Published in final edited form as:

J Phys Chem B. 2019 June 20; 123(24): 5079–5085. doi:10.1021/acs.jpcc.9b03766.

4-Oxoproline as a Site-Specific Infrared Probe: Application to Assess Proline Isomerization and Dimer Formation

Rachel M. Abaskharon^{+,#}, Debopreeti Mukherjee⁺, and Feng Gai^{*}

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

Abstract

Due to its unique structure, proline plays important structural and functional roles in proteins. However, this special amino acid lacks an adequate vibrational mode that can be exploited to probe its local electrostatic and hydration status via infrared spectroscopy. Herein, we show that the C=O stretching vibration of a proline derivative, 4-oxoproline, is sensitive to local environment and hence can be used as a site-specific infrared probe. We further validate this notion by applying this unnatural amino acid to assess the thermodynamics of proline *cis-trans* isomerization in a peptide environment and to examine amino acid dimer formation in concentrated proline and glycine solutions.

Graphical Abstract

1. INTRODUCTION

The functionality of a protein is determined by its structure and dynamics. However, assessing and understanding the underlying structure-dynamics-function relationship of any protein is not easy, due to the intrinsic complexity of the problem as well as the limitation of the currently available experimental techniques. For example, it is difficult, if not impossible, to take snapshots of key atomic/molecular motions during the course of a protein's action (e.g., catalysis or folding) or along a specific 'reaction' coordinate, while in

^{*}Corresponding Author gai@sas.upenn.edu.

⁺These authors made equal contribution

[#]Present Address: 5200 Illumina Way, San Diego, CA 92122, USA

ASSOCIATED CONTENT

Supporting Information: Onsager field-frequency correlation plot for 1M3P, fitting of the C=O stretching bands of KP_{ox}G in water and P_{ox} in 2.0 M proline solution at different temperatures, and the temperature-dependent IR spectra of GP_{ox}G in water. This material is available free of charge at <http://pubs.acs.org>.

the meantime measuring the corresponding changes in relevant physical and/or chemical properties, such as hydration, electric field, force, and energy. Therefore, in practice it is necessary to apply the divide-and-conquer approach, employing multiple techniques and experiments to assess the system of interest from different angles and perspectives. In this regard, spectroscopic studies of protein structure and dynamics have capitalized on a wide variety of intrinsic and extrinsic probes, using them to acquire different types of information. The strategy of utilizing extrinsic probes has been particularly useful for vibrational spectroscopy, since the vibrational signals intrinsic to proteins are often either insufficient or inadequate to offer the information needed. However, because incorporation of an external moiety (except isotopic labeling) will unavoidably perturb the native system in question, effort must be made to minimize such perturbations. For this reason, recent years have seen an increased interest in the utilization of unnatural amino acid-based (UAA-based) vibrational probes,¹⁻⁴ due to their small size and hence potentially less-perturbative nature. In specific applications, however, the perturbation is minimized only when the chosen UAA is structurally similar to the native amino acid it replaces. Therefore, it is necessary to develop a diverse set of UAA-based vibrational probes that are analogs of the twenty natural amino acids. While significant progress has been made in that direction,¹⁻⁴ we still lack an UAA that is suitable to replace proline, a special amino acid, for assessing its local environment via vibrational spectroscopy. Herein, we aim to show that the C=O stretching vibration of 4-oxoproline (P_{ox} , Figure 1) can be used as an infrared (IR) probe of the local electric field of proline.

Proline plays an important role in protein structure, dynamics, and function due to its unique cyclized and rigid structure, where its sidechain is integrated into its backbone unit.⁵ For example, (1) proline is a major component of the triple helices in collagen;⁶ (2) proline is frequently found in transmembrane proteins where it is used to introduce a kink in α -helices⁷ as well as in intrinsically disordered proteins where proline-rich segments are used to disrupt structural formation;⁸ (3) proline is a key component of motifs that specifically recognize signaling proteins;⁹ and (4) proline can serve as a protecting osmolyte, helping cells to deal with various environmental stresses.¹⁰ Additionally, the cyclized proline structure also allows for *cis-trans* isomerization along the peptide bond preceding it, which can dictate the rate of protein conformational transitions.⁵ Given the importance of proline and the fact that it lacks an intrinsic vibrational mode that is adequate to probe its structure and local environment, it would be desirable to find a proline derivative that bears such a vibrational mode. We hypothesize that P_{ox} is such a derivative because (1) the C=O stretching vibration, such as that of ketone, amide, and ester, is sensitive to local electrostatic field and hydration,¹¹⁻¹³ (2) the frequency of this vibrational mode of cyclic ketones is $>1700\text{ cm}^{-1}$,¹⁴ located in a relatively uncongested region of the protein IR spectrum, and (3) Raines and coworkers¹⁵ have shown that in spite of its electrophilic nature, P_{ox} can replace proline in a collagen mimetic peptide without compromising the thermal stability of the corresponding triple helix structure, suggesting that proteins can tolerate a proline-to- P_{ox} substitution. Moreover, P_{ox} exists in natural compounds, such as actinomycin (X_2 and X_{AA}) isolated from the bacterium *Streptomyces chrysomallus*.¹⁶ To establish the potential utility of P_{ox} as an IR probe, we examined (1) the solvent dependence of the C=O stretching vibration of a model compound, 1-methyl-3-pyrrolidinone (1M3P, Figure 1), and a P_{ox} -

containing peptide, (2) the correlation between the C=O stretching frequency of 1M3P and local electric field via molecular dynamics (MD) simulations, and (3) the pH dependence of the C=O stretching vibration of the free amino acid P_{ox} . As expected, our results corroborate the notion that the C=O stretching frequency of P_{ox} is a sensitive reporter of its local environment. Furthermore, to demonstrate its utility, we employed this vibrational mode to assess the thermodynamics of proline *cis-trans* isomerization and to probe amino-acid dimer formation in concentrated proline and glycine solutions.

2. EXPERIMENTAL SECTION

2.1. Materials

All materials were used as received. Specifically, 1-methyl-3-pyrrolidinone (97% purity) was purchased from Sigma-Aldrich, Fmoc-L-Pro(4-keto)-OH was purchased from Chem-Impex International Inc., and 4-oxo-L-proline hydrobromide (90% purity) was purchased from Alfa Aesar. In addition, methanol (MeOH), acetonitrile (ACN), dimethyl formamide (DMF) and dimethyl sulfoxide (DMSO) were purchased from Acros Organics, hexanes and diethyl ether (Ether) were purchased from Fisher Scientific, and deuterium oxide (D_2O ; 99.96% purity) was purchased from Cambridge Isotope Laboratories.

2.2. Peptide Samples

All peptides were synthesized using standard Fmoc protocols on a Liberty Blue microwave peptide synthesizer (CEM, NC) and then cleaved from the rink amide resin using a trifluoroacetic acid (TFA) cleavage cocktail. Crude peptide products were then purified using reverse-phase high performance liquid chromatography (HPLC; Agilent Technologies, CA) and identified using liquid chromatography-mass spectrometry (LC-MS; Water, MA). For each peptide, the purified product was then dissolved in a 0.01 M DCl solution (in D_2O), followed by lyophilization. This process was repeated at least three times, with the purpose of replacing all exchangeable hydrogen atoms with deuterium and removing the residual TFA in the sample. Peptide samples used in the IR measurements were prepared by directly dissolving the lyophilized peptide solids in D_2O , and the final pH of the solution was adjusted by adding an appropriate amount of either a 0.1 M NaOD or 0.1 M DCl solution (in D_2O).

2.3. Fourier Transform Infrared (FTIR) Measurements

FTIR spectra were collected on a Thermo Nicolet Magna 860 FTIR spectrometer equipped with a MCT detector (Nicolet, WI) with 1 cm^{-1} resolution. The detail of the CaF_2 sample cell (56 μm pathlength) was described elsewhere.¹⁷ Sample concentrations were ca. 10 mM for 1M3P, 4–5 mM for free P_{ox} (hydrobromide salt), 10–12 mM for $KP_{ox}G$, and 5–7 mM for $GP_{ox}G$. All reported spectra correspond to an average of 256 scans.

2.4. Gaussian Calculations

The Gaussian 09 software package was used to optimize the geometry of the $GP_{ox}G$ and $KP_{ox}G$ peptide systems and to calculate their respective vibrational frequencies in vacuo. The B3LYP level of theory was used in the 6–31+G(d,p) orbital basis set. The calculated frequencies were scaled by a factor of 0.9632 to yield the reported values.¹⁸

3. RESULTS AND DISCUSSION

3.1. Dependence of the C=O Stretching Frequency of 1M3P on Solvent

In order to assess the feasibility of using the C=O stretching vibration of P_{ox} as an IR probe of its local environment, we first performed vibrational solvatochromic measurements⁴ on a small molecule mimic, 1M3P (Figure 1). The reasons for choosing this small molecule are two-fold: (1) it is soluble in both protic and aprotic solvents whereas P_{ox} is not (or has a low solubility in organic solvents) and (2) it better mimics the P_{ox} residue in a peptide environment than an uncapped version would, as its nitrogen is capped by a methyl group. As shown (Figure 2), the C=O stretching vibrational band (frequency and width) of 1M3P exhibits a strong dependence on solvent. For example, changing the solvent from water to hexane, which has a dielectric constant of 1.89, leads to a blue-shift of its frequency by ca. 20 cm^{-1} , accompanied by a significant narrowing of its bandwidth. This result is consistent with the well-known phenomenon that hydrogen-bonding (H-bonding) interactions between solvent molecules and a C=O moiety could decrease its stretching frequency and also suggests that the C=O stretching vibration of P_{ox} could be used as a local hydration probe of proteins. Interestingly, in methanol, 1M3P gives rise to at least two resolvable C=O stretching bands, suggesting the existence of differently hydrogen-bonded (H-bonded) C=O species.¹¹

For a simple vibration, the solvent induced frequency shift is typically modeled in the framework of Stark effect, which asserts a linear relationship between the vibrational frequency and the external electric field exerted by the solvent molecules onto the oscillator.¹⁹ To show whether the C=O stretching vibration of 1M3P exhibits such a dependence,²⁰ we examined the correlation between its frequency and the Onsager reaction field ($F_{Onsager}$) of the solvent (see detail in SI).²¹ Since the Onsager model does not consider hydrogen bonding interactions, we only analyzed, as previously done for other IR probes,¹¹ the data obtained in aprotic solvents. As shown (Figure S1 in SI), the C=O stretching frequency of 1M3P decreases with increasing the magnitude of $F_{Onsager}$, exhibiting a linear trend. However, the linear correlation is not strong, especially for the frequencies measured in DMSO, DMF, and ACN. We believe that this is due to the molecular structure of 1M3P, which contains a nitrogen (N) atom that, in comparison to other carbon (C) atoms in the ring, can more strongly interact with the solvent molecules. The effect of such interactions, which cannot be captured by the Onsager model, is a change in the net charge on the N atom, which, in turn, can affect the local electric field sensed by the C=O group and hence its frequency. As discussed below, this notion is further corroborated by other experiments and also Gaussian calculations.

3.2. Dependence of the C=O Stretching Frequency of P_{ox} on pH

Because P_{ox} has low solubility in organic solvent, we used pH as a variable to assess whether its C=O stretching frequency depends on local electric field. Since the amine and carboxylic acid groups of an amino acid have very different pK_a values, one can use pH to tune the population of differently charged states (for P_{ox} , these states, referred to as **1**, **2**, and **3**, are given in Figure 1). Therefore, if the C=O stretching vibration of P_{ox} is indeed sensitive to local electric field, we expect its frequency to exhibit a dependence on pH.

As shown (Figure 3), the FTIR spectra of P_{ox} collected at pH 1.0, 3.0, 5.0, 7.0, 9.0 and 11.0 are indeed quite different. Assuming that the pK_a values of P_{ox} are similar to those of proline, which are 1.99 and 10.6 for the carboxyl and amine groups,²² respectively, it is expected that state **1** is the most populous at $pH < 2$, state **2** is the most populous in the pH range of 2 to 8, and state **3** is the most populous at $pH > 9$. This information helps determine the assignment of those peaks in the pH-dependent FTIR spectra of P_{ox} . The spectrum obtained at pH 1.0 contains three distinct peaks at 1625, 1726, and 1772 cm^{-1} , respectively. It is well known that the carbonyl stretching frequency of the protonated carboxylic acid group of amino acids is ca. 1730 cm^{-1} .²³ Therefore, we assign the 1726 cm^{-1} band to the carbonyl stretching vibration of the $-COOH$ group of state **1**. This assignment is supported by the fact that in the spectra obtained at pH 3.0 and 5.0, where the concentration of state **1** becomes significantly smaller, the 1726 cm^{-1} band disappears. Consequently, the 1770 cm^{-1} band must arise from the keto group of state **1**. Furthermore, based on a previous study²⁴ that showed that the deformation vibration of the $-NH_3^+$ group of glycine has a frequency of ca. 1615 cm^{-1} , we attribute the 1625 cm^{-1} band to the $-NH_2^+$ group in state **1**. At pH 3.0 and especially pH 5.0, where state **2** is most populated, we expect to observe a new band corresponding to the asymmetric carbonyl stretching vibration of its $-COO^-$ group. However, the FTIR spectra collected at these pH conditions consist of only the 1625 and 1770 cm^{-1} bands, except that the intensity of the 1625 cm^{-1} band is increased about 3 times. These results indicate that (1) the ionization status of the carboxyl group of P_{ox} has no significant effect on the bands arising from the $-NH_2^+$ and keto moieties, which is expected as this group is not an integral part of the ring and is relatively further away, and (2) the $-COO^-$ band of state **2** is also peaked at ca. 1625 cm^{-1} , making it unresolvable from the $-NH_2^+$ band. As shown (Figure 3), a further increase in pH leads to the formation of two new bands, at 1588 and 1740 cm^{-1} , respectively, with the concomitant decrease of the 1625 and 1770 cm^{-1} bands. Considering the fact that the population of state **3** increases with increasing pH and becomes dominant at pH 11.0, these results indicate that the $-COO^-$ and keto bands of state **3** are shifted to 1588 and 1740 cm^{-1} , respectively. In addition, the complete disappearance of the 1625 cm^{-1} band at pH 11.0 corroborates the notion that the $-NH_2^+$ group absorbs at this frequency. Furthermore, and perhaps more importantly, these results indicate that the charge density on the nitrogen atom in P_{ox} has a significant effect on the $C=O$ stretching frequency of its sidechain, hence supporting the aforementioned dependence of this frequency on local electric field. Moreover, the fact that changing $-NH$ (a neutral moiety) to $-NH_2^+$ (a charged moiety) leads to a blue shift of its stretching frequency is consistent with the aforementioned Onsager field analysis (Figure S1 in SI). This is because adding a positive charge at a point below the carbon atom of the $C=O$ group will effectively reduce the electric field that is exerted on this moiety by the solvent molecules.

3.3. The $C=O$ Stretching Frequency of P_{ox} in a Peptide Environment

To further test the utility of P_{ox} as a local IR probe of proline, we studied two tripeptide systems, $GP_{ox}G$ and $KP_{ox}G$. As shown (Figure 4), the $C=O$ stretching band of P_{ox} in $GP_{ox}G$ is peaked at 1761.4 cm^{-1} in D_2O , which is shifted to 1779.8 cm^{-1} in a more hydrophobic solvent, tetrahydrofuran (THF). These results demonstrate that the environment-dependence of the $C=O$ stretching vibration of P_{ox} is maintained when it is incorporated in a polypeptide chain, further validating its applicability as a site-specific IR probe of proline in proteins. As

indicated (Figure 4), the C=O stretching band of P_{ox} in KP_{ox}G is centered at 1759.5 cm⁻¹ in D₂O, which is red-shifted from that of GP_{ox}G. This frequency shift manifests a change in the local electric field felt by the C=O group, induced by the charged -NH₃⁺ group of Lys, and again highlights the sensitivity of the C=O stretching vibration of P_{ox} to its local environment. In addition, the direction of the shift indicates that the magnitude of the electric field is increased, consistent with the positive nature of the Lys sidechain.

3.4. Application to Assess Proline *cis-trans* Isomerization

To validate the utility of P_{ox} as an IR probe, we used it to assess the thermodynamics of its *cis-trans* isomerization in a peptide environment. We chose KP_{ox}G for this purpose based on the notion that the distance between the positively charged Lys sidechain and P_{ox} will change upon isomerization, which, in turn, will lead to a change in the local electric field sensed by the keto group and hence its stretching frequency. As shown (Figure 5), increasing the temperature, which is expected to change the population ratio between the *cis* and *trans* conformations, indeed results in an apparent shift of the C=O stretching band. A more quantitative analysis indicates that at each temperature the C=O stretching band can be fit by two pseudo-Voigt profiles (Figure S2 in SI). It is known that for prolines in a peptide environment the *trans* form is more populated at low or room temperature.^{25,26} Therefore, we attribute the low frequency band to the *cis* conformer as its intensity increases with increasing temperature. To further substantiate this assignment, we carried out Gaussian frequency calculations in vacuo on KP_{ox}G, where the Lys sidechain is either charged or neutral. As shown (Table 1), in both cases, the carbonyl stretching frequency of P_{ox} in the *cis* form is red-shifted from that in the *trans* form, supporting the above band assignment. Moreover, in support of our hypothesis, the calculation predicts a larger frequency shift (i.e., ca. 5 cm⁻¹) upon P_{ox} isomerization in KP_{ox}G when the Lys sidechain is charged (i.e., K⁺P_{ox}G). As indicated (Table 1 and Figure S3 in SI), both the calculated and experimental results obtained with GP_{ox}G provide additional evidence that a charged sidechain next to P_{ox} will have a more pronounced effect on its local electrostatic environment, since in this case the C=O stretching frequency exhibits less of a dependence on the configuration of P_{ox}. Furthermore, an examination of the charges on the O and N atoms of P_{ox} (Table 1) suggests that the larger frequency shift observed for K⁺P_{ox}G is due to a larger change of the partial charge on the N atom. Moreover, the finding that a larger positive charge on the N atom gives rise to a higher carbonyl stretching frequency is consistent with the aforementioned pH-dependent study on free P_{ox}, which showed that protonation of the -NH group leads to a blue-shift of the frequency of this vibrational mode.

As indicated (Figure 6), the logarithm of the ratio between the integrated areas of the low-frequency and high-frequency bands exhibits a linear dependence on 1/*T*. Assuming that the cross-section of the C=O stretching vibration of P_{ox} is the same for both conformations, this ratio is effectively the equilibrium constant (*K*_{eq}) for the *trans* to *cis* isomerization process. A linear regression of the data in Figure 6 yielded a negative slope of 3.94×10³ K⁻¹ and an intercept of 12.64. Based on the van't Hoff equation, the *H*_{*trans*→*cis*} and *S*_{*trans*→*cis*} values were determined to be 7.8 kcal mol⁻¹ and 25.1 cal K⁻¹ mol⁻¹, respectively. Using these thermodynamic parameters, we further calculated the *G*_{*trans*→*cis*} value to be 0.9 kcal mol⁻¹ at 277 K. A previous study²⁷ on a pentapeptide (sequence: Ac-Ala-Lys-Pro-Ala-Lys-NH₂)

showed that the $G_{trans \rightarrow cis}$ is 1.4 kcal/mol at the same temperature. Therefore, the $G_{trans \rightarrow cis}$ value of 0.9 kcal/mol determined for the KP_{ox}G peptide is in general agreement with those previous studies, indicating the suitability of replacing a proline residue with P_{ox} and the applicability of its C=O stretching vibration as a reporter of proline's *cis-trans* isomerization in peptides or proteins.

3.5. Application to Detect Amino-Acid Dimer Formation in Solution

Free amino acids, such as proline and glycine, are widely used by plants to deal with stresses.²⁸ In particular, they are used as osmolytes to improve the thermal stability of proteins.^{29,30} Therefore, a large number of studies have been dedicated to elucidate the mechanism of action of these two amino acids. While early studies suggested that at a concentration of 2 M, glycine cannot form polymeric aggregates³¹ but proline can,^{32,33} the most recent ones^{34–37} suggest that both glycine and proline are capable of forming small aggregates such as dimers at lower concentrations. To provide insight into this problem and further demonstrate the utility of the C=O stretching vibration of P_{ox}, we use it to probe aggregate formation in concentrated proline and glycine solutions. Our premise is that P_{ox} (in the form of the free amino acid) is able to provide spectroscopic evidence about the nature of the aggregates formed by these amino acids by directly participating in the aggregation process, where, as a probe molecule, the concentration of P_{ox} will be kept low. While it is impossible to determine the structure of the aggregates based on the C=O stretching frequency of P_{ox} alone, this information is nevertheless sufficient to allow identification of the underlying mode of interaction. For example, should proline (glycine) molecules aggregate through dimerization via H-bond formation, as that depicted in several studies,^{34–37} we expect to observe a red-shift in the stretching frequency of the C=O oscillator of P_{ox} in 2.0 M proline (glycine) solution in comparison to that in D₂O. This is because in such dimers the –NH₂⁺ group of P_{ox} would interact with the –COO[–] group of another amino acid (e.g., proline), effectively decreasing the positive charge density on the N atom of P_{ox}. As discussed above, such a decrease would result in a decrease in the C=O stretching frequency of P_{ox}. On the other hand, if the association is initiated through hydrophobic interactions between sidechains, we expect to observe a blue-shift in the C=O stretching frequency of P_{ox} with respect to the C=O band of P_{ox} in D₂O.

As shown (Figure 7), in 2.0 M proline solution (pH = 4.0), the probe molecule, P_{ox} (25 mM), gives rise to two resolvable peaks, at 1768.2 and 1725.2 cm^{–1}, respectively. The result obtained in 2.0 M glycine solution (pH = 4.0) is similar, although the lower frequency peak is shifted to 1734.9 cm^{–1}. In comparison, however, only one peak, at 1769.8 cm^{–1}, is observed for P_{ox} (25 mM) in D₂O (pH = 4.0). Moreover, for the three peaks at ca. 1768 cm^{–1}, the one obtained in D₂O is the most intensive. Therefore, these data indicate that in both 2.0 M proline and glycine solutions the P_{ox} molecules can sample two distinctly different environments (or two states), with one being similar or identical to that in D₂O (i.e., where P_{ox} exists as a D₂O solvated monomer). Since the other band is shifted to a lower frequency, it suggests, as noted above, that in this case the –NH₂⁺ group of P_{ox} engages in H-bonding interaction with the –COO[–] group of another solute molecule. Therefore, these results are in agreement with previous studies indicating that association of proline (glycine) molecules occurs through dimerization.^{34–37}

The temperature dependence of the C=O stretching bands of P_{ox} in 2.0 M proline provides further support of the aforementioned two-state assignment. This is because the intensity of the low-frequency (high-frequency) band decreases (increases) with increasing temperature (Figure 8), as expected for monomer-dimer equilibrium. As shown (Figure S4 in SI), the FTIR spectra can be quantitatively decomposed into two pseudo-Voigt profiles and, similar to the treatment of the *cis-trans* isomerization process, the ratio of the integrated areas of these two pseudo-Voigt bands at a specific temperature is the equilibrium constant of the system at this temperature. As expected (Figure 9), the corresponding van't Hoff plot yielded a straight line with a slope of $1.25 \times 10^3 \text{ K}^{-1}$ and an intercept of -4.34 . Therefore, the H and S values for the dimerization process are calculated to be $-2.6 \text{ kcal mol}^{-1}$ and $-8.6 \text{ cal K}^{-1} \text{ mol}^{-1}$, respectively. To the best of our knowledge, the thermodynamics of proline dimerization in aqueous solution has not been determined before. Therefore, we are unable to compare our result with literature values. Nevertheless, the H value of $-2.6 \text{ kcal mol}^{-1}$ is comparable to those (-1.3 to $-4.9 \text{ kcal mol}^{-1}$) measured for base pairing of various nucleobases in protic solvents, which also involves formation of either two or three H-bonds.^{38–40} Hence, this thermodynamic analysis adds further evidence to support the spectroscopic utility of P_{ox}.

4. CONCLUSIONS

Among the 20 canonical amino acids, proline is unusual because its sidechain is integrated into its backbone unit. As a result, it often plays a unique role in protein structure, dynamics and function and hence requires special attention. However, it is difficult to apply IR spectroscopy, a technique widely used in protein science, to acquire structural or dynamic information specific to a proline residue in proteins because it does not have a convenient and strong vibrational transition that can be used to provide such information. Therefore, we explore the applicability of a proline derivative, 4-oxoproline, as an IR probe. We find that the C=O stretching frequency of this unnatural amino acid, in either free form or a peptide environment, not only is in an uncongested region of the protein IR spectrum (i.e., $1720\text{--}1800 \text{ cm}^{-1}$), but also exhibits a sensitive dependence on the local electrostatic and hydration environment, indicating that this vibrational mode can serve as a site-specific IR probe. This notion is further validated in two proof-of-principle studies. First, we show that the C=O stretching vibration of 4-oxoproline can be used to follow the *cis-trans* isomerization process of proline residues. Second, we demonstrate that this vibrational mode is sensitive enough to allow detection of solute dimers formed in concentrated solution of free amino acids (i.e., proline or glycine).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements.

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978). For the duration of this work, R.M.A. was an NSF Graduate Research Fellow (DGE-1321851). We also thank Dr. Jia Tang for her helpful suggestions.

REFERENCES

- (1). Waagele MM; Culik RM; Gai F Site-Specific Spectroscopic Reporters of the Local Electric Field, Hydration, Structure, and Dynamics of Biomolecules. *J. Phys. Chem. Lett* 2011, 2, 2598–2609. [PubMed: 22003429]
- (2). Ma J; Pazos IM; Gai F Microscopic Insights into the Protein-Stabilizing Effect of Trimethylamine N-Oxide (TMAO). *Proc. Natl. Acad. Sci. U. S. A* 2014, 111, 8476–8481. [PubMed: 24912147]
- (3). Błasiak B; Londergan CH; Webb LJ; Cho M Vibrational Probes: From Small Molecule Solvatochromism Theory and Experiments to Applications in Complex Systems. *Acc. Chem. Res* 2017, 50, 968–976. [PubMed: 28345879]
- (4). Adhikary R; Zimmermann JR; Romesberg FE Transparent Window Vibrational Probes for the Characterization of Proteins With High Structural and Temporal Resolution. *Chem. Rev* 2017, 117, 1927–1969. [PubMed: 28106985]
- (5). Wedemeyer WJ; Welker E; Scheraga HA Proline Cis–Trans Isomerization and Protein Folding. *Biochemistry*, 2002, 41, 14637–14644. [PubMed: 12475212]
- (6). Ramachandran GN Stereochemistry of Collagen. *Int. J. Pept. Protein Res* 2009, 31, 1–16.
- (7). Arkin IT; Brunger AT Statistical Analysis of Predicted Transmembrane α -Helices. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol* 1998, 1429, 113–128.
- (8). Murali MG; Piai A; Bermel W; Felli IC; Pierattelli R Proline Fingerprint in Intrinsically Disordered Proteins. *ChemBioChem* 2018, 19, 1625–1629. [PubMed: 29790640]
- (9). Kay BK; Williamson MP; Sudol M The Importance of Being Proline: The Interaction of Proline-Rich Motifs in Signaling Proteins with Their Cognate Domains. *FASEB J* 2000, 14, 231–241. [PubMed: 10657980]
- (10). Szabados L; Savouré A Proline: a Multifunctional Amino Acid. *Trends Plant Sci* 2010, 15, 89–97. [PubMed: 20036181]
- (11). Pazos IM; Ghosh A; Tucker MJ; Gai F Ester Carbonyl Vibration as a Sensitive Probe of Protein Local Electric Field. *Angew. Chem* 2014, 126, 6194–6198.
- (12). Oh K-I; Fiorin G; Gai F How Sensitive Is the Amide I Vibration of the Polypeptide Backbone to Electric Fields? *ChemPhysChem* 2015, 16, 3595–3598. [PubMed: 26419214]
- (13). Fried SD; Bagchi S; Boxer SG Measuring Electrostatic Fields in Both Hydrogen-Bonding and Non-Hydrogen-Bonding Environments Using Carbonyl Vibrational Probes. *J. Am. Chem. Soc* 2013, 135, 11181–11192. [PubMed: 23808481]
- (14). Brauman JI; Laurie VW Characteristic Vibrational Frequencies of Cyclic Ketones. *Tetrahedron* 1968, 24, 2595–2602.
- (15). Choudhary A; Kamer KJ; Shoulders MD; Raines RT 4-Ketoproline: An Electrophilic Proline Analog for Bioconjugation. *Biopolymers* 2015, 104, 110–115. [PubMed: 25656588]
- (16). Mirau PA; Shafer RH High-Resolution Proton NMR Analysis of the Conformational Properties of Biosynthetic Actinomycin Analogs. *Biochemistry* 1982, 21, 2622–2626. [PubMed: 7093210]
- (17). Huang C-Y; Klemke JW; Getahun Z; DeGrado WF; Gai F Temperature-Dependent Helix–Coil Transition of an Alanine Based Peptide 2001, 123, 9235–9238.
- (18). Irikura KK; Johnson RD III; Kacker RN Uncertainties in Scaling Factors for Ab Initio Vibrational Frequencies. *J. Phys. Chem. A* 2005, 109, 8430–8437 [PubMed: 16834237]
- (19). Boxer SG Stark Realities. *J. Phys. Chem. B* 2009, 113, 2972–2983. [PubMed: 19708160]
- (20). Gunathunge CM; Ovalle VJ; Waagele MM Probing Promoting Effects of Alkali Cations on the Reduction of CO at the Aqueous Electrolyte/Copper Interface. *Phys. Chem. Chem. Phys* 2017, 19, 30166–30172. [PubMed: 29105707]
- (21). Onsager L Electric Moments of Molecules in Liquids. *J. Am. Chem. Soc* 1936, 58, 1486–1493.
- (22). Cohn E; Edsall J Proteins, Amino Acids and Peptides as Ions and Dipolar Ions; Hafner Publishing Company: New York, 1943.
- (23). Barth A Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta - Bioenerg* 2007, 1767, 1073–1101.
- (24). Sebben D; Pendleton P Infrared Spectrum Analysis of the Dissociated States of Simple Amino Acids. *Spectrochim. Acta A Mol. Biomol. Spectrosc* 2014, 132, 706–712. [PubMed: 24910978]

- (25). Wu W-J; Raleigh DP Local Control of Peptide Conformation: Stabilization Of cis Proline Peptide Bonds by Aromatic Proline Interactions. *Biopolymers* 1998, 45, 381–394. [PubMed: 9530015]
- (26). Murray NJ; Williamson MP Conformational Study of a Salivary Proline-Rich Protein Repeat Sequence. *Eur. J. Biochem* 1994, 219, 915–921. [PubMed: 8112343]
- (27). Reimer U; Scherer G; Drewello M; Kruber S; Schutkowski M; Fischer G Side-Chain Effects on Peptidyl-Prolyl Cis/Trans Isomerisation. *J. Mol. Biol* 1998, 279, 449–460. [PubMed: 9642049]
- (28). Stewart GR; Larher F Accumulation of Amino Acids and Related Compounds in Relation to Environmental Stress. *Amin. Acids Deriv* 1980, 609–635.
- (29). Santoro MM; Liu Y; Khan SMA; Hou LX; Bolen DW Increased Thermal Stability of Proteins in the Presence of Naturally Occurring Osmolytes. *Biochemistry* 1992, 31, 5278–5283. [PubMed: 1376620]
- (30). Taneja S; Ahmad F Increased Thermal Stability of Proteins in the Presence of Amino Acids. *Biochem. J* 1994, 303, 147–153. [PubMed: 7945233]
- (31). Sato T; Buchner R; Arka Fernandez S; Chiba A; Kunz W Dielectric Relaxation Spectroscopy of Aqueous Amino Acid Solutions: Dynamics and Interactions in Aqueous Glycine. *J. Mol. Liq* 2005, 117, 93–98.
- (32). Rudolph ASAS; Crowe JH A Calorimetric and Infrared Spectroscopic Study of the Stabilizing Solute Proline. *Biophys. J.*, 1986, 50, 423–430. [PubMed: 3756295]
- (33). S Schobert B; Tschesche H Unusual solution properties of proline and its interaction with proteins. *Biochim. Biophys. Acta, Gen. Subj* 1978, 541, 270–277.
- (34). Bushuev YG; Davletbaeva SV; Koifman OI Molecular Dynamics Simulations of Aqueous Glycine Solutions. *Cryst. Eng. Comm* 2017, 19, 7197–7206.
- (35). Busch S; Lorenz CD; Taylor J; Pardo CL; McLain SE Short-Range Interactions of Concentrated Proline in Aqueous Solution. *J. Phys. Chem. B*, 2014, 118, 14267–14277. [PubMed: 25383939]
- (36). White A; Jiang S Local and Bulk Hydration of Zwitterionic Glycine and Its Analogues through Molecular Simulations. *J. Phys. Chem. B*, 2010, 115, 660–667. [PubMed: 21174438]
- (37). Yani Y; Chow PS; Tan RBH Glycine Open Dimers in Solution: New Insights into α -Glycine Nucleation and Growth. *Cryst. Growth Des* 2012, 12, 4771–4778.
- (38). Spencer JN; Gleim JE; Blevins CH; Garrett RC; Mayer FJ; Merkle JE; Smith SL; Louise Hackman M The N-H Hydrogen Bond. 2. Models for Nucleic Acid Bases. *J. Phys. Chem* 1979, 83, 2615–2621.
- (39). Binford JS; Holloway DM Heats of Base Pair Formation with Adenine and Uracil Analogs. *J. Mol. Biol* 1968, 31, 91–99. [PubMed: 5636881]
- (40). Williams LD; Chawla B; Shaw BR The Hydrogen Bonding of Cytosine with Guanine: Calorimetric and ¹H-NMR Analysis of the Molecular Interactions of Nucleic Acid Bases. *Biopolymers* 1987, 26, 591–603. [PubMed: 3567328]

Figure 1.
Structures of 1M3P and P_{ox} at different ionization states, as indicated.

Figure 2.
Band area normalized C=O stretching bands of 1M3P in different solvents, as indicated.

Figure 3.
FTIR spectra of the free amino acid P_{ox} at different pH values, as indicated.

Figure 4. Normalized C=O stretching bands of GP_{ox}G in D₂O and THF and KP_{ox}G in D₂O, as indicated.

Figure 5.
Temperature dependence of the C=O stretching bands of KP_{ox}G in D₂O (pH 7.0).

Figure 6.

Van't Hoff plot of the ratio between the integrated areas of the low-frequency and high-frequency bands (A_L/A_H) of $KP_{ox}G$ in D_2O (pH 7.0). Fitting these data to a straight line yielded a slope of $-3.94 \times 10^3 \text{ K}^{-1}$ and an intercept of 12.64.

Figure 7.

FTIR spectra of the free amino acid P_{ox} at in D_2O (pH 4.0), 2.0 M glycine solution (pH 4.0), and 2.0 M proline solution (pH 4.0), as indicated. Also shown (dashed lines) are the FTIR spectra of the corresponding 2.0 M proline and glycine solutions.

Figure 8. Temperature dependence of the C=O stretching bands of the free amino acid P_{ox} in 2.0 M proline solution (pH 4.0).

Figure 9.

Van't Hoff plot of the ratio between the integrated areas of the low-frequency and high-frequency bands (A_L/A_H) of $KP_{ox}G$ in D_2O (pH 7.0). Fitting these data to a straight line yielded a slope of $1.25 \times 10^3 \text{ K}^{-1}$ and an intercept of -4.34 .

Table 1.

Calculated C=O stretching frequency of P_{ox} in different peptides and different isomeric forms. Also listed are the Mulliken partial charges on the nitrogen (N) and oxygen (O) atoms of the P_{ox} sidechain.

Peptide, P _{ox} Configuration	Frequency	N Charge	O Charge
K ⁺ P _{ox} G, <i>cis</i>	1795.058	0.106	−0.376
K ⁺ P _{ox} G, <i>trans</i>	1800.096	0.419	−0.364
<i>trans-cis</i>	5.038	0.314	0.013
KP _{ox} G, <i>cis</i>	1787.988	0.126	−0.394
KP _{ox} G, <i>trans</i>	1789.308	0.242	−0.382
<i>trans-cis</i>	1.320	0.116	0.012
GP _{ox} G, <i>cis</i>	1790.126	0.073	−0.381
GP _{ox} G, <i>trans</i>	1790.936	−0.003	−0.376
<i>trans-cis</i>	0.809	−0.076	0.006