Hydrogen Bonding of Carboxyl Groups in Solid-State Amino Acids and Peptides: Comparison of Carbon Chemical Shielding, Infrared Frequencies, and Structures

Zhengtian Gu, Raúl Zambrano, and Ann McDermott*

Contribution from the Department of Chemistry, Columbia University, New York, New York 10027

Received January 7, 1994®

Abstract: We previously reported that the protonation state of the carboxyl group of amino acids and peptides in the solid state can easily be determined by the carbon chemical shielding tensors but not by the isotropic shifts. In this report the substantial variation in the σ_{22} element for both protonated and deprotonated forms is shown to be a result of hydrogen bonding. We have correlated this tensor element with established measures of hydrogen bonding, namely the IR stretching frequencies of the carbonyl and the asymmetric stretching frequency for the protonated and deprotonated carboxy groups, respectively. We also observed a strong correlation between the σ_{22} values and previously reported O···H hydrogen bonding distances from the carbonyl of protonated acids to the nearest proton donor. In the database, we found a fixed geometry for the protonated acids and a variable and complicated geometry for hydrogen bond interaction in deprotonated carboxylate. Correspondingly, the correlation betweeen NMR, IR, and diffraction data is more convincing for the protonated acids.

Introduction

The enormous importance of hydrogen bonding in biological systems has inspired the development of several diffraction and spectroscopic probes for hydrogen bonding strength. Among them, high-resolution solid-state NMR spectroscopy provides useful information about the details of hydrogen bonding and is a general method for observing the motional dynamics of hydrogen bonded groups.[1-3]

We recently reported a study of the ^{13}C chemical shift anisotropy (CSA) of carboxyl groups in amino acids and demonstrated a clear distinction between protonated and deprotonated forms, but we were unable at that point to make a clear conclusion about the signatures for hydrogen bonding in the carbon spectra.[4] The carboxy group is of particular interest because of the strong hydrogen bonds it forms, because of its important roles in modulating the reactivity of other functional groups in proteins, and because of the fact that it is often found in the hydrogen bonded networks with substantial proton mobility.[5] Many studies of the chemical shift anisotropy have been done, and the molecular directions for the chemical shift tensor elements were already known before our studies,[6] but the trends in the tensor elements were not clear. In brief, we reported that the most shielded element, σ_{33}, is invariant among carboxyl groups at 109 ppm vs TMS; this element corresponds to the measurement in which the static field is orthogonal to the plane containing the CO_2 group. The least shielded element, σ_{11}, nicely distinguishes protonated and deprotonated forms, with a value of 238–245 ppm for deprotonated and 251–262 ppm for protonated forms. This element is found along the C–C bond for deprotonated groups and perpendicular to the carbonyl for protonated groups. The remaining element, σ_{22}, varied widely from 168–200 ppm for deprotonated and 138–173 ppm for protonated forms.

In this paper we will address the question of the broad variability in the σ_{22} element. Because the electronic structures of carboxyl groups are greatly affected by hydrogen bonding, we thought it would be reasonable that the ^{13}C anisotropic chemical shift tensors of the carboxyl carbon could be sensitive to hydrogen bonding. For example, systematic trends in CSA are expected if the highest occupied molecular orbitals are perturbed by hydrogen bonding; the so-called paramagnetic, deshielding effects are very sensitive to the energy and symmetry of the highest occupied orbitals when low-lying unoccupied orbitals exist and the magnetic field can mix these orbitals.[7]

In order to establish whether or not hydrogen bonding is the origin of the observed variation in the middle tensor element, we decided to correlate our measurements with other well-established probes of hydrogen bonding. High-quality neutron diffraction can give unambiguous information about the hydrogen bonding strength and geometry, when available. Excellent correlations have been reported previously, relating hydrogen bonding distances determined crystallographically and proton chemical shielding of carboxylic acids,[8] or carbon anisotropic shielding for amides in the solid state.[9]

It has long also been appreciated that the IR frequencies for hydrogen bonded groups correlate strongly with the thermodynamic and structural measurements of hydrogen bonding.[10,11] In the carboxyl group case, the symmetric and asymmetric stretching frequencies of the carboxyl group are extremely sensitive to the hydrogen bonding environment. Since it is difficult to obtain accurate diffraction information about hydrogen bonding for our

* Abstract published in *Advance ACS Abstracts*, June 1, 1994.

(1) Williams, J.; McDermott A. *J. Phys. Chem.* **1993**, *97*, 12393–12398.
(2) Zheng, L.; Fishbein, K. W.; Griffin, R. G.; Herzfeld, J. *J. Am. Chem. Soc.* **1993**, *115*, 6254–6261.
(3) (a) Harris, R. K.; Jackson, P.; Merwin, L. H.; Say, B. J.; Hagele, G. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84*, 3649–3672. (b) McDermott, A.; Rydenour, C. In the *Encyclopaedia of Nuclear Magnetic Resonance*; Grant, D. M., Harris, R. K., Eds.; New York, 1994; in press.
(4) Gu, Z.; McDermott, A. *J. Am. Chem. Soc.* **1993**, *115*, 4282–4285.
(5) Jeffry, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: New York, 1991.
(6) (a) Duncan, T. M. *A Compilation of Chemical Shift Anisotropies*; Farragut Press: Madison, WI, 1990. (b) Veeman, W. *Prog. NMR Spectrosc.* **1984**, *16*, 193–235.
(7) Ebraheem, K. A.; Webb, G. A. *Prog. Nucl. Magn. Reson. Spectrosc.* **1977**, *11*, 149.
(8) (a) Berglund, B.; Linddgren, J.; Tegenfeldt, J. *J. Mol. Struct.* **1978**, *43*, 179. (b) Berglund, B.; Vaughan, R. W. *J. Chem. Phys.* **1980**, *73*, 2037. (c) Jeffrey, G. A.; Yeon, Y. *Acta Crystallogr., Sect. B* **1986**, *42*, 410.
(9) (a) Ando, S.; Ando, I.; Shoji, A.; Ozaki, T. *J. Am. Chem. Soc.* **1988**, *110*, 3380. (b) Asakawa, et al. *J. Am. Chem. Soc.* **1992**, *114*, 3261–3265.
(10) (a) Badger, R. M.; Bauer, S. *J. Chem. Phys.* **1937**, *5*, 839. (b) Badger, R. M. *J. Chem. Phys.* **1940**, *8*, 288.
(11) (a) Nakamoto, K.; Margoshes, M.; Rundle, R. E. *J. Am. Chem. Soc.* **1955**, *77*, 6480. (b) Pimentel G. C.; Sederholm, C. H. *J. Chem. Phys.* **1956**, *24*, 639.

0002-7863/94/1516-6368$04.50/0 © 1994 American Chemical Society

entire database of amino acids and peptides, we decided to first study the correlation between the carbon tensor elements and the vibrational frequency. Some similar studies of hydrogen bonded amides in solution previously indicated a correlation between carbon chemical shifts and IR frequencies both in simple small molecules and in proteins.[12,13] We report in this paper some correlations of carbon chemical shielding tensors with vibrational and diffraction data.

Materials and Methods

The amino acids were purchased from Sigma Chemical Co. and were recrystallized from water. The HCl salts were crystallized from 20–25% HCl solution. The dipeptide crystals were purchased from both Sigma Chemical Co. and BACHEM Bioscience Inc. (Philadelphia, PA) and used without further recrystallization. Those deuterated amino acids used in the infrared spectra assignment were prepared from recrystallization from D_2O twice.

All the SSNMR spectra were taken on a Chemagnetics CMX 400 operating at 99.71 MHz for ^{13}C and 396.5 MHz for ^{1}H. Cross polarization with a background suppressed pulse sequence was used, with contact times of 1–3 ms, a 4.5 μs 90° pulse for protons, and acquisition times of 25 ms. Generally, 600–2400 transients with a 3–15 s recycle delay were accumulated. All spectra were recorded at room temperature and referenced using the adamantane peak at 38.6 ppm (run separately). The principal values of the chemical shift tensors were extracted by computer simulation of the spectrum.[14] The assignment of error bars was based on repetition of the experiment at different spinning speeds. The data presented could be distorted due to the cross polarization. However in one case, DL-aspartic acid, we have compared the static Bloch decay with CPMAS and obtain good agreement within 3 ppm experimental error.

The infrared spectra were obtained on a Perkin-Elmer 1600 series Fourier transform infrared spectrophotometer at 20 °C without further frequency correction. For each spectrum, 64 scans were collected with 2-cm^{-1} resolution and 1-cm^{-1} interval from the 4000–600 cm^{-1} region in a KBr pellet. The assignment of frequencies for the protonated carboxyl group was based on the assumption that C=O and C—O bands are found at 1700 and 1200 cm^{-1}, respectively, while the peptide amide C=O band is near 1650 cm^{-1}.[15] The assignment of the CO_2^- asymmetric stretching frequency is based on comparison with the IR spectrum of the corresponding compound recrystallized from D_2O; under these conditions the NH_3^+ band will shift dramatically but the CO_2^- band is essentially invariant.

Results and Discussion

The present database differs from that used in our previous report in that we are including many peptides and hydrates of peptides in this study. This expansion of our database was intended to include representation from more hydrogen bonding motifs that are typical for proteins, and it allowed us to control the effects of structural features more systematically. However for these compounds, unlike the simple amino acids, a strategy for assigning the various carbonyl resonances was necessary. A typical low-speed ^{13}C CPMAS spectrum for the Gly-Asp monohydrate dipeptide is shown in the Figure 1, with centerbands marked with different groups. Assignment of the carbonyl peaks was based on their typically upfield chemical isotropic shifts in the range 158–173 ppm as compared with carboxy shifts which range from 172 to 183 ppm. In addition, the σ_{33} for amides is generally below 100 ppm, while for acid groups it is in the range 105–112 ppm unless significant motion of the carboxy group or

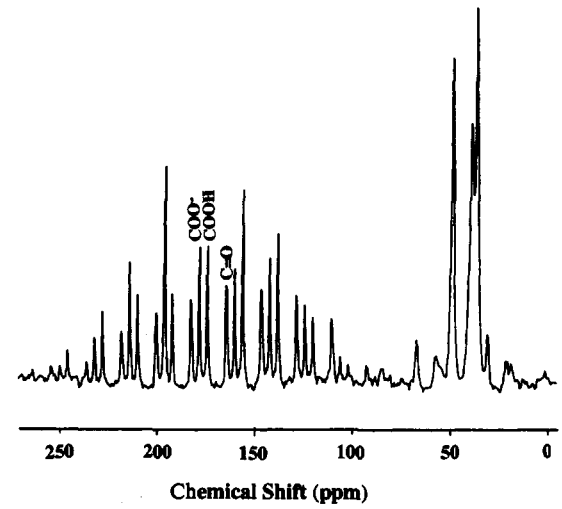

Figure 1, Low spinning speed spectrum of the Gly-Asp monohydrate dipeptide. The spectrum was collected at 1800 Hz. The assignments of the center peaks are indicated, while the arrows above the spectrum indicate the spacing of the sidebands.

carboxy proton occurs, in which case it would be at higher values yet. Distinction between the deprotonated and protonated carboxy groups is based on σ_{11}, which is above 250 ppm for protonated and below 250 ppm for deprotonated carboxy groups.[4] We subsequently confirmed the assignments for several of the compounds using the heteronuclear correlation spectrum; these results will be published separately.[16] We simulated spectra to obtain the principal elements of the shielding tensors which are given in Tables 1 and 2. Parameters and references for carboxyl group stretching frequency and crystallographic structures are also listed.[17–46]

Many of the carboxyl groups we studied do not have high-quality neutron diffraction or X-ray crystal structures. Therefore

(12) Eaton, G.; Symons M. C. R. *J. Chem. Soc., Faraday Trans. I* **1988**, *84*, 3459.
(13) (a) Park, K.; Guo, K.; Adebodun, F.; Chiu, M.; Silgar, S.; Oldfield, E. *Biochemistry* **1991**, *30*, 2333. (b) Augspurger, J.; Dykstra, C.; Oldfield, E. *J. Am. Chem. Soc.* **1991**, *113*, 2447.
(14) (a) Hertzfeld, J.; Berger, E. *J. Chem. Phys.* **1980**, *73*, 6021–6030. (b) deGroot, H.; Harbison, G.; Herzfeld, J.; Griffin, R. *Biochemistry* **1989**, *28*, 3346–3353.
(15) (a) Silverstein, R.; Bassler, G.; Morrill, T. *Spectrometric Identification of Organic Compounds*, 4th ed.; John Wiley & Sons: New York, 1981. (b) Tsuboi, M.; Takenishi, T.; Nakamura, A. *Spectrochimica Acta* **1963**, *19*, 271–284. (c) Salimov, M.; Pachelin, V.; Kerimbekov, A. *Russ. J. Phys. Chem.* **1963**, *37*, 1231–1233. (d) Tomita, N; Shimanouchi, T. *Spectrochim. Acta* **1962**, *18*, 895–905.
(16) Ridenour C. F.; McDermott, A.; Gu, Z.; Iwashita, T. Manuscript in preparation.
(17) Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. *J. Am. Chem. Soc.* **1972**, *94*, 2657–2660.
(18) Eggleston, D. S.; Hodgson, D. J. *Int. J. Peptide Protein Res.* **1983**, *21*, 288–295.
(19) Lehmann, M. S.; Verbist, J. J.; Hamilton, W. C.; Koetzle, T. F. *J. Chem. Soc., Perkin Trans. 2* **1973**, 133–137.
(20) (a) Verbist, J. J.; Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. *Acta Crystallogr. B* **1972**, *28*, 3006–3013. (b) Ramandham, M.; Sikka, S. K.; Chidambaram, R. *Acta Crystallogr. B* **1972**, *28*, 3000–3005.
(21) Eggleston, D. S.; Hodgson, D. J. *Int. J. Peptide Protein Res.* **1985**, *25*, 242–253.
(22) Rao, S. T. *Acta Crystallogr. B* **1973**, *29*, 1718–1720.
(23) Derissen, J. L.; Endeman, H. J.; Peerdeman, A. F. *Acta Crystallogr. B* **1968**, *24*, 1349–1354.
(24) Eggleston, D. S.; Valente, E. J.; Hodgson, D. J. *Acta Crystallogr. B* **1981**, *37*, 1428–1430.
(25) (a) Kerr, K. A.; Ashmore, J. P.; Koetzle, T. F. *Acta Crystallogr. B* **1975**, *31*, 2022–2026. (b) Harding, M. M.; Long, H. A. *Acta Crystallogr. B* **1968**, *24*, 1096–1102.
(26) Oughton, B. M.; Harrison, P. M. *Acta Crystallogr.* **1959**, *12*, 396–404.
(27) Koetzle, T. F.; Frey, M. N.; Lehmann, M. S.; Hamilton, W. C. *Acta Crystallogr. B* **1973**, *29*, 2571–2575.
(28) Lehmann, M. S.; Nunes, A. C. *Acta Crystallogr. B* **1980**, *36*, 1621–1625.
(29) (a) Jonsson, R. G.; Kvick, A. *Acta Crystallogr. B* **1972**, *28*, 1827–1833. (b) Power, L.F.; Turner, K. E.; Moore, F. H. *Acta Crystallogr. B* **1976**, *32*, 11–16.
(30) Eggleston, D. S.; Hodgson, D. J. *Int. J. Peptide Protein Res.* **1982**, *19*, 206–211.
(31) Maddfn, J. J.; McGandy, E. L.; Seeman, N. C. *Acta Crystallogr. B* **1972**, *28*, 2377–2382.
(32) Torii, K.; Iitaka, Y. *Acta Crystallogr. B* **1971**, *27*, 2237–2246.
(33) (a) Coll, M.; Solans, X.; Font-Altaba, M.; Subirana, J. A. *Acta Crystallogr. C* **1986**, *42*, 599–601. (b) Harging, M. M.; Howieson, R. M. *Acta Crystallogr. B* **1976**, *32*, 633–634.
(34) Bhat, T. N.; Vijayan, M. *Acta Crystallogr. B* **1976**, *32*, 891–895.
(35) Torii, K.; Iitaka, Y. *Acta Crystallogr. B* **1973**, *29*, 2799–2807.
(36) (a) Frey, M. N.; Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. *Acta Crystallogr. B* **1973**, *29*, 876–884. (b) Kistenmacher, T. J.; Rand, G. A.; Marsh, R. E. *Acta Crystallogr. B* **1974**, *30*, 2573–2578.

Table 1. Tensor Elements of Deprotonated Carboxyl Groups in Amino Acids, Dipeptides

no.	molecule	σ_{11}, ppm	σ_{22}, ppm	σ_{33}, ppm	σ_{iso}, ppm	d_{C-O}, Å	$d_{C=O}$, Å	$\nu_{C=O}$, cm^{-1}	ν_{C-O}, cm^{-1}	ref
1	alanine	240	184	110	178	1.258	1.242	1594	1413	17
2	alanine (DL)	241	185	112	179			1594	1409	
3	Ala-Ala	234	188	113	178			1556	1397	
4	Ala(*dl*)-Ala(*dl*)	237	187	110	178			1562	1404	
5	Ala-Asp	243	184	107	178	1.258	1.234	1576	1435	18
6	arginine	233	180	107	173	1.254	1.239			19
7	asparagine	243	170	108	174	1.261	1.243			20
8	Arg-Asp (1)[a]	242	191	105	179	1.259	1.232			21
9	Arg-Asp (2)	240	186	108	178	1.253	1.248			
10	aspartic acid (DL)	243	176	107	175	1.251	1.250	1605	1418	22
11	aspartic acid (L)	242	179	106	176	1.252	1.242	1607	1421	23
12	Asp-Ala	243	184	106	178			1588	1398	
13	Asp-Gly	240	180	104	175	1.258	1.246	1599	1398	24
14	cysteine	239	173	109	174	1.251	1.239	1598		25
15	cystine	243	174	108	175	1.250	1.238	1584	1407	26
16	glutamine	241	172	106	173	1.260	1.238	1586	1413	27
17	glutamatic acid	243	185	107	179	1.268	1.242		1420	28
18	Glu-Ala	243	195	106	181			1563	1406	
19	Glu-Thr	241	196	105	180			1563	1406	
20	Glu-Val	242	192	104	179			1560	1406	
21	glycine	241	178	111	177	1.251	1.250	1598	1410	29
22	Gly-Ala	235	200	110	182			1561	1406	
23	Gly-Asp·2H_2O	242	192	106	180	1.253	1.250	1580	1425	30
24	Gly-Glu	244	193	109	182			1583	1394	
25	histidine	245	176	105	175	1.250	1.247			31
26	isoleucine	243	175	108	175	1.254	1.243	1584	1418	32
27	leucine	242	180	108	177	1.260	1.240	1584	1408	33
28	lysine	242	179	109	177			1590	1410	
29	Lys·Asp (1)[a]	238	187	107	177	1.275	1.237			34
30	Lys·Asp (2)	240	185	108	178	1.250	1.247			
31	Lys·Asp (3)	243	176	107	175	1.264	1.250			
32	methionine	240	182	109	177	1.254	1.230	1584	1409	35
33	Met-Glu	237	191	110	179			1588	1393	
34	proline	238	173	113	175					
35	serine	242	174	111	175	1.251	1.248	1598	1411	36
36	Ser-Glu	242	187	107	179			1578	1404	
37	sodium aspartate (DL)[b]	243	185	107	178					
38	sodium glutamate	239	200	107	182					
39	sodium glutamate (δ)	242	202	105	183					
40	threonine	243	168	107	172	1.253	1.236	1630	1418	37
41	valine	245	176	106	175	1.249	1.248	1587	1425	38
42	Val-Ala	239	178	116	178			1589	1403	
43	Val-Glu	238	185	111	178	1.262	1.243	1577	1405	39
	deprotonated groups[c]	241 ± 4	183 ± 18	108 ± 4	177 ± 6					

[a] The assignment of these carboxyl groups is unknown. [b] The two carboxyl groups were unresolved. [c] Average ±2 standard deviations.

in seeking an independent indicator for hydrogen bonding we initially focused on comparing the carbon CSA tensor with vibrational frequencies for the carboxyl groups. Figure 2 shows the elements of the CSA tensors plotted against the carbonyl stretching frequency for the protonated acids and the CO_2^- asymmetric stretching frequency for deprotonated carboxyl groups. We also show typical examples of the type of hydrogen bonding partners that appear in the crystal structures. There is obviously a monotonic correlation between σ_{22} and the stretching frequency. For the protonated carboxyl group, stronger hydrogen bonding is associated with the softening of the carbonyl stretch and decreased shielding for the σ_{22} element. For the deprotonated carboxylates, the reverse is true; stronger hydrogen bonding is associated with stiffening of the asymmetric stretch and increased shielding for the σ_{22} element. Thus the σ_{22} element and the IR frequency have a rather smooth behavior from deprotonated to protonated groups.

The correlation between vibrational frequency and chemical shielding was specific to the carbonyl group vibrational frequency. For example, when the C–O single bond stretching frequency or the CO_2^- symmetric stretching frequency was studied, no correlation was observed with any of the chemical shielding tensor elements. Examination of the dependence of each of the tensor elements also shows that there is no strong correlation between σ_{11}, σ_{33}, and any stretching frequency. There is also no particular relation between the carbon shielding values and hydrogen bonds to the alcohol (C–O–H) group. These facts led us to consider the possibility that the principal perturbation that is manifested in the σ_{22} element of the ^{13}C CSA is the hydrogen bonding to the *carbonyl group* of protonated groups. We therefore examined the crystallographically obtained distances for hydrogen bonding to the carbonyl group. Table 3 lists the available hydrogen bonding structural parameters along with the type of hydrogen bonding of the protonated carboxyl groups. A correlation was observed between the O···H distance and σ_{22} (Figure 3). We have also noted a strong correlation between the C=O (covalent) carbonyl length and the O···H distance (not shown). In many crystallographic analyses the O···O or O···N distances are used as a

(37) (a) Shoemaker, D. P.; Donohue, J.; Schomaker, V.; Corey, R. B. *J. Am. Chem. Soc.* **1950**, *72*, 2328. (b) Ramandham, M.; Sikka, S. K.; Chidambaram, R. *Pramana* **1973**, *1*, 247.
(38) DL-Valine: Mallikarjunan, M.; Rao S. T. *Acta Crystallogr. B* **1969**, *25*, 296–303.
(39) Eggleston, D. S. *Acta Crystallogr. C* **1984**, *40*, 1250–1252.
(40) Blasio, B. D.; Pavone, V.; Pedone, C. *Cryst. Struct. Commun.* **1977**, *6*, 745–748.
(41) Dawson, B. *Acta Crystallogr. B* **1977**, *33*, 882–884.
(42) Sequeira, A.; Rajagopal, H.; Chidambaram, R. *Acta Crystallogr. B* **1972**, *28*, 2514–2519.
(43) Dawson, B. *Acta Crystallogr.* **1953**, *6*, 81–87.
(44) Al-Karaghouli, A. R.; Cole, F. E.; Lehmann, M. S.; Miskel, C. F.; Verbist, J. J.; Koetzle, T. F. *J. Chem. Phys.* **1975**, *63*, 1360–1366.
(45) Kistenmacher, T. J.; Dorrell, T. *J. Cryst. Mol. Struct.* **1974**, *4*, 419–432
(46) Koetzle, T. F.; Golic, L.; Lehmann, M. S.; Verbist, J. J.; Hamilton, W. C. *J. Chem. Phys.* **1974**, *60*, 4690–4696.

Table 2. Tensor Elements of Protonated Carboxyl Groups in Amino Acids, Dipeptides

no.	molecule	σ_{11}, ppm	σ_{22}, ppm	σ_{33}, ppm	σ_{iso}, ppm	d_{C-O}, Å	$d_{C=O}$, Å	$\nu_{C=O}$, cm^{-1}	ν_{C-O}, cm^{-1}	HB[b]	ref
1	alanine hydrochloride	254	163	111	176	1.317	1.177	1723	1199	A	40
2	Ala-Asp	256	153	112	174	1.288	1.195	1727	1227	B	18
3	aspartic acid (DL, γ)	260	166	107	178	1.305	1.219	1690	1212	B	22
4	aspartic acid (γ)	254	162	107	175	1.306	1.202	1692	1250	B	23
5	aspartic hydrochloride (DL)	257	148	106	170	1.320	1.210	1750	1180	B	41
6	aspartic hydrochloride (DL, γ)	260	162	106	176	1.330	1.220	1710	1180	C	
7	aspartic hydrochloride (L)	253	149	106	169			1744			
8	aspartic hydrochloride (L, γ)	256	157	110	174			1720			
9	Asp-Ala	253	166	106	175						
10	Asp-Gly·H_2O	257	156	105	173	1.310	1.210	1713	1227	A	24
11	glutamic acid (L, δ)	259	173	110	180	1.313	1.216	1667	1234	B	28
12	glutamic hydrochloride (L)	256	165	106	176	1.296	1.221	1725	1208	B	42
13	glutamic hydrochloride (L, δ)	252	174	109	178	1.315	1.225	1677	1208	C	
14	glutamic hydrochloride (DL)	259	164	104	176	1.316	1.214	1731	1214	B	43
15	glutamic hydrochloride (DL, δ)	254	174	107	178	1.319	1.194	1678	1214	C	
16	Glu-Ala	262	160	108	177			1698	1218		
17	Glu-Thr	254	171	109	178			1685	1227		
18	Glu-Val	263	157	109	176			1701	1213		
19	glycine hydrochloride	254	155	111	173	1.312	1.202	1738	1230	B	44
20	Gly-Asp·2H_2O	257	161	110	176	1.315	1.199	1717	1196	A	30
21	Gly-Glu	261	162	107	177				1703		
22	histidine·2HCl	254	153	108	172	1.313	1.196	1735	1198	B	45
23	isoleucine hydrochloride	255	147	111	171			1735	1217		
24	leucine hydrochloride	257	153	109	173			1738	1212		
25	lysine·2HCl	254	148	112	171			1738	1204		
26	Met-Glu	255	148	123	176			1726			
27	methyl leucine ester	258	151	111	173						
28	methyl serine ester	257	138	113	170						
29	Ser-Glu	258	164	108	177				1709		
30	valine hydrochloride	257	148	110	172	1.310	1.199	1730	1221	D	46
31	Val-Glu	255	162	110	176	1.315	1.214	1701	1195	B	39
	protonated groups[a]	257 ± 6	158 ± 18	109 ± 8	175 ± 6						

[a] Average ±2 standard deviations. [b] This column indicates the type of hydrogen bonds that carbonyl forms: A, not available; B, C=O···H—N (amine); C, C=O···H—O—CO; D, C=O···H—N but very weak.

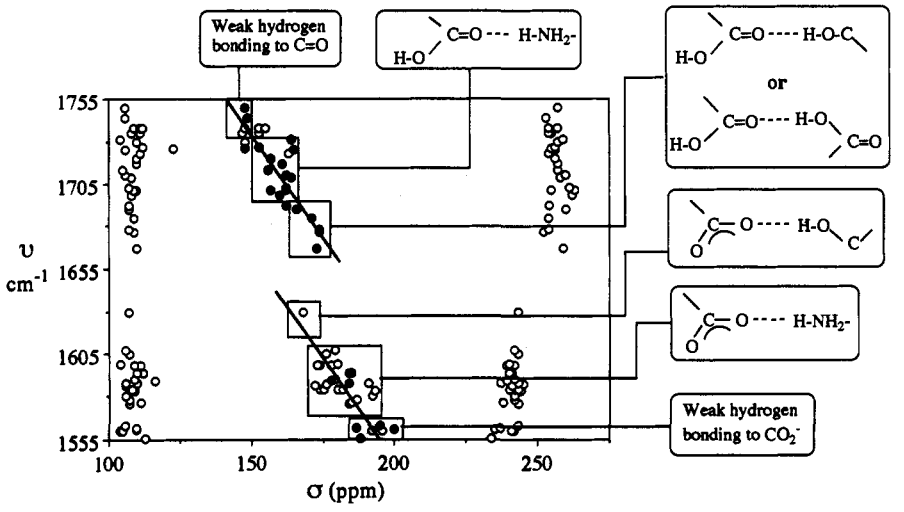

Figure 2. The carbon CSA for 28 protonated and 30 deprotonated carboxyl groups in amino acids and dipeptides is plotted against the stretching frequencies of the carbonyl for protonated acids or the asymmetric mode of the carboxylate group (referenced to adametane at 38.6 ppm). A correlation is apparent between the IR band and the σ_{22} element of the carbon CSA. The other vibrational modes did not exhibit any correlation to the NMR data. The upper cluster of points belongs to the protonated groups and the lower cluster belongs to the deprotonated groups. Specific amino acids in the databank, either simple salts or in peptides, are indicated by the choice of symbol. Filled circles in the protonated groups represent the sidechains of glutamatic acid while open circles with a plus represent aspartatic acid series, and filled circles in the deprotonated groups represent the backbone carboxy of alanine. The hydrogen bonding motif is indicated with the structures in balloons.

measure of hydrogen bond strength. For this database we could not find a significant correlation between σ_{22} and the O···N or O···O distance, presumably because of a variable O···H–X angle.

Although a general trend relating σ_{22} to the IR frequency was seen for deprotonated carboxy groups, the correlation was not as convincing as for the protonated groups. One possible weakness for the deprotonated case is our assignments of the IR bands. In addition, the effect of hydrogen bonding on either NMR or IR is complicated by the presence of three or four proton donors of comparable distances, and the fact that the geometry of the hydrogen bonding is variable for the deprotonated groups. For the protonated groups the geometry is rather fixed: the covalently attached proton is always on the syn side, and the hydrogen bonded proton is near the plane of the carboxy group on the syn side with a C—O···H angle of 120–140° (see Figure 3). This fixed preference for the protonated case and having *apparently* one major hydrogen bonding interaction for each carbonyl probably contribute to the strong correlations that we see in our data. The fact that the geometry is rather fixed for the protonated group but not for the deprotonated case has been discussed in detail

Table 3. Crystallographic Hydrogen Bonding Data of Carbonyl Groups in Amino Acids and Dipeptides

no.	molecule	type of HB[a]	$d_{O\cdots N(O)}$, Å	$d_{O\cdots H}$, Å	∠CO···H, deg	ref	type[b]
1	aspartic acid (DL)	C=O···H—N	2.899	2.07	141.1	20	X
2	Ala-Asp	C=O···H—N	2.828	2.27		16	X
3	glutamic hydrochloride	C=O···H—N	2.904	1.977	120.1	40	N
4	glutamic hydrochloride	C=O···HO—CO	2.637	1.624	139.0		N
5	glutamic acid	C=O···H—N	2.895	1.926	130.9	26	N
6	glycine hydrochloride	C=O···H—N	2.992	2.222	122.5	42	N
7	histidine·2HCl	C=O···H—N	2.90	2.12		43	X
8	Val-Glu	C=O···H—N	2.856	1.98		37	X
9	valine hydrochloride	C=O···H—N (vw)		2.46		44	N

[a] For the protonated carboxyl group, there are basically four kinds of hydrogen bonding motifs (shown in Figure 2): (a) very weak (long) hydrogen bonding in which no major proton donor is identified; (b) C=O···H—N, where N—H is an amine, which is the most common case; (c) C=O···H—O—C; (d) C=O···HO—C=O, which forms the strongest hydrogen bonding. [b] Indicates the type of the crystal structures: X means X-ray crystal structure and N means neutron diffraction crystal structure.

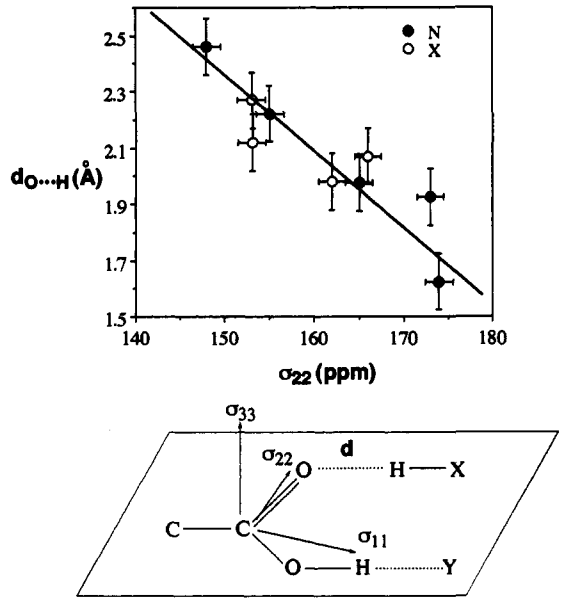

Figure 3. The σ_{22} element of the tensor for protonated carboxyl groups is plotted against the distance between the carbonyl and the nearest proton hydrogen bond donor. A strong correlation is observed, whereas no correlation with crystallographic or spectroscopic properties associated with the C–O–H group of the acid could be found. The error of σ_{22} is ±1.5ppm and that of d_{O-H} is ±0.1 Å. Open circles represent the X-ray data and closed circles represent neutron data. The scheme at the bottom indicates the orientation of the chemical shift tensor and the angle relation with the protons.

earlier.[47] In these earlier studies both the syn and the anti geometries of hydrogen bonding to the deprotonated group were observed to occur frequently. However these two cases have been strongly contrasted in terms of their electronic properties and their respective effects upon the hydrogen bond donor, with the syn hydrogen bonding having a much more pronounced effect upon reactivity. Because of the complicated geometric variation we did not include any analyses on the crystallographic data from the deprotonated groups for the moment.

Data on carboxylic sidechains of peptides involving glutamate and aspartate allowed us to check for the effect of the alkyl sidechain upon the shielding tensor elements. We also measured the IR and NMR of the deprotonated carboxylates from a number of alanine dipeptides. Figure 2 displays these restricted data sets with contrasting symbols. The Asp and Glu sidechains essentially follow the same correlation curves as the HCl salts of amino acids, and carboxylates of a C-terminal alanine fall on the same curve as any other carboxylate of an amino acid zwitterion. By so restricting the choice peptides we did not appreciably reduce the scatter in our correlation curves. Thus we do not see a strong systematic effect of α or β substitutions on any tensor element.

The principal perturbation of the σ_{22} value for the protonated acids appears to be the hydrogen bonding to the carbonyl group which should affect the energy, symmetry, or electron density in the n orbitals of the carbonyl. For the protonated acids the effect of the hydrogen bonding on shielding is qualitatively similar to that which was previously observed for amides, namely a decrease in shielding with increasing hydrogen bond strength.[9] This effect has been rationalized in terms of an n to Π^* mixing which is only allowed for the σ_{22} element.[48] The observed effect of hydrogen bonding on carbon shielding could result from destabilization of n orbitals by hydrogen bonding, consistent with an experimental and theoretical study of the effect of dimerization of acetic acid in the gas phase on the orbital energies.[49] For the deprotonated carboxyl groups the opposite trend occurs: stronger hydrogen bonding is associated with more shielded values for the σ_{22}. In this case a simple argument regarding the stabilization of the n orbitals by electrostatic effects probably holds. Clarification of these points requires high-quality *ab initio* calculations, currently under consideration.

In summary, when analyzing the CSA of a carboxyl group in amino acids, we can identify the protonation state of the group using the σ_{11} value and then determine the approximate hydrogen bonding strength to the carboxyl oxygens from σ_{22}. For the protonated groups hydrogen bonding information concerning the acidic proton is also available from CRAMPS experiments,[3] and a complete picture of the hydrogen bonding around the carboxyl group can be described.

These methods have been used to analyze a transition state analog compound at the active site of triose phosphate isomerase.[50] They can also be used to interpret data previously published concerning the aspartic acids that catalyze proton transfer reactions in bacteriorhodopsin.[51] We are in the process of utilizing this type of analysis to study proton transfer and ionization states in several other enzymes as well.

Acknowledgment. This work was supported by a grant from the International Human Frontiers Research Program. The authors would like to thank Mr. Luis Avila for help with the FTIR measurements.

(47) Görbitz, C. H.; Etter M. C. *J. Am Chem. Soc.* **1992**, *114*, 627–631.

(48) Kempf, J.; Spiess, H.; Haeberlin, U.; Zimmermann, H. *Chem. Phys.* **1974**, *4*, 269.

(49) Carnovale, F.; Gan, T.; Peel, J. *J. Electron Spectrosc. Relat. Phenom.* **1980**, *20*, 53–67.

(50) Tomita, Y.; O'Connor, E.; McDermott, A. *J. Am. Chem. Soc.* Submitted for publication.

(51) (a) Metz, G.; Siebert, F.; Engelhard, M *Biochemistry* **1992**, *31*, 455–462. (b) Engelhard, M.; Hess, B.; Emeis, D.; Metz, G.; Kreutz, W.; Siebert, F. *Biochemistry* **1989**, *28*, 3967–3975.