J|A|C|S
ARTICLES
Published on Web 12/07/2002

Using Nitrile-Derivatized Amino Acids as Infrared Probes of Local Environment

Zelleka Getahun,[‡] Cheng-Yen Huang,[†] Ting Wang,[†] Brenda De León,[†] William F. DeGrado,*[,‡] and Feng Gai*[,†]

Contribution from the Department of Chemistry and Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received September 12, 2002; E-mail: gai@sas.upenn.edu and wdegrado@mail.med.upenn.edu

Abstract: It is well-known that the C≡N stretching vibration in acetonitrile is sensitive to solvent. Therefore, we proposed in this contribution to use this vibrational mode to report local environment of a particular amino acid in proteins or local environmental changes upon binding or folding. We have studied the solvent-induced frequency shift of two nitrile-derivatized amino acids, which are, Ala_{CN} and Phe_{CN}, in H_2O and tetrahydrofuran (THF), respectively. Here, THF was used to approximate a protein's hydrophobic interior because of its low dielectric constant. As expected, the C≡N stretching vibrations of both Ala_{CN} and Phe_{CN} shift as much as $\sim$10 cm^{-1} toward higher frequency when THF was replaced with H_2O, indicative of the sensitivity of this vibration to solvation. To further test the utility of nitrile-derivatized amino acids as probes of the environment within a peptide, we have studied the binding between calmodulin (CaM) and a peptide from the CaM binding domain of skeletal muscle myosin light chain kinase ($MLCK_{579-595}$), which contains a single Phe_{CN}. $MLCK_{579-595}$ binds to CaM in a helical conformation. When the Phe_{CN} was substituted on the polar side of the helix, which was partially exposed to water, the C≡N stretching vibration is similar to that of Phe_{CN} in water. In constrast, when Phe_{CN} is introduced at a site that becomes buried in the interior of the protein, the C≡N stretch is similar to that of Phe_{CN} in THF. Together, these results suggest that the C≡N stretching vibration of nitrile-derivatized amino acids can indeed be used as local internal environmental markers, especially for protein conformational studies.

Introduction

There is a great need for spectroscopic reporters that can monitor conformations, folding, and binding in proteins. For example, spectroscopic study of the process of hydrophobic collapse to bury hydrophobic residues in protein folding requires an optical probe that is sensitive to local environmental changes, for example, from mostly polar to mostly hydrophobic. While fluorophores, such as dye molecules, have been used frequently for this purpose, they tend to be large, relative to the size of amino acid side chains in proteins. As a result, the attachment of an external fluorophore to a protein can therefore cause undesirable perturbation to the native structure. By contrast, infrared probes such as NO, CO, or CN are much smaller and may serve as an atomic substitution within an amino acid side chain. Here, we reported the use of nitrile-derivatized amino acids (Figure 1) as internal local environment markers. Because of its small size (Figure 2) and intermediate polarity, a nitrile group (CN) should minimally perturb the native structure of a protein.

For a vibrational mode to be useful as a monitor of the local environment it should be a simple transition, largely decoupled from the rest of the molecule, and it should also have a relatively intense narrow absorption band that is separated from other absorption bands of the molecule. For example, the CO and NO stretching vibrational modes meet these requirements, and these compounds have been extensively used as infrared probes for investigating the structure, dynamics, and function of heme proteins.[1] These groups have the advantage of having high extinction coefficients in a region of the infrared (IR) spectrum that is free of competing absorptions from other functional groups found in proteins. Thus, a single CO or NO bound to a heme protein can be measured with an excellent signal-to-noise ratio. The primary disadvantage of these probes, however, is that they are applicable to only certain heme and metalloproteins where they can bind to metals as ligands. Thus, it would be

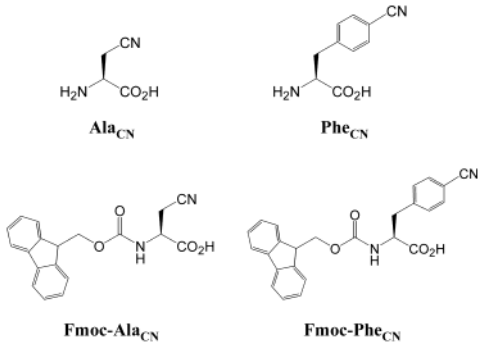

Figure 1. Nitrile-derivatized amino acids used in this study, where Fmoc stands for 9-fluorenylmethoxycarbonyl.

[†] Department of Chemistry.
[‡] Department of Biochemistry and Biophysics.

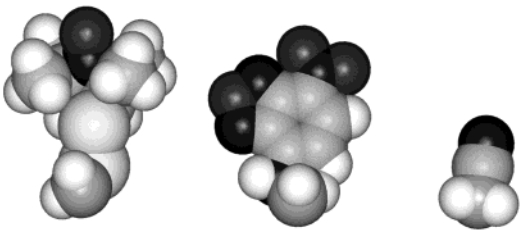

Figure 2. Nitriles are smaller than commonly used spin and fluorescent labels. Space-filling models are shown for a proxyl sulfide spin label (left), an NBD fluorophore (middle), and a nitrile. In each case, the probe is shown connected to a methylene at the bottom of the figure.

quite advantageous to devise a probe that is as small as CO or NO but could be placed *virtually anywhere* within a protein.

Alkyl and aryl nitriles have the potential to meet all the above criteria but do not have the limitation of the necessity to bind as ligands to metal ions. Furthermore, using modern synthetic and biological methods[2] now it is possible to place an unnatural amino acid in particular regions of a polypeptide chain or a protein matrix. Thus, we aimed to develop nitriles as infrared probes for structural and dynamic studies. The nitrile stretching mode is well understood experimentally and theoretically. Reimer and Hall[3] as well as others[4,5] have examined solvation effects of acetonitrile, and their results show the sensitivity of the C≡N vibrational mode to the environment (or solvent). Caughey and co-workers demonstrated that cyanide could be used as infrared probes for investigating the ligand binding sites of hemeproteins.[6] In a continuation of their elegant work on vibrational Stark effects, Andrews and Boxer have measured the Stark effect of a series of nitriles.[7] They found a linear Stark tuning rate of 0.01/*f* to 0.04/*f* Debye, indicating that nitriles might be used as local probes of electric field in chemical and biological systems.

Compared to other site-directed probes, such as the widely used nitroxide spin labels[8] and fluorescent dye molecules, the nitrile moiety is much smaller (Figure 2). Thus, a nitrile can serve as an atomic substitution within an amino acid side chain, whereas fluorescence and electron spin resonance (ESR) probes have been used as substitutions for individual residues within a protein. Therefore, the perturbation to the native conformation caused by introducing a nitrile group into the side chain of an amino acid should be minimal. Further, a nitrile has a polarity between that of an amide group and a methylene, and thus it should be readily accommodated in the hydrophobic interior of a protein or a membrane as well as on the hydrophilic surface of a water-soluble protein. Thus, nitrile-derivatized amino acids can potentially be used as versatile infrared probes for investigating a wide variety of structural and dynamic questions, provided the C≡N stretching vibration is sufficiently sensitive to environment.

Experimental Section

Materials. Tetrahydrofuran (Aldrich, 99.9%), acetonitrile (Fisher, HPLC grade), *p*-tolunitrile (Aldrich, 98%), Ala_{CN} and Phe_{CN} (Bachem), CaM (Sigma) were used without further purification. Millipore water was used to prepare aqueous solutions. The $MLCK_{CN}$ peptides were synthesized using standard Fmoc-protocol employing Pal resin and were purified to homogeneity and characterized by electrospray-ionization mass spectroscopy. Fmoc-Ala_{CN} and Fmoc-Phe_{CN} were prepared from Ala_{CN} and Phe_{CN}, respectively, by Fmoc protection of the amine using standard methods.[9] Briefly, a solution of Fmoc-*N*-hydroxysuccinimide ester in CH_3CN was added to Phe_{CN} or Ala_{CN} dissolved in aqueous 10% Na_2CO_3; upon completion of the reaction the product was isolated by sequential basic and acidic extractions with diethyl ether.

Equilibrium Infrared and Fluorescence. FTIR spectra were collected on a Magna-IR 860 spectrometer (Nicolet) using 1 cm^{-1} resolution. A CaF_2 sample cell that is divided into two compartments using a Teflon spacer (52 μm) was employed to allow the separate measurements of the sample and reference (solvent) under identical conditions. The temperature was maintained at 20 °C for all measurements. To correct for slow instrument drift, an automated translation stage was used to move both the sample and reference side in and out of the IR beam alternately, and each time a single-beam spectrum corresponding to an average of 8 scans was collected. The final result was usually an average of 32 such spectra, both for the sample and the reference. The spectra shown correspond to the net absorbance of the solutes, and the resulting C≡N stretching vibrational band was modeled by either a Lorentzian or a Gaussian line shape function plus a linear background using the solver function in MS Excel. Peptide−CaM binding was studied using fluorescence spectroscopy at room temperature. Fluorescence spectra were collected on a Fluorolog3 (SPEX, NJ) fluorometer with a standard 1-cm quartz cuvette with 1-nm excitation and emission resolutions.

Results and Discussion

The FTIR spectra of two nitrile-derivatized amino acids, Ala_{CN} and Phe_{CN} (Figure 1), and their amine-protected forms used in peptide synthesis, which are, Fmoc-Ala_{CN} and Fmoc-Phe_{CN}, were measured in H_2O, THF or THF−H_2O mixtures. THF was used to approximate the hydrophobic interior of proteins because it is an aprotic solvent and its dielectric constant is roughly 10 times smaller than that of H_2O. Because of the low solubility, the IR spectra of Ala_{CN} and Phe_{CN} in pure THF were not obtained. Instead, Fmoc-Ala_{CN} and Fmoc-Phe_{CN} were used in this case. For comparison, the IR spectra of two analogues of Ala_{CN} and Phe_{CN}, acetonitrile and *p*-tolunitrile, were also measured. It was found that the C≡N stretching vibration of these molecules in both H_2O and THF can be modeled reasonably well by a Lorentzian function (Figure 3) although deviation from a symmetric line shape is clearly visible in some cases, especially for those of acetonitrile and *p*-tolunitrile in THF (Supporting Information). The fitting param-

(1) Park, E. S.; Boxer, S. G. *J. Phys. Chem. B* **2002**, *106*, 5800. Stavrov, S. S.; Wright, W. W.; Vanderkooi, J. M.; Fidy, J.; Kaposi, A. D. *Biopolymers* **2002**, *67*, 255. Rubtsov, I. V.; Zhang, T.; Nakajima, H.; Aono, S.; Rubtsov, G. I.; Kumazaki, S.; Yoshihara, K. *J. Am. Chem. Soc.* **2001**, *123*, 10056. Ding, X. D.; Weichsel, A.; Andersen, J. F.; Shokhireva, T. K.; Balfour, C.; Pierik, A. J.; Averill, B. A.; Montfort, W. R.; Walker, F. A. *J. Am. Chem. Soc*. **1999**, *121*, 128. Phillips, G. N., Jr.; Teodoro, M.; Li, T.; Smith, B.; Gilson, M. M.; Olson, J. S. *J. Phys. Chem. B* **1999**, *103*, 8817. Reddy, K. S.; Yonetani, T.; Tsuneshige, A.; Chance, B.; Kushkuley, B.; Stavrov, S. S.; Vanderkooi, J. M. *Biochemistry* **1996**, *35*, 5562.
(2) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498. Döring, V.; Mootz, H. D.; Nangle, L. A.; Hendrickson, T. L.; de Crécy-Lagard, V.; Schimmel, P.; Marlière, P. *Science* **2001**, *292*, 501. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 19.
(3) Reimers, J. R.; Hall, L. E. *J. Am. Chem. Soc.* **1999**, *121*, 3730.
(4) Pace, E. L.; Noe, L. J. *J. Chem. Phys.* **1968**, *49*, 5317. Eaton, G.; Pena-Nuñez, A. S.; Symons, M. C. R. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84*, 2181. Fawcett, W. R.; Liu, G. J.; Kessler, T. E. *J. Phys. Chem*. **1993**, *97*, 9293.
(5) Nyquist, R. A. *Appl. Spectrosc.* **1990**, *44*, 1405.
(6) Yoshikawa, S.; O'Keeffe, D. H.; Caughey, W. S. *J. Biol. Chem.* **1985**, *260*, 3518.
(7) Andrews, S. S.; Boxer, S. G. *J. Phys. Chem. A* **2000**, *104*, 11853. Andrews, S. S.; Boxer, S. G. *J. Phys. Chem. A* **2002**, *106*, 469.
(8) Hubbell, W. L.; Gross, A.; Langen, R.; Lietzow, M. A. *Curr. Opin. Struct. Biol.* **1998**, *8*, 649.
(9) Carpino, L. A.; Han, G. Y. *J. Org. Chem*. **1972**, *37*, 3404. Milton, R. C. de L.; Becker, E.; Milton, S. C. F.; Baxter, J. E. J.; Elsworth, J. F. *Int. J. Pept. Protein Res.* **1987**, *30*, 431.

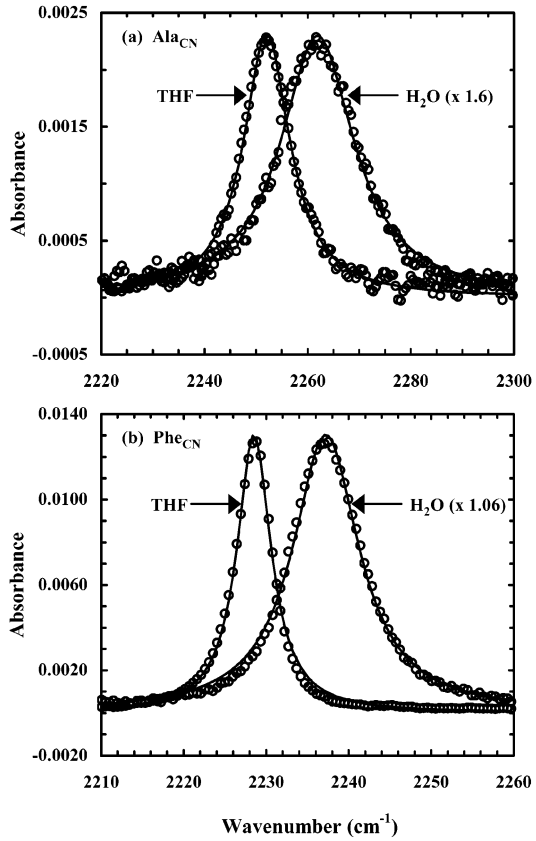

Figure 3. FTIR spectra (in C≡N stretching region) of Ala_{CN} and Phe_{CN} in H_2O and Fmoc-Ala_{CN} and Fmoc-Phe_{CN} in THF, as indicated. A linear background has been subtracted. The solid lines are fits to a Lorentzian function. The fitting parameters are given in Table 1. The concentrations, estimated by weight, were Ala_{CN} ~5 mM, Phe_{CN} ~10 mM, Fmoc-Ala_{CN} ~12 mM, and Fmoc-Phe_{CN} ~10 mM, respectively.

Table 1. Band Position (ν) and Full Width at Half Maximum ($\Delta\nu$) of the Stretching Vibration of C≡N in Different Solvents and Molecules

solvent→	H_2O		THF	
solute↓	ν(cm^{-1})	$\Delta\nu$(cm^{-1})	ν(cm^{-1})	$\Delta\nu$(cm^{-1})
Ala_{CN}	2261.9	18.2	-	-
Fmoc-Ala_{CN}	-	-	2252.1	11.0
acetonitrile	2260.1	9.0	2251.8	7.9
Phe_{CN}	2237.2	9.8	-	-
Fmoc-Phe_{CN}			2228.5	5.0
p-tolunitrile	2234.4	8.4	2228.8	5.0
$MLCK_{3CN}$	2235.2	12.4	-	-
$MLCK_{3CN}$+CaM	2228.4	8.3	-	-
$MLCK_{5CN}$	2233.9	14.4	-	-
$MLCK_{5CN}$+CaM	2235.1	17.4	-	-

eters are summarized in Table 1. These results indicate that the C≡N stretching vibrational modes of both Ala_{CN} and Phe_{CN} are sensitive to solvent, as expected.

As one of the commonly used organic solvents, the IR spectra of acetonitrile have been studied extensively in a wide variety of solvents. It is well-known that the C≡N stretching vibration (ν_2) in acetonitrile is sensitive to environment. For example, a large blue shift of the ν_2 band has been observed for acetonitrile aqueous solutions containing inorganic salts, due to coordination of a cation with the CN moiety. Organic solvents were also found to influence the frequency of the ν_2 band. A pseudolinear empirical correlation was observed between the Gutmann's accept number (AN)[10] of the solvent and the frequency shift of the ν_2 band.[5] Recently, Reimer and Hall[3] have discussed in detail the solvation of acetonitrile in 33 solvents using a theoretical model that takes into account both specific and nonspecific interactions between the solute and solvent molecules. We found that the C≡N stretching vibration of acetonitrile has a maximum at 2260.1 cm^{-1} in H_2O and 2251.8 cm^{-1} in THF. These numbers match well with previously reported values.[3] Compared to its analogue, acetonitrile, Ala_{CN} has a much broader C≡N stretching vibration in H_2O, indicative of a greater inhomogeneous broadening effect due to the zwitterionic character of Ala_{CN}. In THF, however, the C≡N stretching band of Fmoc-Ala_{CN} is almost identical to that of acetonitrile (Table 1), suggesting that the solvent−solute interaction in this case is largely nonspecific in nature. Even under the very dilute conditions used in current study (~10 mM), the C≡N stretching band of acetonitrile is still asymmetric, especially when THF was used as solvent. The nature of this asymmetry has been a subject of intensive studies, and a number of explanations, including microheterogeneity,[11] dimer formation,[12] and hot band,[13] have been proposed. Our results seemed to be in agreement with the microheterogeneity model.

For Phe_{CN} and its analogues, the C≡N stretching vibration affords similar characteristics as those observed for Ala_{CN}. For example, in H_2O, Phe_{CN} has a CN band centered at 2237.2 cm^{-1} with a width of 9.8 cm^{-1}, whereas in THF this band shifts to 2228.5 cm^{-1} and concomitantly narrows its width to 5.0 cm^{-1}. A noticeable difference between Ala_{CN} and Phe_{CN} is in their absorption cross sections. On the basis of the measured absorbance as well as the estimated concentration, the oscillator strength of the C≡N stretching vibration in Phe_{CN} is roughly four times larger than that of Ala_{CN}. This enhancement is likely due to the phenyl ring which facilitates the creation of a larger charge separation accompanied with the nitrile vibrational transition.

Compared to THF, the C≡N stretching vibration in H_2O gives rise to a broader absorption profile with a higher frequency. An increased width suggests that H_2O can support a larger number of solvent−solute complexes whose structures remain static on the time scale of the vibrational transition, a likely result for hydrogen-bonded solvents. Furthermore, the formation of a direct hydrogen bond between a H_2O molecule and the CN group, also quite likely for this solvent, would cause the net bond order along the CN axis to increase due to the alleviation of the antibonding character of the CN bond upon complexation,[14] resulting in a higher vibrational frequency.[15] Consistent with this picture, the ν_2 band of acetonitrile was found to depend on solvent acidity.[5]

To further test how the nitrile stretching vibration behaves in a more complicated environment, for example, in mixed solvent system, the C≡N stretching vibrations of Phe_{CN} in THF−H_2O mixtures were also measured (Figure 4). It is found that three Lorentzians, which are labeled as water-like (WL), THF-like (TL), and water−THF (WT), respectively, are required to describe these vibrational bands. Globally fitting all the

(10) Gutmann, V.; Resch, G.; Linert, W. *Coord. Chem. Rev.* **1982**, *43*, 133.
(11) Moreau, C.; Douhéret, G. *J. Chim. Phys.* **1974**, *71*, 1313. Kovacs, H.; Laaksonen, A. *J. Am. Chem. Soc.* **1991**, *113*, 5596. Marcus, Y.; Migron, Y. *J. Phys. Chem.* **1991**, *95*, 400.
(12) Loewenschuss, A.; Yellin, N. *Spectrochim. Acta* **1975**, *31A*, 207.
(13) Fini, G.; Mirone, P. *Spectrochim. Acta* **1976**, *32A*, 439.
(14) Cho, H.-G.; Sheong, B.-S. *J. Mol. Struct. (THEOCHEM)* **2000**, *496*, 185.
(15) Purcell, K. F.; Drago, R. S. *J. Am. Chem. Soc.* **1966**, *88*, 919.

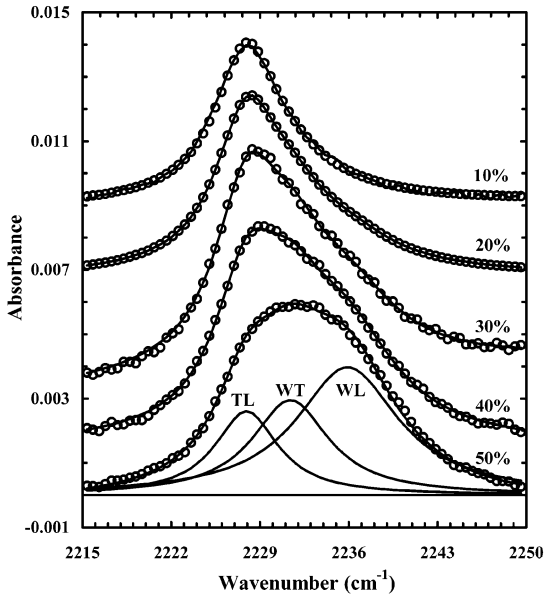

Figure 4. FTIR spectra of Phe_{CN} (in C≡N stretching region) in THF−H_2O solutions. The vol % of H_2O in these solvents is listed in the plot. These spectra were modeled (solid lines) globally by three Lorentzian functions, which are labeled as water-like (WL), THF−like (TL), and water-THF (WT), respectively. The bands that make up the fit for the 50% H_2O solution are shown. The concentration of Phe_{CN} in these solvents was ∼5−7 mM, estimated by weight.

spectra yields the following parameters, WL: $\nu = 2235.9\ cm^{-1}$, $\Delta\nu = 9.8\ cm^{-1}$; TL: $\nu = 2227.9\ cm^{-1}$, $\Delta\nu = 6.0\ cm^{-1}$; WT: $\nu = 2231.4\ cm^{-1}$, $\Delta\nu = 7.4\ cm^{-1}$. The fact that three bands can fit all the spectra suggests that at least three major species (or solvation states) exist in this mixed-solvent system. On the basis of the band position and width, we speculate that the WL band arises from Phe_{CN} molecules that are solvated directly by H_2O molecules and the TL band arises from Phe_{CN} molecules that are solvated directly by THF molecules, whereas the WT band is due to Phe_{CN} molecules that are simultaneously solvated by both THF and H_2O molecules or by THF−H_2O complexes. Indeed, previous studies have shown that in a mixture of THF and H_2O, some THF molecules are associated with H_2O molecules. Although other effects, such as long-range electrostatic interactions, also influence the state of solvation, the current results suggest that the specific solute−solvent interactions play a much more important role in controlling the C≡N stretching vibrational frequency.

The above picture is further supported by the fact that the band intensities change as a function of H_2O content in the mixed solvent (Figure 5). It is clear that the WT component reaches a maximum at a point where the solution contains roughly 40% H_2O (by volume), while the other two bands vary monotonically with increasing of H_2O, indicating that the solvation of the Phe_{CN} molecules can indeed be divided into three different regions. This conclusion is in agreement with the results of Katz et al.,[16] who also suggested that a THF−H_2O mixture could be treated as a ternary system where free THF molecules, free H_2O molecules, and THF−H_2O complexes are present. Recent studies by Scott as well as Mistry et al.[17] further indicated that at low H_2O content a THF−H_2O mixture

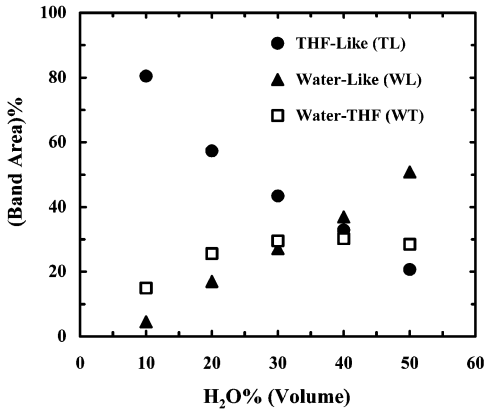

Figure 5. Percentage of the integrated area of the TL, WL, and WT bands as a function of H_2O% (by volume) in THF.

can be regarded largely as a binary system, with mostly free THF molecules and THF−H_2O complexes. This is also consistent with the results from current study. As indicated (Figure 5), when water's volume fraction decreases to ∼10%, the IR band associated with free H_2O molecules becomes negligible. In mixed-solvent systems, however, solvation can well be complicated, and the final result is an equilibrium that depends on the complex web of interactions among the same and different solvent molecules and also interactions between solvent and solute molecules. Sometimes preferential solvation can occur where the solute molecules interact strongly with only one type of solvent molecules. For the later case, using a solute to probe the properties of a mixed-solvent system is apparently inadequate. For example, it is not clear whether the WT component is due completely to a THF−H_2O complex described in other studies or THF-induced distortion to the hydrogen bond formed between water and the CN group, because of the existence of THF. Although the details of the solvation of the CN group in these solvents are interesting and merit further studies, these results nevertheless support our view that the C≡N stretching can be used as an environment sensor.

It is well-known that the C≡N stretching frequencies of aryl nitriles are lower than those of alkyl nitriles.[18] Our results on nitrile-derivatized amino acids also confirmed this trend. As indicated (Table 1), in both H_2O and THF the C≡N stretching vibration of Phe_{CN} is lower in frequency but narrower, by roughly a factor of 2, than that of Ala_{CN}, indicating that the structure of the amino acid is also important in determining the solvation of the CN group. Compared to that of Ala_{CN}, the C≡N stretching band of Phe_{CN} is red-shifted by as much as $\sim 25\ cm^{-1}$ in H_2O, due presumably to the π electrons of the phenyl ring. It has been observed that an increase in the electron density on the CN group shifts its stretching frequency to lower wavenumbers. For example, the gas-phase $C{\equiv}N^-$ has a fundamental vibrational frequency[19] of ca. $2035\ cm^{-1}$, whereas the Q-branch absorption of the gas-phase acetonitrile[3] exhibits a C≡N stretching frequency of $2266.5\ cm^{-1}$. Indeed, the electron-rich phenyl ring would effectively increase the electron density of the CN group and consequently weaken the CN bond because of the antibonding character of the highest occupied molecular orbital state. As a result, the C≡N stretching vibration of Phe_{CN}

(16) Katz, E. D.; Ogan, K.; Scott, R. P. W. *J. Chromatogr.* **1986**, *352*, 67.
(17) Scott, R. P. W. *J. Liq. Chromatogr. Relat. Technol.* **2000**, *23*, 3083. Mistry, K.; Cortes, H.; Meunier, D.; Schmidt, C.; Feibush, B.; Grinberg, N.; Krull, I. *Anal. Chem.* **2002**, *74*, 617.
(18) Skinner, M. W.; Thompson, H. W. *J. Chem. Soc.* **1955**, 487.
(19) Bradforth, S. E.; Kim, E. H.; Arnold, D. W.; Neumark, D. M. *J. Chem. Phys.* **1993**, *98*, 800. Reimers, J. R.; Hush, N. S. *J. Phys. Chem. A* **1999**, *103*, 10580.

shifts to lower frequencies and is largely separated from that of Ala_{CN}. As a matter of fact, there is virtually no overlap between the two bands. This result may turn out to be quite useful, because it might allow the simultaneous use of two or more nitrile moieties in a protein.

In many protein conformational studies, such as temperature-induced folding/unfolding experiments,[20] temperature is subjected to change. To investigate whether the intrinsic temperature dependence of the C≡N stretching frequency would obscure its dependence on environment, we studied the C≡N stretching vibration of Phe_{CN} as a function of temperature (Figure 6). As indicated, the peak frequency of the C≡N stretching vibration depends linearly on temperature, with a slope of -0.048 cm^{-1}/°C. This result shows that the C≡N stretching vibration exhibits only a relatively weak temperature dependence, compared to that induced by environment changes. Therefore, such a probe can be used in conformational studies involving changes in temperature. It is noteworthy, however, that the bandwidth of the C≡N stretching vibration becomes narrower when temperature is increased, suggestive of a more homogeneous environment at higher temperatures.

To test the utility of Phe_{CN} as a probe of the environment within a peptide, we prepared two derivatives (the $MLCK_{CN}$ peptides) of the calmodulin binding domain of skeletal muscle myosin light chain kinase (MLCK):[21]

$MLCK_{579\text{-}595}$: RRWKKNFIAVSAANRFK-$CONH_2$

$MLCK_{3CN}$: RR Phe_{CN} KKNFIAVSAANRFK-$CONH_2$

$MLCK_{5CN}$: RRWK Phe_{CN} NFIAVSAANRFK-$CONH_2$

Peptides derived from this region of MLCK are known to bind to CaM in a Ca^{2+}-dependent manner with very high affinity. Upon binding to CaM, Trp581 is sequestered into a very hydrophobic pocket within the C-terminal lobe of CaM,[22] which can accommodate a number of other aromatic groups,[23] whereas Lys583 remains exposed to a polar environment, as indicated by the NMR structure of the MLCK−CaM complex determined by Bax and co-workers.[24] Thus, it was expected that if Phe_{CN} were substituted for Trp581, the nitrile stretching of the $MLCK_{3CN}$−CaM complex would occur at a position similar to that observed in THF. Similarly, if Phe_{CN} were substituted for Lys583, the $MLCK_{5CN}$−CaM complex would display a C≡N stretching band similar to that observed in H_2O.

The binding of the $MLCK_{CN}$ peptides to CaM was investigated by fluorescence spectroscopy (Figure 7 and Supporting Information), using methods described previously.[25] Upon addition of CaM into the solution containing $MLCK_{5CN}$, the fluorescence emission of Trp581 increases in intensity and concomitantly shifts its maximum to lower wavelength (Figure 7). This result is similar to that observed for binding between

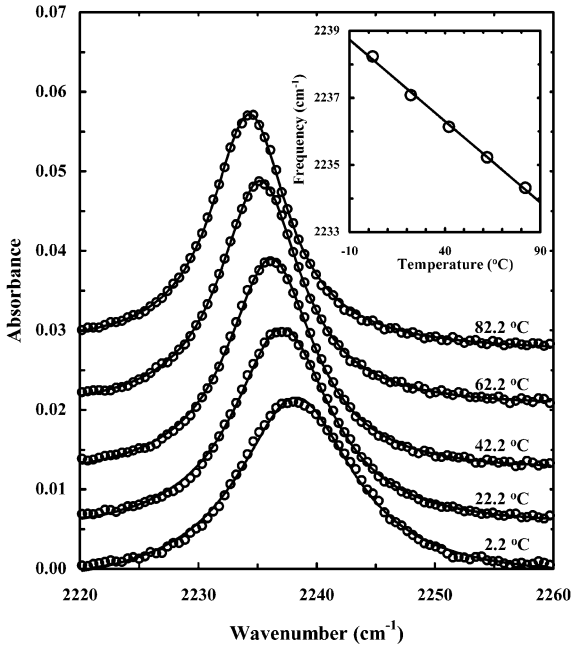

Figure 6. FTIR spectra of Phe_{CN} (in C≡N stretching region) in H_2O at different temperatures, as indicated. The solid lines are fit to a Lorentzian function plus a linear background. The band position (ν, in cm^{-1}) and full width at half-maximum ($\Delta\nu$, in cm^{-1}) for each temperature are summarized below. 2.2 °C: $\nu = 2238.2$, $\Delta\nu = 11.9$; 22.2 °C: $\nu = 2237.1$, $\Delta\nu = 10.1$; 42.2 °C: $\nu = 2236.1$, $\Delta\nu = 9.0$; 62.2 °C: $\nu = 2235.2$, $\Delta\nu = 8.3$; 82.2 °C: $\nu = 2234.3$, $\Delta\nu = 7.8$. (Inset) Temperature-dependent band positions. The solid line is the linear regression to the data points, which yields a slope of -0.048 cm^{-1}/°C.

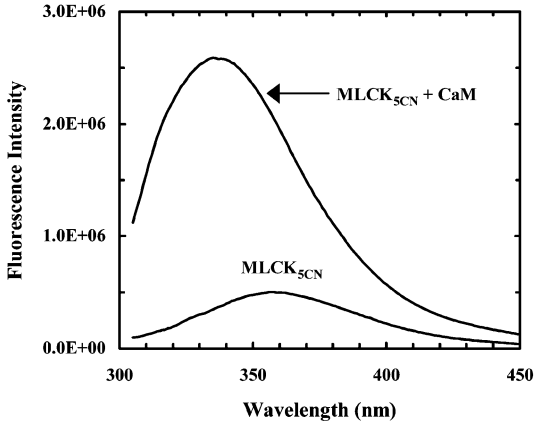

Figure 7. Fluorescence spectra of $MLCK_{5CN}$ and $MLCK_{5CN}$ + CaM, as indicated. The concentration of both CaM and $MLCK_{5CN}$ peptide was 5 μM, in a buffer solution containing 1 mM Tris-HCl, pH 7.0, 0.5 mM $CaCl_2$. These spectra were collected with an excitation wavelength of 295 nm and excitation and emission resolutions of 1 nm, respectively. The $MLCK_{5CN}$ peptide contains a Trp, whereas CaM lacks Trp, allowing fluorescence to be used to monitor binding. Compared to that of free $MLCK_{5CN}$, the fluorescence spectrum of $MLCK_{5CN}$ + CaM system has an emission maximum shifted to the blue and a higher intensity, indicating that the $MLCK_{5CN}$ peptide binds to CaM under these conditions.

$MLCK_{579-595}$ and CaM, indicating that Phe_{CN} is an acceptable replacement of Lys583 and that $MLCK_{5CN}$ can indeed form a tightly bound complex with CaM. The binding of $MLCK_{3CN}$ to CaM was investigated by a competition assay[23] because it lacks a Trp side chain. A titration of $MLCK_{3CN}$ into CaM−MLCK complex led to dissociation of the MLCK peptide, indicating that the $MLCK_{CN}$ peptide binds to CaM with a dissociation constant within a factor two of that of MLCK (which is in the

(20) Huang, C.-Y.; Getahun, Z.; Zhu, Y. J.; Klemke, J. W.; DeGrado, W. F.; Gai, F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2788.
(21) Blumenthal, D. K.; Takio, K.; Edelman, A. M.; Charbonneau, H.; Titani, K.; Walsh, K. A.; Krebs, E. G. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 3187.
(22) Ikura, M.; Barbato, G.; Klee, C. B.; Bax, A. *Cell Calcium* **1992**, *13*, 391.
(23) O'Neil, K. T.; DeGrado, W. F. *Proteins* **1989**, *6*, 284. Changenet-Barret, P.; Choma, C. T.; Gooding, E. F.; DeGrado, W. F.; Hochstrasser, R. M. *J. Phys. Chem. B* **2000**, *104*, 9322.
(24) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632.
(25) Erickson-Viitanen, S.; DeGrado, W. F. *Methods Enzymol.* **1987**, *139*, 455. Ghirlanda, G.; Lear, J. D.; Lombardi, A.; DeGrado, W. F. *J. Mol. Biol.* **1998**, *281*, 379.

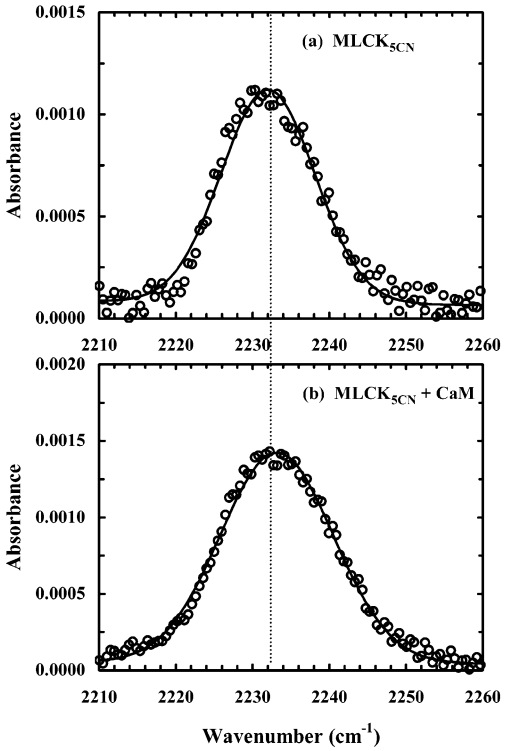

Figure 8. FTIR spectra of the $MLCK_{5CN}$ peptide (in C≡N stretching region) with and without CaM present, as indicated in the plot. The solid lines are fits to a Gaussian line shape function plus a linear background. The fitting parameters are summarized in Table 1. The peptide concentration was ∼1 mM in (a), ∼1 mM in (b); the CaM concentration was also ∼1 mM in (b), estimated by weight.

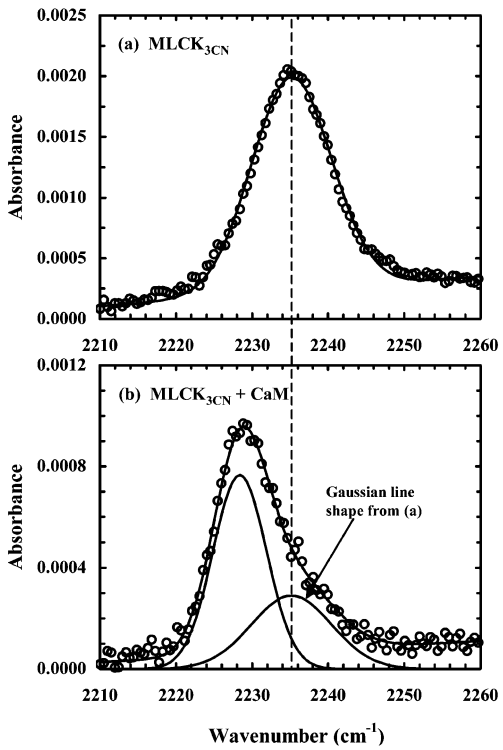

Figure 9. FTIR spectra of the $MLCK_{3CN}$ peptide (in C≡N stretching region) with and without CaM present, as indicated in the plot. The solid lines passing through the data (open circle) are fits to (a) a Gaussian line shape function plus a linear background and (b) a Gaussian line shape function plus the Guassian function recovered from (a), as indicated, and a linear background. The fitting parameters are summarized in Table 1. The peptide concentration was ∼3 mM in (a), ∼2 mM in (b); the CaM concentration was also ∼2 mM in (b), estimated by weight.

low nM range[25]). Thus, Phe_{CN} is also a good functional replacement of Trp581.

As expected, binding[26] to CaM also resulted in a prominent change in the nitrile stretching of $MLCK_{3CN}$ but not $MLCK_{5CN}$. With or without CaM present, the C≡N stretching band of $MLCK_{5CN}$ is almost identical, and its position is close to that observed of Phe_{CN} in H_2O (Figure 8 and Table 1). Thus, the Phe_{CN} of $MLCK_{5CN}$ experiences a similar environment in both the free and bound state, as expected from its polar environment in the CaM−MLCK peptide complex.[24] It is noticed, however, that the nitrile stretching band of the complex is slightly broader than that of the free peptide, due perhaps to either the binding inhomogeniety or uncomplexed peptides whose CN groups may experience a slightly different environment than that of the bound peptides. Another possibility is due to the conformational change of MLCK associated with complexation. It is known that upon binding with CaM the MLCK peptide undergoes a change in conformation, from random to helical. In the absence of CaM, $MLCK_{3CN}$ showed a peak at 2235.2 cm^{-1} (Figure 9 and Table 1), similar to the value of 2237.2 cm^{-1} for the free amino acid in H_2O. Upon addition of approximately 1.0 equiv of CaM, the peak shifted to 2228.4 cm^{-1} (Figure 9), which compares well to the value of 2228.5 cm^{-1} observed for the free Fmoc-Phe_{CN} in THF. In the spectrum of the complex, a small shoulder appears at higher wavenumbers, which is probably due to a small amount of uncomplexed peptide or a heterogeneity of the environment of the nitrile in the complex or both. Further experiments with higher CaM concentrations, compared to that of $MLCK_{3CN}$ peptide, should help to clarify this ambiguity. Such an experiment is currently underway. Regardless of the origin of this small shoulder, however, these results again support our view that the CN stretching vibration can indeed be used as an internal marker for conformational studies.

By far, the most commonly used infrared probe in protein conformational studies is the amide I band of polypeptides, which arises mainly from the amide C═O stretching vibration and is sensitive to conformation.[27] However, the limitation of the amide I band is that it cannot be employed to report on the local environment of amino acid side chains other than Gln and Asn. Thus, the use of nitrile-derivatized amino acids as local protein environmental sensors is complementary to infrared methods that involve the use of amide bands as conformational reporters. Because of the intrinsic sensitivity of the CN stretching vibration to many factors, however, a straightforward interpretation of the observed shift in the CN stretching frequency may not be easy or even obtainable in some cases. Nevertheless, such a probe would be particular useful to monitor the kinetics of conformational changes whose starting and end points are well defined, for example, protein folding. In addition, this method is in spirit equivalent to that of using Trp fluorescence

(26) The MLCK−CaM binding study was carried out in a buffer solution containing 0.6 M Tris-HCl, pH 7.0, + 0.3 M $CaCl_2$.

(27) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181. Kalnin, N. N.; Baikalov, I. A.; Venyaminov, S. Y. *Biopolymers* **1990**, *30*, 1273. Hamm, P.; Lim, M.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2036. Manas, E. S.; Getahun, Z.; Wright, W. W.; DeGrado, W. F.; Vanderkooi, J. M. *J. Am. Chem. Soc.* **2000**, *122*, 9883. Huang, C.-Y.; Klemke, J. W.; Getahun, Z.; DeGrado, W. F.; Gai, F. *J. Am. Chem. Soc.* **2001**, *123*, 9235.

as probes in protein conformational studies. Although the photophysics of Trp are not fully understood, it still serves as one of the most commonly used fluorescent probes in binding and folding studies.

In conclusion, nitrile-derivatized amino acids can be used as local environment sensors. It is especially useful to employ such a probe to monitor processes in which a particular amino acid undergoes large environmental changes, for example, from polar to hydrophobic or vice versa. When combined with modern chemical and biological[2] synthetic methods and infrared spectroscopy, nitrile-derivatized amino acids may be used to study in great detail a variety of structural and dynamic questions, such as side chain packing, side chain−side chain interaction, peptide−protein binding, and protein−protein interaction. Currently, we are employing such a probe to investigate the initial formation and subsequent consolidation of the hydrophobic core in protein folding.

Acknowledgment. F.G. gratefully acknowledges financial support from Research Corporation, the University of Pennsylvania Research Foundation, the NSF (CHE-0094077), and the NIH (GM-065978). B.D.L. is an NSF-REU summer student. We also thank Dr. M. J. Therien and his group for assistance with fluorescence measurements.

Supporting Information Available: Fluorescence binding study of $MLCK_{3CN}$ to CaM; infrared spectra of acetonitrile and *p*-tolunitrile in H_2O and THF (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0285262