Modeling Vibrational Spectra of Amino Acid Side Chains in Proteins: Effects of Protonation State, Counterion, and Solvent on Arginine C−N Stretch Frequencies†

Mark S. Braiman,*,‡ Deborah M. Briercheck,§ and Kerry M. Kriger

University of Virginia Health Sciences Center, Biochemistry Department # 440, Charlottesville, Virginia 22908

Received: July 14, 1998; In Final Form: December 23, 1998

The vibrational spectrum of arginine's side chain in various protein environments is modeled by measuring IR spectra of ethylguanidinium (EG) salts under varying conditions. Characteristic ν_{C-N} stretch vibrational frequencies of monoalkylguanidinium are assigned to observed IR bands at 1655−1685 cm^{-1}, 1615−1635 cm^{-1}, and 1170−1180 cm^{-1}. Each of these bands is observed to downshift by 4−9 cm^{-1} upon $[^{15}N]_2$ substitution at the terminal nitrogens. Additional weaker bands from vibrations involving nitrogen motion are also discernible at ∼920, ∼1085, and ∼1440 cm^{-1}. Deprotonation of EG−Cl is accompanied by an overall decrease in IR absorption intensity and substantial changes in the ν_{C-N} vibrational bands. The ∼1670 and ∼1180 cm^{-1} bands appear to shift substantially in the deprotonated state. A strong band near 1635 cm^{-1} remains in ethylguanidine, but this is interpreted as being due to a $\delta(NH_2)$ scissor mode based on previously published assignments of the guanidine IR spectrum. New bands attributable to C−N stretch vibrations of deprotonated EG are observed at 1600, 1566, and 1208 cm^{-1}. The ν_{C-N} bands of EG cation also show characteristic shifts of up to ∼40 cm^{-1} depending on solvent and counterion conditions. The biggest effects are seen for counterion variations in the solid state, where the highest ν_{C-N} frequency ranges from 1695 cm^{-1} for EG−carbonate down to 1652 cm^{-1} for EG−bromide. In nonpolar solvent, ion pairing occurs, as evidenced by reproducible differences seen for vibrational frequencies of different EG salts, e.g., 1185 cm^{-1} for the acetate vs 1179 cm^{-1} for the iodide. In polar solvents (methanol, ethanol, dimethyl sulfoxide, or water), however, there is little if any difference in the vibrational frequencies of EG acetate, chloride, bromide, iodide, or phenolate, indicating independently solvated anion and cation. We conclude that when arginine's side chain is buried as part of an ion pair within a hydrophobic region of a protein, its strongly IR-absorbing ν_{C-N} frequencies are likely to be sensitive to perturbations such as changing the nature or position of the counterion or altering the hydrogen-bonding capacity of nearby neutral amino acids.

Introduction

Vibrational spectroscopy can be a useful tool for analyzing enzyme mechanisms because of its intrinsically fast time resolution and its potential sensitivity to individual chemical groups within a protein. IR measurements have modeled various amino acid side chain functionalities and the effects of different protein environments on them, in greatest detail for the carboxylic acid resides (Asp and Glu)[1] and to a lesser extent for other ionizable residues such as His, Cys, Tyr, and Lys (see, for example, ref 2).

Arginine side chains play important, but as yet poorly defined, roles in a number of very important membrane proteins for which no high-resolution crystal structures are available. For example, motions of arg-82 in bacteriorhodopsin (and its homologue arg-108 in halorhodopsin) are thought to play a key role in the mechanism of H^+ transport[3] (or, in the case of halorhodopsin, Cl^- transport[7]). Yet the four high-resolution models of bacteriorhodopsin that have been published using data from high-resolution cryoelectron microscopy and X-ray crystallography indicate different positions for the side chain of arg-82, with sigificantly different nearest-neighbor contacts.[3−6] It may be some time before crystallographic methods define an unambiguous resting position for arg-82, much less the changes it undergoes during the photoreactions of bacteriorhodopsin or halorhodopsin.

As in halorhodopsin, arginines are likely to be involved in membrane-buried chloride binding sites in mammalian chloride transport systems such as the CFTR protein, defects in which are responsible for cystic fibrosis.[8] In G-protein-coupled receptors, activation of G-protein has been found to be dependent on the presence of a specific arginine in the receptor (the central residue in the highly conserved ERY sequence found at the cytoplasmic end of helix III).[9] These are just a few examples of arginines which play active roles in protein functions important in human physiology, yet whose environments and interactions are at the moment almost completely unknown.

IR spectroscopy has begun to provide some methods for testing the roles of specific arginines in such membrane proteins. However, detailed interpretation of such spectra is limited in part by an absence of adequate modeling of the vibrational spectrum of the arginine side chain. To analyze IR spectra for evidence of such possible structural changes as deprotonation, counterion substitution, and desolvation, we have examined alkylguanidino model compounds in which such effects can be studied systematically.

* To whom correspondence should be addressed at Syracuse University Chemistry Department, Room 1-014, Center for Science and Technology, Syracuse, NY 13210. Phone: (315) 443-4691. Fax: (315) 443-4070. E-mail: mbraiman@syr.edu.

† This is the third paper in a series. See ref 1 and 13 for the preceding papers.

‡ Current address: Chemistry Department, Syracuse University.

§ Current address: Department of Biological Sciences, Carnegie Mellon University, 4400 Fifth Ave., Pittsburgh, PA 15213.

10.1021/jp983011b CCC: $18.00
Published on Web 05/06/1999

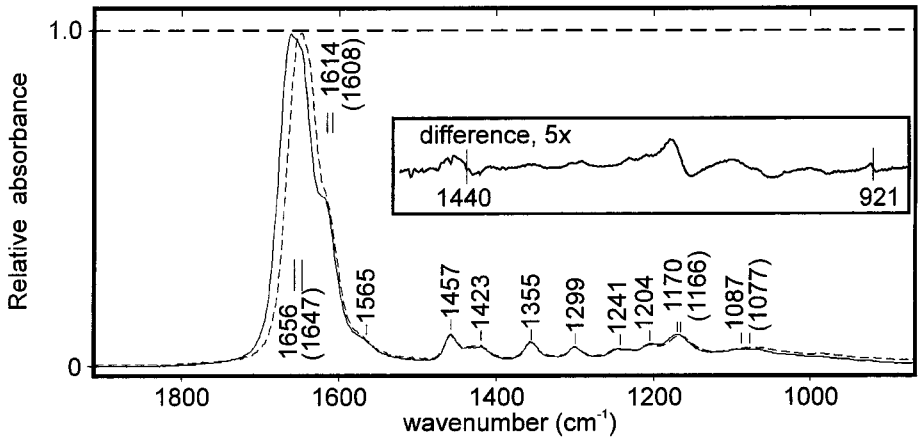

Figure 1. IR spectra of thioethylguaniginium bromide, TEG−Br, measured as a film dried down on an 3M IR card using a Bruker IFS66 spectrometer. Spectral resolution was 1 cm^{-1}. The dotted line represents an isotopically substituted version of the same molecule with ^{15}N incorporated at >95% at the two terminal nitrogens. Both spectra were normalized relative to the maximum absorbance in this spectral range, which occurs near 1650 cm^{-1}. Labeled frequencies and tick marks correspond to centers of band shapes in the unsubstituted spectrum, as obtained from a band fitting program. All labeled bands had identical frequencies in the $[^{15}N]_2$-substituted spectra, with the exception of the four for which downshifted frequencies are given in parentheses. (Inset) difference spectrum between the natural-abundance and $[^{15}N]_2$-labeled samples, displayed at a ∼5×-magnified vertical scale. The labeled marks indicate two isotope-shifted difference bands that were not apparent in the individual spectra.

While environmental changes are expected to affect the N−H stretching vibrations of the guanidino group most strongly, measurement of such vibrations in proteins (in the region near 3000−3500 cm^{-1}) can be difficult due to their breadth and the presence of numerous strongly overlapping OH and NH bands from water and other amino acid residues. More important, however, is the fact that such N−H stretch bands cannot easily be assigned to individual residues, e.g., by means of residue-specific isotope labeling. For example, it is impossible to deuterium-label the N−H bonds of arginine specifically within a protein; such labeling of arginine would be accompanied by equal labeling of all exchangeable protons, including many other N−H bonds with overlapping vibrational frequencies. Residue-specific ^{15}N labeling is achievable, but such isotope substitution is expected to generate only small shifts in the relatively broad N−H stretch bands, too small to detect reliably.

Thus, we focus our attention on understanding the C−N stretch vibrations of the alkylguanidino group. These vibrations have strong IR absorptivity as well as preresonant Raman intensity due to coupling with the lowest lying $\pi^* \leftarrow \pi$ transition.[10] The C−N bonds all have partial double-bond character due to the presence of multiple resonance structures for the guanidino group. The relative stability of the resonance structures is expected to be influenced by counterion and solvent interactions. Additionally, coupling of the C−N stretch vibrations to in-plane NH bending motions means that changes in H-bond interactions will lead to ν_{C-N} frequency shifts. These bands are expected to be assignable in many cases to individual amino acids via residue-specific isotope labeling. The strong IR bands arising from C−N stretch vibrations are therefore expected to serve as sensitive spectroscopic indicators of the state of the arginine group at various protein sites.

Here we use ^{15}N isotope labeling to determine empirically the involvement of the two terminal nitrogens in C−N vibrational modes of the of the alkylguanidino side chain of arginine and then examine effects of deprotonation, counterion, and solvent on those ν_{C-N} frequencies.

Materials and Methods

Except as noted in figure legends, all FTIR absorbance spectra were measured using a Nicolet 740 FTIR spectrometer. Subsequent spectral manipulations were carried out using GRAMS software (Galactic Industries, NH).

All solvents were reagent grade or better. Solvents used for obtaining spectra in nonpolar media ($CHCl_3$/MeOH 97:3 and isotope derivates thereof) were kept dry by storage over 4-A molecular sieves. Ethylguanidinium chloride (EG−Cl) from Aldrich was used without further purification, except for comparisons with other EG salts. For such comparisons, all salts were prepared by first converting EG−Cl to the carbonate using Ag_2CO_3, then treating an aqueous solution of EG carbonate with the conjugate acid of the desired anion (chloride, bromide, iodide, acetate, nitrate, or phenolate), as described previously.[11]

Thioethylguanidinium bromide (TEG−Br) was synthesized from thiourea and 2-bromoaminopropane (Aldrich) using a published procedure[12] and was recrystallized from ethanol/diethyl ether. TEG−Br with ^{15}N substitution at the two terminal $-NH_2$ groups was prepared by using ^{15}N-labeled thiourea (Cambridge Isotope Laboratories).

Results

The IR spectra of alkylguanidinium salts (e.g., Figure 1) are in general very similar to that of the side chain of arginine hydrochloride measured in water.[2] The strongest features in the latter spectrum, a band with a maximum at 1673 cm^{-1} and a distinct shoulder at 1635 cm^{-1}, have been ascribed to C−N vibrations because they are in a frequency range characteristic of C═N bonds. As discussed further below and in a related paper describing ab initio electronic and force field calculations on ethylguanidinium,[13] these bands can now be quantitatively assigned to two specific linear combinations of stretching motions of the three C−N bonds of the guanidino group. The "third" linear combination of CN_3 stretches is more strongly admixed with other in-plane motions (such as in-plane hydrogen bending and C_{alkyl}−N stretching), as shown below.

Empirical Assignment of C−N Vibrations based on Isotopic Shifts. Because alkylguanidinium compounds have an extended π orbital system involving three nearly equivalent nitrogens bonded to a central carbon, the three C−N stretch internal coordinates are strongly coupled to one another. In the case of guanidinium itself, which has D_{3h} or C_{3v} symmetry depending on its environment, the C−N stretches are easily

classified as a symmetric stretch mode near 1005 cm^{-1} and a 2-fold degenerate asymmetric mode near 1665 cm^{-1}.[10]

Monosubstituted guanidinium species such as arginine generally exhibit a related pattern of C−N stretch vibrations. In methylguanidinium, for example, the "asymmetric" CN_3 stretch character was assigned to a broad Raman band at 1670 cm^{-1}.[10] However, the substituent breaks the 3-fold symmetry of guanidinium, precluding simple symmetry group descriptions of the C−N stretch vibrations. Thus, it was concluded that, for monosubstituted guanidinium, the "symmetric" CN_3 stretching vibration could not be unambiguously identified and that it might not exist as a single vibrational band. Vibrations at 920 cm^{-1} (in methylguanidine) and near 1180 cm^{-1} (in methylguanidinium and in arginine itself) were attributed as having significant C−N stretch character, based on their Raman intensity enhancement as the excitation wavelength was brought into resonance with the lowest lying electronic excited state.[10]

More recently, high-resolution IR data on various EG salts in organic solvent were used to support the assignment of vibrational bands near 1670, 1635, and 1180 cm^{-1} as likely having C−N stretch character.[11] This work showed that these three bands of alkylguanidinium compounds show significant shifts depending on counterion pairing (see also below for a further discussion).

However, until now, no detailed normal-mode calculations or heavy-atom (^{13}C or ^{15}N) isotope substitutions of alkylguanidinium molecules appear to have been performed. Such data are necessary to define accurately the C−N stretch character of observed vibrations. Heavy-atom isotope substitution is generally superior to deuteration for defining C−N stretch frequencies because overall it induces a much smaller perturbation on the vibrations of the molecule, making it unlikely to lead to significant remixing of the normal modes. Here, we provide such measurements using ^{15}N labeling at two of the three nitrogens of EG.

Thiolated ethylguanidinium (TEG) was used for this work principally because it greatly simplified the isotope labeling, making it possible to use a simple two-step published synthesis and commercially available starting materials. Selective ^{15}N labeling of the two terminal $-NH_2$ groups was achieved by using commercially available isotope-labeled thiourea, $S{=}C(^{15}NH_2)_2$, as a starting material. (The other starting material, 2-bromoaminoethane, is not commercially available with ^{15}N label). The thiol group is not expected to perturb significantly the IR spectrum of the ethylguanidinium moiety. The SH stretch vibration, observed at 2510 cm^{-1} with an integrated intensity similar to that of the 1355 cm^{-1} band (spectral region not shown), is expected to be the strongest additional vibration contributed by the thiol group.

However, the presence of the thiol group limits somewhat the conditions under which the isotope shifts could be monitored, since TEG was found to be much less soluble than EG itself in most organic solvents. Furthermore, the thiol group's relatively low pK_a of 9−10 interfered with making analogous isotope-shift measurements on deprotonated alkylguanidine (see below). Thus, we limited our measurements to the protonated form and measured the IR spectrum only of solid samples.

In Figure 1 we present a comparison of the IR spectra of unsubstituted and ^{15}N-substituted thioethylguanidinium bromide, $HS{-}(CH_2)_2{-}NH{-}C{-}(NH_2)_2^+{\cdot}Br^-$. These spectra provide for the first time a direct empirical measure of the relative contributions of nitrogen motion to the IR bands at 1656, 1614, and 1170 cm^{-1}. Each of these bands downshifts significantly (by 4−9 cm^{-1}) upon isotope labeling.

The spectra in Figure 1 provide clear support, however, for the idea that the C−N stretches are significantly mixed with each other as well as with other vibrations. A mass substitution of ^{15}N for ^{14}N would be expected to downshift the frequency of a single unmixed C−N stretch vibration by a factor of $(15-(14+12)/14(15+12))^{1/2}$, or 1.0157. The observed shift factors of the 1656, 1614, and 1170 cm^{-1} bands are in the range 1.003−1.005, i.e., substantially smaller. Furthermore, there are clearly more than three bands in the 800−1800 cm^{-1} frequency range that are sensitive to ^{15}N substitution. The biggest isotope-induced downshift, in fact, appears to occur in a weak, broad band centered at 1087 cm^{-1}. Additional downshifts, in bands near 1440 and 921 cm^{-1} that are too weak to be discerned in the raw absorbance spectra, become readily apparent in a vertically expanded $^{14}N-^{15}N$ difference spectrum (Figure 1, inset).

The observed 921 cm^{-1} band (Figure 1, inset) likely corresponds quite closely to the Raman-active band previously measured at 920 cm^{-1} for methylguanidinium.[10] Its IR intensity is low, as expected of a symmetric C−N stretch. However, the relatively small ^{15}N isotope shift of this band (∼6 cm^{-1} according to band fitting of the difference spectrum, or a shift factor of 1.006) suggests that this mode cannot be described as a pure symmetric CN_3 stretch.

Instead, much of the remaining CN_3 stretch character appears to be associated with the bands near 1170 and 1087 cm^{-1}. Based on the greater IR and resonance Raman intensity of the former, it is more likely to play a significant role in analysis of arginine environments within proteins. In the remainder of this paper, we will generally include the 1170 cm^{-1} vibration as one of three characteristic IR CN_3 group stretch frequencies in alkylguanidinium compounds. However, vibrational normal-mode calculations show that this mode includes substantial contributions from other internal coordinates, especially the $N-C_{alkyl}$ stretch at the substituted nitrogen (ref 13; see also Discussion).

IR Spectral Signals due to Arginine Side Chain Deprotonation. Very little has been published previously regarding the vibrational spectrum of the deprotonated arginine side chain. This is possibly because arginine, with a pK_a typically above 11, is unlikely to be found deprotonated under most biological conditions. An additional reason may be that, like other guanidinium compounds, aqueous arginine has a tendency to decompose at very high pH values. Previous workers have noted base-catalyzed hydrolysis to both citrulline[14] and ornithine.[15]

Recent IR measurements on the effects of pH on the most commonly ionized amino acids[2] did include a comparison of IR spectra of arginine at pH 5.6 and 11.2 in an attempt to determine the spectroscopic consequences of arginine deprotonation. However, this comparison was complicated by the presence of other titratable IR-active groups (the amino and carboxylate moieties). Correction for this (by subtracting the spectrum of alanine at corresponding pH values) was only approximate. Furthermore, the uppermost pH value of 11.2 was probably inadequate to bring about more than ∼50% deprotonation.

To avoid hydrolysis under conditions expected to induce more complete deprotonation of EG, we measured its spectrum in methanolic sodium methoxide. Such conditions are expected to cause even more deprotonation of EG than 1 M aqueous KOH (i.e., pH 14) because methanol, being less polar than water, is less capable of solvating and stabilizing the ionized state. The resulting spectrum of mostly deprotonated EG (Figure 2B) may be taken as a qualitative representation of the IR absorbance signal expected from a deprotonated arginine side chain buried inside a protein.

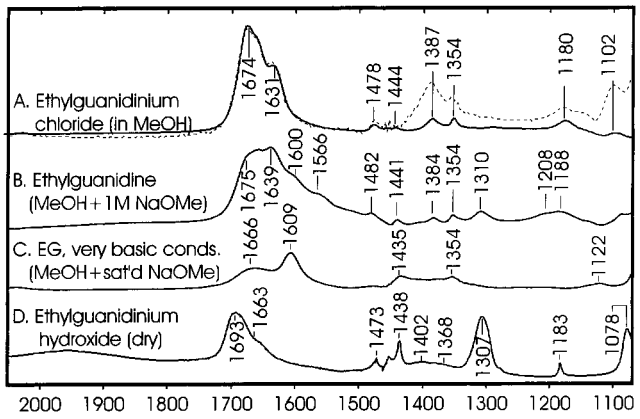

Figure 2. IR spectra of ethylguanidinium chloride, EG−Cl, and its alkaline products, measured at 1 cm^{-1} resolution. The strong methanol band near 1450 cm^{-1} precluded accurate solvent subtraction in some spectra. The indicated frequencies and tick marks correspond to centers of fitted band shapes. EG−Cl was (A) dissolved in MeOH, (B) dissolved in MeOH containing 1 M sodium methoxide, (C) dissolved in MeOH containing saturated sodium methoxide, and (D) measured as a film dried down from 1 M KOH/MeOH. Spectrum D, measured using a 3M IR card, was similar to that measured using a film dried down on a CaF_2 window (not shown). In (A), the solid line represents data from a crystalline sample freshly dissolved in reagent-grade methanol. The dotted line represents an identical sample first incubated for 5 h in 1 M NaOMe/MeOH (during which time an aliquot was taken to measure a spectrum similar to that of (B)), then acidified by gently bubbling HCl gas through the solution for 1 min. The liquid spectra were all measured using a 50 μm path length and similar concentrations of EG−Cl (~20 mM). However, the ethylguanidine spectra (B and C) were rescaled by a factor of ~5 relative to the ethylguanidium spectra in (A), to compensate for the weaker IR absorbance of the deprotonated species.

In the deprotonated state, there appears to be a significant decrease in the integrated intensity of the C−N stretch vibrations. (Note the scale change between parts A and B of Figure 2). The strongest IR absorption band shifts down in frequency, from ~1675 cm^{-1} in the protonated state to ~1639 cm^{-1} in the unprotonated state. In addition, new bands are seen at 1600 and 1566 cm^{-1}. The latter two bands are also observable as weak shoulders appearing only at high pH in the previously published comparison of arginine at pH 5.6 and 11.2.[2] However, it is now apparent that the latter spectrum of arginine at pH 11.2 corresponded to substantially less than 100% deprotonation of the arginine side chain. The previously published spectrum was approximately matched by a computerized addition of our spectra of protonated EG (Figure 2A) and deprotonated EG (Figure 2B) in an approximate ratio of 1:3, suggesting that 75% deprotonation of the arginine side chain had been attained at pH 11.2. This could be an overestimate, however, since the persistence of a band near ~1675 cm^{-1} in our spectrum of deprotonated EG (Figure 2B) probably signifies that our sample also contained a small amount of protonated EG.

We checked for reversibility of the deprotonation of EG in 1 M sodium methoxide by measuring its spectrum after reacidification (Figure 2A, broken line). While the spectrum of the reacidified species shows some measurable changes, most spectral features correspond to those of the fresh EG−Cl, indicating that the deprotonation was largely reversible.

We attempted to bring about more complete deprotonation of EG by using higher concentrations of sodium methoxide in methanol. The results with a saturated solution, shown in Figure 2C, suggest that there are more than two species involved in this titration. While a ~1600 cm^{-1} band attributable to ethylguanidine continues to increase in intensity at the higher NaOMe concentration, bands at 1566 and 1310 cm^{-1} do not. Our attempts at reversing this species (Figure 2C) back to ethylguanidine itself were unsuccessful. These results suggest that it may be difficult to disentangle the pure ethylguanidine spectrum from data obtained from varying concentrations of NaOMe in MeOH.

When we attempted to measure the spectrum of deprotonated EG as a thin film, by dissolving it first in aqueous or methanolic 1 M KOH and then removing solvent as rapidly as feasible (<1 min), we obtained spectra like that in Figure 2D. This spectrum would appear to indicate that the strongest ν_{C-N} band characteristic of the alkylguanidino group upshifts from ~1675 to ~1690 cm^{-1} upon deprotonation in a dried film. However, previously published spectra of arginine at pH 11.6[2] show no indication of upshifted ν_{C-N} vibration(s). Furthermore, upon redissolution and reacidification we were not able to recover material that gave a spectrum matching the original EG−Cl (Figure 2A, dotted line). It is unlikely that rapid air-drying from methanolic KOH immediately hydrolyzed EG to products analogous to citrulline (ethylurea) or ornithine (ethylamine), since the ~1690 cm^{-1} band in Figure 2D is >40 cm^{-1} higher in frequency than any comparably intense band in standard reference spectra of methylurea or ethylamine.

The aforementioned facts suggest that films made by air-drying alkaline solutions of EG probably contain principally EG^+OH^- (as well as KOH and KCl), rather than deprotonated EG or a hydrolysis product of EG. We conclude that true deprotonated alkylguanidine species exhibit only downshifted, not upshifted, ν_{C-N} frequencies, as exemplified in parts B and C of Figure 2, and that the IR spectrum shown in Figure 2D corresponds principally to (solid) ethylguanidinium hydroxide.

Our results and interpretation are supported by what has been observed with deprotonated guanidine itself.[16] In this case, simple air-drying from an ethanolic hydroxide solution was found to yield principally solid guanidinium hydroxide. Guanidine itself was generated only by intensively vacuum-drying a pure stoichiometric preparation of guanidinium hydroxide in the presence of P_2O_5, a strong desiccant. The published IR spectrum of the resulting ultra-dry guanidine film[16] was quite analogous to our spectrum of MeOH-dissolved ethylguanidine (Figure 2B). The degenerate asymmetric ν_{C-N} mode of guanidinium near 1650 cm^{-1} drops in frequency upon deprotonation, giving rise to two strong guanidine IR bands at 1603 and 1469 cm^{-1} that were assigned as ν_{C-N} modes. A weaker band remaining near 1650 cm^{-1} was assigned as a $\delta(NH_2)$ in-plane scissor vibration. By analogy with these assignments, we assign the new 1600 and 1566 cm^{-1} bands in the ethylguanidine spectrum (Figure 2B) to downshifted ν_{C-N} modes and a significant portion of the residual intensity at 1639 cm^{-1} to the $\delta(NH_2)$ scissor mode. The remaining intensity at 1639 cm^{-1}, as well as that near 1675 cm^{-1}, are likely due to residual protonated EG in the saturated sodium methoxide solution.

It might be supposed that deprotonation of the guanidino group should give rise to an upshift in the highest ν_{C-N} vibrational frequency, due to the increased bond order expected for the C−N bond where deprotonation occurs. However, this argument neglects the effect of coupling with the C−NH in-plane vibrations, which lie in the range 1100−1600 cm^{-1}. Such coupling tends to raise the frequency of the higher frequency asymmetric ν_{C-N} vibrations. Deprotonation eliminates this coupling of the two asymmetric ν_{C-N} vibrations to an N−H bend vibration, lowering the frequency of the former. This effect was analyzed at greater length for guanidinium.[16] A completely analogous effect is also well-known for Schiff bases, which

TABLE 1: Observed Frequencies for EG Salts in the Solid State (as Films)

counterion	ν_{C-N} frequencies, cm^{-1}[a]		
Br^{-1}	1652	1618	1176
Cl^{-2}	1652	1623	1173
phenolate[b]	1680	1640	1198
acetate[c]	1686	1614	1192
carbonate	1695	1643	1202
OH^{-d}	1693	1663	1183

[a] Frequencies are reproducible to 2 cm^{-1}. [b] Spectra not shown. [c] Spectrum presented in SI-1 of the Supporting Information. [d] Taken from the spectrum in Figure 2C.

likewise show deprotonated ν_{C-N} frequencies that are typically 10−30 cm^{-1} lower in frequency than those for the corresponding protonated species.[17]

Effects of Counterion in Ethylguanidinium Salts. It was noted previously that for guanidinium compounds in the solid state there are small changes in ν_{C-N} frequencies depending on counterion.[10,19] We observed even greater effects for ethylguanidinium in the solid state (Table 1). Even limiting ourselves to several of the counterions that are likely to be found as arginine counterions in proteins (bromide, chloride, and carbonate, plus phenolate and acetate as models for the side chains of tyrosine and asparate or glutamate), we find a substantial variation. The highest frequency ν_{C-N} vibration, for example, is nearly 45 cm^{-1} higher in frequency in the carbonate than in the halide salts, with the phenolate at an intermediate position.

From the results in Table 1, one can generalize that ions forming stronger H-bonds produce higher ν_{C-N} frequencies, at least for the most intense ν_{C-N} vibration (1650−1700 cm^{-1}). The highest values for this frequency in ethylguanidinium are seen with the oxyanions (phenolate, acetate, or carbonate). Furthermore, oxyanions with a capability of simultaneously forming multiple O···HN hydrogen bonds seem to produce the very highest ethylguanidinium frequencies. (However, this last generalization may be contradicted by the case of the hydroxide salt, if that is indeed what the spectrum in Figure 2C represents).

However, solid-state spectra are of limited value because each EG cation is surrounded by multiple anions in an arrangement that may vary considerably due to crystal form. The effect of a single counterion on EG in a nonpolar solvent is of considerable interest for several reasons. First, this system is expected to model closely the vibrations of an isolated ion pair, which is more amenable to calculation than systems involving more molecules. Second, an ion pair in a nonpolar solvent may be expected to resemble rather closely a hydrogen-bonded salt bridge buried within a hydrophobic region of a protein. Such ion-paired arginines within hydrophobic protein regions are known from protein structural work and are potentially expected to play a role in important processes, e.g., in binding anions as they are transported across a hydrophobic transmembrane channel.

Previously, it was demonstrated that a solvent system of $CHCl_3$/MeOH 97:3 is capable of dissolving EG ion pairs at a concentration sufficient to obtain good IR spectra.[11] Effects of different anions on the EG bands near 1670, 1630, and 1180 cm^{-1} were noted. However, the accuracy of spectral frequencies for the ∼1180 cm^{-1} vibration was limited by the presence of nearby solvent absorption bands.

To confirm that this vibration is indeed counterion-sensitive, we have obtained additional spectra from the same salts, using C^2HCl_3 instead of C^1HCl_3 in the solvent mixtures (Figure 3). The nonexchangeable deuterium on the chloroform does not interfere with measuring accurate spectra separately of ethylguanidinium containing ^1H or ^2H at its five exchangeable proton positions.

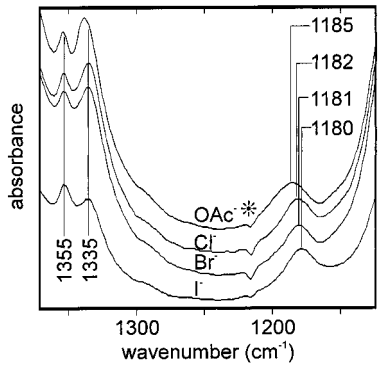

Figure 3. IR spectra of EG salts measured in nonpolar solvent (C^2HCl_3/MeOH) showing the counterion dependence of the ν_{C-N} vibration near 1170 cm^{-1}. Indicated band positions and tick marks correspond to the center of the best fit of the 1200−1160 cm^{-1} range to a single Lorentzian peak shape on top of a sloping baseline.

TABLE 2 [a]

counterion	observed ν_{C-N} frequencies, cm^{-1}					
	^1H-exchanged			^2H-exchanged		
I^-	1670.4	1626.5	1179.7	1606.4	1581.4	1135.1
Br^-	1672.5	1627.0	1181.4	1605.9	1581.2	1136.2
Cl^-	1673.3	1632.7	1182.2	1606.0	1581.1	1136.5
acetate	1683.1	1636.3	1185.5	1605.8	1585.9	1139.9

[a] All wavenumber values are taken from bandfits to spectra shown in the Supporting Information. A representative portion of some of these spectra is shown in Figure 3. Fits were performed generally as described previously,[11] i.e., to 4 Voigtian bands, labeled A−D, plus a sloping baseline in the 1725−1570 cm^{-1} region of SI-1; to 1 band, labeled E, plus a sloping baseline in the 1170−1180 cm^{-1} region of SI-1; to 2 bands, labeled A and B, in the 1725−1570 cm^{-1} region of SI-2, and to 1 band, labeled C, in the 1130−1140 cm^{-1} region of SI-2. Band centers in this table are estimated to be accurate to 0.5 cm^{-1} or better based on repeated fits with variations in the fitted region, number of peaks, etc.

The results of our measurements of the anion dependence of all three ν_{C-N} modes are presented in Table 2. (The spectra themselves are presented as Supporting Information.) These data, which were obtained on samples completely independent of those presented earlier,[11] confirm the anion dependence of the ν_{C-N} frequencies and provide more accurate values for them, especially for the vibrations in the 1100−1200 cm^{-1} range.

It should be noted that the frequency of the strongest absorbing band in the EG−OAc spectrum listed in Table 2 is nearly 20 cm^{-1} lower than that given in the earlier reference.[11] The cause for this is the different spectral regions fitted (1725−1540 cm^{-1} in the earlier fits; 1725−1570 cm^{-1} in the current fits), along with the absence of a return to baseline of the EG−OAc spectrum anywhere in the region near 1550 cm^{-1}. The fits of the acetate salts in this region are therefore much more susceptible than the other salts to choices in number of bands, baseline, etc. The earlier choice for the number of bands (five) led to the fitting program finding several nearly equal bands associated with a single intensity maximum near 1606 cm^{-1}. This led to a fitted frequency which, although it was very reproducible under the stated conditions of the earlier fits, we now consider not to be directly comparable to those of the other salts, which never gave fits with multiple bands under the 1606 cm^{-1} maximum. Other smaller discrepancies in frequencies between Table 2 and the earlier table are likewise explained by minor changes in fitting conditions. In every case, the current choices are justifiable based on the spectra in the Supporting Information.

Interestingly, in the ^{2}H-exchanged sample, only the lowest frequency ν_{C-N} mode, near 1135 cm^{-1}, shows a reliable monotonic counterion dependence, i.e., a frequency that consistently increases with increasing H-bond capacity of the anion. From this, we conclude that the anion dependence of the ν_{C-N} modes in the nondeuterated samples is largely a reflection of their coupling with in-plane hydrogen bending motions in the 1200−1600 cm^{-1} range. These frequencies are expected to be directly affected by the H-bonding force field of the anion. The stiffer the H-bond, the higher the in-plane hydrogen bending frequency, and the higher the coupled ν_{C-N} frequencies. In the deuterium-exchanged ethylguanidinium, however, the in-plane deuterium motions are significantly downshifted, to below 1000 cm^{-1}.[10] At this low frequency, they are capable of effectively coupling only to the lowest frequency ν_{C-N} mode.

Effects of Solvent Environment on the C═N Stretch Frequency. In addition to counterion effects, it would be desirable to be able to use measured vibrational frequencies of arginine residues in proteins to determine the nature of nearby neutral molecules (other amino acid side chains and/or solvent). To this end, we compare EG frequencies in several different solvents.

The three C−N stretching vibrations of EG−Cl are at 1673, 1633, and 1182 cm^{-1} in $CHCl_3$/MeOH 97:3 (Table 1 and ref 11); at 1673, 1635, and 1180 cm^{-1} in methanol (Figure 2A); at 1673, 1648, and 1183 cm^{-1} in dimethyl sulfoxide (DMSO; data not shown), and at 1674, 1635, and 1180 cm^{-1} in H_2O (data not shown). The last set of frequencies are very close to those measured for aqueous arginine hydrochloride.[2] These results would seem to indicate that the three ν_{C-N} bands of arginine, at least as the chloride salt, should be quite insensitive to solvent environment. (The relatively weak middle band near 1630 cm^{-1} shows the largest solvent dependence, about 8 cm^{-1} lower in frequency in nonpolar solvent.) This result appears consistent with an earlier conclusion from ab initio calculations that H-bonding to solvent (H_2O) should have little effect on ν_{C-N} frequencies.[18]

However, EG exhibits ν_{C-N} frequencies that are anion-independent in the polar solvents (MeOH, DMSO, and H_2O; spectra not shown), whereas they are anion-dependent in a nonpolar solvent (Table 1; see also spectra in SI-1). Thus, significant differences are observed between the ν_{C-N} frequencies of EG acetate measured in $CHCl_3$/methanol (97:3) and in a polar solvent such as methanol. Specifically, the strongest IR band of EG acetate is found at 1682 cm^{-1} in nonpolar solvent (Table 1), but at 1673 cm^{-1} in methanol (data not shown).

As discussed above, the widest variation in ν_{C-N} frequencies of EG salts is actually seen in the solvent-free solid state. Compared to the large effects of the counterion, there is no simple relation that can yet be discerned between solvation and C−N stretch frequencies of the arginine side chain. In general, however, it appears that solvation by a strong H-bond-accepting solvent may compete with counterion pairing in raising the highest of the ν_{C-N} frequencies. The lowest value that we have observed for this frequency, 1652 cm^{-1}, occurs for the "weakest"counterion, Br^-, without solvent present. This frequency is raised either by adding solvent (even one as nonpolar as $CHCl_3$/MeOH 97:3) or by switching to a stronger paired counterion. The effect of both switching the counterion to Cl^- and solvating with $CHCl_3$/MeOH 97:3, which raises the frequency to 1673 cm^{-1}, just about matches the effect of dissolving in a single-component H-bond-acceptor solvent such as H_2O, DMSO, or MeOH. The added counterion strength of a single paired oxyanion such as phenolate or acetate, on the other hand, apparently exceeds the solvation effect of neutral H-bond acceptors such as H_2O, DMSO, or MeOH, even though multiple solvent molecules undoubtedly form H-bonds with the guanidino group.

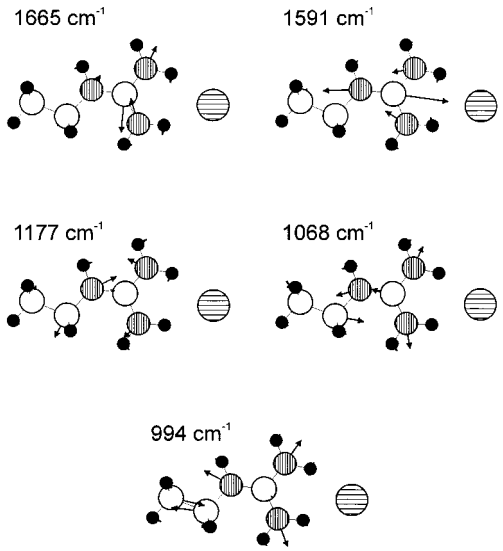

Figure 4. Mass-weighted atomic displacements for five normal modes discussed in the text. The frequencies and displacements were obtained directly from the output of published GAUSSIAN94 calculations on ethylguanidinium chloride[13] and converted to a graphical format by using ANIMOL software.

To test such effects with greater assurance, however, it will be necessary to find other solvent systems, especially solvents even less polar than $CHCl_3$/MeOH 97:3 or DMSO, in which pairing between alkylguanidinium and its counterion can reliably be shown to occur.

Discussion

Our IR spectra of [^{15}N]$_2$-thioethylguanidinium bromide allow a fairly detailed empirical assignment of the C−N stretching vibrations of arginine's side chain. The observed isotope shifts confirm the commonly accepted view that two strong characteristic IR bands near 1620 and 1670 cm^{-1} correspond largely to C−N stretching modes. Contrary to the conclusions of previous studies, however, a comparison of ^{15}N isotope shifts suggests that the majority of the remaining C−N stretch character is at ∼1170 and ∼1085 cm^{-1}, rather than 922 cm^{-1}.

Recent ab initio density functional calculations of the geometry and force field of ethylguanidinium chloride[13] predict vibrations having significant C−N stretch character at 1665, 1591, 1177, 1068, and 994 cm^{-1}, remarkably close to the frequencies of 1656, 1614, 1170, 1087, and 922 cm^{-1} that we have observed to downshift upon [^{15}N]$_2$ isotopic substitution of thioethylguanidinium bromide (Figure 1). Ab initio calculations also reproduce, to within ∼30%, the observed isotope shifts of each of these bands upon either [^{15}N]$_2$ or [^{2}H]$_5$ substitution (Figure 1 and Table 2, respectively). They thus confirm the empirical assignments described above and provide a quantitative description of the character of each mode. Conversely, the observed bandshifts in the current work provide a useful test of the accuracy of the ab initio calculations. Together, the close agreement between theory and experiment indicates that the energy-minimized geometry and force field of the arginine side chain complexed with halide ions have now been modeled with considerable accuracy.

Figure 4 presents a graphical representation of the mass-weighted displacements in each of these five normal modes,

based on the ab initio calculation.[13] The displacement demonstrates that the symmetric C−N stretch character is divided between the vibrations at 1068 and 994 cm^{-1}, where it is strongly mixed with other internal coordinates, especially the C−C and C−N stretch motions of the alkyl side chain. The displacement vectors in Figure 4 also indicate that, in the 1177 cm^{-1} mode, the C−N bond to the alkyl-substituted nitrogen accounts for most of the guanidino C−N stretch character. In this mode, the motion of the two terminal nitrogens, which reflects for the sensitivity of this mode to isotope substitution at these positions, involves N−C−N bending as well as C−N stretching.

Spectra presented here also provide a model for what is expected to happen to the strongest arginine IR absorption bands when this residue is subjected to various perturbations within a protein, e.g., during transmembrane anion transport or during an enzymatic reaction. Deprotonation of arginine's side chain is expected to lead to an overall decrease in absorption strength, along with 50−100 cm^{-1} downshifts in the strong ν_{C-N} modes to ∼1600 cm^{-1}. If the arginine participates in a salt bridge within a nonpolar region of a protein, substitution with a different counterion can lead to shifts of up to ∼40 cm^{-1} in the strongest ν_{C-N} vibration in the range 1650−1700 cm^{-1}. In general, the weaker the H-bonding tendency of the counteranion(s), the lower the expected frequency. An arginine surrounded by a number of weakly H-bonding I^- ions would be expected to give rise to a ν_{C-N} frequency near 1655 cm^{-1}. On the other hand, surrounding the guanidino group with multiple strongly H-bonding anions (e.g., glutamate or aspartate side chains or hydroxide) is expected to lead to a very high ν_{C-N} frequency, near 1695 cm^{-1}. The intermediate H-bonding provided, e.g, by −OH-containing side chains of neutral amino acids (e.g., Ser, Thr) or water is expected to produce intermediate frequencies.

It should be noted, however, that such environmental dependence of C−N frequencies is expected only for arginines containing ^1H. As a result of weaker couplings with C−N−D bending motions, the strong C−N stretch bands of deuterated arginine are expected to be much less sensitive to the hydrogen-bonding environment of its surroundings (for example, see spectra in SI-2).

The unusually high published frequency of 1695 cm^{-1} for arg-108 of halorhodopsin[11,20] can now be explained qualitatively based on empirical results with model compounds. Such a high frequency is most consistent with this arginine being H-bonded principally to multiple carboxylate (i.e., aspartate or glutamate) anions, rather than to polar neutral side chains (Ser, Thr) and/or H_2O. In addition to the multiple carboxylates, however, this arginine likely makes at least one H-bond with a bound halide ion, since the frequency of the ∼1695 cm^{-1} band shifts down in frequency (progressively by 1 cm^{-1} increments) if the protein is washed from medium containing Cl^- into medium containing Br^- or I^-.[11]

For bacteriorhodopsin, in which arg-82 is known to play a key role in the transient light-induced proton release, the model compound spectra presented here suggest a spectroscopic means to test whether deprotonation of arg-82 or a nearby water molecule occurs. Recently, a negative band at 1670 cm^{-1} and some of the positive intensity near 1615 cm^{-1} in the bR → M difference spectrum have been assigned to arg-82, based on site-directed $^{14}N/^{15}N$ substitution of the bacteriorhodopsin mutant arg82 → cys with thioethylguanidinium attached to position 82 (U. Alexiev, M. S. Hutson, and M. S. Braiman, unpublished data). The assigned C−N vibrational frequencies do not support the proposal that arg-82 forces the release of a proton from a nearby water molecule to form arginine hydroxide. However, deprotonation of arg-82 itself is not ruled out by these data.

Acknowledgment. This work was supported by NIH Grant GM46854. D.M.B. was supported in part by Molecular Biophysics Training Grant GM08323. We are grateful to Carl Trindle for many useful discussions and to Richard Krebs for obtaining confirmatory IR spectra of thioethylguanidinium bromide.

Supporting Information Available: IR spectra over the entire 1800−1100 cm^{-1} region of the acetate, chloride, bromide, and iodide salts of EG in chloroform/[^1H]-methanol (97:3) are included as Figure SI-1. In the region from 1540−1100, deuterated chloroform was used to eliminate interfering solvent vibrations; vertical scales of spectra in this region were carefully scaled to correspond to show correct absorption strength compared to the 1800−1540 cm^{-1} region. Also included in Figure SI-1 is the spectrum of a dried film of EG acetate. Figure SI-2 shows spectra of the same set of EG salts but measured using solvent containing MeO^2H, resulting in deuteration of five exchangeable positions on the EG ion. Deuterated chloroform was used for obtaining the entire spectral region. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Dioumaev, A. K.; Braiman, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 10572−10574.
(2) Venyaminov, S. Y.; Kalnin, N. N. *Biopolymers* **1990**, *30*, 1243−1257.
(3) Grigorieff, N.; Ceska, T. A.; Downing, K. H.; Baldwin, J. M.; Henderson, R. *J. Mol. Biol.* **1996**, *259*, 393−421.
(4) Kimura, Y., Vassylyev, D. G.; Miyazawa, A.; Kidera, A.; Matsushima, M.; Mitsuoka, K.; Murata, K.; Hirai, T.; Fujiyoshi, Y. *Nature* **1997**, *389*, 206−211.
(5) Pebay-Peyroula, E.; Rummel, G.; Rosenbusch, J. P.; Landau, E. M. *Science* **1997**, *277*, 1676−1681.
(6) Luecke, H.; Richter, H.-T.; Lanyi, J. K. *Science* **1998**, *280*, 1934−1937.
(7) Oesterhelt, D. *Israel J. Chem.* **1995**, *35*, 475−494.
(8) Anderson, M. P.; Gregory, R. J.; Thompson, S.; Souza, D. W.; Paul, S.; Mulligan, R. C.; Smith, A. E.; Welsh, M. J. *Science* **1991**, *253*, 202−205.
(9) Baldwin, J. M. *Curr. Opin. Cell Biol.* **1994**, *6*, 180−190.
(10) Sension, R. J.; Hudson, B.; Callis, P. R. *J. Phys. Chem.* **1990**, *94*, 4015−4025.
(11) Braiman, M. S.; Walter, T. J.; Briercheck, D. M. *Biochemistry* **1994**, *33*, 1629−1635.
(12) Doherty, D. G.; Shapira, R. H.; Burnett, W. T., Jr. *J. Am. Chem. Soc.* **1957**, *79*, 5667−5671.
(13) Trindle, C.; Braiman, M. S.; Prager, A. B. *Int. J. Quantum Chem.*, in press.
(14) Smith, J. N. *Biochem. J.* **1958**, *69*, 509.
(15) Chang, M. L. W.; Johnson, B. C. *J. Biol. Chem.* **1957**, *226*, 799.
(16) Jones, W. J. *Trans. Faraday Soc.* **1958**, *55*, 524−531.
(17) Smith, S. O.; Braiman, M. S.; Myers, A. B.; Pardoen, J. A.; Courtin, J. M.; Winkel, C.; Lugtenburg, J.; Mathies, R. A. *J. Am. Chem. Soc.* **1987**, *109*, 3108−3125.
(18) Magalhães, A. L.; Gomes, J. A. N. F. *Int. J. Quantum Chem.* **1997**, *61*, 725−739.
(19) Bonner, O. D.; Jordan, C. F. *Spectrochim. Acta* **1976**, *32A*, 1243−1246.
(20) Rüdiger, M.; Haupts, U.; Gerwert, K.; Oesterhelt, D. *EMBO J.* **1995**, *14*, 1599−1606.