Published in final edited form as:

Phys Chem Chem Phys. 2016 March 14; 18(10): 7027–7034. doi:10.1039/c5cp04413h.

C≡N Stretching Vibration of 5-Cyanotryptophan as an Infrared Probe of Protein Local Environment: What Determines Its Frequency?

Wenkai Zhang^{1,2,#}, Beatrice N. Markiewicz^{1,#}, Rosalie S. Doerksen¹, Amos B. Smith III¹, and Feng Gai^{1,2,*}

¹Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104

²The Ultrafast Optical Processes Laboratory, University of Pennsylvania, Philadelphia, PA 19104

Abstract

Recently it has been suggested that the C≡N stretching vibration of a tryptophan analog, 5-cyanotryptophan, could be used as an infrared probe of the local environment, especially the hydration status, of tryptophan residues in proteins. However, the factors that influence the frequency of this vibrational mode are not understood. To determine these factors, herein we carried out linear and nonlinear infrared measurements on the C≡N stretching vibration of the sidechain of 5-cyanotryptophan, 3-methyl-5-cyanoindole, in a series of protic and aprotic solvents. We found that while the C≡N stretching frequencies obtained in these solvents do not correlate well with any individual Kamlet-Taft solvent parameter, i.e., π^* (polarizability), β (hydrogen bond accepting ability), and α (hydrogen bond donating ability), they do however, collapse on a straight line when plotted against $\sigma = \pi^* + \beta - \alpha$. This linear relationship provides a firm indication that both specific interactions, i.e., hydrogen-bonding interactions with the C≡N (through α) and indole N-H (through β) groups, and non-specific interactions with the molecule (through π^*) work together to determine the C≡N stretching frequency, thus laying a quantitative framework for applying 5-cyanotryptophan to investigate the microscopic environment of proteins in a site-specific manner. Furthermore, two-dimensional and pump-probe infrared measurements revealed that a significant portion (~31%) of the ground state bleach signal has a decay time constant of ~12.3 ps, due to an additional vibrational relaxation channel, making it possible to use 5-cyanotryptophan to probe dynamics occurring on a timescale on the order of tens of picoseconds.

Keywords

Infrared probe; 5-Cyanotryptophan; Nitrile vibration; 2D IR spectroscopy; Hydrogen bond; Kamlet-Taft parameters

*Corresponding Author: gai@sas.upenn.edu.

#Equal contribution

SUPPORTING INFORMATION

Electronic Supplementary Information (ESI) available: Details of 3M5CI synthesis, kinetic model, and all FTIR spectra. See DOI: XXX

1. INTRODUCTION

Vibrational spectroscopy is a powerful tool for assessing the structure and conformational dynamics of proteins. While many intrinsic vibrational modes, such as the amide I mode of the protein backbone, have been used for this purpose, they often lack the ability to reveal site-specific information due to spectral overlapping and/or vibrational coupling. To overcome this limitation, the past decade has seen an increased effort towards the development of unnatural amino acid-based infrared (IR) probes that can be used to site-specifically interrogate various structural and environmental properties of proteins, such as the local electrostatic field, hydrogen-bonding (H-bonding) interactions, and degree of hydration.^{1–5} One of those unnatural amino acids is 5-cyanotryptophan (Trp_{CN}). Waegle *et al.*⁶ found that the C≡N stretching vibrational mode of Trp_{CN}, especially its bandwidth, is highly dependent on the percentage of water in water and tetrahydrofuran (THF) mixtures and, thus, suggested that this unnatural amino acid could be used as an IR probe of the local hydration status of proteins. By performing a combined QM/MM study on 5-cyanoindole, they further showed that, besides the H-bonding interactions between water and the C≡N group, the interactions between water and the aromatic indole ring, especially the H-bonding interactions between water and the indole amine (N-H) group, also influenced the C≡N stretching frequency.⁷ More specifically, they found that the H-bonding interactions occurring at the C≡N and N-H groups had different effects on the C≡N stretching frequency, with one (C≡N) shifting the frequency toward higher wavenumbers and the other (N-H) toward lower wavenumbers. This finding is particularly interesting as it suggests that the indole N-H group of Trp_{CN} could be used to sense local H-bonding dynamics through measurement of the C≡N stretching vibration, especially under conditions where the nitrile group is not directly involved in hydrogen bond (HB) formation. To further verify this notion, herein we carried out static and ultrafast IR studies on 3-methyl-5-cyanoindole (3M5CI), which is the sidechain of Trp_{CN} (Figure 1), in different solvents.

Tryptophan (Trp) plays an important role in defining the folding, structure and function of many proteins; it is frequently found at or near sites that are responsible for protein-protein interaction,⁸ ligand binding,^{9,10} protein-DNA interaction,^{11,12} and enzyme catalysis.^{13,14} For example, the M2 proton channel of the influenza A virus uses a Trp tetrad to gate and control proton conduction across the viral membrane in an asymmetric manner after endocytosis, which enables uncoating and release of the viral RNA into the host cell for viral replication.¹⁵ In addition, it is well recognized that Trp plays a key role in anchoring membrane proteins and peptides in lipid bilayers as it is preferentially located at the water-membrane interface.¹⁶ For these reasons, Trp, which fluoresces upon excitation with ultraviolet light, has become one of the most frequently utilized fluorophores in the study of the structure-dynamics-function relationship of proteins via fluorescence spectroscopy. On the other hand, the study of the role of Trp in protein structure and dynamics using IR spectroscopy is scarce. This is because none of the intrinsic IR active vibrational modes of Trp is particularly ideal for being used as a site-specific IR probe of proteins, due to spectral congestion, low extinction coefficients, or insensitivity to local environment.^{17,18} Since IR spectroscopy is capable of offering a higher temporal and sometimes structural resolution than fluorescence spectroscopy, it would be quite useful to confer distinct IR utility to Trp

by adding a relatively non-perturbing exogenous moiety to the indole ring that displays a strong, localized, and environmentally sensitive absorption band located in a non-congested region of the IR spectrum of proteins. While the C≡N stretching vibration of Trp_{CN} seems to meet these requirements, further experimental study is required to delineate the factors that determine its frequency and to lay a quantitative foundation to use this vibrational mode to investigate various biophysical problems, such as the changes in local hydration environment and H-bonding dynamics. Our results showed that the frequency of the C≡N stretching mode of 3M5CI is not simply dictated by the immediate environment of the nitrile group, but instead, it depends on the microscopic surroundings of the entire molecule. In other words, specific and/or non-specific interactions with the C≡N group, the aromatic ring, and the pyrrole N-H group combined determine the position and width of this vibrational band. In addition, we found that when the nitrile group is buried in an aprotic environment, the C≡N stretching frequency exhibits a simple and linear dependence on the polarizability and H-bonding ability of the solvent. We believe that this finding is particularly encouraging and useful in applying Trp_{CN} to probe local HB dynamics in cases where the benzene ring is situated in a dehydrated environment. Further time-resolved measurements indicated that an ultrafast process at the vibrationally excited state produces a long-lived ground state bleach signal ($\tau \approx 12.3$ ps), making it possible to use Trp_{CN} to probe dynamic events occurring on the timescale of tens of ps.

MATERIALS AND METHODS

Materials and sample preparation

The details of the 3M5CI synthesis are given in the ESI. N-methyl-5-cyanoindole (NM5CI) was purchased from Thermo Fischer Scientific (Loughborough, UK) and used as received. The following solvents (spectroscopic grade) were purchased from Acros Organics: methanol (MeOH), 2-propanol, dichloromethane (DCM), dimethyl sulfoxide (DMSO), dimethylformamide (DMF), pyridine, acetophenone, cyclopentanone, tetrahydrofuran (THF), 1,4-dioxane, toluene, tetrachloromethane, and 2,2,2-trifluoroethanol (TFE); and hexafluorobenzene was purchased from Oakwood Products. We checked the water content in two representative solvents, DMSO and THF, using the OH stretching band at ~ 3500 cm^{-1} . As indicated (Figure S1), in both cases the amount of water is negligible. Samples were freshly prepared before use by directly dissolving either 3M5CI or NM5CI in the desired solvent and the final concentration of the solute was approximately 33 mM for static and 100 mM for time-resolved measurements.

Static and time-resolved IR measurements

All static IR measurements were carried out on a Thermo Nicolet 6700 FTIR spectrometer at a resolution of 1 cm^{-1} . A solvent background has been subtracted for each spectrum shown. 2D IR spectra were obtained on a heterodyne-detected photon-echo setup with a boxcar geometry that has been described in detail elsewhere.¹⁹ IR pump-probe data were obtained using a transient absorption spectrometer derived from the 2D IR setup. Briefly, the local oscillator and one of the three pump beams were blocked; one of the remaining two beams was used as the pump, and the other was attenuated and directed to the monochromator to act as the probe. The polarization of the probe was set at the magic angle

with respect to that of the pump. For both the static and time-resolved measurements, the sample solution was placed between two 2 mm CaF₂ windows separated by either a 25 μm (for time-resolved experiment) or 50 μm (for static measurement) spacer.

RESULTS AND DISCUSSION

C $\equiv$ N vibrational bands of 3M5CI in different solvents

It is well known that protic solvents, such as water and simple alcohols, can form HBs with the C $\equiv$ N group of aryl and alkyl nitriles.²⁰ For 3M5CI, such solvents are expected to also interact with the indole N-H group via H-bonding interactions. Thus, it is impossible to distinguish the effects of these two types of H-bonding interactions based on measurements of the C $\equiv$ N stretching vibrational band of 3M5CI in such protic solvents alone. Therefore, we have chosen mostly solvents that can only form HBs with a solute through their H-bonding accepting groups, based on the Kamlet-Taft solvent parameters (Table 1).^{21,22} As expected (Figure 2), the C $\equiv$ N band of 3M5CI shows a clear dependence on solvent. Furthermore, the band shape in each case can be described satisfactorily by a Voigt profile (Figure S2, ESI) and the corresponding spectral parameters are listed in Table 1. A cursory inspection of the results suggests, as predicted by the computational study of Waegle *et al.*,⁷ that H-bonding interactions between solvent and the indole N-H group can have a measurable influence on the stretching frequency of the C $\equiv$ N group located on the other side of the aromatic ring. This can be seen by comparing the spectra obtained in TFE, DMSO and MeOH. In TFE, which is a strong HB donor according to its Kamlet-Taft parameters ($\beta = 0$, $\alpha = 1.51$), the C $\equiv$ N stretching band is centered at 2230.7 cm^{-1} , whereas in DMSO, which is a strong HB acceptor ($\beta = 0.76$, $\alpha = 0$), the C $\equiv$ N stretching band is shifted to 2216.1 cm^{-1} . On the other hand, in MeOH ($\beta = 0.62$, $\alpha = 0.93$), which is capable of forming HBs with both the C $\equiv$ N and N-H groups, the C $\equiv$ N stretching band is located between those obtained in TFE and DMSO and centered at 2223.6 cm^{-1} .

It is noticeable that the C $\equiv$ N bands of 3M5CI obtained in protic solvents, especially in TFE, are much broader than those obtained in aprotic solvents (Table 1), indicating that in protic solvents the inhomogeneity sensed by the C $\equiv$ N stretching vibration becomes larger. Previously, Cho and coworkers²³ have shown that the C $\equiv$ N stretching frequency of acetonitrile is dependent on the HB configuration, leading to inhomogeneous broadening of the band. As shown (Figure S3, ESI), the bandwidth of the C $\equiv$ N stretching vibration of acetonitrile in TFE is approximately 14 cm^{-1} , which is broader than that ($\sim 8 \text{ cm}^{-1}$) in THF.²⁴ This is consistent with the study of Cho and coworkers.²³ In comparison, however, the bandwidth of this vibration in 3M5CI shows a more significant increase upon changing the solvent from THF to TFE (i.e., 8 to 24 cm^{-1}). Taken together, these results indicate that in the case of 3M5CI, besides an inhomogeneous distribution of HB configurations at the C $\equiv$ N site, other solvent interactions, likely H-bonding interactions with the indole ring also contribute significantly toward the broadening of the bandwidth. This is consistent with a recent study of McLain and coworkers,²⁵ which showed that in a methanol-water mixture, instead of being solvated by the methyl groups of methanol, the indole ring is in fact forming HBs with the -OH groups of water and methanol. Further evidence supporting this picture is, as shown (Figure S4, ESI), that for 3M5CI the dependence of the C $\equiv$ N bandwidth

on the overall solvent property, as judged by $\tau = \pi^* + \beta + \alpha$, are different for protic and aprotic solvents. It is worth noting that the bandwidth of the C $\equiv$ N stretching vibration of 3M5CI in hexafluorobenzene was found to be concentration dependent (Figure S5, ESI), which indicates that 3M5CI may oligomerize in this case and, as a result, the measured frequency may not be entirely representative of the monomeric form. Therefore, we excluded this solvent in the following analysis.

This difference could be exploited to probe preferential interactions between the indole ring and a specific solvent component. To illustrate this point, we measured the C $\equiv$ N stretching band of 3M5CI in a binary solvent composed of DMSO and TFE. We chose this mixture because it is well known that at a relatively high molar fraction DMSO will replace the protic solvent molecules, such as water, that are initially solvating the indole ring.²⁶ As shown (Figure 3), the results obtained at different volume ratios of TFE and DMSO are indeed consistent with this picture. It is clear that at a 50:50 ratio the C $\equiv$ N stretching band is already similar to that obtained in pure DMSO, which indicates, as expected, exclusion of TFE molecules from the indole ring.

Quantifying the C $\equiv$ N stretching frequency of 3M5CI in aprotic solvents

While the simple comparison discussed above revealed a picture that is consistent with the notion that HB formations at both the nitrile and N-H sites affect the C $\equiv$ N stretching frequency, a more quantitative assessment of the FTIR data is needed in order to extract the exact contribution of the H-bonding interactions at the N-H site. To do so, we first considered solvents that can only form HBs with the N-H group but not the nitrile moiety. In other words, these solvents all have a Kamlet-Taft α parameter of zero (Table 1). A previous study by Moog and coworkers²⁷ indicated that the C $\equiv$ N stretching frequency of benzonitrile shows a linear dependence on the Kamlet-Taft π^* parameter of solvents with $\alpha = 0$. Since the π^* parameter is a relative scale measuring the solvent's polarizability, this result indicates that the C $\equiv$ N stretching frequency of benzonitrile, which lacks any HB donating groups, is determined only by the local electrostatic field exerted by the surrounding solvent molecules when the C $\equiv$ N group is not involved in any direct H-bonding interactions.²⁷ As indicated (Figure 4A), however, the C $\equiv$ N stretching frequency of 3M5CI does not show a strong linear correlation with the π^* parameter as observed for benzonitrile. Similarly, a strong linear correlation is not observed when other solvent parameters (i.e., β and ϵ) are used alone (Figure 4B and 4C). Thus, these simple analyses indicate that in this case other factors are also at play. Since the most pronounced difference between benzonitrile and 3M5CI, in the context of the current study, is that the latter can interact with the solvent through its N-H group, one needs to consider the effect arising not only from the solvent's polarizability (π^*) but also its HB accepting ability (β). Indeed, we found that the C $\equiv$ N stretching frequency of 3M5CI obtained in solvents with $\alpha = 0$ exhibits a strong linear dependence on $\gamma = \pi^* + \beta$ (Figure 5). This finding provides a strong corroboration for the aforementioned notion that HB formations involving the indole N-H group of Trp_{CN} can be sensed by the C $\equiv$ N stretching vibration. In addition, and perhaps more importantly, this linear relationship makes it possible to use the C $\equiv$ N stretching frequency to characterize changes in the H-bonding interactions between the N-H group of a Trp_{CN} residue and a

neighboring solvent molecule or sidechain in proteins, especially under conditions where the 5-cyano position is immersed in an aprotic environment.

To further validate the conclusions reached above, we measured the C≡N stretching modes of N-methyl-5-cyanoindole (NM5CI, Figure 1) in DMSO since NM5CI is incapable of forming HBs with DMSO due to the added methyl group. As shown (Figure 6), the C≡N band of NM5CI in DMSO is centered at 2217.1 cm⁻¹, which is blue-shifted from that of 3M5CI. In addition, its bandwidth (8.4 cm⁻¹) is narrower than that (10.6 cm⁻¹) of 3M5CI. This blue-shift and band narrowing is once again consistent with the idea that H-bonding interactions with the indole N-H group is an important determinant of the C≡N stretching frequency.

Finally, we attempted to derive a simple relationship, using the Kamlet-Taft parameters alone, to describe the C≡N stretching frequencies obtained in all solvents. As discussed above, direct H-bonding interactions with the nitrile group shifts its stretching frequency to higher wavenumbers. Thus, considering the fact that the α parameter is a measure of the HB donating ability of the solvent and the results presented in Figure 5 are for solvents with $\alpha = 0$, we hypothesized that the simplest surrogate variable that can reasonably capture the overall effect of a solvent on the C≡N stretching frequency of 3M5CI is $\sigma = \pi^* + \beta - \alpha$. As shown (Figure 7), this parameter (σ), to our surprise, proves to be an excellent scale to quantify the C≡N stretching frequencies measured in all solvents, including water. The significant linear correlation exhibited between ω_0 and σ provides a simple and quantitative way to interpret the C≡N stretching frequency of 3M5CI or Trp_{CN}. Despite this success, however, it is worth noting that this linear relationship is obtained based on empirical solvent parameters. Hence, it is impossible to use it to directly yield a microscopic interpretation of the environment of a specific Trp_{CN} residue in proteins. Nonetheless, this relationship validates the notion that, besides the direct interaction with the nitrile group, interactions with other parts of the indole ring can also be sensed by the C≡N stretching vibration, making Trp_{CN} a more versatile IR probe in this regard.

Time-resolved IR measurements

While the bandwidth of a linear IR spectrum is informative about the degree of inhomogeneous broadening, the underlying dynamics can only be assessed by nonlinear spectroscopic techniques, such as 2D IR spectroscopy.²⁸ To further test the feasibility of using Trp_{CN} to probe such dynamics, we carried out 2D IR measurements on the C≡N stretching mode of 3M5CI in a 50:50 DMSO:TFE mixture. The reason that we chose this particular system as the testbed is based on the following considerations: (1) our results indicated, as discussed above (Figure 3), that under this condition the solute experiences a DMSO-like environment, indicating preferential accumulation of DMSO molecules near the indole ring; (2) a previous simulation study by Bagchi and coworkers²⁶ on Trp solvated by DMSO-water mixtures indicated that when the fraction of DMSO is larger than 15%, preferential interactions between the indole ring and DMSO occur wherein the solvent molecules form a distinct network or cluster surrounding the solute enhanced by favorable solvent-solute hydrophobic interactions; (3) we hypothesized that such a scenario also happens in DMSO-TFE mixtures, which can be tested by measuring the spectral diffusion

dynamics²⁸ of the C≡N stretching vibrations via 2D IR spectroscopy as cluster formation has been shown to result in a prolonged spectral diffusion time.^{29,30} As shown (Figure 8), 2D IR spectra of the C≡N band of 3M5CI obtained at 4 different waiting times (T) clearly indicate spectral diffusion dynamics, as manifested by changes in the contour of the 2D peak corresponding to the 0–1 transition. However, even at the longest waiting time of the experiment, 20 ps, this 2D peak still shows an appreciable tilt towards the diagonal direction, signifying the slowness of the spectral diffusion process. Since the spectral diffusion dynamics of the C≡N stretching vibration in simple, pure liquids typically occur on a few ps timescale,^{29,31} this 2D IR result is therefore consistent with our hypothesis that DMSO can form clusters surrounding the indole ring, making its microscopic environment fluctuate at a slower time scale.

Interestingly, the 2D IR spectra also revealed the presence of a second 1–2 transition. Since this spectral signature is not detectable at $T = 0$, it most likely corresponds to a dark state (i.e., the corresponding 0–1 transition is forbidden). A similar phenomenon was also observed for the C≡N stretching vibration of cyanophenol in methanol by Cho and coworkers,³² which was attributed to a combination band arising from the combination of two vibrational modes of the parent molecule. To better characterize the impact of this dark state, we conducted IR pump-probe measurements under the magic angle polarization condition on 3M5CI in DMSO. As shown (Figure 9), the time-resolved spectra clearly reveal the existence of three distinguishable spectral features: one negative band centered at $\sim 2216\text{ cm}^{-1}$, which corresponds to contributions from the ground state bleach (GSB) and stimulated emission (SE) signals, and two positive bands, centered at ~ 2190 and $\sim 2205\text{ cm}^{-1}$, respectively. Because the 2190 cm^{-1} band is more intense at earlier delay times and the 2205 cm^{-1} band grows in with time, the simplest model capable of explaining the decay kinetics of these features is a competing relaxation process from the excited state, $(\text{C}\equiv\text{N})_{v=1}$, to the ground state, $(\text{C}\equiv\text{N})_{v=0}$, (through two channels: one goes through the $v = 1$ state of the dark state and the other takes the initially prepared excited state population directly to the ground state, as indicated in the following kinetics scheme:

where k_{ij} is the rate constant of the corresponding kinetic step. To determine the key rate constants, we analyzed the transient absorption kinetics at two representative probing frequencies, $\omega_A = 2190\text{ cm}^{-1}$ and $\omega_B = 2223\text{ cm}^{-1}$. Because the transient signal at ω_A contains contribution mostly from the excited state absorption from the $v = 1$ state of the

C≡N stretching vibration, the relaxation kinetics at this frequency should reveal $k_1 = k_{10} + k_{11}$. As shown (Figure 10), the signal at ω_A can be satisfactorily described by a single-exponential function with a time constant of 1.3 ± 0.1 ps, indicating that $k_1 = (1.3 \pm 0.1 \text{ ps})^{-1}$. On the other hand, the transient signal at ω_B contains contributions from both channels and, thus, should decay in a double-exponential manner. Indeed, as shown (Figure 10), the transient kinetics at this frequency can be fit to a double-exponential function with $\tau_1 = 1.2 \pm 0.1$ ps and $\tau_2 = 12.3 \pm 1.6$ ps. Based on the rate equations of the above kinetic scheme (see details in ESI), it is easy to show that $\tau_1 = (k_1)^{-1}$ and $\tau_2 = (k_{20})^{-1} = 12.3 \pm 1.6$ ps. It is clear that the τ_1 value is consistent with the k_1 value determined above from the ω_A data. To further determine k_{11} , we took advantage of the fact the ratio (R) between the population going through the two channels is $k_{11}/(2k_{10} - 2k_{20} + k_{11})$. It is straightforward to show (ESI) that $R = A_2/A_1 = 0.18$, where A_1 and A_2 are the amplitudes of the τ_1 and τ_2 components obtained from the double-exponential fit of the ω_B kinetics. Using this ratio and the above determined value of k_1 , we calculated k_{10} and k_{11} to be $(1.8 \text{ ps})^{-1}$ and $(4.7 \text{ ps})^{-1}$, respectively. A previous study has shown that the vibrational lifetime of the C≡N stretching vibration of benzonitrile in DMSO is about 4 ps,³³ which is longer than that (1.8 ps) of 3M5CI. However, due to the second decay channel, a significant population (~31%) is transferred to another excited state, leaving a significant portion of the GSB signal of 3M5CI long lived. This prolonged GSB recovery time (12.3 ps) could be useful in the study of protein dynamic events occurring on the timescale of tens of ps. However, we note that the solute concentration used in the current 2D IR experiments is much higher than that of any typical protein solutions and, hence, the feasibility of this statement requires further test. In addition, future work is needed in order to provide a more comprehensive understanding of the factors that determine the onset and percentage of the aforementioned dark state, especially in aqueous solutions or under biological conditions.

Applying Trp_{CN} to measure local H-bonding dynamics in proteins

In many regards, the indole ring of Trp makes it a unique amino acid: it has a large hydrophobic surface area, a permanent dipole moment, a large aromaticity, and a HB donating group (i.e., the pyrrole N-H). As such, it can interact with the surroundings via different forces. For example, it can form π -hydrogen bonds^{34–36} with backbone and/or sidechain H-bond donating groups and also hydrophobic clusters with other aromatic amino acids. In addition, it tends to immerse itself in an environment where different interactions can occur; for instance, the benzene ring buries in a dehydrated and hydrophobic environment whereas the pyrrole N-H undergoes HB formation. One such example is the aforementioned M2 proton channel, wherein the N-H ends of the key Trp41 residues, according to a recent high-resolution crystal structure,^{37,38} point toward the aqueous pore, leaving the benzene rings facing the hydrocarbons of the lipid. Another example is the gramicidin A proton channel,^{39,40} where H-bonding interactions involving the N-H groups of several Trp residues are believed to be crucial for stabilizing the functional channel conformation. A third example is transmembrane peptides and proteins, where Trp residues are often located at the water-membrane interface with their N-H groups H-bonded to either water or a lipid headgroup.^{41,42} The results obtained in the current study, we believe, provide a foundation to use Trp_{CN} to study such H-bonding interactions using various linear and nonlinear IR spectroscopic methods. Finally, our findings suggest that the C≡N stretching

vibration of Trp_{CN} could be also used to study cation- π interactions involving Trp, which are prevalent in many biological systems.

CONCLUSIONS

Trp residues are frequently found at locations that are crucial for structure, interaction and functions of proteins. However, assessment of the microscopic environment of a specific Trp sidechain in proteins using linear and/or nonlinear IR spectroscopy has been hampered by the fact that none of its intrinsic IR active vibrational modes are ideally suited for this purpose. While a Trp analog, Trp_{CN}, has been suggested to be useful in this regard because the bandwidth of its C $\equiv$ N stretching vibration is sensitive to hydration, an experimental delineation of the factors that affect this vibration is lacking. Herein, we studied the C $\equiv$ N stretching vibration of 3M5CI, which is the sidechain of Trp_{CN}, in a series of solvents, aiming to provide a better understanding and characterization of these factors. Our results revealed that the C $\equiv$ N stretching frequency of 3M5CI depends not only on solvent interactions with the nitrile group, but also on interactions with the indole ring, making Trp_{CN} a versatile IR probe of proteins. Specifically, we found that a single solvent parameter, $\sigma = \pi^* + \beta - \alpha$, where π^* , β , and α are the Kamlet-Taft parameters characterizing the polarizability, HB accepting ability, and HB donating ability of the solvent, respectively, is sufficient to describe the C $\equiv$ N stretching frequencies obtained in all solvents via a simple linear function. This relationship thus confirms the possibility of using the C $\equiv$ N stretching vibration of Trp_{CN} to sense the dynamics of H-bonding interactions with the pyrrole N-H group, especially under conditions, as often seen in membrane proteins, where the N-H group is involved in HB formation but the benzene ring is buried in a dehydrated or hydrophobic environment. Linear and 2D IR measurements on 3M5CI in a binary solvent consisting of DMSO and TFE provide further evidence that its C $\equiv$ N stretching vibration is a sensitive probe of the microscopic environment of the molecule. The results, which are consistent with the literature, show preferential accumulation of DMSO molecules around the indole ring when its mole fraction reaches about 0.5. In addition, the 2D IR data indicated the presence of a second excited state species, leading to a long-lived ground state bleach signal. Kinetic measurements using IR pump-probe spectroscopy allowed us to further determine the formation time constant (~ 5.0 ps) of this additional excited state as well as the decay time constant (~ 12.3 ps) of the long-lived ground state bleach signal.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The ultrafast IR experiments were performed on instruments that were developed under an NIH Research Resource Grant (P41-GM104605). B.N.M is supported by an NIH Ruth Kirschstein National Research Service Award Predoctoral Fellowship (F31AG046010).

References

1. Fafarman AT, Sigala PA, Herschlag D, Boxer SG. J Am Chem Soc. 2010; 132:12811–12813. [PubMed: 20806897]
2. McMahon HA, Alfieri KN, Clark CAA, Londergan CH. J Phys Chem Lett. 2010; 1:850–855. [PubMed: 20228945]
3. Hu WH, Webb LJ. J Phys Chem Lett. 2011; 2:1925–1930.
4. Waagele MM, Culik RM, Gai F. J Phys Chem Lett. 2011; 2:2598–2609. [PubMed: 22003429]
5. Ma J, Pazos IM, Zhang W, Culik RM, Gai F. Annu Rev Phys Chem. 2015; 66:357–377. [PubMed: 25580624]
6. Waagele MM, Tucker MJ, Gai F. Chem Phys Lett. 2009; 478:249–253. [PubMed: 20161057]
7. Waagele MM, Gai F. J Phys Chem Lett. 2010; 1:781–786. [PubMed: 20436926]
8. Bogan AA, Thorn KS. J Mol Biol. 1998; 280:1–9. [PubMed: 9653027]
9. Spier AD, Lummis SCR. J Biol Chem. 2000; 275:5620–5625. [PubMed: 10681544]
10. Shimazaki Y, Yajima T, Takani M, Yamauchi O. Coord Chem Rev. 2009; 253:479–492.
11. Saikumar P, Murali R, Reddy EP. Proc Natl Acad Sci U S A. 1990; 87:8452–8456. [PubMed: 2236054]
12. Hsu CH, Chen C, Jou ML, Lee AYL, Lin YC, Yu YP, Huang WT, Wu SH. Nucleic Acids Res. 2005; 33:4053–4064. [PubMed: 16034027]
13. Samanta U, Pal D, Chakrabarti P. Proteins: Struct, Funct, Bioinf. 2000; 38:288–300.
14. Bartlett GJ, Porter CT, Borkakoti N, Thornton JM. J Mol Biol. 2002; 324:105–121. [PubMed: 12421562]
15. Hong M, DeGrado WF. Protein Sci. 2012; 21:1620–1633. [PubMed: 23001990]
16. de Jesus AJ, Allen TW. Biochim Biophys Acta, Biomembr. 2013; 1828:864–876.
17. Fabian H, Schultz C, Backmann J, Hahn U, Saenger W, Mantsch HH, Naumann D. Biochemistry. 1994; 33:10725–10730. [PubMed: 8075073]
18. Barth A. Prog Biophys Mol Biol. 2000; 74:141–173. [PubMed: 11226511]
19. Asplund MC, Zanni MT, Hochstrasser RM. Proc Natl Acad Sci U S A. 2000; 97:8219–8224. [PubMed: 10890905]
20. Allerhand A, Schleyer PV. J Am Chem Soc. 1963; 85:866.
21. Kamlet MJ, Abboud JLM, Abraham MH, Taft RW. J Org Chem. 1983; 48:2877–2887.
22. Marcus Y. Chem Soc Rev. 1993; 22:409–416.
23. Oh KI, Choi JH, Lee JH, Han JB, Lee H, Cho M. J Chem Phys. 2008; 128:10.
24. Getahun Z, Huang CY, Wang T, De Leon B, DeGrado WF, Gai F. J Am Chem Soc. 2003; 125:405–411. [PubMed: 12517152]
25. Johnston AJ, Zhang Y, Busch S, Carlos Pardo L, Imberti S, McLain SE. J Phys Chem B. 2015; 119:5979–5987. [PubMed: 25893741]
26. Roy S, Bagchi B. J Chem Phys. 2013; 139:034308. [PubMed: 23883028]
27. Aschaffenburg DJ, Moog RS. J Phys Chem B. 2009; 113:12736–12743. [PubMed: 19711975]
28. Hamm, P., Zanni, M. Concepts and Methods of 2D Infrared Spectroscopy. Cambridge University Press; 2011.
29. Ma JQ, Pazos IM, Gai F. Proc Natl Acad Sci U S A. 2014; 111:8476–8481. [PubMed: 24912147]
30. Hilaire MR, Abaskharon RM, Gai F. J Phys Chem Lett. 2015; 6:2546–2553. [PubMed: 26266732]
31. Chung JK, Thielges MC, Lynch SR, Fayer MD. J Phys Chem B. 2012; 116:11024–11031. [PubMed: 22909017]
32. Lee KK, Park KH, Choi JH, Ha JH, Jeon SJ, Cho M. J Phys Chem A. 114:2757–2767. [PubMed: 20141151]
33. Banno M, Kotani A, Ohta K, Tominaga K. Bull Chem Soc Jpn. 2014; 87:470–478.
34. Steiner T, Koellner G. J Mol Biol. 2001; 305:535–557. [PubMed: 11152611]
35. Steiner T. Biophys Chem. 2002; 95:195–201. [PubMed: 12062379]
36. Perutz MF. Philos Trans R Soc, A. 1993; 345:105–112.

37. Acharya R, Carnevale V, Fiorin G, Levine BG, Polishchuk AL, Balannik V, Samish I, Lamb RA, Pinto LH, DeGrado WF, Klein ML. *Proc Natl Acad Sci U S A*. 2010; 107:15075–15080. [PubMed: 20689043]
38. Williams JK, Zhang Y, Schmidt-Rohr K, Hong M. *Biophys J*. 2013; 104:1698–1708. [PubMed: 23601317]
39. Maruyama T, Takeuchi H. *Biochemistry*. 1997; 36:10993–11001. [PubMed: 9283091]
40. Chaudhuri A, Haldar S, Sun H, Koeppe RE II, Chattopadhyay A. *Biochim Biophys Acta, Biomembr*. 2014; 1838:419–428.
41. Vostrikov VV, Daily AE, Greathouse DV, Koeppe RE. *J Biol Chem*. 2010; 285:31723–31730. [PubMed: 20667827]
42. Pogozheva ID, Mosberg HI, Lomize AL. *Protein Sci*. 2014; 23:1165–1196. [PubMed: 24947665]
43. Haynes, WM. *CRC Handbook of Chemistry and Physics*. 96. CRC Press; Boca Raton, FL: 2015.

Highlight

This study shows that the $\text{C}\equiv\text{N}$ frequency of 5-cyanotryptophan depends, and hence reports, on multiple interactions with the solvent.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1. Structures of (A) 3-methyl-5-cyanoindole (3M5CI) and (B) 1-methyl-1H-indole-5-carbonitrile (NM5CI).

Figure 2.
C≡N stretching bands of 3M5CI in different solvents, as indicated.

Figure 3.
C≡N stretching bands of 3M5CI in TFE and DMSO mixtures with different volume ratios, as indicated.

Figure 4. Center frequency (ω_0) of the C $\equiv$ N stretching band of 3M5CI versus π^* (A), β (B), and ϵ (C). Only frequencies obtained in solvents with $\alpha = 0$ were used.

Figure 5.

Center frequency (ω_0) of the C $\equiv$ N stretching band of 3M5CI versus the solvent γ parameter, where $\gamma = \pi^* + \beta$. The solid line represents the linear regression of this data, yielding a slope of -5.6 ± 0.2 cm⁻¹ and an intercept of 2226.1 ± 0.2 cm⁻¹.

Figure 6.
Comparison of the C≡N stretching bands of NM5CI and 3M5CI, as indicated, in DMSO.

Figure 7.

Center frequency (ω_0) of the C $\equiv$ N stretching band of 3M5CI versus the solvent σ parameter, where $\sigma = \pi^* + \beta - \alpha$. The solid line represents the best fit of this data to a line with a slope of $-5.6 \pm 0.2 \text{ cm}^{-1}$ and an intercept of $2226.2 \pm 0.2 \text{ cm}^{-1}$.

Figure 8.
2D IR spectra of 3M5CI in a 50:50 TFE:DMSO mixture, at various waiting times (T), as indicated.

Figure 9.
Time-resolved absorption spectra of 3M5CI in DMSO.

Figure 10.

Transient absorption kinetics obtained at 2190 cm^{-1} (red circle) and 2223 cm^{-1} (blue cross). The transient data at 2190 cm^{-1} can be fit to a single exponential function (red line) with a time constant of $1.3 \pm 0.1 \text{ ps}$. On the other hand, the kinetics at 2223 cm^{-1} can be best described by a double-exponential function (blue line) with the following time constants (amplitude): $1.2 \pm 0.1 \text{ ps}$ ($3.8 \pm 0.1 \text{ mOD}$) and $12.3 \pm 1.6 \text{ ps}$ ($0.7 \pm 0.1 \text{ mOD}$).

Table 1

The center frequency (ω_0) and full-width at half maximum (FWHM) of the C $\equiv$ N stretching band of 3M5CI in different solvents. Also listed for each solvent are its Kamlet-Taft parameters, π^* (polarizability), β (hydrogen bond donor), α (hydrogen bond acceptor), as well as its dielectric constant (ϵ).⁴³ Unless otherwise indicated, all Kamlet-Taft parameters are from ref. 21.

Solvent	ω_0 , cm ⁻¹	FWHM, cm ⁻¹	π^*	β	α	ϵ
Water	2224.0 ^d	18.0 ^d	1.09	0.47 ^b	1.51	80.1
MeOH	2223.6	14.2	0.60	0.62	0.93	33.0
2-propanol	2223.6	14.7	0.48	0.95	0.76	20.2
DCM	2221.7	10.1	0.82 ^b	0.10 ^b	0.13 ^b	8.9
DMSO	2216.1	10.6	1.00	0.76	0.00	47.2
DMF	2217.2	9.7	0.88	0.69	0.00	38.2
Pyridine	2217.9	10.3	0.87	0.64	0.00	13.3
Acetophenone	2218.5	10.6	0.90	0.49	0.00	17.4
Cyclopentanone	2218.7	8.9	0.76	0.52	0.00	13.6
THF	2219.8	8.0	0.58	0.55	0.00	7.5
1,4-dioxane	2221.1	8.7	0.55	0.37	0.00	2.2
Toluene	2222.4	8.7	0.54	0.11	0.00	2.4
Tetrachloromethane	2224.5	11.5	0.28	0.00	0.00	2.2
Hexafluorobenzene	2226.9	10.7	0.33 ^b	0.02 ^b	0.00 ^b	2.0
TFE	2230.7	24.1	0.73	0.00	1.51	27.7

^aIR data in water were from ref. 6 and measured for 5-cyanoindole in a 95/5 (v/v) water/MeOH mixture.

^bRef. 22