Origins of Fluorine Chemical Shifts in Proteins

Sarah E. Chambers, Edmond Y. Lau, and J. T. Gerig*

Department of Chemistry
University of California
Santa Barbara, California 93106

Received December 1, 1993

Incorporation of fluorinated amino acids into proteins provides a valuable tool for study of the structural changes in these molecules that result from ligand binding or interaction with other macromolecules.[1] As a result of tertiary structure-specific NMR shielding effects, fluorine resonances from a fluoroamino acid-containing protein are typically dispersed over a 5–15-ppm range and cleanly resolved lines, assignable to specific amino acid residues of the structure, are often observed even when the protein is quite large.[2] The origins of these tertiary structure-specific chemical shift effects remain obscure. However, a full understanding of the relations between structure and shift would be of value in making assignments and interpreting observed changes in shielding in these systems.

Sources of tertiary structure-specific proton shielding effects in proteins include magnetic anisotropies associated with the peptide bond system and with the circulation of π-electrons in aromatic rings.[3-5] Calculations using parameters available in the literature show that such anisotropies typically could account for at most about 2 ppm of the range of fluorine shifts noted above. Thus, other sources of fluorine shielding effects in proteins and nucleic acids must be considered; these sources include electric fields, van der Waals interactions, and hydrogen bonding.

Building on the theoretical developments of Stephen[6] and Buckingham,[7] Pearson *et al.* have recently reported theoretical studies of electrostatic field effects on fluorine shielding in proteins.[8] Using a multipole expansion to describe the effect of an electric field and electric field gradients on the fluorine shielding tensor and empirical molecular dynamics simulations to take into account averaging by molecular motions, these authors demonstrate a reasonable linear correlation between calculated tertiary structure-specific shielding and the experimental fluorine shielding in a 5-fluorotryptophan-containing analog of galactose-binding protein (*Escherichia coli*). Unambiguous assignments of all fluorine signals from this protein have been made using site-directed mutations.[9] As a result of this correlation, Pearson *et al.* concluded that fluorine shielding in proteins "appears to be dominated by weak electrical interactions" and that van der Waals interactions are not significant.[8]

The van der Waals effect on fluorine shielding has been extensively studied,[10] and there have been indications that van der Waals interactions are a significnat component of tertiary structure-induced fluorine shielding effects in fluorine-containing alkaline phosphatase,[11] rat cellular retinol-binding protein,[12] and rhodopsin.[13] In its simplest form, the van der Waals contribution to shielding is given by

$$\sigma_{vdW} = -B\sum_i \frac{3P_iI_i}{r_i^6} \qquad (1)$$

where r_i is the distance between the fluorine nucleus under observation and an interacting atom i, I_i is the first ionization potential of that atom, P_i is the polarizability of the interacting atom, and B is a parameter that includes the ionization potential and polarizability of the fluorine.[7,14] Experimental values for B are in the range (38–380) $\times$ 10^{-6} A^3 eV^{-1} depending on the chemical nature of the fluorine under consideration.

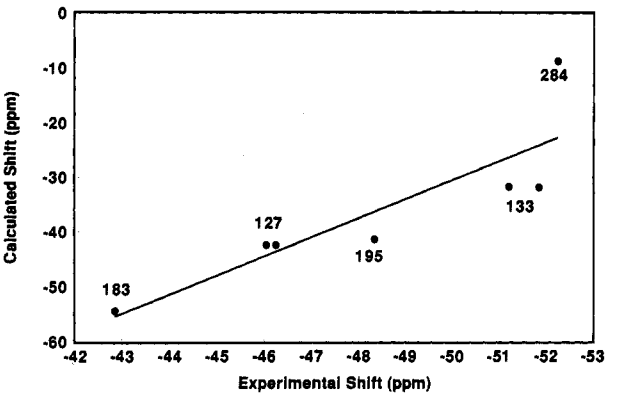

Figure 1. Comparison of experimental fluorine shifts for 5-fluorotryptophan-containing galactose-binding protein from *E. coli* with van der Waals shielding contribution calculated by means of eq 1. The experimental data are from ref 9. The experimental data are referenced to trifluoroacetate; on this scale free 5-fluorotryptophan appears at –49.5 ppm. Algebraically smaller numbers for the experimental data correspond to larger fluorine shielding parameters. Two fluorine signals each are observed for tryptophans-127 and -133, presumably because of structural isomerism due to nearby proline residues.

In continuation of our previous efforts to understand the origins of fluorine shielding in proteins,[15] we have examined the application of eq 1 to shielding in galactose-binding protein and other proteins. Coordinates for all atoms of the fluorotryptophan-containing proteins were developed from the available heavy atom coordinates (Brookhaven Protein Databank) by attaching protons and fluorines to the structure using standard procedures (QUANTA, Version 3.2, Molecular Simulations, Inc.), followed by extensive conformational energy minimization (CHARMM, Version 21.3, provided by M. Karplus, Harvard University). Standard ionization potentials[16] and polarizabilities[17,18] and the internuclear distances from the models so developed were used to compute the van der Waals shielding contribution according to eq 1, using a value for B of 67.7 $\times$ 10^{-6} A^3 eV^{-1} previously developed for aromatic fluorine.[19] Figure 1 compares computed van der Waals shielding contributions with experimental data for 5-fluorotryptophan-containing galactose-binding protein. It is clear that a strong correlation exists, one equally as good as

(1) For a recent review, see: Gerig, J. T. *Prog. Nucl. Magn. Reson. Spectrosc.*, in press.
(2) For a recent example, see: Sun, Z.-Y.; Truong, H.-T. N.; Pratt, E. A.; Sutherland, D. C.; Kulig, C. E.; Homer, R. J.; Groetsch, S. M.; Hsue, P. Y.; Ho, C. *Protein Sci.* **1993**, *2*, 1938–1947.
(3) Asakura, T.; Nakamura, E.; Asakawa, H.; Demura, M. *J. Magn. Reson.* **1991**, *93*, 355–360.
(4) Williamson, M. P.; Asakura, T.; Nakamura, E.; Demura, N. *J. Biomol. NMR* **1992**, *2*, 83–98.
(5) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647–1651.
(6) Stephen, M. J. *Mol. Phys.* **1957**, *1*, 223.
(7) Buckingham, A. D. *Can. J. Chem.* **1960**, *38*, 300–307.
(8) Pearson, J. G.; Oldfield, E.; Lee, F. S.; Warshel, A. *J. Am. Chem. Soc.* **1993**, *115*, 6851–6862.
(9) Luck, L. A.; Falke, J. J. *Biochemistry* **1991**, *30*, 4248–4256.
(10) Rummens, F. H. A. *NMR: Basic Princ. Prog.* **1975**, *10*, 1–118.
(11) Hull, W. E.; Sykes, B. D. *J. Mol. Biol.* **1975**, *98*, 121–153.
(12) Winter, N. S.; Bratt, J. M.; Banaszak, L. J. *J. Mol. Biol.* **1993**, *230*, 1247–1259.
(13) Mirzadegan, T.; Humblet, C.; Ripka, W. C.; Colmenares, L. U.; Liu, R. S. H. *Photochem. Photobiol.* **1992**, *56*, 883–893.
(14) Bothner-By, A. A. *J. Mol. Spectrosc.* **1960**, *5*, 52–61.
(15) Gregory, D. H.; Gerig, J. T. *Biopolymers* **1991**, *31*, 845–858.
(16) Levin, R. D.; Lias, S. G. *Ionization Potential and Appearance Potential Measurements (1971–1981)*; NSRDS-NBS-71; U.S. National Bureau of Standards, U.S. Government Printing Office: Washington, DC, 1982.
(17) Pasto, D. J.; Johnson, C. R. *Organic Structure Determination*; Prentice-Hall: New York, 1969; p 71.
(18) Dreisback, R. R. *Physical Properties of Organic Compounds II*; Advances in Chemistry, 23; American Chemical Society: Washington, DC, 1959; p 1–128.
(19) Boden, N.; Emsley, J. W.; Feeney, J.; Suttcliffe, L. H. *Mol. Phys.* **1964**, *8*, 467–471.

0002-7863/94/1516-3603$04.50/0

the one demonstrated by Pearson *et al.* Fluorinated tryptophan-284 is near the surface of the protein where its shielding is likely influenced by interactions with water molecules; the effects of solvent molecules were not taken into account in our calculations. If the datum for residue 284 is excluded from both calculations, a correlation coefficient of 0.95 is observed for shifts predicted either by eq 1 or by the consideration of electrostatic fields. This observation suggests, in agreement with the indications from the studies previously cited, that van der Waals interactions *can* be a significant determinant of tertiary structure-induced shielding effects in this system.

We have carried out similar calculations with 6-fluorotryptophan-containing rat cellular retinol-binding protein, for which site-specific assignments are available,[20] and we find a reasonable correlation between the computed van der Waals shielding term and experimental shifts. The approach outlined also predicts correctly the direction of the tertiary structure effect on fluorine shielding in a 4-fluorophenylalanine-containing analog of Xfin-31, a 25-member, zinc coordinating peptide that has a well-defined tertiary structure[21] (E. Lau, work in progress).

While strong correlation with experimental results is apparent, the magnitudes of the van der Waals shielding effects calculated using eq 1 are appreciably larger than those observed experimentally. This observation was also made in previous work with a fluorinated ribonuclease analog.[15] It must be borne in mind, however, that eq 1 is an approximation and that the value for B could well be better chosen. Moreover, the previous work demonstrated that including molecular motions leads to an appreciable reduction of the van der Waals shielding term. It is our opinion that the results of a more complete study, one that includes the anisotropy of the fluorine polarizability and extensive conformational averaging, while changing the magnitude of computed shift effects, will not appreciably alter the ordering of the computed residue-specific shifts indicated in Figure 1.

The work of Pearson *et al.* represents an important step in understanding the role of electrostatic fields and their time variation in determining the shielding parameter for fluorine within biological macromolecules. However, we suggest that the conclusion that electrostatic fields are the *only* factors defining the dependence of fluorine shielding on the tertiary structure of proteins is premature, and that van der Waals interactions should also be considered in this context.

Acknowledgment. This work has been supported by Grant GM-25975 from the National Institutes of Health.

(20) Li, E.; Qian, S.-J.; Winter, N. S.; D'Avignon, A.; Levin, M. S.; Gordon, J. I. *J. Biol. Chem.* **1991**, *266*, 3622–3629.

(21) Lee, M. S.; Gippert, G. P.; Soman, K. V.; Case, D. A.; Wright, P. E. *Science* **1989**, *245*, 635–637.