Carbon-13 Nuclear Magnetic Resonance of Biologically Important Aromatic Acids. I. Chemical Shifts of Benzoic Acid and Derivatives

Katherine Násfay Scott*

Contribution from the Departments of Chemistry and Radiology,
University of Florida, Gainesville, Florida 32601,
and the Veterans Administration Hospital, Gainesville, Florida 32601.
Received March 3, 1972

Abstract: The aim of the present work was to determine if empirically predicted ^{13}C chemical shifts of biologically important aromatic acids are sufficiently reliable for structural investigation and compound identification. Precise values of the ^{13}C chemical shifts were determined for phenol, anisole, benzoic acid, methyl benzoate, and 12 hydroxy- and methoxy-substituted benzoic acid derivatives. The additivity of substituent effects on the aromatic ^{13}C chemical shifts was evaluated in detail in order to establish the reliability of the chemical shifts predicted from additivity of substituent effects. The larger deviations from additivity of the chemical shifts of the 2-hydroxy-substituted benzoic acids are attributed to intramolecular hydrogen bonding between the hydroxyl proton and the carbonyl oxygen. The empirical scheme for predicting ^{13}C chemical shifts obtained in the present work is sufficiently reliable for the characterization of the benzoic series of compounds.

Carbon-13 nmr has lagged behind proton nmr as a tool for structural investigation and compound identification mainly because of the experimental difficulties in obtaining ^{13}C spectra of compounds with natural abundance of ^{13}C. However, recent advances in instrumental techniques, notably heteronuclear lock, heteronuclear broadband decoupling, and long-time computer averaging of spectrometer signal, have greatly reduced the difficulty of obtaining ^{13}C spectra of readily soluble compounds available in gram quantities. Finally, the very recent commercial availability of Fourier transform nmr spectrometers has made ^{13}C spectroscopy of these materials almost as routine a task as ^{1}H spectroscopy. More importantly, Fourier transform nmr makes it possible to obtain ^{13}C spectra of sparingly soluble materials or of biological compounds, the latter being often available only in milligram quantities. These compounds are inaccessible to ^{13}C nmr without the Fourier transform technique. Thus, ^{13}C nmr promises to be a very useful method for structural investigation and compound identification because of the large range of ^{13}C chemical shifts and the ease of interpretation of ^{13}C nmr spectra.

Structural determination by nmr depends in large measure on the ability to predict chemical shifts for a given compound. For several substituted benzenes, the aromatic carbon chemical shifts have been predicted from additivity of substituent effects.[1-12] The deviations between observed and predicted chemical shifts have generally been 2 ppm or better; however, ortho-substituted benzenes showed larger deviations.[1-4,7,8,10-12]

Thus there are successful empirical schemes available for predicting the ^{13}C shifts for a proposed substituted benzene. However, the available literature data on ^{13}C nmr of biologically important aromatic acids are very scanty, and the available data are unsystematic, having been obtained under different experimental conditions of magnetic field strength, solvent, concentration, and chemical shift reference. Such data are quite unsatisfactory for the efficient identification of unknowns. Thus we have undertaken a compilation and evaluation of precisely determined chemical shifts obtained under the same experimental conditions, in order to provide for greater efficiency and a higher confidence level in the identification of unknowns, whether of biological or synthetic origin. In the systematic study of biologically important aromatic acids, benzoic acid and its derivatives were chosen as the first group of compounds to be investigated. The additivity of substituent effects on the chemical shifts was evaluated in detail in order to establish the reliability of the predicted chemical shifts.

Experimental Section

All the compounds were commercially available and were used as received. The molecular weight and structure of each compound were confirmed by mass spectrometry using a CEC 491 spectrometer and proton nmr spectroscopy using a Varian HA 100 spectrometer.[13] Carbon-13 nmr spectra were obtained of solid samples as saturated solutions in perdeuterioacetone or in perdeuteriodimethyl sulfoxide, and of liquid materials as 2:1 (v/v) sample-perdeuterioacetone solutions. Both the solvent and the concentration selected represent a carefully considered compromise. For chemical-shift comparisons, solutions approaching infinite dilution in the same nonpolar solvent are preferred; however, accumulation of ^{13}C spectra of dilute solutions is a prohibitively time-consuming task. For most of the compounds in the present series, solubility considerations required the use of a polar solvent. Thus, in order to minimize solvent effects and make chemical-shift comparisons within the series more meaningful, a polar solvent was used for all the compounds.

Carbon-13 nmr spectra were obtained on natural abundance compounds using a Varian XL-100 spectrometer at 25.2 MHz. The deuterium of the solvent was used as lock and primary refer-

* Address correspondence to the Department of Radiology, University of Florida.
(1) P. C. Lauterbur, J. Amer. Chem. Soc., 83, 1846 (1961).
(2) P. C. Lauterbur, J. Chem. Phys., 38, 1406 (1963).
(3) P. C. Lauterbur, ibid., 38, 1415 (1963).
(4) P. C. Lauterbur, ibid., 38, 1432 (1963).
(5) G. B. Savitsky, J. Phys. Chem., 67, 2723 (1963).
(6) G. E. Maciel and J. Natterstad, J. Chem. Phys., 42, 2427 (1965).
(7) K. S. Dhami and J. B. Stothers, Can. J. Chem., 43, 479 (1965).
(8) K. S. Dhami and J. B. Stothers, ibid., 43, 498 (1965).
(9) K. S. Dhami and J. B. Stothers, ibid., 43, 510 (1965).
(10) K. S. Dhami and J. B. Stothers, ibid., 44, 2855 (1966).
(11) K. S. Dhami and J. B. Stothers, ibid., 45, 233 (1967).
(12) K. S. Dhami and J. B. Stothers, Tetrahedron Lett., 631 (1964).

(13) K. N. Scott, J. Magn. Resonance, 2, 361 (1970).

ence. In order to show enough fine structure arising from carbon–hydrogen coupling to permit spectral assignment, the spectral region of interest for each compound was obtained in two or more sections of 250- or 500-Hz width. Carbon-13 spectra were obtained by computer averaging the signal from the XL-100 spectrometer with a C-1024 computer for 4–12 hr for each of the spectral sections. For all the compounds other than the disubstituted benzoic acids, precise values of chemical shifts were then determined by broadband decoupling all the protons in the compound and running spectra of 100-Hz sweep width for each of the carbons. The sweep-offset frequency of the peaks from the deuterium lock of the solvent was counted with a frequency counter and this value served as the primary chemical shift. The decoupling conditions for each carbon were slightly varied to give the narrowest peak. The peak positions of either repeat spectra of the same solution or of spectra from different solutions of the same compound could be determined with better than 1 Hz precision. The chemical shifts relative to the deuterium lock were then converted to shifts relative to tetramethylsilane (TMS) by subtracting the TMS sweep-offset value of a 2:1 (v/v) TMS–perdeuterioacetone solution. The externally referenced chemical shifts thus obtained are believed to be accurate to 0.1 ppm. Precise values of chemical shifts of the disubstituted benzoic acids were obtained from the computer output of Fourier transform spectra obtained at the Varian Associates Applications Laboratory. The solutions of these compounds contained TMS as internal reference. The internally referenced chemical-shift values thus obtained agreed to 0.1 ppm with the externally referenced values. The addition of 5–10% by volume TMS to the perdeuterioacetone decreased the solubility of the solid compounds in the present series to such an extent that internal referencing of the non-Fourier transform spectra was not feasible.

The sample of 3,4-dimethoxybenzoic acid was insufficiently soluble in acetone, and spectra for this compound were obtained of a perdeuteriodimethyl sulfoxide solution. The ^{13}C chemical shifts of 3,4-dimethoxybenzoic acid relative to the perdeuteriodimethyl sulfoxide deuterium lock were then converted to shifts relative to TMS by subtracting the TMS sweep-offset value of a 2:1 (v/v) TMS–perdeuterioacetone solution. These chemical shifts are believed to be accurate to 0.4 ppm, since for 3-methoxy-4-hydroxybenzoic acid in perdeuteriodimethyl sulfoxide solution, the sweep-offset values of the ^{13}C peaks from the deuterium lock differ on the average by ± 8 Hz from the values obtained in perdeuterioacetone solution.

Results

The aromatic region of the spectra of phenol, anisole, benzoic acid, and methyl benzoate could be assigned without difficulty from the carbon-to-hydrogen coupling patterns and the intensities of the peaks. Thus the carbon with the substituent did not show the 160-Hz directly bonded carbon–hydrogen coupling. The para carbon peaks had only half the intensity of either the ortho or the meta carbons. The ortho and the meta carbons could be distinguished from each other by the difference in spectral pattern arising from long-range carbon–hydrogen coupling. The carbon–hydrogen couplings in benzene are known to be:[14] two bond, 1.0 Hz; three bond, 7.4 Hz; and four bond, 1.1 Hz. Three-bond carbon–hydrogen couplings have been found to be generally larger in magnitude than two-bond couplings.[15–20] Thus, in a monosubstituted benzene, the ortho carbon resonances are expected to show two three-bond carbon–hydrogen splittings of about 7 Hz, approximating a triplet pattern, whereas
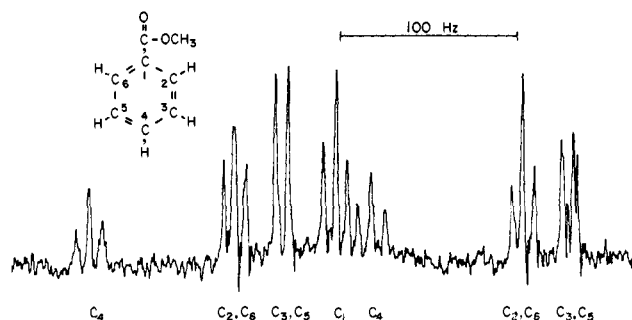

Figure 1. Aromatic region of the ^{13}C spectrum of methyl benzoate.

the meta carbon resonances show only one three-bond carbon–hydrogen coupling of about 7 Hz, resulting in a doublet pattern. Figure 1 shows that the aromatic region of the ^{13}C spectrum of methyl benzoate can easily be assigned by inspection, as outlined above, whereas the proton spectrum of methyl benzoate required complete computer simulation[13] before chemical-shift and coupling constant values could be obtained. The spectra of the mono- and disubstituted benzoic acids could similarly be assigned from the carbon–hydrogen coupling patterns and the expected substituent effects on the chemical shifts.

The ^{13}C chemical shifts are summarized in Table I. In o-hydroxybenzoic acid, methyl o-hydroxybenzoate, m-hydroxybenzoic acid, and 2,4-dihydroxybenzoic acid, sufficiently similar chemical shifts and carbon–hydrogen splitting patterns were observed for two carbons each, so that alternate assignments are possible, as indicated in Table I. The numbering of the carbons in Table I and throughout this paper is as in Figure 1. Carbon-13 shifts for phenol, anisole, benzoic acid, and methyl benzoate have been reported previously in the literature,[6,10,11,21] but under different experimental conditions. The literature values agree with the values listed in Table I to ± 0.6 ppm, except for the C-1 and C-4 shifts of phenol, which deviate by 1.8 and 1.2 ppm, respectively. These differences are undoubtedly due to hydrogen bonding differences caused by changes in solvent and concentration.

Discussion

The hydroxy and methoxy substituents in phenol and anisole show a substantial alternating shielding and deshielding effect on the even- and the odd-numbered carbons, respectively. This alternating chemical shift substituent effect mirrors the effect of these substituents on the π-electron charge densities of the carbon nuclei. There is a general correlation between the ^{13}C chemical shifts observed in this work and the calculated π-electron charge densities reported in the literature.[22–28] Comparison of the ^{13}C chemical shifts of benzoic acid

(14) F. J. Weigert and J. D. Roberts, J. Amer. Chem. Soc., 89, 2967 (1967).
(15) G. J. Karabatsos, ibid., 83, 1230 (1961).
(16) G. J. Karabatsos, J. D. Graham, and F. M. Vane, ibid., 84, 37 (1962).
(17) F. J. Weigert and J. D. Roberts, ibid., 90, 3543 (1968).
(18) K. Tukahashi, T. Sone, and K. Fujieda, J. Phys. Chem., 74, 2765 (1970).
(19) A. R. Tarpley, Jr., and J. H. Goldstein, J. Mol. Spectrosc., 37, 432 (1971).
(20) A. R. Tarpley, Jr., and J. H. Goldstein, J. Amer. Chem. Soc., 93, 3573 (1971).

(21) A. M. Ihrig and G. L. Marshall, ibid., 94, 1756 (1972).
(22) K. Nishimoto and R. Fujishiro, Bull. Chem. Soc. Jap., 31, 1036 (1958).
(23) A. Julg and M. Bonnet, Theor. Chim. Acta, 1, 6 (1962).
(24) S. Forsen and T. Alm, Acta Chem. Scand., 19, 2027 (1965).
(25) D. T. Clark and J. W. Emsley, Mol. Phys., 12, 365 (1967).
(26) G. G. Dvoryantseva, V. P. Lezina, V. F. Bystrov, T. N. Ul'yanova, G. P. Syrova, and Y. N. Sheinker, Izv. Akad. Nauk SSSR, Ser. Khim., 994 (1968).
(27) M. M. Krayshkin, S. G. Gagarin, and V. V. Sevost'yanova, ibid., 33 (1970).
(28) P. Lazzeretti and F. Taddei, Tetrahedron Lett., 805 (1970).

Table I. Carbon-13 Chemical Shifts[a]

Sample[b]	C=O	C-1	C-2	C-3	C-4	C-5	C-6
Benzoic acid	168.5	130.8	130.2	129.1	133.7	129.1	130.2
Methylbenzoate[c]	167.0	131.0	130.0	129.0	133.4	129.0	130.0
o-Hydroxybenzoic acid	172.3	113.0	162.6	117.8[d]	136.2	119.5[d]	131.0
Methyl o-hydroxybenzoate[e]	171.1	113.4	162.4	118.4[f]	136.3	120.0[f]	130.6
m-Hydroxybenzoic acid	168.6	132.2	116.9	157.9	120.8[g]	130.2	121.6[g]
p-Hydroxybenzoic acid	169.0	121.9	132.7	115.8	162.5	115.8	132.7
2,3-Dihydroxybenzoic acid	172.9	113.2	151.1	146.6	121.4	119.6	121.4
2,4-Dihydroxybenzoic acid	172.6	105.2	164.9[h]	103.3	164.8[h]	108.7	133.0
2,5-Dihydroxybenzoic acid	172.2	112.7	156.2	118.7	124.9	149.9	115.6
2,6-Dihydroxybenzoic acid	172.1	101.1	161.6	108.4	136.7	108.4	161.6
3,4-Dihydroxybenzoic acid	169.4	122.4	117.5	145.3	150.8	115.7	123.9
3,5-Dihydroxybenzoic acid	168.9	132.7	109.0	159.0	108.3	159.0	109.0
3-Methoxy-4-hydroxybenzoic acid[i]	167.9	122.8	113.5	148.0	152.0	115.5	124.9
3,4-Dimethoxybenzoic acid[j]	167.6	123.3	112.5	148.8	153.1	111.2	123.7
Phenol		157.6	116.1	130.1	120.4	130.1	116.1
Anisole[k]		160.5	114.6	130.1	121.2	130.1	114.6

[a] In ppm downfield from TMS; numbering of carbons as in Figure 1. [b] Perdeuterioacetone solution. [c] Methyl, 52.3 ppm. [d] Or C-3 = 119.5 and C-5 = 117.8 ppm. [e] Methyl, 52.4 ppm. [f] Or C-3 = 120.0 and C-5 = 118.4 ppm. [g] Or C-4 = 121.6 and C-6 = 120.8 ppm. [h] Or C-2 = 164.8 and C-4 = 164.9. [i] Methyl, 56.3 ppm. [j] Perdeuteriodimethyl sulfoxide solution; methyl, 55.8 and 55.9 ppm. [k] Methyl, 55.3 ppm.

and methyl benzoate with those of phenol and anisole shows a much smaller substituent effect for the carboxy and carbomethoxy groups than for the hydroxy and methoxy groups. For benzoic acid and methyl benzoate, no marked alternating shielding and deshielding of alternating carbons is observed, and no correlation between chemical shifts and calculated π-electron charge densities[26,27,29] could be found. Undoubtedly the larger neighbor anisotropy effect of the carboxy and carbomethoxy groups masks their mesomeric effect.

Shielding parameters of the aromatic carbons by the CO_2H group were obtained by subtracting from the various carbon chemical shifts of benzoic acid the chemical shift of benzene, which was 128.9 ppm under the present conditions. These shielding parameters are given in Table II, which also contains the shielding parameters of the aromatic carbons by the CO_2CH_3, OH, and the OCH_3 groups, which were obtained from methyl benzoate, phenol, and anisole, respectively. Positive shielding parameters indicate deshielding by the substituent. It should be emphasized that these shielding parameters are for concentrated acetone solutions and are expected to show some variation with concentration and solvent.

Assuming additivity of substituent effects, the chemical shift of carbon "a" in a polysubstituted benzene can be predicted using the relationship

$$\delta_a = \delta^0 + \sum_i S_{x,o,m,\,or\,p;\,X_i} \qquad (1)$$

where δ_a is the chemical shift of carbon "a," δ^0 is the chemical shift of benzene, $S_{x,o,m,\,or\,p}$ is the shielding parameter of substituent X_i on the substituted carbon or on the carbon ortho, meta, or para to it, and where the summation is carried over all of the substituents on the ring.

The aromatic carbon chemical shifts in the mono- and disubstituted benzoic compounds were calculated by using eq 1 and the appropriate shielding parameters in Table II. For the monohydroxybenzoic compounds, the average of the absolute deviations between observed and predicted chemical shifts was 1.4 ppm or 3% of the 50-ppm range observed for the monosubstituted benzoic

(29) R. M. Aminova, I. D. Morozova, and T. A. Zyablikova, *Izv. Akad. Nauk SSSR, Ser. Khim.*, 1709 (1969).

Table II. Aromatic Carbon-13 Shielding Parameters

Substituent	Shielding parameters[a]			
	S_x	S_o	S_m	S_p
CO_2H	+1.9	+1.3	+0.2	+4.8
CO_2CH_3	+2.1	+1.1	+0.1	+4.5
OH	+28.7	-12.8	+1.2	-8.5
OCH_3	+31.6	-14.3	+1.2	-7.7

[a] In ppm from benzene. Positive value indicates deshielding by substituent. Designation of parameters: S_x is carbon attached to the substituent; S_o, S_m, and S_p are carbons ortho, meta, and para, respectively, to the substituent.

shifts. None of the deviations exceeded 5.0 ppm or 10% of the range. For disubstituted benzoic acids the average of the absolute deviations was 2.0 ppm or 3% of the 64-ppm range observed for the shifts of disubstituted benzoic acids. None of the deviations exceeded 6.5 ppm or 10% of the range. Empirical schemes that reproduce chemical shifts to 5% of the observed range are usually quite satisfactory for the purpose of predicting chemical shifts for compound identification. Thus, most of the aromatic carbon shifts of the mono- and disubstituted benzoic compounds could be predicted satisfactorily by eq 1. However, there were individual large deviations of up to 10% of the observed chemical shift range.

In an attempt to minimize these larger deviations, the aromatic shifts were predicted by a further empirical scheme, which had given very satisfactory results for the proton chemical shifts of some polysubstituted benzenes,[30] and the proton chemical shifts of the benzoic series.[13] Reed[30] could predict the ring proton chemical shifts of a large number of variously methyl-, chloro-, and hydroxy-polysubstituted benzenes to within 0.02 ppm by employing an additivity scheme, which in addition to the ortho, meta, and para shielding parameters of eq 1 uses additional "crowding" constants for two or three substituents mutually ortho to each other. These additional constants are obtained from the appropriately di- or trisubstituted benzenes and affect the shielding of only the proton ortho to the substituents.

(30) J. J. R. Reed, *Anal. Chem.*, **39**, 1586 (1967).

In ortho-disubstituted benzenes the crowding of substituents is expected to affect most of the chemical shifts of the carbons bearing the substituents. Indeed, for both the mono- and disubstituted benzoic series, the largest deviations from additivity occur for the C-1 and C-2 carbons of the 2-hydroxy-substituted compounds. The observed shifts for C-1 are always to higher field and for C-2 are to lower field than the predicted ones. Thus, for the 2-hydroxy-substituted benzoic compounds, the averaged crowding parameter was −5.1 ppm for C-1 and +4.1 ppm for C-2. The inclusion of these crowding parameters greatly reduces the deviations between observed and predicted chemical shifts. For the monosubstituted benzoic series the averaged absolute deviation is 0.7 ppm and the maximum deviation is 2.4 ppm. For the disubstituted benzoic series the average deviation is 1.4 ppm and the maximum deviation is 4.8 ppm. Figure 2 shows a plot of predicted vs. observed ^{13}C chemical shifts of the aromatic carbons; the predicted values shown were obtained with the inclusion of the crowding parameters.

Larger deviations from additivity for ortho-disubstituted benzene carbon chemical shifts have been observed previously[4-6,9-12,15] and have been attributed to steric interactions of the substituents ortho to each other. For the o-hydroxy-substituted benzoic acids and for methyl o-hydroxybenzoate, in addition to the general steric interaction, there is a possibility of intramolecular hydrogen bonding interaction between the hydroxyl proton and the carbonyl oxygen. Thus, the larger deviations from additivity of the C-1 and C-2 carbons in the o-hydroxybenzoic compounds could arise either from steric interactions or from intramolecular hydrogen bonding, or from both.

Examination of the carbonyl carbon chemical shifts shows that in benzoic acid, m- and p-hydroxybenzoic acid, and in 3,4- and 3,5-dihydroxybenzoic acid the carbonyl chemical shifts fall in the range from 168.5 to 169.4 ppm. Whereas in o-hydroxybenzoic acid and the dihydroxybenzoic acids with substituents in the 2 position, the carbonyl chemical shifts are in the range from 172.1 to 172.9 ppm. Thus, there is an average deshielding effect of 3.5 ppm on the carbonyl chemical shift by an o-hydroxy substituent. A downfield shift of 3-7 ppm of the carbonyl carbon chemical shift has been observed previously[7,31,32] for several o-hydroxy- or o-amine-substituted benzyl compounds and has been attributed to intramolecular hydrogen bonding between the carbonyl carbon and the hydroxyl or amine proton. The carbonyl chemical shifts are not further affected by the second o-hydroxy grouping in either 2,6-dihydroxy-substituted acetophenones[7] or in 2,6-dihydroxybenzoic acid. This fact is consistent with the intramolecular hydrogen bonding hypothesis. If the deshielding of the carbonyl carbons were caused mainly by steric inhibition of conjugation, then a further deshielding would be expected on the introduction of a second o-hydroxy substituent. For o-alkyl-substituted acetophenones, for which the deshielding is attributed to steric inhibition of conjugation, a very pronounced substituent effect is observed on the introduction of a second ortho substituent.
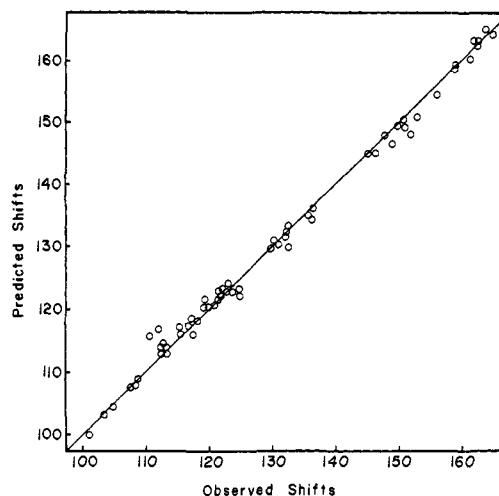
(31) P. C. Lauterbur, Ann. N. Y. Acad. Sci., 70, 841 (1958).
(32) G. E. Maciel and G. B. Savitsky, J. Phys. Chem., 68, 437 (1964).

Figure 2. Plot of predicted vs. observed aromatic ^{13}C chemical shifts in ppm from TMS. The curve drawn represents perfect (theoretical) correlation.

stituent.[7] Since the deshielding of the carbonyl carbons in o-hydroxy-substituted benzoic acids can be attributed to intramolecular hydrogen bonding, it seems likely that the greater deviations from additivity of the C-1 and C-2 carbons in these compounds arise also mainly from intramolecular hydrogen bonding. This explanation is supported by the fact that the C-1 carbon in 2,6-dihydroxybenzoic acid shows no increased deviation from additivity over the C-1 carbon of o-hydroxybenzoic acid.

Conclusions

Because of recent advances in instrumentation, ^{13}C nmr is becoming an increasingly useful tool in structural investigation and compound identification. However, although the number of compounds for which ^{13}C nmr spectra have been cataloged is increasing, it is still fortuitous if a given compound is one for which a ^{13}C spectrum has been obtained previously. For this reason, empirical schemes are most important in the prediction of structure from the ^{13}C spectrum of the compound. The additivity scheme of substituent effects on the ^{13}C chemical shifts outlined in the present paper could have unequivocally identified all but one of the mono- and disubstituted benzoic acid derivatives, had these been compounds of unknown structure. The only compound in the present series which the additivity scheme cannot identify unambiguously is 3-methoxy-4-hydroxybenzoic acid, for which 3-hydroxy-4-methoxybenzoic acid is an equally likely formula from the observed ^{13}C spectrum.

It is significant that the substituent effects on the ^{13}C aromatic chemical shifts were found to be sufficiently additive, so that the ^{13}C chemical shifts varied in a systematic and predictable manner in the benzoic acid derivatives. For all previously studied substituents with additive effects,[1-12] the substituent effects were much larger than those of either the carboxy or the carbomethoxy group. For aromatic systems of biological importance, the substituents encountered show smaller substituent effects, so that deviations from additivity could have sufficiently masked the smaller differences in chemical shift and could have

made the ^{13}C spectra unreliable for the prediction of structure. Often biologically quite different compounds are sufficiently similar chemically and spectroscopically that it is difficult to distinguish them. For example, if given the mass spectrum of a tri- or higher-substituted benzene, one cannot select among the several positional isomers. Nmr, however, can identify such positional isomers. In particular, ^{13}C nmr, because of the larger range of chemical shifts observed and the ease of interpretation of spectra, is ideally suited for the characterization of biological unknowns. If predictability of substituent effects on ^{13}C chemical shifts can be shown to be generally true for other aromatic systems of biological importance,

this fact would be of great significance in the characterization of biological unknowns.

Tertiary Structure in Carboxypeptidase

I. D. Kuntz

Contribution from the Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, San Francisco, California 94122. Received March 6, 1972

Abstract: Intramolecular hydrogen bonding and hydrophobic interactions in carboxypeptidase are analyzed. The hydrogen bonding in the helices and the β sheet is well described by the classical models. The ends of each helix and the ends and edges of the β sheet are not hydrogen bonded within the protein molecule. Hairpin turns do not fit the bonding pattern suggested in the literature. Few internal hydrogen bonds are made in the turns, so that the turns appear to be weak links in the peptide chain. For the entire protein about 42% of the potential donor or acceptor sites are not occupied by intramolecular bonds. It is likely that water molecules are bound at most of these sites. The eight helices are held to each other and to the β sheet by a maximum of 17 hydrogen bonds, less than two bonds per helix. This level of interconnection appears insufficient to be the dominant term in determining protein tertiary structure. To explore the role of hydrophobic bonding, we define all side-chain carbon atoms as "hydrophobic." We find large regions (of the order of 1000 Å^3) of carboxypeptidase that contain *ca.* 97% hydrophobic atoms. These regions are roughly channel shaped and hold two-thirds of the hydrophobic atoms of the molecule. The channels are located between the helices and the β sheet, sometimes extending to the molecular surface. The density of the hydrophobic regions is surprisingly low being close to that of liquid benzene.

The importance of intramolecular hydrogen bonding and hydrophobic bonding in determining the tertiary structure of globular proteins is still a matter of some controversy. The difficulty with a facile assessment of the role of hydrogen bonding is the small free energy of formation of amide hydrogen bonds in aqueous solutions.[1] Cooperative secondary features such as α helices or β pleated sheet structures clearly involve some substantial contribution of hydrogen bonding to the overall potential energy. What is not clear is the importance of hydrogen bonding in connecting these secondary elements together. Kauzmann's discussion of hydrophobic interactions[2] has led to a very simple model for the structure of globular proteins which does not require direct reference to hydrogen bonding. In this model the protein is represented as a hydrophilic shell enclosing a hydrophobic core. Qualitative assessments of the known protein structures have supported this picture in that relatively few polar residues are found in the protein interior.[3] Recently investigations have

raised some interesting difficulties. Klotz[3] and Lee and Richards[4] have pointed out that a number of hydrophobic residues are on the protein surface. The latter authors, particularly, question whether hydrophobic interactions have much importance in determining protein structure. Klapper,[5] in a different approach, has made use of scaled particle theory to estimate the atom density, free volume, and compressibility of the protein interior. He finds that the average density is so high as to make a "wax-like" rather than an "oil-like" description more useful. In sum, both of the widely discussed approaches to simple models for protein structure have yet to gain complete acceptance. In this paper we will make use of the availability of atomic coordinates for globular proteins to investigate further intramolecular hydrogen bonding and hydrophobic interactions. Our initial questions are: (1) do hydrogen bonds provide direct links between secondary structural elements; (2) can hydrophobic regions be quantitatively identified; and (3) if so, what are the

(1) I. M. Klotz and J. S. Franzen, *J. Amer. Chem. Soc.*, **84**, 3461 (1962).
(2) W. Kauzmann, *Advan. Protein Chem.*, **14**, 1 (1959).

(3) I. M. Klotz, *Arch. Biochem. Biophys.*, **138**, 704 (1970).
(4) B. Lee and F. M. Richards, *J. Mol. Biol.*, **55**, 379 (1971).
(5) M. H. Klapper, *Biochim. Biophys. Acta*, **229**, 557 (1971).