Correlating Hydration Shell Structure with Amino Acid Hydrophobicity

David Hecht, Lema Tadesse, and Lee Walters*

The Scripps Research Institute
Department of Molecular Biology, MB-15
LaJolla, California 92037

Received January 8, 1993

Historically, valuable insights have resulted from successful correlations between molecular structure and resulting macroscopic thermodynamic properties that are, by definition, structure independent. This communication provides data, obtained by infrared spectroscopy, on the degree of solvent–solvent hydrogen bonding in the hydration shells surrounding eight "hydrophobic" amino acids. These structural data are correlated with thermodynamic criteria for amino acid hydrophobicity.

Hydrophobicity has traditionally been defined in classical thermodynamic terms that are based upon macroscopic properties of solutes and their solutions. These include (a) the free energy of transfer of compounds from a nonaqueous to an aqueous phase[1-3,4a-h] and for amino acids (b) estimates of the solvent-accessible surface area that is lost as an amino acid is buried inside a fully folded protein[4i-p,5,6] and (c) the effect of amino acid substitutions on protein stability.[7]

The data presented below demonstrate systematic patterns in the vibrational transitions of water molecules that are associated with eight L-amino acids in aqueous solution. Shifts in vibrational frequencies and intensities of solvating water molecules are consistent with an increase in the degree of hydrogen-bonded structure[8,9] that is generated as the hydrocarbon moieties of amino acid side chains increase in size. The measured vibrations result from a complex convolution of the numbers and the enthalpies of solvent–solvent intermolecular hydrogen bonds.[8]

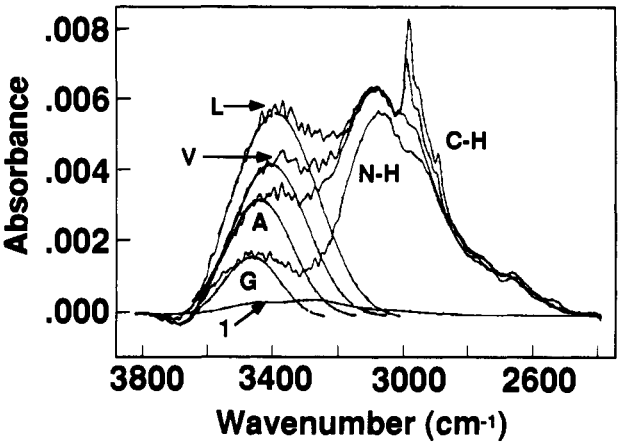

Figure 1. Individual difference spectra, i.e., (spectrum of zwitterionic amino acid solution) – (spectrum of pure water), of aqueous 100 mM amino acid solutions (prepared from ≥98% pure compounds, Calbiochem, with deionized and filtered water, $\leq 5.6 \times 10^{-8}$ S cm^{-1}, pore size = 0.22 μm; Millipore). Each difference spectrum results from the average of two or more independently measured sample spectra, with spectra exhibiting a variation of less than 4×10^{-4} AU (absorbance units) prior to averaging. The spectrum labeled 1 is the difference spectrum that results from subtracting two independently measured water samples. The regions where hydration shell water vibrations occur for each of the amino acid difference spectra are labeled with the appropriate single letter code (i.e., glycine, G; alanine, A; valine, V; leucine, L). Also shown in this region are Gaussian peaks representing water vibrations. These peaks resulted from deconvolutions that were performed using a curve fitting, iterative, linear least-squares algorithm (Fit; Galactic Industries Corp.). Regions where NH vibrations and CH vibrations occur are labeled accordingly. Difference spectra were smoothed using a maximum likelihood algorithm[13] (ESmooth; Galactic Industries Corp.). As described previously,[9] each difference spectrum was corrected to account for the exclusion of water molecules from a given volume of the solution. All spectra were measured on a Nicolet 740 FTIR as described previously.[9] Samples were examined in internal reflectance cells in duplicate experiments (pathlength ~7.5 μm, SpectraTech) containing ZnSe or GeAsSe crystals (45° conical ends). All samples were measured at pH 7.0 ± 0.1 at 24.0 ± 0.1 °C.

The spectra of four representative 100 mM aqueous amino acid solutions are shown in Figure 1. N–H vibrations occur in the 3200–2600-cm^{-1} region; C–H vibrations occur at wavenumbers <3000 cm^{-1}.[10]

The 3800–2800-cm^{-1} region of the absorbance spectrum of liquid water consists of three major families of infrared transitions:[10] at the high-energy end of this region are the asymmetric stretch ν_3 (←H–O→←H, ~3630 cm^{-1}) and the symmetrical stretch ν_1 (H→O←H, ~3450 cm^{-1}); at the low-energy end of the spectrum is the ν_2 overtone bending vibration (←H–O–H→, ~3225 cm^{-1}). Previous spectroscopic studies have

(1) Kauzmann, W. *Adv. Protein Chem.* **1959**, *14*, 1–63.
(2) Nozaki, Y.; Tanford, C. *J. Biol. Chem.* **1971**, *246*, 2211–2217.
(3) (a) Wolfenden, R. V.; Cullis, P. M.; Southgate, C. C. F. *Science* **1979**, *206*, 575–577. (b) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849–855.
(4) (a) Leo, A.; Hansch, C.; Elkins, D. *Chem. Rev.* **1971**, *71*, 525–616. (b) Yunger, L. M.; Cramer, R. D. *Mol. Pharmacol.* **1981**, *20*, 602–608. (c) Zaslavsky, B. Y.; Mestechkina, N. M.; Miheeva, L. M.; Rogozhin, S. V. *J. Chromatogr.* **1982**, *240*, 21–28. (d) Fauchère, J. L.; Pliska, V. *Eur. J. Med. Chem.* **1983**, *18*, 369–375. (e) Damodaran, S.; Song, K. B. *J. Biol. Chem.* **1986**, *261*, 7220–7222. (f) Abraham, D. J.; Leo, A. J. *Proteins: Structure, Function and Genetics*; Alan R. Liss, Inc.: New York, 1987; pp 130–152. (g) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664–1670. (h) Bhat, R.; Kishore, N.; Ahluwalia, J. C. *J. Chem. Soc., Faraday Trans. I* **1988**, *84*, 2651–2665. (i) Chothia, C. *Nature* **1974**, *248*, 338–339. (j) Chothia, C. *J. Mol. Biol.* **1976**, *105*, 1–14. (k) Wertz, D. H.; Scheraga, H. A. *Macromolecules* **1978**, *11*, 9–15. (l) Manavalan, P.; Ponnuswamy, P. K. *Nature* **1978**, *275*, 673–674. (m) Janin, J. *Nature* **1979**, *277*, 491–492. (n) Rose, G. D.; Roy, S. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 4643–4647. (o) Rose, G. D.; Geselowitz, A. R.; Lesser, G. J.; Lee, R. H.; Zehfus, M. H. *Science* **1985**, *229*, 834–838. (p) Chothia, C. *Annu. Rev. Biochem.* **1984**, *53*, 537–572.
(5) Kyte, J.; Doolittle, R. F. *J. Mol. Biol.* **1982**, *157*, 105–132.
(6) (a) Sweet, R.; Eisenberg, D. *J. Mol. Biol.* **1983**, *171*, 479–488. (b) Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C.; Wilcox, W. *Faraday Symp. Chem. Soc.* **1982**, *17*, 109–120. (c) Eisenberg, D. *Annu. Rev. Biochem.* **1984**, *53*, 595–623. (d) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199–203.
(7) (a) Matsumura, M.; Becktel, W. J.; Matthews, B. W. *Nature* **1988**, *334*, 406–410. (b) Kellis, J. T.; Nyberg, K.; Fersht, A. R. *Biochemistry* **1989**, *28*, 4914–4922. (c) Lim, W. A.; Sauer, R. T. *Nature* **1989**, *339*, 31–35. (d) Sandberg, W. S.; Terwilliger, T. C. *Science* **1989**, *245*, 54–57. (e) Karpusas, M.; Baase, W. A.; Matsumura, M.; Matthews, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 8237–8241. (f) Shortle, D.; Stites, W. E.; Meeker, A. K. *Biochemistry* **1990**, *29*, 8033–8041. (g) Bowie, J. U.; Reidhaar-Olson, J. F.; Lim, W. A.; Sauer, R. T. *Science* **1990**, *247*, 1306–1310. (h) Pakula, A. A.; Sauer, R. T. *Nature* **1990**, *344*, 363–364. (i) Pace, C. N. *TIBTECH* **1990**, *8*, 93–98. (j) Creighton, T. E. *Curr. Opin. Struct. Biol.* **1991**, *1*, 5–16. (k) Moult, J.; Unger, R. *Biochemistry* **1991**, *30*, 3816–3824. (l) Green, S. M.; Meeker, A. K.; Shortle, D. *Biochemistry* **1992**, *31*, 5717–5728. (m) Inoue, M.; Yamada, H.; Yasukochi, T.; Kuroki, R.; Miki, T.; Horiuchi, T.; Imoto, T. *Biochemistry* **1992**, *31*, 5545–5553.
(8) (a) Pimentel, G. C.; McClellan, A. L. *The Hydrogen Bond*; W. H. Freeman and Company: San Francisco, CA, 1960. (b) Pauling, L. *J. Am. Chem. Soc.* **1936**, *58*, 94–98. (c) Badger, R. M.; Bauer, S. H. *J. Chem. Phys.* **1937**, *5*, 605–608. (d) Badger, R. M.; Bauer, S. H. *J. Chem. Phys.* **1937**, *5*, 839–851. (e) Badger, R. M. *J. Chem. Phys.* **1940**, *8*, 288–289. (f) Herman, R. C. *J. Chem. Phys.* **1940**, *8*, 252–258. (g) Lord, R. C.; Merrifield, R. E. *J. Chem. Phys.* **1953**, *21*, 166–167. (h) Nakamoto, K.; Margoshes, M.; Rundle, R. E. *J. Am. Chem. Soc.* **1955**, *77*, 6480–6486. (i) Becker, E. D.; Pimentel, G. C. *J. Chem. Phys.* **1956**, *25*, 224–228. (j) Pimentel, G. C.; Sederholm, C. H. *J. Chem. Phys.* **1956**, *24*, 639–641. (k) Huggins, C. M.; Pimentel, G. C. *J. Phys. Chem.* **1956**, *60*, 1615–1619. (l) Van Thiel, M.; Becker, E. D.; Pimentel, G. C. *J. Chem. Phys.* **1957**, *27*, 486. (m) Van Thiel, M.; Pimentel, G. C. *J. Chem. Phys.* **1957**, *27*, 95–99. (n) Glew, D. N.; Rath, N. S. *Can. J. Chem.* **1971**, *49*, 837–856. (o) Lindgren, J.; Tegenfeldt, J. *J. Mol. Struct.* **1974**, *20*, 335–342.
(9) Hecht, D.; Tadesse, L.; Walters, L. *J. Am. Chem. Soc.* **1992**, *114*, 4336–4339.
(10) (a) Herzberg, G. *Infrared and Raman Spectra of Polyatomic Molecules*; D. Van Nostrand Company: Princeton, NJ, 1945; Vol. 2. (b) Colthup, N. B.; Daly, L. H.; Wiberley, S. E. *Introduction to Infrared and Raman Spectroscopy*, 3rd ed.; Academic Press, Inc.: San Diego, CA, 1990.

0002-7863/93/1515-3336$04.00/0

Table I. Peak Areas and Maximum Absorbance Values (λ_{max}) for the Water Peaks Obtained from Deconvolutions of the Amino Acid Solution Difference Spectra

solute (0.1 M, pH 7)	λ_{max} (cm^{-1})	area (AU[a]/cm)	fwhm (cm^{-1})
glycine	3437	0.390	232
methionine	3412	1.285	259
proline	3402	1.147	265
alanine	3401	0.943	265
phenylalanine	3397	1.633	289
valine	3386	1.221	263
isoleucine	3384	1.684	276
leucine	3383	1.737	278

[a] AU = absorbance units.

established that both ν_3 and ν_1 shift to lower frequencies and that the ν_2 overtone shifts to higher frequencies as *either* the enthalpies *or* the number of hydrogen bonds increases.[8]

Assuming that component peaks were well represented by Gaussian distributions, all spectra were analyzed and deconvolved with an iterative curve-fitting procedure.[9] The frequency of maximum absorbance for each water peak (λ_{max}) reflects the average energy for the overlapping ν_1, ν_2, and ν_3 transitions. Net shifts of λ_{max} to lower frequencies in the high-energy 3800–3300-cm^{-1} region of the spectrum are reliable indicators of an overall increase in the degree of hydrogen bonding in water.[8] With the strategies used in this communication, the waters being examined are in the hydration shells of the indicated amino acids and represent a direct determination of solute-induced hydrogen bonding.

Thus the ranking of amino acids, in order of increasing hydration shell structure, is as follows (Table I): G < M < P,A < F < V,I,L.[11]

We have compared this ranking (Table I) to five commonly used thermodynamic hydrophobicity scales (Table II). With only two exceptions (glycine in Wolfenden's[3] hydration potentials and leucine in Kyte and Doolittle's hydrophobicity index[5]), for amino acids containing only alkyl side chains the ordering is the same using either the newly presented hydration shell *structural* data or classical *thermodynamic* criteria for hydrophobicity. This is consistent with previous proposals that as the relative hydrophobicity of a solute increases so does the extent of hydrogen bonding in its solvation shell.[9,12]

There is general disagreement (Table II) regarding the ranking of methionine and phenylalanine. The solvation shell transitions induced by sulfides and aromatic rings are currently being examined to gain greater insight into the origins of the relative hydrophobicities of Met and Phe.

(11) As the data were acquired at 2-cm^{-1} resolution, it is not yet clear whether the λ_{max} differences between P and A as well as those between V, I, and L (Table I) are significant. This issue is now being examined.

Table II. Ordering of the Amino Acids Based upon the Structure of the Hydration Shells, λ_{max}, Compared to Several Representative Preexisting Hydrophobicity Scales

evaluation criteria	relative ranking or hydrophobicity									
λ_{max} (cm^{-1})	G	M	P/A		F		V/I/L			
$\Delta G_{tr\ (ETOH\text{-}HOH)}$ [a]	G		A	M			V		L	F
Hydration Potential [b]		M	F A				V	I	L	G
$\Delta G_{Surface\ Tension}$ [c]	G		A	M			V	I	F L	
Hydrophobicity Index [d]	G		A	M	F	L	V	I		
Hydrophobicity Parameter [e]	G		A				V		M	F

[a] Nozaki and Tanford.[2] [b] Wolfenden et al.[3] [c] Bull and Breese.[14] [d] Kyte and Doolittle.[5] [e] Eisenberg.[6]

In this report, solute-induced perturbations of solvent *structure* have been measured and are observed to correlate with *thermodynamic* criteria for amino acid hydrophobicity. The strategy described is currently being extended to include other amino acids, amino acid analogues, and small peptides. The information gained from these and future studies will lead to a better understanding of the relationship between the molecular structure of hydration shells and the origins of classical thermodynamic measurements of hydrophobicity.

Acknowledgment. The authors wish to thank Jeffrey Skolnick and Cris Lewis for stimulating discussions. We also wish to acknowledge the expertise and talented craftsmanship of Chris Fish and Ward Coppersmith of the Instrumentation and Design Laboratory at The Scripps Research Institute. Lastly we extend our special appreciation to Richard Lerner, William Beers, and Peter Wright for their support in establishing a vibrational spectroscopy facility at The Scripps Research Institute.

(12) (a) Frank, H. S.; Evans, M. W. *J. Chem. Phys.* **1945**, *13*, 507. (b) Mirejovsky, D.; Arnett, E. M. *J. Am. Chem. Soc.* **1983**, *105*, 1112–1117. (c) Gill, S. J.; Dec, S. F.; Olofsson, G.; Wadso, I. *J. Phys. Chem.* **1985**, *89*, 3758–3761. (d) Muller, N. *Acc. Chem. Res.* **1990**, *23*, 23–28. (e) Goldammer, E. V.; Hertz, H. G. *J. Phys. Chem.* **1970**, *74*, 3734. (f) Némethy, G.; Scheraga, H. A. *J. Chem. Phys.* **1962**, *36*, 3401–3417. (g) Némethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1962**, *66*, 1773–1789. (h) Pottel, R.; Kaatze, U. *Ber. Bunsen-Ges. Phys. Chem.* **1969**, *73*, 437–446. (i) Dashevsky, V. G.; Sarkisov, G. N. *Mol. Phys.* **1974**, *27*, 1271–1290. (j) Geiger, A.; Rahman, A.; Stillinger, F. H. *J. Chem. Phys.* **1979**, *70*, 263. (k) Rossky, P. J.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 1913. (j) Rossky, P. J.; Karplus, M.; Rahman, A. *Biopolymers* **1979**, *18*, 825. (l) Pangali, C.; Rao, M.; Berne, B. J. *J. Chem. Phys.* **1979**, *71*, 2982–2990.

(13) Freiden, B. R. *Deconvolution, With Applications in Spectroscopy*; Academic Press: New York, 1984.

(14) Bull, H. B.; Breese, K. *Arch. Biochem. Biophys.* **1974**, *161*, 665.