band of p-Ac-Phe with λ_{max} at 255 nm (ε_{255} = 9.8 mM^{-1} cm^{-1}), which masks any changes in the envelope at 280 nm that accompany the titration.

(B) Spectra from panel A, except subtracting the contribution from [S-peptide] for each one. The resulting absorption traces reveal that the underlying 280-band consistently narrows (the full-width half-maximum decreases from 27 to 18 nm) and red-shifts (from 277 to 279 nm) – both of which are suggestive of the formation of a structured folded protein. The largest spectral change occurs for [S-peptide] between 10 μM and 20 μM, which suggests this as the approximate range for the dissociation constant.[59]

Figure S5. Circular Dichroism (CD) spectroscopy of [p-Ac-Phe]RNase S

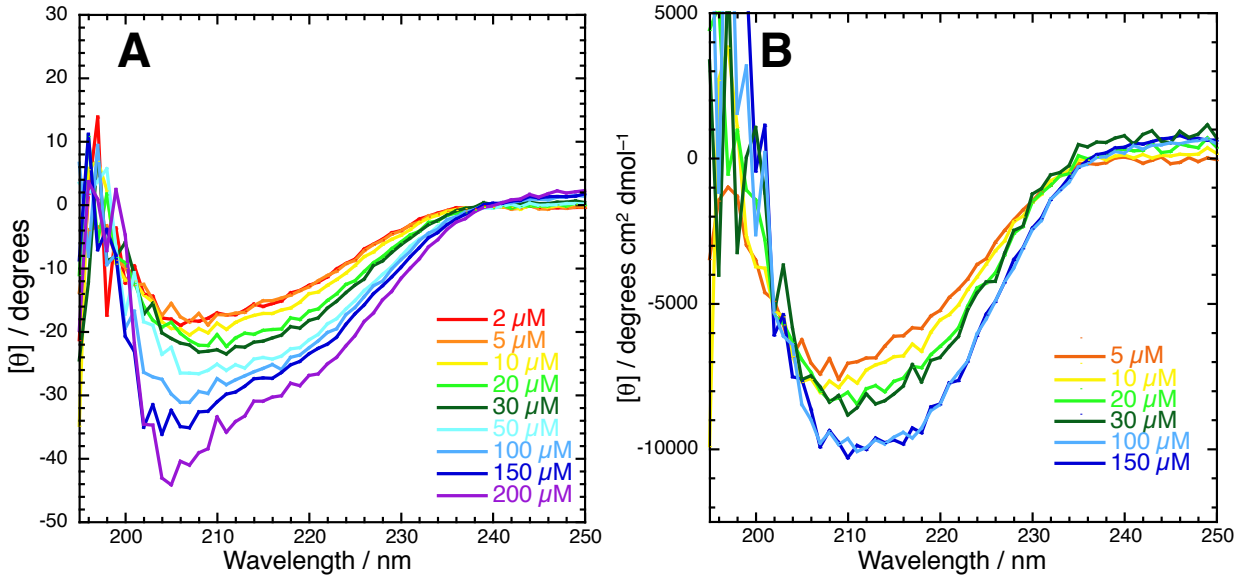


Figure S5. CD spectra of the ribonuclease S-protein equilibrated with [p-Ac-Phe]S-peptide, all in 50 mM sodium acetate, 100 mM sodium chloride, pH 6.0. [S-protein] was kept constant at 20 μM, and [S-peptide] varied between 2 μM and 200 μM. The spectra were baseline corrected by subtracting the CD spectrum of pure buffer. Spectra were obtained using a 1-mm path length cuvette, 1 nm resolution, and averaging signal over 10 s/nm.

(A) Traces correspond to baseline corrected but otherwise unmodified CD spectra. Increasing the concentration of S-peptide results in a consistent increase in the broad negative peak around 211 nm. The changes observed result from a superposition of (1) the increasing dichroism of S-protein as it assembles into RNase S, and (2) the intrinsic dichroism of the S-peptide itself (which has some helical character).

(B) To isolate effect (1), spectra were corrected by subtracting the contribution from [S-peptide] for each one, and then normalizing the ellipticity to the mean residue ellipticity (using the molar mass of RNase S). Fewer traces are shown for clarity, but the same trend is observed for the S-peptide concentrations not shown. The corrected spectra clearly reveal two changes in the optical activity when unstructured S-protein is converted to folded RNase S. First, the negative band increases in intensity from 6,000 to 10,000 $^{\circ}$cm^2/dmol; and second, the negative band becomes sharper on the blue edge. The broad band with a max around 210 nm is characteristic of RNase, which has a mixture of alpha and beta secondary structure.[60] In particular, the CD spectrum at 150 μM S-peptide is very similar to that of RNase A, including the intensity of the ellipticity at 210 nm.[60] Moreover, the similarity between the 100 μM and 150 μM traces

suggest that the S-protein is fully folded and resembles RNase A at these high [S-peptide]'s.

Based on these observations, one can say roughly that half of S-protein is folded for an [S-peptide] between 20 μM and 30 μM, because these have mean residue ellipticities approximately half of the full range from zero to saturation. This observation suggests that the dissociation constant is between 20 μM and 30 μM.

Figure S6. Isothermal Titration Calorimetry (ITC) of [p-Ac-Phe]RNase S

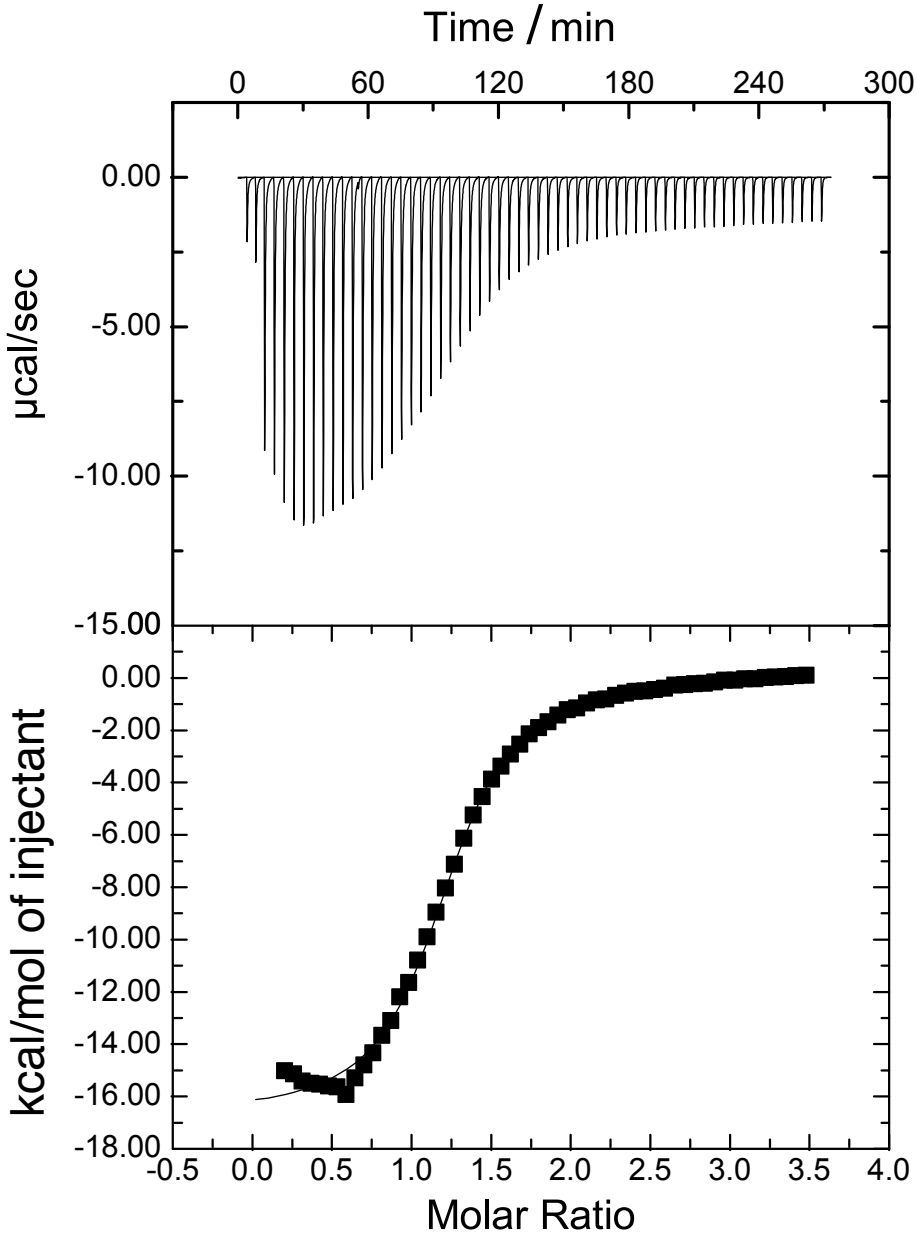


Figure S6. Isothermal titration calorimetry of [p-Ac-Phe]RNase S to measure the binding thermodynamics of [p-Ac-Phe]S-peptide to S-protein. The raw data (top) were integrated, baseline-subtracted, and fit (bottom) using MicroCal LLC's Origin package to determine the thermodynamic parameters for the RNase S binding/self-assembly process. The thermodynamic parameters extracted from the fit are found in Table 3.

The relatively high dissociation constant of 12.7 μM reflects a 120-fold weaker affinity of [p-Ac-Phe]S-peptide compared to that of the unmodified S^{15}-peptide. Overall, the process is highly favored energetically but costly entropically. Both of these features

are consistent with a folding process. Note that the value of K_D measured is qualitatively consistent with the estimates from UV-Vis and CD titrations.

Moreover, combining this K_D with the relatively mutation-insensitive value of k_{on} ($4.4 \times 10^5\ M^{-1}\ s^{-1}$),[61] we can calculate an approximate k_{off} of this system of $5\ s^{-1}$.

Figure S7. Field-frequency analysis using the Onsager model

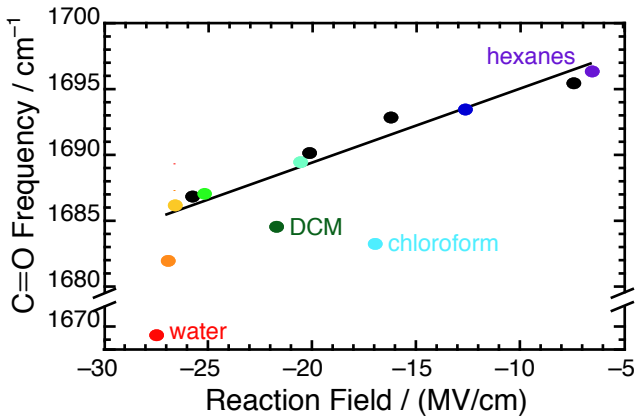


Figure S7. Plot of acetophenone's peak C=O frequency in 13 solvents compared against the electric field of those solvents, as calculated by the Onsager reaction equation (Eq. S1) using μ_0 = 2.95 D, n^2 = 2.35, and a^3 = 46.3 Å^3. Excluding H-bond-donating solvents, the trend line is $\bar{\nu}_{C=O} = 0.62\left|F_{reaction}\right| + 1701.9$ with R^2 = 0.91. Including the three H-bond-donating solvents, R^2 falls to 0.57.

While the present work has focused on MD-based approaches to calculate solvation fields, it is important to compare these approaches to continuum models as these were employed in earlier work.[22,31,41] The simplest such example is the reaction field equation devised by Onsager:[62]

$$\left|F_{Onsager}\right| = \frac{\mu_0}{a^3}\left[\frac{2(\varepsilon - 1)(n^2 + 2)}{3(2\varepsilon + n^2)}\right] \qquad \text{(S1)}$$

The solute is treated as a point dipole in a spherical cavity and is represented by its dipole moment μ_0, the cavity radius a, and its refractive index n. The solvent enters into Eq. S1 by way of its static dielectric constant, ε. By employing the experimentally-determined dipole moment[63] and index of refraction of acetophenone,[64] we used the Onsager formalism to model the electric field in 13 different solvents (Table S1), and plotted the C=O stretching frequencies of acetophenone against them. Figure S7 displays a qualitative linear correlation; however, dichloromethane and chloroform fall considerably off the trend and water is grossly misrepresented. It is quite clear that the reason for these three outliers is that the Onsager model only describes the effect of dielectric polarization, but not specific interactions such as H-bonds. On the other hand, MD simulations captured these effects properly.

It is interesting to point out that C≡N vibrations generally correlated *better* with the Onsager reaction field.[31,41] In hindsight we can say this was due to a cancellation of errors: the reaction field's inability to capture the effect of H-bonds was compensated for by the nitrile's non-linear behavior when H-bonded.[12,13]

As an aside, the cavity radius cubed, a^3, was calculated by dividing acetophenone's formula weight (120 g/mol) by its density (1.028 g/mL) to obtain the cavity volume (194 $Å^3$/molecule). The cavity volume was then multiplied by 3/(4π) to obtain a^3. We must report that in previous papers,[22,31,41] we had mistakenly used the cavity volume in place of the cavity radius cubed in Eq. S1, resulting in electric fields that were uniformly too small by a factor of ca. 4.2.

3. Simulation Methods

1. Parameterization of p-Ac-Phe

The following procedure was used to add p-Ac-Phe (FAC) to the AMBER99SB-ILDN forcefield (ff99SB-ILDN).

Step 1. The molecule p-(tolyl)ethanone was modeled in GaussView and then geometry optimized by DFT at the B3LYP/6-31G* level in Gaussian 09. The output coordinates were converted to the pdb format.

Step 2. The ANTECHAMBER utility from AmberTools12 was run on the p-(tolyl)ethanone pdb and parameterized into the GAFF forcefield. Charges were obtained using the AM1-BCC method. Afterwards, the PARMCHK utility was run on the resulting mol2 files to obtain missing parameters.

Step 3. Using the tLEaP program from AmberTool12, the mol2 file and frcmod file were inputted to form a new Amber unit. Amber parameter (prmtop) files and coordinate (inpcrd) files were saved.

Step 4. The AcPype utility was called in order to convert from the Amber prmtop and inpcrd files into GROMACS topology (top) and coordinate (gro) files.

Step 5. The charges for FAC were determined as follows. The charges that ff99SB-ILDN uses for the backbone atoms (N, H, CA, C, O) and HB of Tyrosine were retained. The AM1-BCC charges determined for p-(tolyl)ethanone starting with the “CG” atom and moving to the end of the molecule were mapped onto CG, CD1, HD1, etc. atoms of FAC. The charge of CB was set equal to whatever value was needed to neutralize the residue fragment. This simplified protocol is justifiable given the similarity of FAC to natural amino acids and the fact that ff99SB-ILDN uses identical charges for the backbone atoms in Phe and Tyr, despite the large difference in polarity of their side chains.

Step 6. In general, the van der Waals parameters and valence bonding parameters used for FAC were maintained from ff99SB-ILDN. The parameters were loaded into the system’s data files ($GMXDATA/gromcas/top/amber99sb-ildn.ff), so that FAC could be parameterized simply by calling the PDB2GMX utility. The full list of parameters placed into amber99sb-ildn.ff files for FAC are given below.

added to aminoacids.rtp

```
[ FAC ]
 [ atoms ]
     N    N           -0.41570     1
     H    H            0.27190     2
```

```
    CA     CT           -0.00140     3
    HA     H1            0.08760     4
    CB     CT            0.05440     5
   HB1     HC            0.02950     6
   HB2     HC            0.02950     7
    CG     CA           -0.04530     8
   CD1     CA           -0.14750     9
   HD1     HA            0.13600    10
   CE1     CA           -0.07800    11
   HE1     HA            0.14450    12
    CZ     CA           -0.18360    13
    CH     C             0.57970    14
    OH     O            -0.53910    15
    CQ     CT           -0.19810    16
   HQ1     HC            0.06370    17
   HQ2     HC            0.06370    18
   HQ3     HC            0.06370    19
   CE2     CA           -0.07800    20
   HE2     HA            0.14500    21
   CD2     CA           -0.14750    22
   HD2     HA            0.13600    23
     C     C             0.59730    24
     O     O            -0.56790    25
[ bonds ]
     N      H
     N     CA
    CA     HA
    CA     CB
    CA      C
    CB    HB1
    CB    HB2
    CB     CG
    CG    CD1
    CG    CD2
   CD1    HD1
   CD1    CE1
   CE1    HE1
   CE1     CZ
    CZ     CH
    CZ    CE2
    CH     OH
    CH     CQ
    CQ    HQ1
    CQ    HQ2
    CQ    HQ3
   CE2    HE2
   CE2    CD2
   CD2    HD2
```

```
     C      O
    -C      N
 [ impropers ]
    -C     CA      N      H
    CA     +N      C      O
    CG    CE2    CD2    HD2
    CZ    CD2    CE2    HE2
   CD1     CZ    CE1    HE1
    CG    CE1    CD1    HD1
   CD1    CD2     CG     CB
   CE1    CE2     CZ     CH
    CQ     CZ     CH     OH
```

added to aminoacids.hdb

```
FAC      8
1        1        H        N        -C       CA
1        5        HA       CA       N        CB       C
2        6        HB       CB       CA       CG
1        1        HD1      CD1      CG       CE1
1        1        HE1      CE1      CD1      CZ
1        1        HE2      CE2      CZ       CD2
1        1        HD2      CD2      CG       CE2
3        4        HQ        CQ      CH       CZ
```

under [angletypes]

```
CA  C    O            1   123.440    574.630 ; from GAFF (for ca-c-o)
CA  C    CT           1   118.540    523.750 ; from GAFF (ca-c-c3)
```

Step 7. The missing dihedrals were taken from GAFF and hard-coded into the protein's topology file. This was simply done by manually editing the top files after running PDB2GMX. We chose to do this is because GAFF uses a different rule for dihedrals than ff99SB-ILDN. For instance, the line in the topology file **under [dihedrals]** that read

```
116   121   118   119     4
```

was replaced with

```
116   121   118   119     1   180.00   4.60240   2
```

using information found in the topology file from step 4. This replacement is repeated for each missing dihedral parameter, which can be easily found by examining the errors raised upon calling GROMPP.

2. Python scripts used in electric field calculations. Note that separate itp files with charge-neutral TIP3P water and ions must also be prepared.

A. discharger.py

```
#The purpose of this script is to prepare a charge-neutral topology file of a
particular protein.
#Call it with the topology file as the first argument; i.e.,

#python discharger.py myprotein.top

import sys
import re


top_fn = sys.argv[1]
top0q_fn = '%s_0q.%s' %(top_fn.split('.')[0] , top_fn.split('.')[1])
top = open( top_fn , 'r')
top0q = open( top0q_fn, 'w')

inAtoms = 0
for line in top:
    #Discharge all of the [ atoms ] records
    if line.startswith('[ bonds ]'):
        inAtoms = 0
    if inAtoms == 1:
        atomInfo = line.split()
        if line.startswith(';') or line == '\n':
            top0q.write( line )
        elif atomInfo[3] == 'FAC': #the resn of the probe-bearing residue
            top0q.write( line )
        else:
            atomInfo[6] = '0.00000'
            whitespace = re.split(r'\S+', line)
            newline = ''
            for i in range( len( atomInfo ) ):
                newline += whitespace[i]
                newline += atomInfo[i]
            newline += '\n'
            top0q.write( newline )
    #Point to a new discharged solvent and ions topology
    elif re.search('tip3p.itp' , line):
        top0q.write( line.replace('amber99sb-ildn.ff/tip3p.itp' ,
'./tip3p_0q.itp') )
    elif re.search('ions.itp' , line):
        top0q.write( line.replace('amber99sb-ildn.ff/ions.itp' ,
'./ions_0q.itp') )
    else:
        top0q.write( line )
    if line.startswith('[ atoms ]'):
        inAtoms = 1

top0q.close()
top.close()
```

B. protein_fields.py

```
#The purpose of this script is to calculate electric fields in a protein
#Call it with the topology file as the first argument
#And the force+coordinate trajectory file as the second argument; i.e.,

#python protein_fields.py myprotein.top myprotein_md.trr

#As a pre-requisite, you need to have an index file called probe.ndx that has
#the index-numbers of the two probe atoms

import sys
import re
import itertools
import os
import numpy as np

def isnumber(string):
    try:
        float(string)
        return True
    except ValueError:
        return False

top_fn = sys.argv[1]
top0q_fn = '%s_0q.%s' %(top_fn.split('.')[0] , top_fn.split('.')[1])

#Step 2, rerun the trajectory with the discharged topology
traj_fn = sys.argv[2]                  #i.e., myprotein_md.trr
sys_n = traj_fn.split('_')[0]          #i.e., myprotein

command = 'grompp -f md-langevin.mdp -p %s -c %s_npt2.gro -t %s_npt2.cpt -o
%s_md0q.tpr' %(top0q_fn, sys_n, sys_n, sys_n)
print "Executing command:",command
os.system(command)

command = 'mdrun -s %s_md0q.tpr -rerun %s -o %s_md0q.trr' %(sys_n, traj_fn,
sys_n)
print "Executing command:",command
os.system(command)


#Step 3, perform gtraj to get coordinates of probe, forces, and discharged
forces
command = 'g_traj -s %s_md.tpr -f %s -n probe.ndx -ox %s_probeX.xvg' %(sys_n,
traj_fn, sys_n)
print "Executing command:",command
os.system(command)

command = 'g_traj -s %s_md.tpr -f %s -n probe.ndx -of %s_probeF.xvg' %(sys_n,
traj_fn, sys_n)
print "Executing command:",command
os.system(command)

command = 'g_traj -s %s_md0q.tpr -f %s_md0q.trr -n probe.ndx -of
%s_probeF0q.xvg' %(sys_n, sys_n, sys_n)
print "Executing command:",command
os.system(command)
```

```
#Step 4, get the relevant forces and coords
f_x = open('%s_probeX.xvg' %sys_n ,'r')
f_f = open('%s_probeF.xvg' %sys_n , 'r')
f_f0q = open('%s_probeF0q.xvg' %sys_n , 'r')
#And finally calculate the fields
fields = open('%s_FIELDS.txt' %sys_n , 'w')
for line_x, line_f, line_f0q in itertools.izip(f_x, f_f, f_f0q):
    if isnumber( line_x.split()[0] ):
        #get the time
        time_x = float( line_x.split()[0] )
        time_f0q = float( line_f0q.split()[0] )

        #check that the info matches!
        if time_x != time_f0q:
            print "time indeces do not match.  Error!"
            sys.exit()

        #get the C & O coordinates
        #calculate the CO bond length and CO unit vector
        x_C = np.array([float(x) for x in line_x.split()[1:4]])
        x_O = np.array([float(x) for x in line_x.split()[4:7]])
        COvec = x_O - x_C
        COlen = np.sqrt((COvec**2).sum())
        COunitvec = COvec / COlen
        #get the C & O forces
        f_C = np.array([float(x) for x in line_f.split()[1:4]])
        f_O = np.array([float(x) for x in line_f.split()[4:7]])
        #get the C & O discharged forces
        f0q_C = np.array([float(x) for x in line_f0q.split()[1:4]])
        f0q_O = np.array([float(x) for x in line_f0q.split()[4:7]])
        #calculate the electrostatic force
        fe_C = f_C - f0q_C
        fe_O = f_O - f0q_O
        #calculate the electrostatic force projection onto CO
        feproj_C = np.dot( fe_C , COunitvec )
        feproj_O = np.dot( fe_O , COunitvec )
        #calculate the electric field and convert
        Eproj_C = (feproj_C / charge_C) * 0.1036427
        Eproj_O = (feproj_O / charge_O) * 0.1036427
        Eproj = (Eproj_C + Eproj_O)/2
        EprojDrop = (Eproj_O - Eproj_C)


        #print to file
        fieldInfo = str(time_x) + '\t' + str(Eproj_C) + '\t' + str(Eproj_O) +
'\t' + str(Eproj) + '\t' + str(EprojDrop) + '\n'
        fields.write ( fieldInfo )
f_x.close()
f_f.close()
f_f0q.close()
fields.close()
```

4. References

(58) Bagchi, S.; Kim, Y. S.; Charnley, A. K.; Smith III, A. B.; Hochstrasser, R. M. *J. Phys. Chem. B* **2007**, 111, 3010–3018.
(59) Schalley, C.; Hirose, K. *Analytical Methods in Supramolecular Chemistry*; Wiley-VCH: Weinheim, 2007. Chapter 2: Determination of Binding Constants.
(60) Catanzano, F.; Graziano, G.; Cafaro, V.; D'Alessio, G.; Di Donato, A.; Barone, G. *Int. J. Biol. Macromol.* **1998**, 23, 277-285.
(61) Bachmann, A.; Wildemann, D.; Praetorius, F.; Fischer, G.; Kiefhaber, T. *Proc. Natl. Acad. Sci.* **2011**, *108*, 3952–3957.
(62) Onsager, L. *J. Am. Chem. Soc* **1936**, *58*, 1486–1493.
(63) Granzhan, V. A.; Manole, S. F.; Laktionova, S. K.; Filippov, M. P. *J. Struct. Chem.* **1971**, *12*, 397–402.
(64) ChemSpider: The free chemical database. URL: http://www.chemspider.com/Chemical-Structure.7132.html