Studies toward this end are now in progress in this Laboratory.

Acknowledgments.—The authors are pleased to acknowledge valuable discussions with Drs. E. R. Garrett and M. L. Bender and would like also to thank Dr. C. Zioudrou for her very kind assistance in the determination of some of the rate data and also for the synthesis of the amide XI. This research was aided by grants to T. C. Bruice by the U. S. Public Health Service (A-980-C2) and to J.M. Sturtevant from the National Science Foundation (G-2855) and the U. S. Public Health Service (RG-4725).

NEW HAVEN, CONN.

[CONTRIBUTION NO. 148 FROM THE CENTRAL RESEARCH DEPARTMENT, MINNESOTA MINING AND MANUFACTURING CO.]

Proton N.S.R. Spectroscopy. V. Studies of Amino Acids and Peptides in Trifluoroacetic Acid

BY F. A. BOVEY AND G. V. D. TIERS

RECEIVED SEPTEMBER 2, 1958

The proton spin resonance spectra of trifluoroacetic acid solutions of all the common and several less common amino acids and of many of their glycyl peptides and N-acetyl derivatives are reported. The use of trifluoroacetic acid as solvent has the advantage that an internal reference standard, tetramethylsilane, may be employed, enabling peak positions to be very accurately determined; in addition, it is an excellent solvent for these compounds. By measurement and comparison of peak positions, widths and multiplicities, conclusions of considerable interest may be drawn concerning charge distribution, inductive effects of polar groups and positive charges, rates of proton exchange with solvent, base strength and (in one case) molecular conformation. In addition, the tabulation of peak positions is very useful in the interpretation of the n.s.r. spectra of proteins.

Introduction

The proton nuclear spin resonance (n.s.r.) spectra of glycine, alanine, β-alanine, cysteine, proline, hydroxyproline, glycylglycine and alanylalanine have been reported by Takeda and Jardetzky,[1] who examined 2–3 *M* solutions in strongly acidic, neutral and strongly basic aqueous medium. These studies have recently been extended to include 22 amino acids.[2] The protons of the $-NH_3^+$ group can be distinguished in 6 *N* HCl, in concentrated H_2SO_4 and in $H_2SO_4 \cdot H_2O$; in less strongly acidic solutions the peak due to these protons coalesces with that of the water, in the fashion typical for rapid exchange between the species. In the spectra of the dipeptides the peak due to peptide hydrogen likewise appears only in strongly acid solutions. Takeda[3] has extended these studies to β-alanine, α-aminobutyric acid and γ-aminocaproic acid.

We have studied the proton n.s.r. spectra in trifluoroacetic acid solvent of all the common and of several less common amino acids and of many of their glycyl peptides and N-acetyl derivatives. Trifluoroacetic acid is an excellent solvent for these substances, 20% (weight/volume) solutions of nearly all of them being readily prepared. Glycine is soluble to the extent of about 15% and cysteine and cystine are still less soluble but sufficiently so (5–8%) to give excellent spectra. In trifluoroacetic acid, for which the Hammett acidity function H_0 is -4.4,[4] those groups acquire protons which would normally do so in strongly acidic aqueous medium. Peptide and amide protons show n.s.r. peaks, as does the sulfhydryl group of cysteine, but those of hydroxyl and carboxyl exchange too rapidly with the solvent to be observable; small concentrations of water likewise give no n.s.r. peak distinguishable from that of the solvent. The present study provides information useful in the interpretation of the n.s.r. spectra of proteins, as will be demonstrated in a later publication.

Referencing.—In D_2O or H_2O solutions the n.s.r. peak for water protons in the solution is variable in position and hence not satisfactory as a reference point for a scale of shielding values. The use of an "external reference," for example, H_2O or toluene in a capillary tube, would be satisfactory, *provided that the required*[5–7] *extrapolation of the solute peak position to infinite dilution were carried out.* Unfortunately this is virtually never done, unreliable values being reported instead.[8] Careful temperature control and measurement also are necessary because of the differing magnetic susceptibilities and coefficients of expansion of the solvent and reference. On the basis of work done in organic liquids[8] we suspect that an internal reference might be preferable even in aqueous media; however, the proper choice of such a reference for aqueous solutions has not been discussed in the scientific literature, and is not dealt with in this research.

One of the significant advantages provided by the use of trifluoroacetic acid as solvent is that tetramethylsilane may be employed as the internal reference.[8] It does not appear to be sufficiently soluble in purely aqueous solutions. The usual concentration of tetramethylsilane is 1% by volume. Its n.s.r. peak is very sharp and falls beyond the usual range of proton resonances. It is nonassociative, magnetically isotropic, and is usually

(1) M. Takeda and O. Jardetzky, *J. Chem. Phys.*, **26**, 1346 (1957).

(2) O. Jardetzky and C. D. Jardetzky, *J. Biol. Chem.*, **233**, 383 (1958).

(3) M. Takeda, paper presented to Div. of Biol. Chem., 131st Meeting of American Chemical Society, April 7-12, 1957.

(4) (a) G. V. D. Tiers, THIS JOURNAL, **78**, 4165 (1956), and (b) E. L. Mackor, P. J. Smit and J. H. Van der Waals, *Trans. Faraday Soc.*, **53**, 1309 (1957).

(5) A. A. Bothner-By and R. E. Glick, THIS JOURNAL, **78**, 1071 (1956).

(6) A. L. Allred and E. G. Rochow, *ibid.*, **79**, 5361 (1957).

(7) J. S. Waugh and R. W. Fessenden, *ibid.*, **79**, 846 (1957).

(8) G. V. D. Tiers, *J. Phys. Chem.*, **62**, 1151 (1958).

chemically unreactive, although its concentration is occasionally observed to decrease on long standing under circumstances where evaporation is impossible.

In the tables which follow, the dimensionless spectral positions reported for the peaks are shielding values (sometimes called "chemical shifts") which become more positive with increasing shielding of the proton. The numerical values (in parts per million) are entitled "τ-values," and are referred to the "τ-scale,"[8] on which the tetramethylsilane peak is taken as +10.000 p.p.m. by definition. Nearly all the τ-values observed have been found to be positive; in this study only a very few (*e.g.*, NH of the imidazole ring of histidine) show negative values. The τ-values here tabulated may be compared directly with the large number of τ-values (over 2000) measured for organic compounds in carbon tetrachloride solution.[9] The τ-value measurements are believed to be accurate to *ca.* ± 0.01 p.p.m. standard deviation except for certain broad or ill-defined peaks. A large number of the peaks observed are multiplets. The *observed* degree of multiplicity is designated by d (doublet), t (triplet) and q (quadruplet). The listed τ-value corresponds to the center of the multiplet. The separation of the observable components of the multiplet, *i.e.*, the coupling constant *J*, is expressed in cycles-sec.$^{-1}$ (c.p.s.). Both the peak positions and the coupling constants thus should be independent of the observing frequency.

When a multiplet is expected but its components are unresolved, the peak is designated "m"; the position of its center is recorded but no value for *J* is assigned. However, the peak width at half-height, designated "*W*" and recorded in cycles-sec.$^{-1}$, may be used to estimate *J*. An unresolved peak which is apparently broader than 10 c.p.s., for whatever reason, is designated "b." In most cases, the peak width can be measured and is recorded, but in some cases, as when two peaks are closely fused together or resolution is very poor, this is not possible.

Tetramethylsilane at 1% concentration in trifluoroacetic acid has its n.s.r. peak at +5.35 p.p.m. with respect to water, the latter being present at 25.0° as a separate phase in a capillary tube. Conversion to a scale on which external water is zero *does not* make data obtained in one solvent comparable to that obtained in another solvent of differing magnetic susceptibility, though this fact apparently is ignored in nearly all scientific papers in which n.s.r. data are cited.[8]

The spectra presented in Figs. 1–6 of this paper have been limited to those which illustrate important classes of compounds, or particular spectral features.

Experimental

The common amino acids were obtained from Schwarz Laboratories, Inc., Mount Vernon, N. Y. Betaine hydrochloride, DL-dihydroxyphenylalanine, glycocyamine, hydroxy-L-proline, allo-hydroxy-D-proline, DL-1-methylhistidine, L-thiolhistidine, D-alloisoleucine, DL-allothreonine and the N-acetylamino acids were obtained from the California Corporation for Biochemical Research, Los Angeles. N-Acetyltyrosine ethyl ester and all peptides were obtained from Mann Research Laboratories, New York 6, N. Y.

(9) G. V. D. Tiers, paper presented at the 133rd Meeting of the American Chemical Society, San Francisco, Calif., April 13–18, 1958, and currently being prepared for publication. Preprints are available on request.

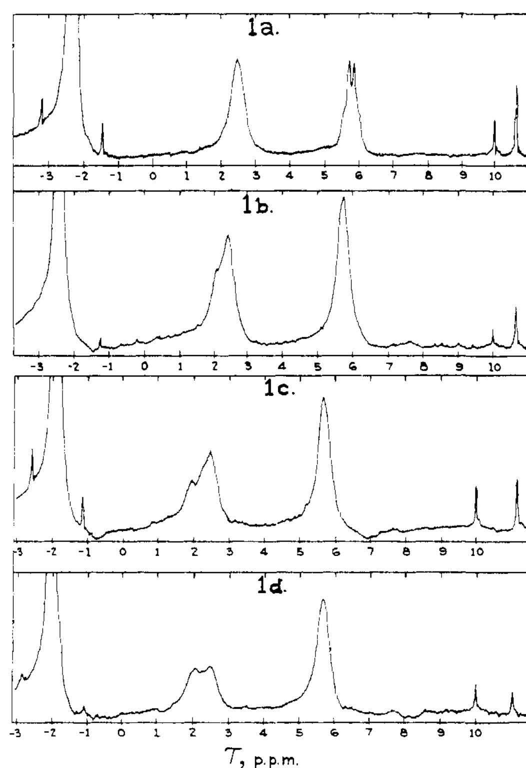


Fig. 1.—N.s.r. spectra of glycine and polyglycines in trifluoroacetic acid: a, glycine; b, diglycine; c, triglycine; d, tetraglycine.

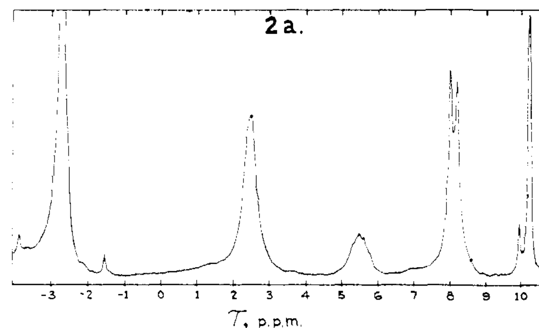


Fig. 2a. N.s.r. spectrum of DL-alanine in trifluoroacetic acid.

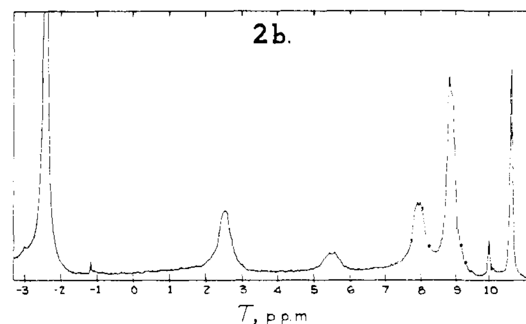


Fig. 2b.—N.s.r. spectrum of L-leucine in trifluoroacetic acid.

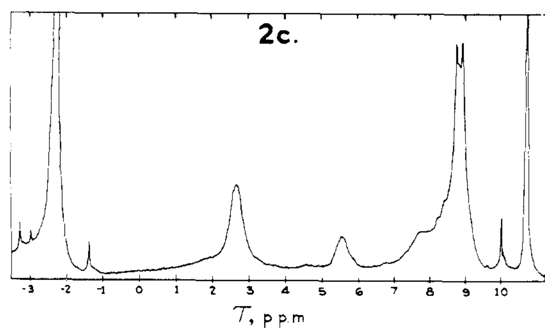


Fig. 2c.—N.s.r. spectrum of DL-isoleucine in trifluoroacetic acid.

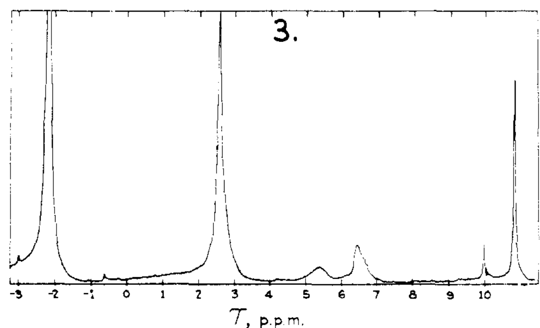


Fig. 3.—N.s.r. spectrum of DL-phenylalanine in trifluoroacetic acid.

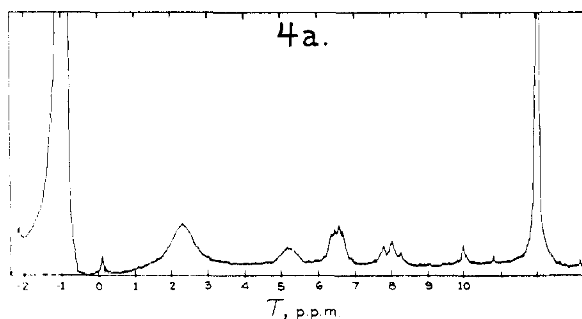


Fig. 4a.—N.s.r. spectrum of L-cysteine in trifluoroacetic acid.

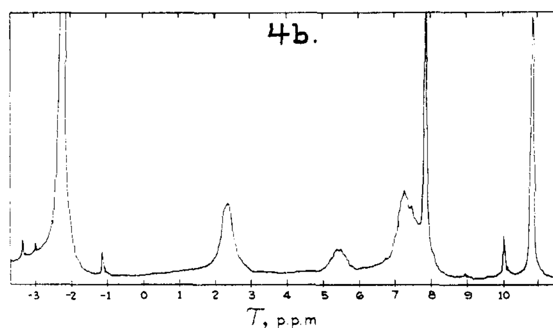


Fig. 4b.—N.s.r. spectrum of DL-methionine in trifluoroacetic acid.

All compounds were used as received without further testing or purification. In all cases, the n.s.r. spectra were consistent with the known structure of the compound and showed no evidence of any substantial amounts of impurities. It is probable, however, that as much as 3–5% of impurity would not be detectable under the experimental conditions employed. Optical isomers of course give identical spectra,

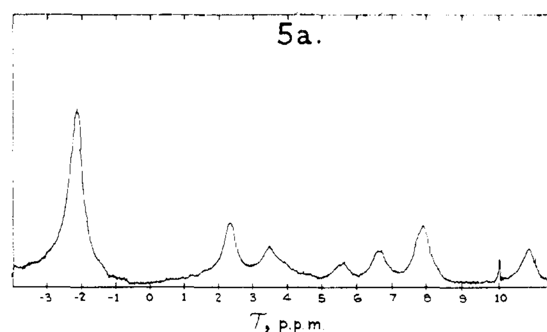


Fig. 5a.—N.s.r. spectrum of L-arginine in trifluoroacetic acid.

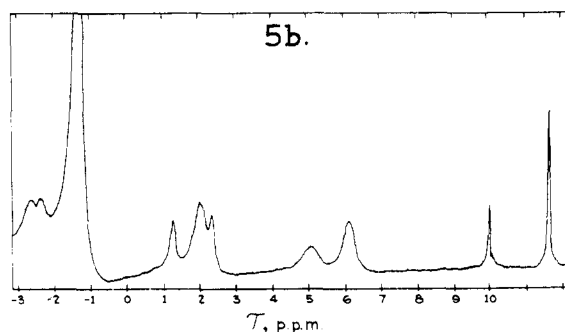


Fig. 5b.—N.s.r. spectrum of L-histidine in trifluoroacetic acid.

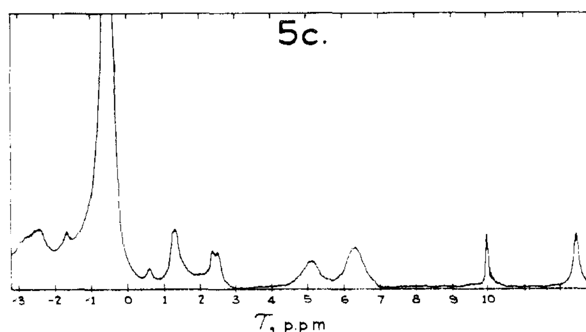


Fig. 5c.—N.s.r. spectrum of L-histidyl-L-histidine in trifluoroacetic acid.

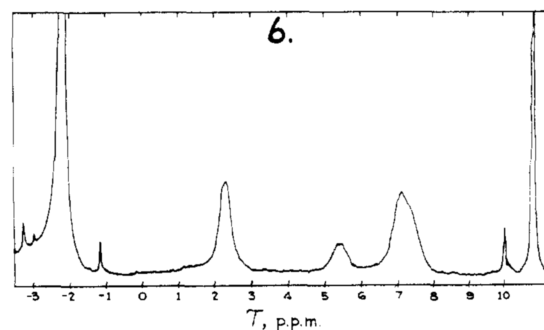


Fig. 6.—N.s.r. spectrum of L-glutamic acid in trifluoroacetic acid.

so that varying proportions of these would give no observable effect.

Trifluoroacetic acid was obtained from the Fluorochemical Division of Minnesota Mining and Manufacturing Co., and was redistilled through an all-glass fractionating column. Tetramethylsilane, the "internal reference," obtained from

Anderson Laboratories, Inc., Weston, Mich., was added to the extent of 1.0% by volume to the trifluoroacetic acid solvent.

Solutions were prepared by dissolving 100 mg. of sample in 0.50 ml. of "referenced" trifluoroacetic acid in a 0.5-dram screw-cap vial fitted with a Teflon cap liner. In some cases, noted above, the substance was not soluble to this extent and so a saturated solution was prepared, care being taken not to retain solid particles in the solution, since these produce erratic effects in the spectrum. The solution was then transferred to a 5 mm. o.d. Pyrex tube which in many cases was not quite concentric and modulated the magnetic field while spinning; this effect is responsible for the weak, sharp peaks observable near the solvent peak in the spectra. Throughout the preparation of the samples, the trifluoroacetic acid, which is very hygroscopic, was exposed to air as little as possible, although small quantities of moisture do not have any perceptible effect on the spectrum.

A Varian V-4300-2 40.00 Mc/s. spectrometer was employed, together with a sample spinner, audio-oscillator, Hewlett Packard 522-B frequency counter and Varian recorder. The sweep was such that 13 p.p.m. (corresponding to 520 c.p.s. side-bands) occupied about 220 mm. on the chart. The 520 c.p.s. side-band peak corresponding to the solvent appears at the extreme right in all spectra. Just to the left of it at 10.00 τ appears the smaller peak due to the tetramethylsilane reference, while on the extreme left is the solvent peak. The τ-values were measured by linear interpolation in these spectra, *ca.* ± 0.02 p.p.m. standard deviation being observed. In many cases the results from 3 or 4 spectra were averaged; thus the resultant standard deviation in τ-values is closer to ± 0.01 p.p.m.

Results and Discussion

Glycine and Polyglycines

The experimentally determined n.s.r. spectra of glycine, diglycine, triglycine and tetraglycine are shown in Fig. 1 and the peak positions are summarized in Table I. The peak

TABLE I

PROTON N.S.R. SHIELDING VALUES, τ, FOR GLYCINE AND POLYGLYCINES IN TRIFLUOROACETIC ACID SOLUTION

Compound	CONH	NH_3^+	CH_2, CH_3
Glycine	..	2.46 m, b; $W = 20$	5.80 q, $J = 6$
N-Acetylglycine	1.65 t; $J = 5$	..	CH_2: 5.54 d $J = 5$ CH_3: 7.53
Diglycine	2.05 d, $J = 6$	2.46 m, b; $W = 17$	5.71 m, b; $W = 16$
Triglycine	1.91	2.44 m, b; $W = 17$	5.69 m, b; $W = 13$
Tetraglycine	2.05, m 2.25, m	2.49 m, b; $W = 17$	5.68 m, b; $W = 16$
Diketopiperazine	1.72 $W = 9$	..	5.56 $W = 4$

at 2.46 τ in the glycine spectrum (Fig. 1-a) is that of the charged amino group. In the absence of interfering effects, this would be split into a triplet of triplets by spin-spin interaction with the three spin states of the N^{14} nucleus ($I = 1$, $J \sim 50$ c.p.s.) and with the spins of the two neighboring methylene protons.[10] Here, however, as will be discussed later, the half-height width of the ammonium n.s.r. peak is only 20 c.p.s. This narrowing cannot be due to proton exchange, for the methylene absorption shows clearly a quadruplet splitting by the three ammonium protons.[10]

In the polyglycines the peptide hydrogen n.s.r. peak appears near 2.0 τ. For diglycine this peak appears split by the neighboring methylene hydrogens, but only part of the expected triplet is discernible. In triglycine and tetraglycine the peptide hydrogen absorption becomes larger, the latter showing a peak equal to that for the ammonium group, as expected. The methylene τ-value remains nearly constant for the polyglycines but is 0.1 p.p.m. less than that for glycine. Takeda and Jardetzky[1] report two methylene peaks in the diglycine spectrum, which they attributed to the two different methylene groups. In trifluoroacetic acid only one peak is found. Diketopiperazine has markedly lower τ-values for both peptide and methylene protons; this effect is apparently general in cyclic compounds.

Alanine, Valine, Leucine, Isoleucine, Proline, Hydroxyproline, Betaine, Their Peptides and N-Acetyl Derivatives

In Table II are presented data for this group of substances. In Fig. 2 are shown the spectra of DL-alanine, L-leucine and DL-isoleucine. When incorporated into proteins (except of course at a chain end) glycine, alanine, valine, leucine, isoleucine and proline contribute no "functional" groups. The high field (most shielded) portion of the n.s.r. spectra of proteins is attributable[11,12] to the side-chain protons of these amino acids, which constitute about half of the non-exchangeable protons of most globular proteins. Protons farther away from the charged ammonium group show peaks at higher τ-value because of the decreasing field strength of the positive charge.[2] The uncharged carboxyl group, a dipole, exerts a marked but shorter-range field (or inductive) effect in the same direction. The n.s.r. peaks due to α-hydrogens in these amino acids appear at τ-values 0.11 to 0.26 p.p.m. lower (0.54 for proline) than those of glycine despite the presence of alkyl groups on the same carbon atom; similar effects have been noted for esters and halo-alkanes.[9,13] For the glycyl peptides, τ-values for the α-protons are lower by 0.3 to 0.4 unit, and the N-acetyl derivatives show a still further decrease by about 0.1 p.p.m. The peaks for side-chain protons behave oppositely, moving to higher τ-values in the glycyl peptides and N-acetyl derivatives, this effect being seen most clearly for the methyl group of alanine and its derivatives. These observations suggest that the peptide linkage exerts a very powerful but short-range electrostatic anti-shielding[14] or electron-withdrawing effect (as would be expected for the C=O dipole) which is felt strongly at the α-protons but very little two carbons away. It should be noted that magnetic anisotropy effects[13] may be important in this case. The electron-withdrawing effect of the ammonium group, on the other hand, is clearly measurable three carbons away (*e.g.*, in the methyl groups of valine) but not four carbons away (*e.g.*, in the methyl groups of leucine); the field due to a pole falls off less rapidly than that of a dipole.

The coupling constants for the peptide hydrogen and the β-hydrogen with the α-hydrogen are usually about equal in glycyl peptides and N-acetyl derivatives, being about 5–6 c.p.s. For N-acetylvaline, however, a quadruplet is observed.

An important feature of the spectra of the N-acetylamino acids and of several of the glycyl peptides is the narrow multiplet resonance observed for the peptide protons; this effect has been discussed elsewhere.[15]

Table II presents three cases of isomeric pairs of dipeptides: glycylvaline and valylglycine; glycylleucine and leucylglycine; and glycylproline and prolylglycine. In all three cases the spectra are appreciably different. Most of the difference arises from the electrostatic effects just discussed. In a protein or long polypeptide chain, interchanges of non-terminal residues would probably produce at most only very slight shifts.

The spectrum of betaine consists of two simple peaks; τ-values are included in Table II. By way of comparison, the peak for $(CH_3)_4N^+$ is at 6.67 τ, and the peaks for the CH_3 and CH_2 groups in $(C_2H_5)_4N^+$ are at 8.59 τ and 6.60 τ, respectively.

Aromatic Amino Acids

The protons of and near the benzene ring of the aromatic amino acids and their peptides have peaks at the relatively low τ-values to be expected for aromatic compounds. This behavior is a result of the reinforcement of the applied

(10) E. Grunwald, A. Lowenstein and S. Meiboom, *J. Chem. Phys.*, **27**, 630 (1957).

(11) O. Jardetzky and C. D. Jardetzky, THIS JOURNAL, **79**, 5332 (1957).

(12) F. A. Bovey, G. V. D. Tiers and S. S. Yanari, paper presented at the 133rd Meeting of the American Chemical Society, San Francisco, Calif., April, 1958.

(13) A. A. Bothner-By and C. Naar-Colin, THIS JOURNAL, **80**, 1728 (1958).

(14) T. W. Marshall and J. A. Pople, *Molecular Phys.*, **1**, 199 (1958).

(15) G. V. D. Tiers and F. A. Bovey, *J. Phys. Chem.*, **63**, 302 (1959).

Table II

Proton N.S.R. Shielding Values, τ, for Alanine, Valine, Leucine, Isoleucine, Proline, Hydroxyproline and Betaine, and Their Peptides and N-Acetyl Derivatives in Trifluoroacetic Acid Solution

Compound	CONH	NH_3^+	α-CH	CH. CH_2	CH_3
DL-Alanine	..	2.52 m, b; $W = 18$	5.54 m; $J \cong 5$		8.15 d; $J = 7$
N-Acetyl-DL-alanine	1.65 d; $J = 7$	..	5.11 d; $J = 7$		Ac: 7.50 8.30 d; $J = 7.5$
Glycyl-DL-alanine	2.20 d; $J = 8$	2.47 m, b; $W = 18$	a: 5.24 t; $J = 7$ g: 5.76 d; $J = 5$		8.35 d; $J = 7.5$
DL-Valine	..	2.66 m, b; $W = 17$	5.69 t; $J = 5$	7.56 m; $W \cong 30$	8.81 d; $J = 6$
N-Acetyl-DL-valine	1.72 d; $J = 8.5$	..	5.24 q $J = 6.5, 8.5$	(Under Ac)	Ac: 7.54 8.89 d; $J = 6.5$
Glycyl-DL-valine	2.16	2.44 m, b; $W = 19$	v: 5.34 t, $J = 7$ g: 5.71 d; $J = 5$	7.70 m; $W \cong 30$	8.92 d; $J = 6.5$
DL-Valylglycine	1.97	2.58 m, b; $W = 17$	v, g: 5.64 m, b; $W = 17$	7.55 m; $W \cong 30$	8.75 d; $J = 6$
L-Valyl-L-valine (2) (1)	2.15 d; $J = 4$	2.52 m, b; $W = 19$	(1): 5.64 m, b; $W = 18$ (2): 5.26 m, b; $W \cong 18$	7.61 m; $W \cong 30$	(1): 8.91 d; $J = 7$ (2): 8.52 d; $J = 7$
L-Leucine	..	2.64 m, b; $W = 16$	5.64 d; $J = 6$	CH_2: 8.02 d; $J \cong 5$	8.93 d; $J = 3.6$
N-Acetyl-DL-leucine	1.72 d; $J = 8$	..	5.10 q; $J = 8$	CH_2: 8.13 d; $J = 5$	Ac: 7.56 8.92 d; $J = 3$
Glycyl-L-leucine	2.22	2.48 m, b; $W = 24$	1: 5.26 t; $J = 6$ g: 5.75 m; $J = 5$	CH_2: 8.18 m; $W = 16$	8.96
DL-Leucylglycine	2.01 m, b; $W = 17$	2.60 m, b; $W = 16$	l, g: 5.62 m, b; $W = 14$	CH_2: 8.07 m, b; $W = 18$	8.92
DL-Isoleucine	..	2.70 m, b; $W = 17$	5.60 m, b; $W = 13$	CH_2: 8.36 m CH_2 ~7.8 m, b; $W \cong 20$	β: 8.85 d; $J \sim 6$ ω: 8.93 t; $J \sim 6$
D-Alloisoleucine	..	2.70 m, b; $W = 17$	5.58 m, b; $W = 13$	CH_2: 8.33 m CH, ~7.8 m, b; $W \cong 20$	β: 8.79 ω: 8.86
L-Proline	..	2.17 m, b; $W = 36$	5.26 t; $J = 6$	CH_2N: 6.30 t, $J = 6$ $(CH_2)_2$: 7.64 m	..
Glycyl-L-proline	..	2.50 m, b; $W = 17$	p: 5.25 t; $J = 6$ g: 5.73 d; $J = 5$	CH_2N: 6.30 m, b; $W = 12$ $(CH_2)_2$: 7.72	..
L-Prolylglycine		2.05 m, b; $W = 17$	p: 5.24 t; $J = 6$ g: 5.78 d; $J = 5$	CH_2N: 6.33 t; $J = 5$ $(CH_2)_2$: 7.64 m, b; $W = 17$	..
L-Hydroxyproline	..	1.36, b (?) 1.86, b (?)	OCH + α: 4.94 m, b; $W = 14$	CH_2N: 6.12 m, b; $W = 14$ CH_2: 7.20 m, b; $W = 22$	..
Betaine	..	..	5.55	..	6.49

magnetic field by that arising from the circulation of the π-electrons of the benzene ring.[7,16–19] However, for protons occupying (average) positions near the face of the benzene ring, as may occur when the compound has a folded struc-

(16) J. A. Pople, *J. Chem. Phys.*, **24**, 1111 (1956).

(17) A. A. Bothner-By and R. E. Glick, *ibid.*, **26**, 1651 (1957).

(18) L. W. Reeves and W. G. Schneider, *Can. J. Chem.*, **35**, 251 (1957).

TABLE III

PROTON N.S.R. SHIELDING VALUES, τ, FOR AROMATIC AMINO ACIDS IN TRIFLUOROACETIC ACID SOLUTION

Compound	CONH	Aryl	NH_3^{+} [a]	α-H	CH, CH_2, CH_3
DL-Phenylalanine	..	2.60		5.40 m, b; $W = 19$	6.56 d; $J = 7$
Glycyl-DL-phenylalanine	2.35 d; $J = 7.5$	2.77		p: 4.93 q: $J = 5$ g: 5.88 q; $J = 5$	6.76 d; $J = 7$
DL-Tyrosine	..	2.82	2.60	5.37 m, b; $W = 17$	6.50 m, b; $W = 15$
Glycyl-DL-tyrosine	2.36 d; $J = 7.5$	3.00	2.69	ty: 5.06 q; $J = 5$ g: 5.84 q; $J = 5$	6.79 d; $J = 6$
N-Acetyl-L-tyrosine ethyl ester	1.63 d; $J = 7.5$	3.00	..	5.03 q; $J = 7$	Ac: 8.01 CH_2: 6.80 d; $J = 6$ CH_2O: 5.63 q; $J = 7.5$ CH_3: 8.66 t; $J = 7.5$
D-Leucyl-L-tyrosine	2.24 d; $J = 8.5$	2.87	2.65	ty: 4.84 m, b; $W \cong 20$ l: 5.71 m, b; $W \cong 20$	l-CH_2: 8.45 m, b; $W \cong 20$ CH_3: 9.05 ty-CH_2: 6.82 d; $J = 11$
L-Leucyl-L-tyrosine	2.26 d; $J = 8.5$	2.94	2.66	ty: 4.94 m, b; $W \cong 20$ l: 5.59 m, b; $W \cong 20$	l-CH_2: 8.16 m, b; $W \cong 19$ CH_3: 8.90 ty-CH_2: 6.72 $W = 17$
DL-Dihydroxyphenylalanine	..	2.98	2.60 m, b; $W = 24$	5.32 m, b; $W = 19$	6.47 m, b; $W = 19$
L-Tryptophan	..	Aryl: 2.72 CH: 2.82	2.98 m, b; $W = 19$	5.46 d; $J = 5$	6.40 m, b; $W = 14$
Glycyl-L-tryptophan	1.91 b, (poor)	2.70	2.41	t: 4.86 m, b; $W \cong 20$ g: 5.95 d; $J = 6$	6.45

[a] NH_3^+ resonance peaks are frequently so close to those of the aromatic protons that it is impossible to measure their half-height widths; the general appearance of the spectrum suggests that they may be rather narrow in many cases.

ture of the proper conformation, anomalously high shielding will be observed, for now the protons are in the ring's region of oppositely-directed magnetic field.[7,16-19]

The spectrum of phenylalanine (Fig. 3) shows the prominent narrow peak at 2.60 τ, arising from the aromatic protons, which in this compound happens to coincide exactly with that of the ammonium protons. The methylene proton peak is appreciably displaced to lower τ-values, but the effect becomes small at the α-proton. The peak for the aromatic protons of glycyl-DL-phenylalanine (Table III) is at somewhat lower τ-value than that of toluene (2.91)[8] because of the proximity of peptide and carboxyl groups, and is displaced to a still lower field in phenylalanine by the charged ammonium group. In tyrosine and glycyltyrosine the resonance of the ring protons is displaced about +0.23 p.p.m., compared to phenylalanine and glycylphenylalanine; a similar effect has been noted for *p*-cresol as compared to toluene.[9] The spectrum of dihydroxyphenylalanine shows a still greater positive shift, as might be expected; it should be noted that the electron-donating property of oxygen is exhibited in aromatic systems, while its electron-withdrawing behavior is made evident by large negative displacements of peaks for adjacent protons in aliphatic compounds. Both types of hydroxyl protons exchange too rapidly with the solvent to give a separate peak.

The benzylic proton peak should be split to a doublet by the adjacent α-proton. In acetyltyrosine ethyl ester this is clearly seen; in most of the other compounds the doublet is poorly resolved or unresolved.

A problem of some interest is provided by a comparison of the spectra of the diastereoisomers D-leucyl-L-tyrosine and L-leucyl-L-tyrosine. The first of these is considerably more soluble in water than the second; its amino group has a pK' of 8.3, whereas that of the LL-isomer has the appreciably lower pK' of 7.8.[20] In water solution, the configuration of the DL-isomer presumably is such that its carboxylate group is somewhat closer to the ammonium group than is the case in the LL-isomer. The n.s.r. spectra of these compounds in trifluoroacetic acid are qualitatively very similar. However, as can be seen from Table III, there is a *significant* difference in the τ-values corresponding to the side-chain of the leucyl residue. The methylene peak of the DL-isomer is displaced by +0.43 p.p.m., compared to the corresponding peak of leucine itself, an indication that these protons are in the opposing-field region of the benzene ring. Thus they are markedly closer to the face of the ring than in the LL-isomer, for which the displacement, though still noticeable, is much smaller. Molecular models provide no plausible reason for this difference. In trifluoroacetic acid the tyrosine carboxyl group is uncharged so that a direct comparison with the observations made in water solution is not possible; nevertheless, the evidence suggests a more compact structure for D-leucyl-L-tyrosine in both solvents.

In the tryptophan and glycyl-L-tryptophan spectra, the benzene ring peak appears at a τ-value close to that for the

(19) C. E. Johnson, Jr., and F. A. Bovey, *J. Chem. Phys.*, **29**, 1012 (1958).

(20) S. S. Yanari, private communication.

Table IV

Proton N.S.R. Shielding Values, τ, for Serine and Threonine and Their Glycyl Peptides in Trifluoroacetic Acid Solution

Compound	CONH	NH_3^+	α-H	CH, CH_2, CH_3
DL-Serine	..	2.25	5.46	CH_2O: 5.46
		m, b; W = 14	m, b; W = 12	m, b; W = 12
Glycyl-DL-serine	1.70	2.43	s: 5.07	CH_2O: 5.66
	d; J = 7	m, b; W = 17	m, b; W = 20	m, b; W = 15
			g: 5.66	
			m, b; W = 15	
Threonine	..	2.34	5.26	OCH: 5.53
		m, b; W = 14	d; J = 5 (?)	m, b; $W \leqq$ 19
				CH_3: 8.34
				d; J = 5
Glycyl-DL-threonine	1.75	2.50	th: 5.16	OCH: 5.66(?)
	d; J = 7	m, b; W = 17	d; J = 7	CH_3: 8.56
			g: 5.66	d; J = 5
			d; J = 5	
DL-Allothreonine	..	2.35	5.31	OCH: 5.31
		m, b; W = 15		m, b; W = 22
				CH_3: 8.42
				d; J = 5

Table V

Proton N.S.R. Shielding Values, τ, for Sulfur-containing Amino Acids and Peptides in Trifluoroacetic Acid

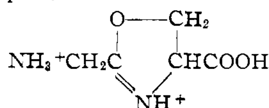
Compound	CONH	NH_3^+	α-H	CH_2, CH_3
L-Cysteine	(SH: 8.06	2.34	5.22	6.54
	t; J = 9)	m, b; W = 26	q; J = 5	d, d; J = 5 & 9
L-Cystine	..	2.25	5.20	6.39
		m, b; W = 17	m, b; W = 19	m, b; W = 14
DL-Methionine	..	2.33	5.40	$(CH_2)_2$: 7.22
		m, b; W = 14	d; J = 5	m, b; $W \leqq$ 27
Glycyl-DL-methionine	2.05	2.47	m: 5.00	SCH_3: 7.81
	d; J = 7	m, b; W = 19	d; J = 6	$(CH_2)_2$: 7.28
				m, b
			g: 5.72	SCH_3: 7.75
			d; J = 5.5	
Glutathione, L-γ-glutamyl-L-cysteinylglycine	2.10		c, gly: 5.65	CH_2S: 6.36
	m, b; W = 19		m, b	m, b
	(one peak)		glut: ~5.3	$(CH_2)_2$: 7.22
			m, b	m, b

aromatic protons in glycylphenylalanine; the two protons of the pyrrole ring appear at slightly higher τ-values. Tryptophan solutions in trifluoroacetic acid are stable for several weeks at room temperature, but ultimately undergo darkening with the formation of insoluble products.

Serine and Threonine

A solution of DL-serine made up in trifluoroacetic acid shows only two peaks when freshly prepared: ammonium at 2.25 τ and α-hydrogen plus methylene hydrogens together at 5.46 τ (Table IV). Upon standing at room temperature for only 30 minutes, the solution begins to exhibit a new peak at 4.97 τ, which becomes more prominent as aging continues. If the temperature of the solution is raised, serine undergoes a complex and drastic rearrangement or condensation. After 1 hour at 100°, the α-hydrogen–methylene pattern disappears, and peaks at lower τ-values appear which may correspond to olefinic hydrogen. The ammonium peak decreases and shifts to lower τ-value, and methyl and ethyl groups at much higher τ-values.

Glycylserine is more stable than serine in trifluoroacetic acid, but likewise undergoes reaction after several days at room temperature, the serine α-hydrogen peak apparently increasing at the expense of the methylene–glycyl–α–hydrogen peak, while the amide peak shifts 0.17 p.p.m. to higher τ-values. These changes seem to indicate the formation of the oxazoline peptide



Threonine is more stable than serine in trifluoroacetic acid, the spectrum undergoing no marked changes after several weeks at room temperature. After 1 hour at 100°, however, the spectrum is altered in a manner consistent with a partial dehydration to form α-aminocrotonic acid. The glycyl peptide undergoes similar reactions, and also perhaps partial transformation to the oxazoline peptide.

These changes in the spectra of serine and threonine and their glycyl peptides are not surprising, for it is known that these amino acids, like tryptophan, are partially destroyed during the hydrolysis of proteins with concentrated hydrochloric acid.

The spectrum of DL-allothreonine is very similar to that of threonine; however there are differences in the positions of the peaks of the side-chain protons which far exceed experimental error.

Sulfur-containing Amino Acids and Peptides

Cysteine and cystine are much less soluble in trifluoroacetic acid than are the other amino acids. Their solubilities were not measured exactly but appeared to be 5 to 8% at 25° and proved quite sufficient to obtain very satisfactory spectra. The spectrum of cysteine (Fig. 4-a and Table V) shows a well-resolved triplet (J = 9 c.p.s.) for the sulfhydryl proton, indicating that this SH proton cannot exchange rapidly with the solvent. Exchange with H_2O is reported to be rapid.[2] The methylene proton peak is a well resolved doublet of doublets arising from splitting by the SH proton (J = 9 c.p.s.) and by the α-proton (J = 5 c.p.s.). The sulfhydryl group is considerably less electronegative than the hydroxyl group,[2] the neighboring methylene protons appearing at a

τ-value nearly 1 p.p.m. higher than that of the corresponding protons of serine.

Cystine shows a spectrum very similar to that of cysteine, but which of course lacks the sulfhydryl peak. The disulfide linkage is slightly more electronegative than the sulfhydryl group, the methylene protons being displaced to a τ-value lower by 0.15 p.p.m.

The spectra of methionine (Fig. 4b) and glycylmethionine[15] are distinguished by a very narrow and prominent peak at 7.8 τ arising from the methyl group. This methyl peak is unique among all those occurring in the natural amino acids in that there are no neighboring protons to split it. As will be shown in a subsequent publication, this peak can be clearly seen in the spectra of many proteins even though the number of protons involved is a very small proportion of the total in the molecule.

The methyl peak of glycylmethionine is at a lower τ-value than that of methionine; this is the only known case in which side-chain proton peaks are displaced in this direction for the glycyl derivative.

The spectrum of glutathione is consistent with its structure, but the resolution of individual types of protons proved to be rather poor. The triplet corresponding to the sulfhydryl proton does not appear, probably as a result of poor resolution of this low peak.

TABLE VI

PROTON N.S.R. SHIELDING VALUES, τ, FOR BASIC AMINO ACIDS AND PEPTIDES IN TRIFLUOROACETIC ACID

Compound	$>NH_2^+$, $\geqslant NH^+$, $-NH_3^+$	α-H	CH, CH_2, CH_3
L-Lysine	(α): 2.33 m, b; W = 19 (ϵ): 2.96 m, b; W = 24	5.58 m, b; W = 24	$(CH_2)_3$: 8.03 m, b; W = 19 CH_2N: 6.62 m, b; W = 26
Glycyl-L-lysine	CONH: 2.03 b gly. NH_3^+: 2.47 m, b; $W \simeq 19$ (ϵ): 3.06 m, b; $W \simeq 24$	1: 5.25 m, b; (poor) g: 5.65 m, b; $W \simeq 24$	$(CH_2)_3$: 8.19 m, b; W = 22 CH_2N: 6.71 d; J = 6
L-Ornithine	(α): 2.25 m, b; W = 24 (δ): 2.86 m, b; W = 24	5.52 m, b; W = 22	$(CH_2)_2$: 7.70 m, b; W = 22 CH_2N: 6.60 m, b; W = 24
L-Arginine	α-NH_3^+: 2.27 m, b; W = 17 NH_2: 3.37 m, b; $W \simeq 30$ NH: 3.88 (poor)	5.50 m, b; W = 22	$(CH_2)_2$: 7.79 m, b; W = 19 CH_2N: 6.53 d; J = 5
Glycocyamine	3.14 t; J = 6 3.60 m, b; W = 20	5.71 d; J = 6	...
L-Histidine	NH (1,3): −2.38, −2.62, m, b $-NH_3^+$: 2.04 m, b; W = 16	5.09 m, b; W = 19	NCHN: 1.27 CH: 2.36 CH_2: 6.12 m, b; W = 16
Imidazole	−2.71 m, b; W = ?	..	NCHN: 1.17 $(CH)_2$: 2.35
L-2-Thiolhistidine	$-NH_3^+$: 2.02 m, b; W = 24	5.24 m, b; W = 19	CH: 2.45 CH_2: 6.27 m, b; W = 19
DL-1-N-Methylhistidine	$-NH_3^+$: 2.00 m, b; $W \simeq 30$	5.22 m, b; $W \simeq 20$	NCHN: 1.38 CH: 2.40 CH_2: 6.12 m, b; W = ? CH_3: 6.00
L-Histidine methyl ester	NH {−2.20, −2.60} $-NH_3^+$: 2.04 m, b; W = 17	5.04 q; J = 5	NCHN: 1.17 CH: 2.30 CH_2: 6.24 m, b; W = ? CH_3: 5.96
L-Histidyl-L-histidine*	NH {−2.42, −2.85} b CONH & NH_3^+: 1.35 m, b; W = 12	both 5.12 m, b; J = 20	NCHN: 1.35 CH: 2.54 CH*: 2.39 both CH_2: 6.34 m, b; W = 20

Basic Amino Acids and Peptides

The resonance which is shown by the ammonium protons does not appear as the "expected" widely spaced triplet ($J \backsimeq 50$ c.p.s.) but rather as a single peak having a width of about 20 c.p.s.; this is true for most amino acids and peptides, with certain exceptions such as proline. Collapse of the triplet to a broad singlet might in some cases be partly due to proton exchange with solvent, but this cannot be the chief reason in those cases where a well-resolved splitting of the α-proton peak by the ammonium protons can be observed, and it is probable that quadrupole relaxation of the nitrogen is always important. In many compounds, the peptide proton resonance is collapsed to a very narrow singlet from this cause alone.[15]

The spectrum of lysine (Table IV) shows well separated peaks for the α- and ϵ-ammonium protons, the ϵ-peak appearing at relatively higher τ-value, probably chiefly because of its greater distance from the carboxyl group. The ϵ-peak is somewhat broader and lower than the α-, but the difference is not great. In ornithine, both the α- and δ-peaks are at lower τ-values than in lysine, probably because of the greater mutual inductive effect of their positive charges. In water, the pK' of the α-amino group of lysine is 8.95[21]; the ϵ-amino group is a markedly stronger base, its pK being reported as 10.5.[21] Its greater strength is of course a consequence of its greater distance from the carboxyl group. In trifluoroacetic acid, the difference in base strengths must be considerably greater, for the carboxyl group is uncharged and can exert its full "weakening" effect on the α-amino group. In ornithine, the δ-amino group is reported to be somewhat stronger (pK' 10.76) than the ϵ-amino group in lysine.[21] This is surprising, since one would expect it to be weaker. Perhaps one of the reported values is in error. Ellenbogen[22] reports a pK' of 10.80 for the ϵ-amino group of lysine, and this seems to be more reasonable.

In arginine (Fig. 5-a), the peak width for protons of the guanidino group is much greater than that of α-amino groups. Quadrupole relaxation of the nitrogen nucleus is probably at least as great as in the ammonium group, for in the closely related glycocyamine[15] the resonance of the proton on the nitrogen linked to the CH_2 group is very narrow, and in the arginine spectrum there is an indication of this same narrow resonance (poorly resolved) at 3.88 τ. Greater breadth would be expected if the average residence time of the protons on the guanidino nitrogen were greater than that for the charged ammonium group, for then they would be able to experience a greater splitting by the three spin states of the nitrogen. A longer average residence time is not a necessary consequence of the greater base strength of the guanidino group (pK' 12.48 in arginine[21]), as the guanidino proton resonance in glycocyamine is narrow,[15] having the same width as that of most α-amino groups, 20 c.p.s. In trifluoroacetic acid, the guanidino group of glycocyamine would be expected to be weaker than that of arginine because of its closer proximity to the carboxyl group.

The position of the guanidino peak, although probably mainly determined by proximity to the electronegative carboxyl or peptide groups, may also be influenced by the rate of proton exchange with the solvent.[2] When exchange is rapid, the guanidino peak and the solvent peak will broaden (as observed); if exchange were rapid enough the two peaks would disappear and reappear, merged, at a common intermediate point. Presumably the guanidinium cation is weakly amphoteric.[2] The peaks for other weakly basic groups may be expected to be somewhat broadened for this reason also. In a subsequent publication we shall deal in more detail with this question.

The imidazole group of histidine and its derivatives is aromatic in character even when uncharged, but may be expected to have yet lower τ-values in trifluoroacetic acid because of its acquisition of a positive charge. Histidine (Fig. 5b) shows peaks at −2.38 and −2.62 τ, even lower than the carboxyl protons of the solvent. These are due to the two nitrogen-bound protons of the ring; it is not certain which is which, but it is probable that the proton in the 1-position is at the lower field. The proton in the 2-position appears at 1.27 τ and the remaining ring proton at 2.36 τ. These assignments are confirmed by the peak heights and positions of the spectrum of imidazole (Table VI). The assignment of the 1.27 τ peak is also confirmed by the spec-

(21) E. J. Cohn and J. T. Edsall, "Proteins, Amino Acids, and Peptides," Reinhold Publishing Corp., New York, N. Y., 1943, pp. 84–85.

(22) E. Ellenbogen, THIS JOURNAL, 74, 5198 (1952).

TABLE VII

PROTON N.S.R. SHIELDING VALUES, τ, FOR ACIDIC AMINO ACIDS AND PEPTIDES IN TRIFLUOROACETIC ACID

Compound	$CONH_2$	NH_3^+	α-H	CH_2
L-Aspartic acid	..	2.24 m, b; $W = 16$	5.31 d; $J = 5$	6.45 m, b; $W = 16$
L-Asparagine	2.68 m, b; $W \cong 17$	2.15 m, b; $W \cong 18$	5.33 m, b; $W = 17$	6.51 m, b; $W = 12$
Glycyl-L-asparagine	CONH: 1.73 d; $J = 7$ $CONH_2$: 2.48 (in NH_3^+ peak)	2.48 m, b; $W = 21$	a: 4.77 m, b; $W = 19$ g: 5.67 m, b; $W = 19$	6.73 m, b; $W = 17$
L-Glutamic acid	..	2.32 m, b; $W = 15$	5.45 d; $J = 5$	β: 7.18 γ: ~7.45 (poor) both m, b
Glycyl-L-glutamic acid	2.12 d; $J \cong 7$	2.39 m, b; $W = 24$	glut.: 5.12 m, b; $W = ?$ gly.: 5.76 m, b; $W \cong 22$	$(CH_2)_2$: 7.42 m, b; $W = 24$
L-Glutamine	1.74	2.20 m, b; $W = 14$	5.47 m, b; $W = 22$	$(CH_2)_2$: 7.26 m, b; $W = 18$

trum of 2-thiolhistidine, in which this peak is missing. (In this compound, the sulfhydryl hydrogens, unlike those of cysteine, evidently exchange rapidly with the solvent and give rise to no distinguishable peak.)

The spectral peak positions of 1-N-methylhistidine and L-histidine methyl ester are in accordance with expectation. The interpretation of the spectrum of L-histidyl-L-histidine (Fig. 5-c) shows the α-ammonium resonance to be at an unusually low τ-value, as compared to other dipeptides. The pK' for the α-amino group of histidylhistidine in water is 7.80,[21] appreciably lower than the value of 8.2 ± 0.1 characteristic of dipeptides having one basic group and one carboxyl.[21] In trifluoroacetic acid this difference may be expected to be considerably exaggerated because the α-amino group is being observed in a medium of low dielectric constant and under conditions where both imidazole groups are probably positively charged. The displacement of the peak due to the α-amino group is too great to be attributed to the direct inductive effect of the nearest charged imidazole group, but may well result from π-electron ring currents.[7,16-19] The peaks of the α-amino groups of histidine and its other one-ring derivatives (Table VI) are also at somewhat lower τ-values than those of most other compounds reported in this study, but the displacement is much less than found for histidylhistidine.

Acidic Amino Acids and Peptides

The peak positions of the acidic amino acids, aspartic and glutamic, and of asparagine and glutamine, together with two glycyl peptides, are given in Table VII; the spectrum of glutamic acid in Fig. 6. The amide proton peak for asparagine is at considerably higher τ-value than that of glutamine; in glycylasparagine, a similar deviation is seen, the peak due to the peptide proton (identifiable by virtue of its splitting) appearing at considerably lower τ-values than that for the $CONH_2$ protons, which is evidently included under the ammonium peak.

Acknowledgment.—The authors gratefully acknowledge the important contributions to this work made by Mr. George Filipovich, who maintained and operated the n.s.r. equipment, and by Miss Lucetta Stifter, who prepared the solutions for measurement.

ST. PAUL 6, MINN.

[CONTRIBUTION FROM THE DEPARTMENT OF ORGANIC CHEMISTRY, UNIVERSITY OF ATHENS, AND THE LABORATORY OF BIOCHEMISTRY, NATIONAL CANCER INSTITUTE, NATIONAL INSTITUTES OF HEALTH]

Studies on Arginine Peptides. II. Synthesis of L-Arginyl-L-arginine and other N-Terminal Arginine Dipeptides

BY LEONIDAS ZERVAS, THEODORE T. OTANI, MILTON WINITZ AND JESSE P. GREENSTEIN

RECEIVED NOVEMBER 1, 1958

N^α,N^ω-Dicarbobenzoxy-L-arginine and tricarbobenzoxy-L-arginine have been utilized for the preparation of a variety of dipeptides containing an N-terminal arginine residue. The former compound, either as its acid chloride–hydrochloride derivative or under the condensing action of dicyclohexylcarbodiimide, was coupled with diethyl L-glutamate and the condensation product subsequently converted to the corresponding free peptide *via* successive saponification and catalytic hydrogenolysis; the formation of N^α,N^ω-dicarbobenzoxyanhydro-L-arginine accompanied the coupling reaction, an occurrence presumably attributable to the fact that the basicity of the guanido moiety of N^α,N^ω-diacylated arginines is not completely masked. No analogous evidence of intramolecular cyclization was revealed in the case of the more exhaustively protected tricarbobenzoxy-L-arginine upon like condensation, *via* its mixed carbonic–carboxylic acid anhydride derivative, with the benzyl ester derivatives of L-alanine, L-aspartic acid, L-glutamic acid, glycine, L-isoleucine, D-alloisoleucine, L-leucine, L-phenylalanine, L-tyrosine and L-valine. Catalytic hydrogenolysis of the tricarbobenzoxy-L-arginylamino acid benzyl esters so procured led to the corresponding dipeptides, which were isolated in high over-all yield. A comparable condensation between tricarbobenzoxy-L-arginine and benzyl N^ω-carbobenzoxy-L-argininate permitted the ultimate synthesis of L-arginyl-L-arginine, which was isolated and characterized as its diflavianate and dipicrolonate derivatives.

Introduction

The successful union of an arginine residue with another amino acid residue in peptide linkage requires that the basicity of the former be adequately masked prior to its implication in the peptide-forming step. For such purpose, the basicity of the