BBA 4154

ELECTRONIC STRUCTURE OF THE α-AMINO ACIDS OF PROTEINS

I. CHARGE DISTRIBUTIONS AND PROTON CHEMICAL SHIFTS

G. DEL RE*, B. PULLMAN AND T. YONEZAWA
Université de Paris, Institut de Biologie Physico-Chimique, 13 rue Pierre Curie, Paris (France)

(Received March 26th, 1963)

SUMMARY

The method of molecular orbitals of quantum chemistry has been used, in the approximation developed by DEL RE, for the study of the distribution of the electronic charges in the α-amino acids of proteins, in their neutral, dipolar, cationic and anionic forms. The results of the calculation are correlated with the chemical shifts observed in nuclear magnetic resonance, in different solvents.

A satisfactory over-all parallelism can be observed between the proton shifts (δ_H) and the σ-electronic charges of the hydrogens (Q_H), as well as between the shifts and the charges of the neighbouring carbons (Q_C). The results may be best presented by the general equation: $\delta_H = -AQ_C - BQ_H + C$, the method of least squares leading to the values $A = 9.92$, $B = 133.93$ and $C = 9.67$ for data relating to alkaline and acidic solutions. This indicates in agreement with general theory, that the contribution of Q_C to the proton chemical shift is much smaller than that of Q_H.

INTRODUCTION

The application of quantum theory to biochemistry has been essentially concerned with the π-electron systems of conjugated molecules[1]. Although this is justified to a large extent by the preponderant role played by these systems in the processes of life[2,3], the situation is nevertheless not entirely satisfactory. In fact, it is essentially a consequence of the technical difficulties involved in the application of quantum chemistry to σ-electron systems, Recently, solutions to these difficulties have been suggested by different authors[4-11]. In this paper, the method proposed by one of us[10] for the determination of the energy levels and the charge distributions in saturated molecules (hydrocarbons and heterocompounds) is used for the study of the electronic structure of the α-amino acids of proteins.

The method, the details of which were presented previously[10], may be considered as the counterpart, for σ-bonds, of the HÜCKEL method as used for π-electrons. In fact, it employs the same formalism as the general HÜCKEL procedure, but treats all the σ-bonds as strictly localized units, and, instead of delocalized orbitals, introduces an inductive effect (in a way similar to that originally proposed by WHELAND AND PAULING[12]) as the main feature describing the influence of the neighbouring atoms

---

[*] Universita de Napoli, Instituto di Fisica Teorica, Naples (Italy).

on each particular bond of a given molecule. The method, like the HÜCKEL method, is a semi-empirical one, and requires therefore the choice of a suitable set of parameters. Those suggested in the original paper[10] lead to charge distributions which correspond to dipole moments in satisfactory agreement with experiment; they enable also to account satisfactorily for a number of physico-chemical properties of molecules (*e.g.*, dissociation constants) and for their chemical reactivities. The procedure represents thus a satisfactory starting point for the first investigation of the properties of saturated biomolecules.

The choice of α-amino acids as the first group of such compounds to be studied quantum mechanically, was motivated by their importance in biology and by the fact that they form a large series of related substances: it is in such series that a semi-empirical method is at its best.

In the present paper, we shall describe the results concerning the distribution of the σ-electronic charges in the α-amino acids, and discuss them in connection with the proton chemical shifts, recently measured for these compounds in nuclear magnetic resonance studies[13-15]. As will be mentioned in more detail later, these shifts are strictly related to the charges of the hydrogen atoms and of their neighbours. Therefore, they constitute excellent experimental data to be used for testing the validity of the theoretical calculations of charge distributions. As will be seen, a simple relation does indeed provide a good fitting of the experimental shifts with the calculated charges. The few instances where any serious disagreement between the two sets of data has been found correspond to compounds where an important role is played by effects not directly related to σ-charges. Among such effects are *e.g.* those due to the ring currents in aromatic molecules.

An approximate estimate of the corresponding corrections can nevertheless be made sometimes (*vide infra*).

THE CALCULATIONS

As the method used is fully described in ref. 10 we shall here make only a few remarks concerning the choice of the parameters, and give a numerical example of the actual calculations.

The parameters used in the calculations are of three kinds. The bond parameters, ε_{XY}, correspond strictly to the exchange integrals of the usual HÜCKEL method[16]. The two other parameters (δ_X^0 and $\gamma_{X(Y)}$) are used to determine the Coulomb integrals δ_X of the various atoms, according to the formula:

$$\delta_X = \delta_X^0 + \sum_{Y \text{ adj. } X} \gamma_{X(Y)} \delta_Y \qquad (1)$$

Here the inductive parameters $\gamma_{X(Y)}$ indicate the fraction of the Coulomb integral of Y that has to be included in the Coulomb integral of X, in order to take into account the inductive effect of Y on X.

The numerical values of the parameters are reported in Table I. Most of them are taken from Ref. 10, the only novelties being the parameters for bonds containing S and $O^{-\frac{1}{2}}$. The value of δ_S^0 was taken equal to δ_C^0, because the electronegativities of sulfur and carbon are the same; the inductive parameters $\gamma_{C(S)}$ and $\gamma_{S(C)}$ were taken equal to $\gamma_{C(Cl)}$ and $\gamma_{Cl(C)}$, in accordance with the rule, introduced in ref. 10, that they should be the same for atoms of the same rows of the periodic table. Finally,

ε_{SC} was estimated by extrapolating the data for ε_{CO}, ε_{CF}, and ε_{CCl}. An entirely similar procedure was followed for the SH and SS bonds, the inductive parameters being estimated from dipole moments. The parameters for $O^{-\frac{1}{2}}$ were determined, as had been done in ref. 10 for the parameters of N^{+}, by an interpolation taking into account the change in electronegativity due to the presence of the formal charge.

TABLE I

PARAMETERS ADOPTED

Bond	*C-H*	*C-C*	*C-N*	*C-N+*	*C-O*	*C-O-½*
ε_{AB}	1.00	1.00	1.00	1.33	0.95	0.80
$\gamma_{A(B)}$	0.30	0.10	0.10	0.10	0.10	0.10
$\gamma_{B(A)}$	0.40	0.10	0.10	0.10	0.10	0.10
δ_A^0	0.07	0.07	0.07	0.07	0.07	0.07
δ_B^0	0.00	0.07	0.24	0.31	0.40	0.33
Bond	*N-H*	*N+-H*	*C-H*	*C-S*	*S-H*	*S-S*
ε_{AB}	0.45	0.60	0.45	0.75	0.70	0.60
$\gamma_{A(B)}$	0.30	0.30	0.30	0.20	0.30	0.10
$\gamma_{B(A)}$	0.40	0.40	0.40	0.40	0.40	0.10
δ_A^0	0.24	0.31	0.40	0.07	0.07	0.07
δ_B^0	0.00	0.00	0.00	0.07	0.00	0.07

As an example of the calculations, we describe explicitly here the calculation for glycine, the simplest amino acid, in its neutral form.

The Coulomb parameters are determined from a system of equations corresponding to Eqn. 1. With the parameters given in Table I, and the numbering of the atoms shown in Fig. 1, this system takes the explicit form (2).

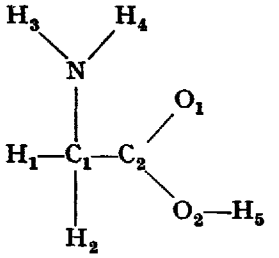


Fig. 1. Numbering of atoms in glycine.

$$\begin{aligned}
&\delta_{H1} - 0.40\,\delta_{C1} = 0.00\\
&\delta_{H2} - 0.40\,\delta_{C1} = 0.00\\
&\delta_{C1} - 0.30\,\delta_{H1} - 0.30\,\delta_{H2} - 0.10\,\delta_N - 0.10\,\delta_{C2} = 0.07\\
&\delta_N - 0.10\,\delta_{C1} - 0.30\,\delta_{H3} - 0.30\,\delta_{H4} = 0.24\\
&\delta_{H3} - 0.40\,\delta_N = 0.00\\
&\delta_{H4} - 0.40\,\delta_N = 0.00\\
&\delta_{C2} - 0.10\,\delta_{C1} - 0.10\,\delta_{O1} - 0.10\,\delta_{O2} = 0.07 \qquad (2)\\
&\delta_{O1} - 0.10\,\delta_{C2} = 0.40\\
&\delta_{O2} - 0.10\,\delta_{C2} - 0.30\,\delta_{H5} = 0.40\\
&\delta_{H5} - 0.40\,\delta_{C2} = 0.00
\end{aligned}$$

The values obtained for the nine unknowns are:

$$\begin{aligned} \delta_{H_1} &= \delta_{H_2} = 0.064 \\ \delta_{C_1} &= 0.159 \\ \delta_{N} &= 0.337 \\ \delta_{H_3} &= \delta_{H_4} = 0.135 \\ \delta_{C_2} &= 0.175 \\ \delta_{O_1} &= 0.418 \\ \delta_{O_2} &= 0.474 \\ \delta_{H_5} &= 0.190 \end{aligned} \tag{3}$$

The atomic charges can now be obtained as follows[16]: for each bond XY one solves the secular equation:

$$\begin{bmatrix} \alpha_X - E & \beta_{XY} \\ \beta_{XY} & \alpha_Y - E \end{bmatrix} = 0 \tag{4}$$

where $\alpha_X = \alpha + \delta_X\beta$; $\beta_{XY} = \varepsilon_{XY}\beta$; its lower root E_B, when inserted into the linear system of equations:

$$\begin{aligned} (\alpha_X - E)C_1 + \beta_{XY}C_2 &= 0 \\ \beta_{XY}C_1 + (\alpha_Y - E)C_2 &= 0 \end{aligned} \tag{5}$$

gives the coefficients C_1 and C_2 of the bonding orbital,

$$\psi = C_1\chi_X + C_2\chi_Y$$

(here χ_X and χ_Y are appropriate atomic orbitals of X and Y). The net charge appearing on X as a result of the formation of the XY bond is therefore:

$$q_{X(Y)} = 1 - 2\,C_1^2 \tag{6}$$

Solving Eqns. 4 and 5 and substituting into Eqn. 6, one gets the final formula:

$$q_{X(Y)} = \frac{\delta_Y\delta_X}{2\varepsilon_{XY}} \left(\frac{1}{\sqrt{1 + \left(\frac{\delta_Y - \delta_X}{2\delta_{XY}}\right)^2}} \right) \tag{7}$$

This formula has been actually used in our calculations. It gives, for instance, the following contributions to the net charge of N:

$$\begin{aligned} q_{N(H_3)} &= -0.218 \\ q_{N(H_4)} &= -0.218 \\ q_{N(C_1)} &= -0.089 \end{aligned}$$

Clearly, the total net charge of an atom is given by:

$$Q_X = \sum_Y q_{X(Y)} \tag{8}$$

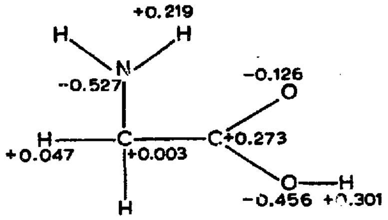
In the case of the nitrogen atom of glycine, we get,

$$Q_N = -0.525$$

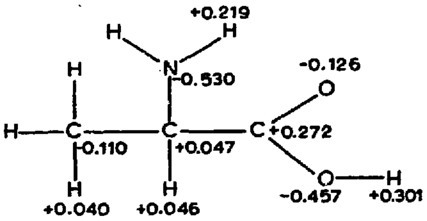
An entirely similar procedure gives the results for all the other atoms (see Fig. 2).

FIGURE 2

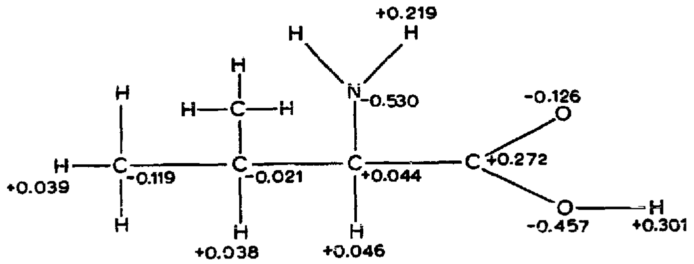
Distribution of electronic charges in α-amino acids (neutral form).



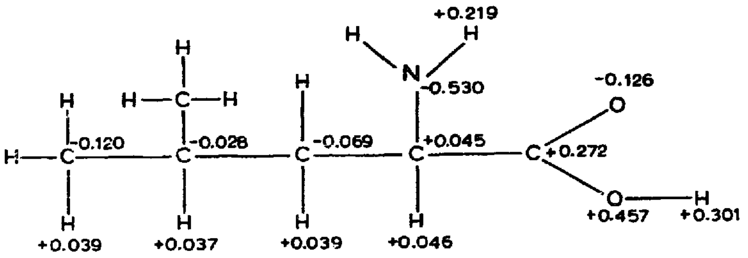
1.Glycine



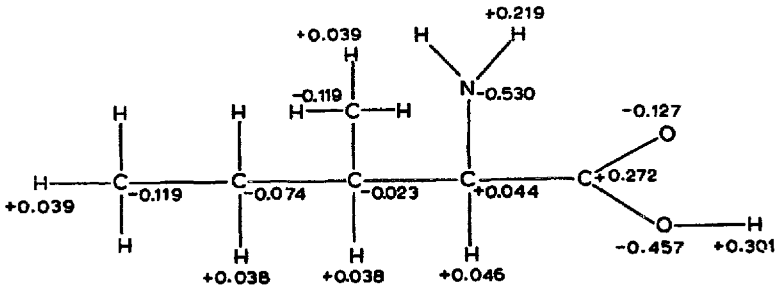
2.Alanine



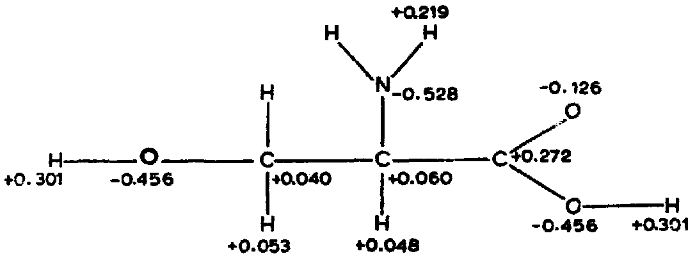
3.Valine



4. Leucine

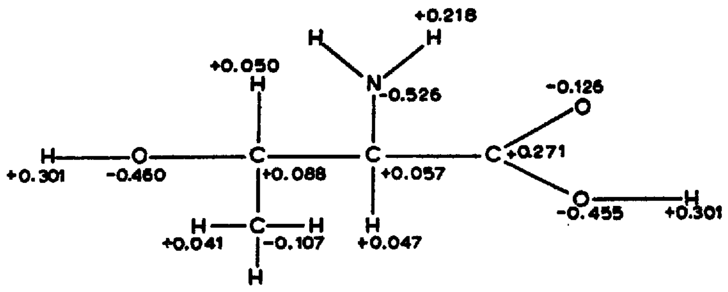


5. Isoleucine

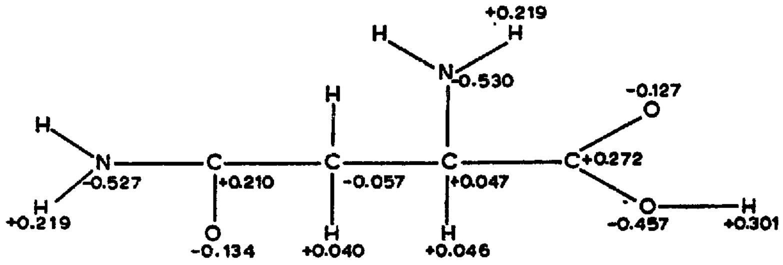


6. Serine

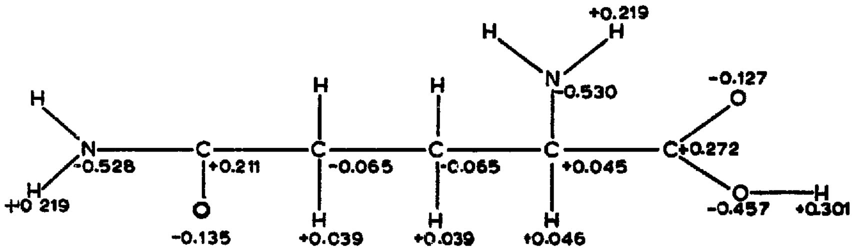
Fig. 2. (1–6).



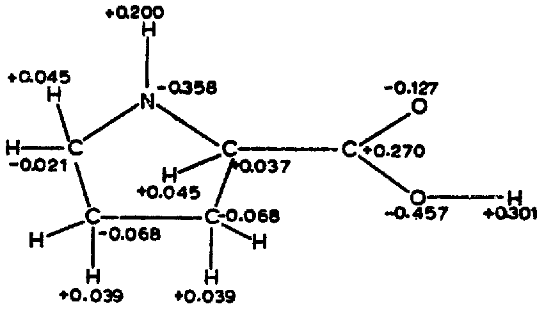
7. Threonine



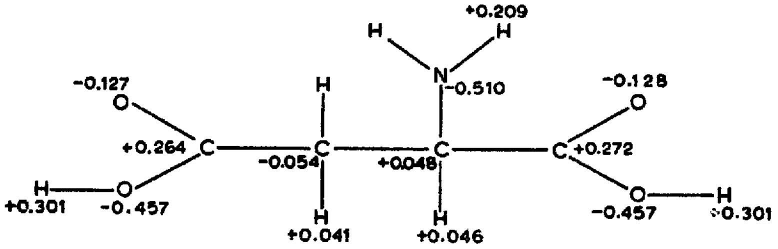
8. Asparagine



9. Glutamine

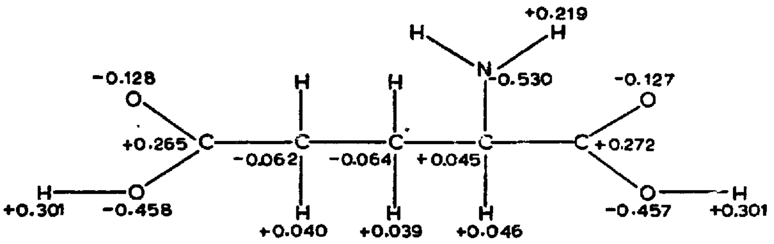


10. Proline

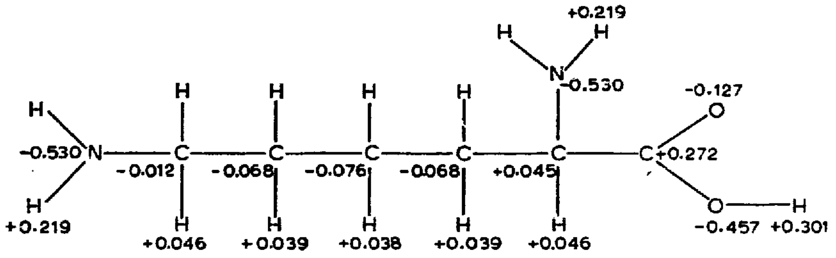


11. Aspartic acid

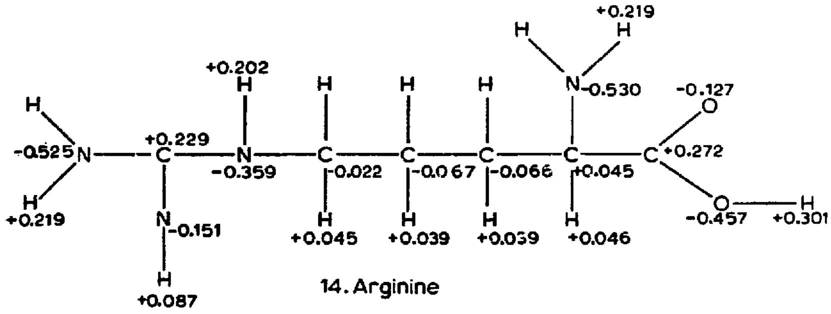
Fig. 2. (7–11)



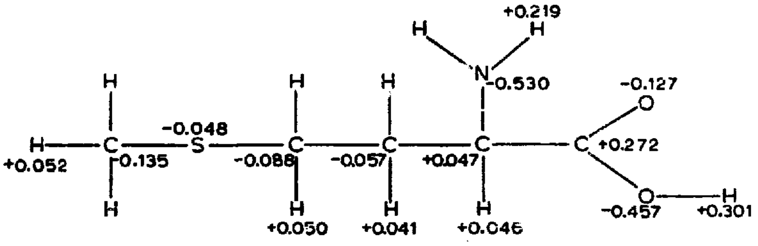
12. Glutamic acid



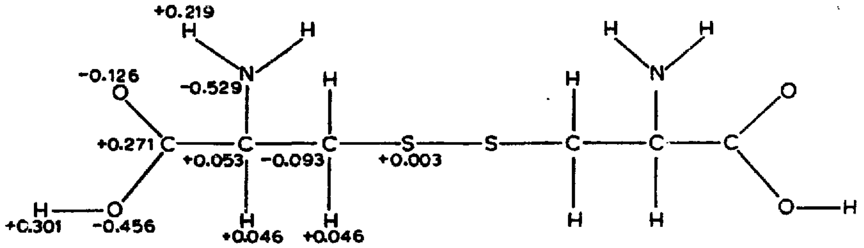
13. Lysine



14. Arginine

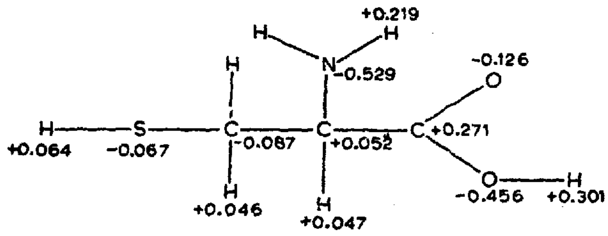


15. Methionine

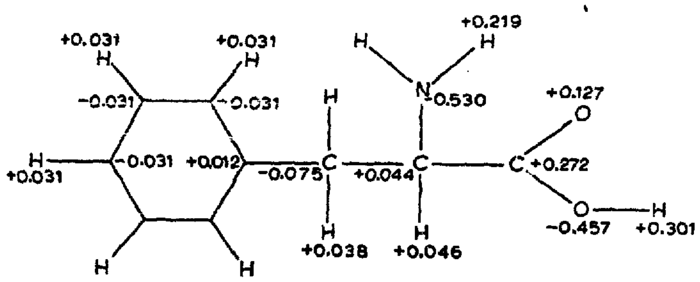


16. Cystine

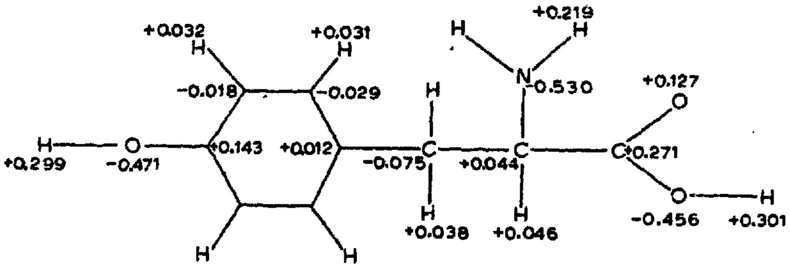
Fig. 2. (12–16)



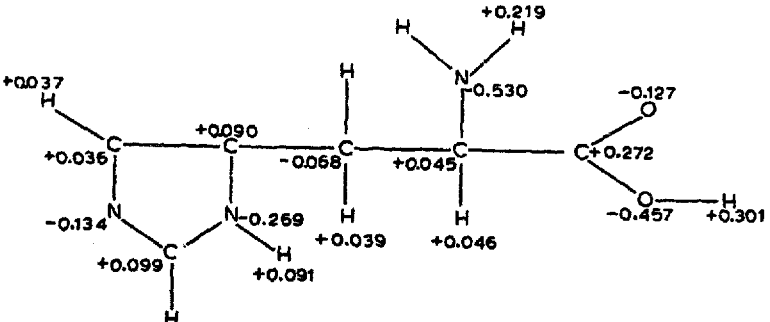
17. Cysteine



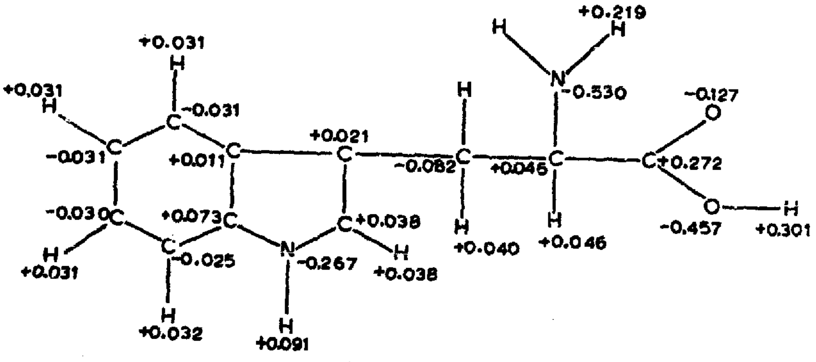
18. Phenylalanine



19. Tyrosine

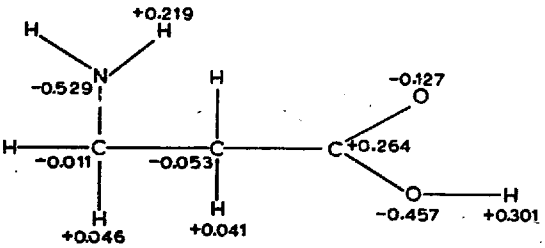


20. Histidine

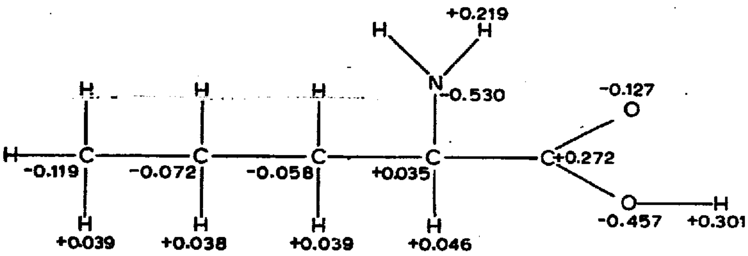


21. Tryptophan

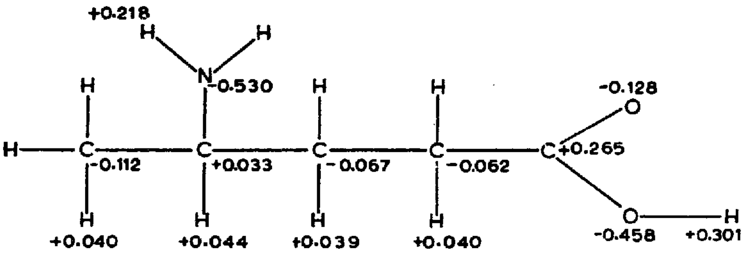
Fig. 2. (17–21)



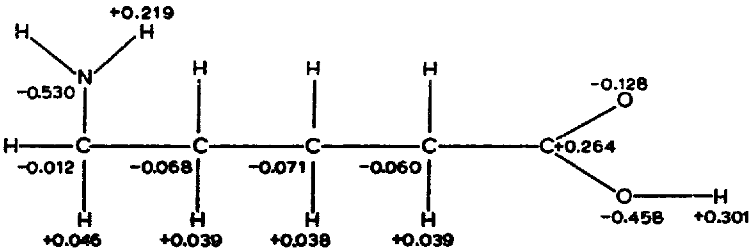
22. β-alanine



23. α-aminovaleric acid



24. γ-aminovaleric acid



25. γ-aminonorvaline

Fig. 2. (22–25)

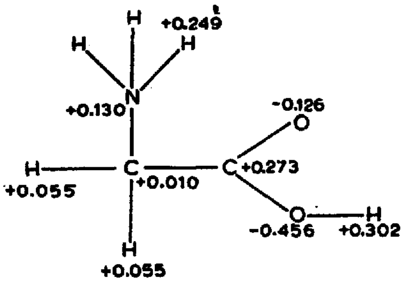
σ-ELECTRON DISTRIBUTIONS

The calculations of σ-electron distributions have been carried out for four different forms of 25 amino acids of biochemical importance, 21 of which are constituents of proteins. The forms considered are the uncharged (RNH_2COOH), the cation (RN^+H_3COOH), the anion (RNH_2COO^-), and the dipolar ion ($RN^+H_3COO^-$) forms.

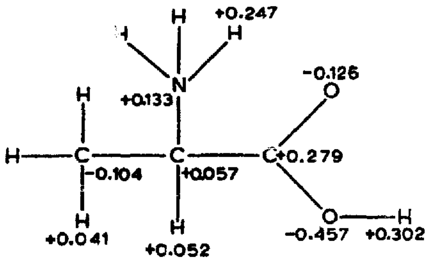
In some amino acids, *e.g.* aspartic acid, glutamic acid, lysine, arginine, histidine, etc., other ionic structures are undoubtedly present. These structures, however, have been omitted from the present investigation. We stress also that as regards the aromatic amino acids, phenylalanine, tyrosine, histidine and tryptophan, only the σ-electrons of their aromatic rings have been taken into account in the calculations;

FIGURE 3

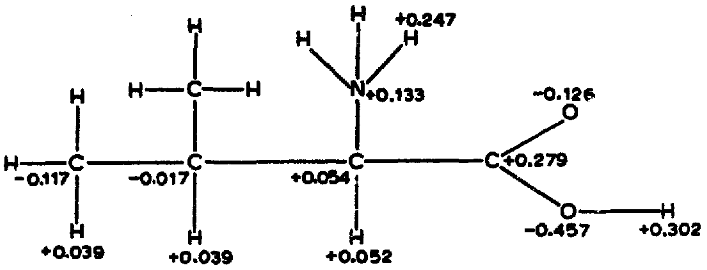
Distribution of electronic charges in α-amino acids (cations).



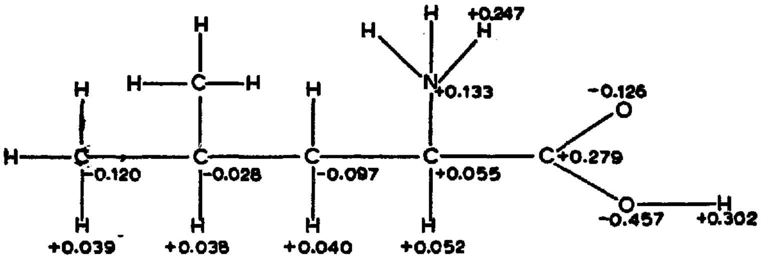
1. Glycine



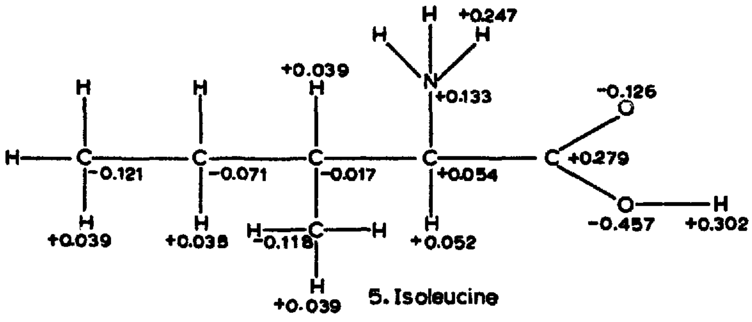
2. Alanine



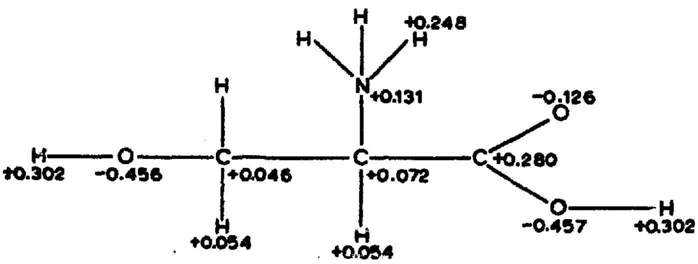
3. Valine



4. Leucine

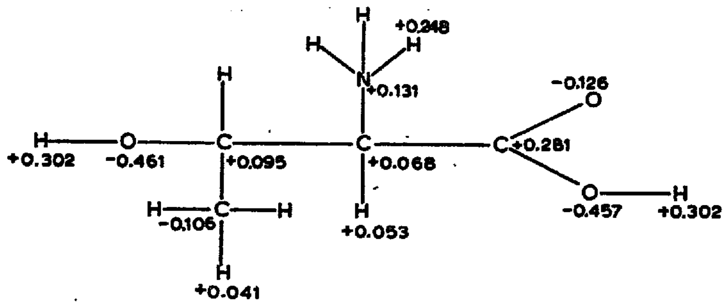


5. Isoleucine

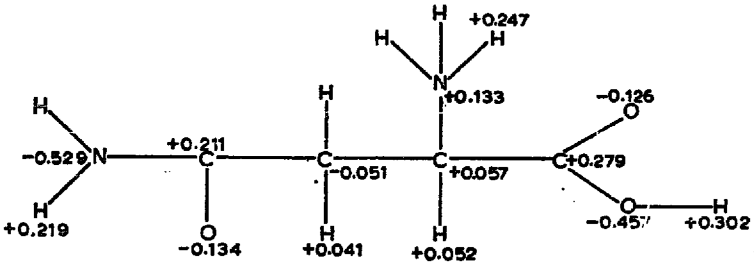


6. Serine

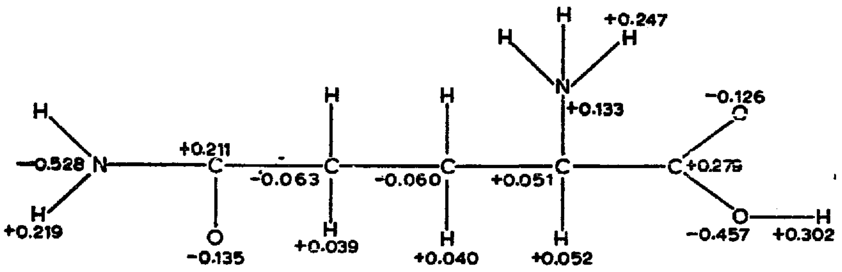
Fig. 3. (1–6)



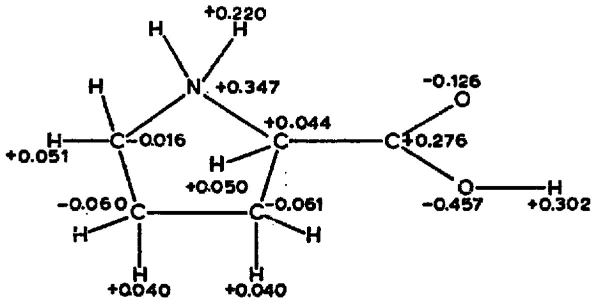
7. Threonine



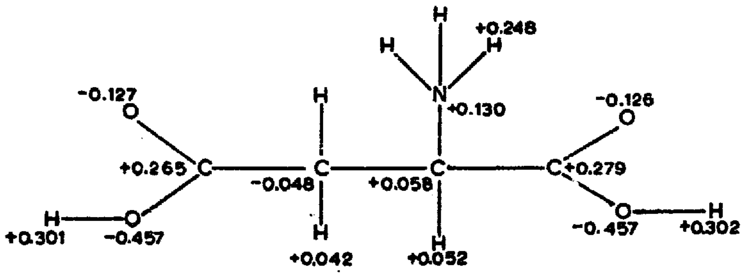
8. Asparagine



9. Glutamine

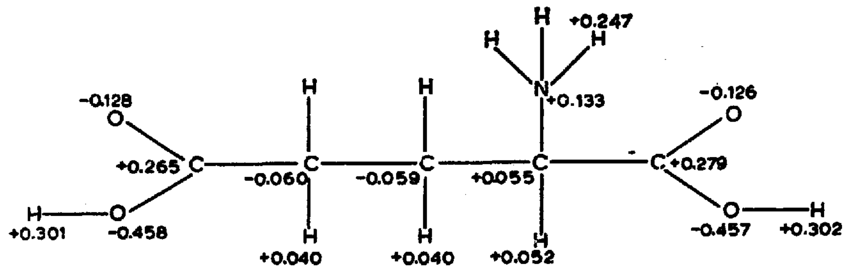


10. Proline

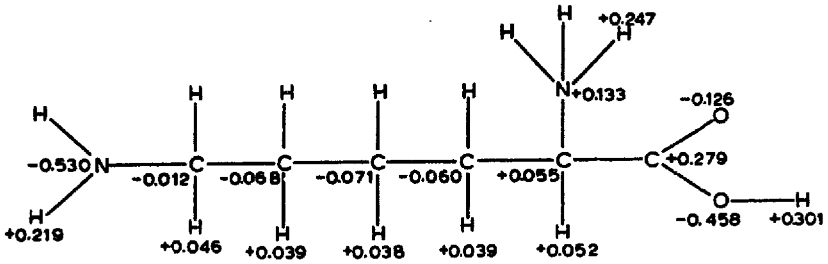


11. Aspartic acid

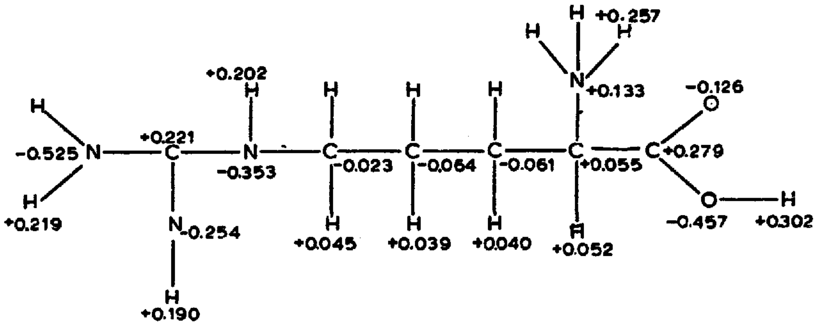
Fig. 3. (7–11)



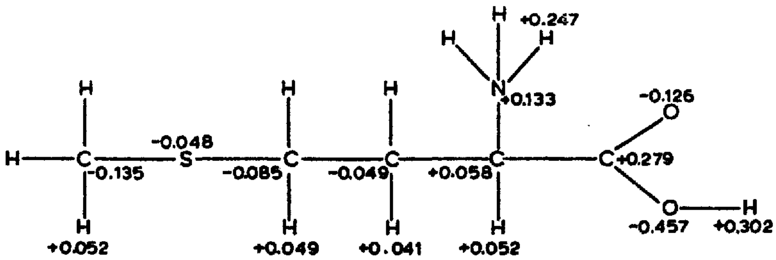
12. Glutamic acid



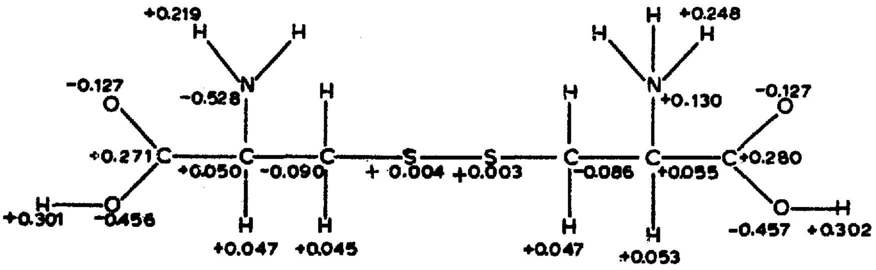
13. Lysine



14. Arginine

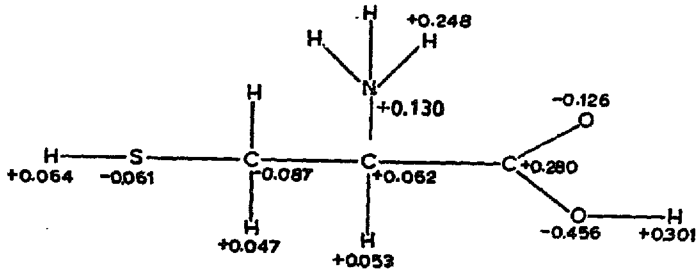


15. Methionine

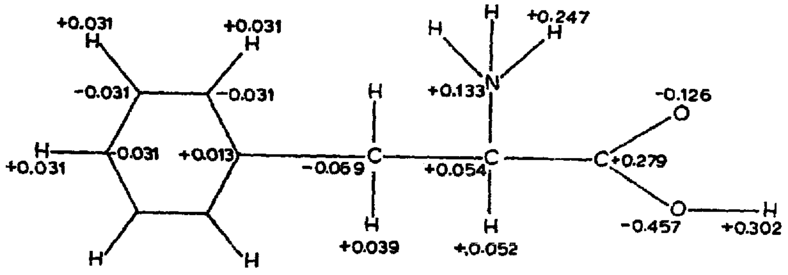


16. Cystine

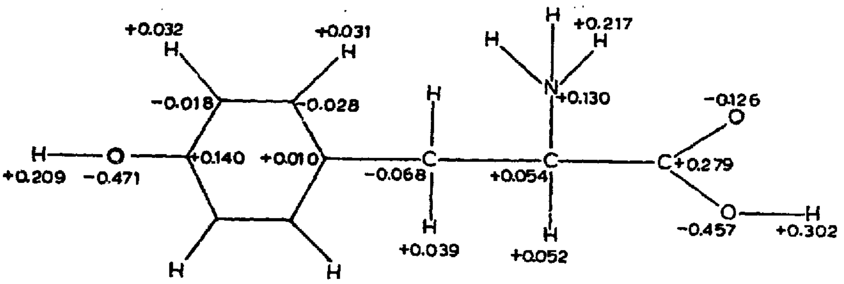
(Fig. 3. 12–16)



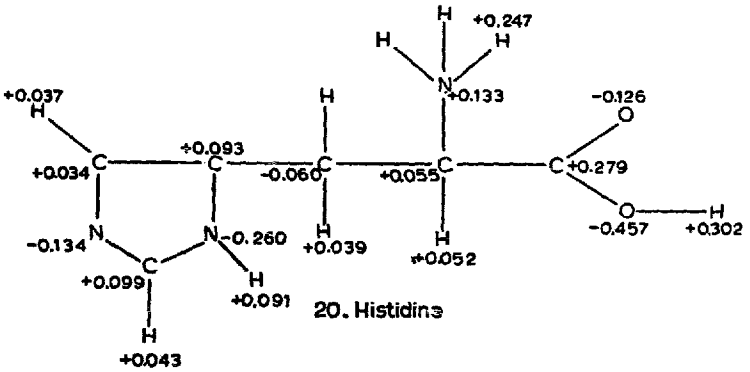
17. Cysteine



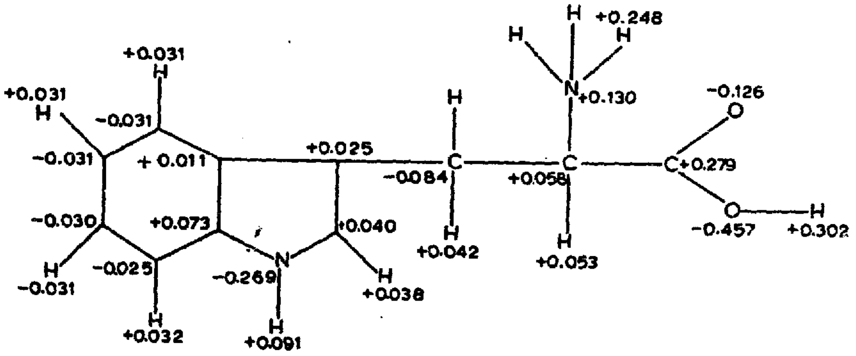
18. Phenylalanine



19. Tyrosine

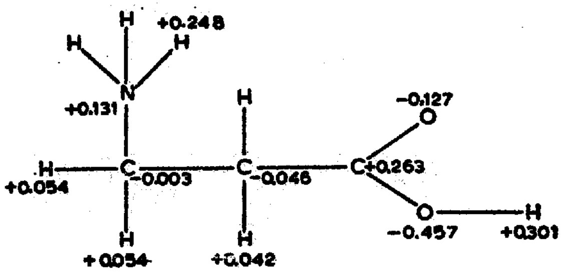


20. Histidine

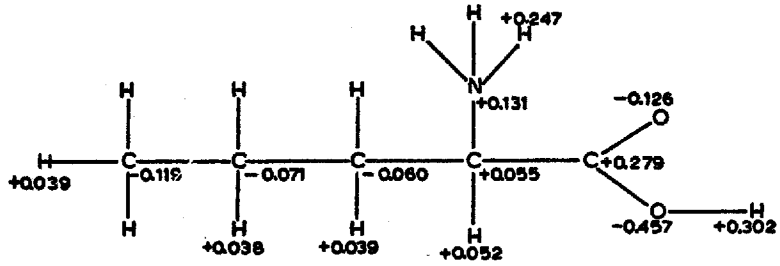


21. Tryptophan

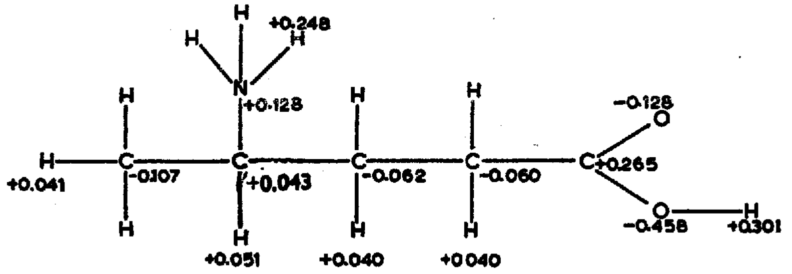
Fig. 3. (17–21)



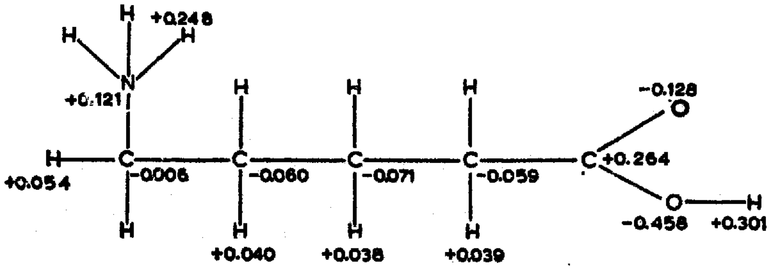
22.β-alanine



23. α-aminovaleric acid



24.γ-aminovaleric acid



25. γ-aminonorvaline

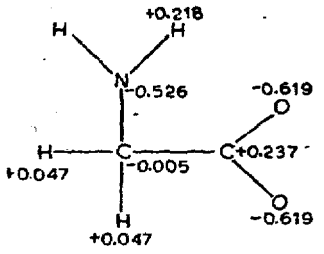
Fig. 3. (22–25)

the difference in hybridization between the carbon atoms in the two parts of these molecules has also been neglected.

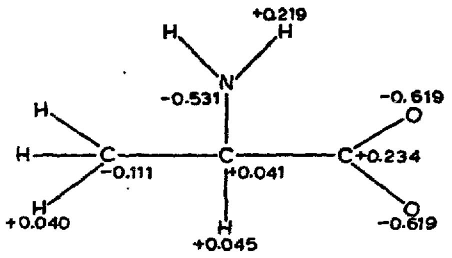
The results of the calculations are shown in Figs. 2–5. The numbers written on the diagrams are the calculated net charges of the various atoms in atomic units (charges of equivalent atoms are given only once). It may be useful to remark that, as a charge of $-0.5e$ was assumed to be localized originally on each oxygen atom in the COO^- group, the net charge of an oxygen atom of this group in Figs. 4 and 5 is $-0.5 + Q_O$, Q_O being the net charge induced on the oxygen by the inductive effect of the environment. Similarly, the charge of the nitrogen atom in the N^+H_3 group in Figs. 3 and 5 is actually equal to $1 + Q_N$.

FIGURE 4

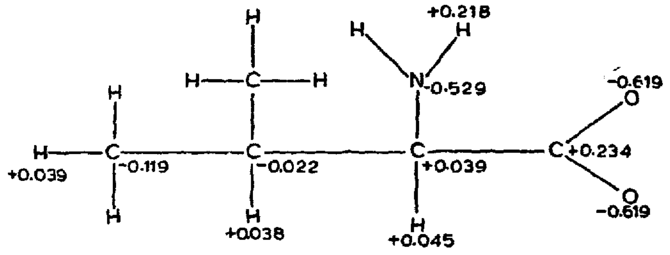
Distribution of electronic charges in α-amino acids (anions).



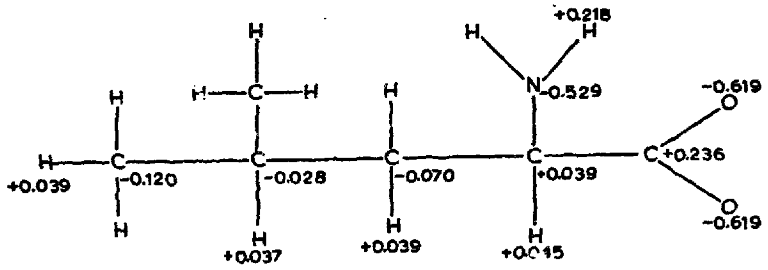
1. Glycine



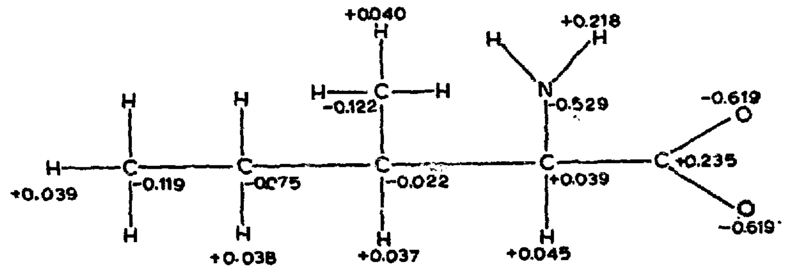
2. Alanine



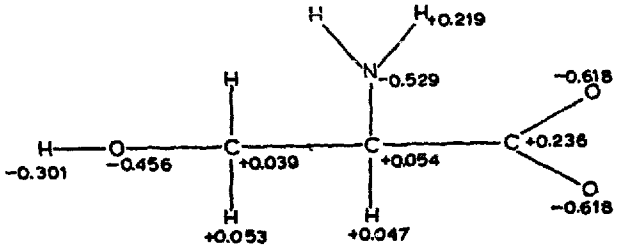
3. Valine



4. Leucine

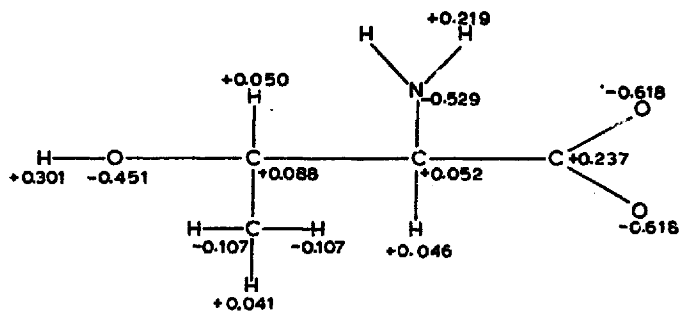


5. Isoleucine

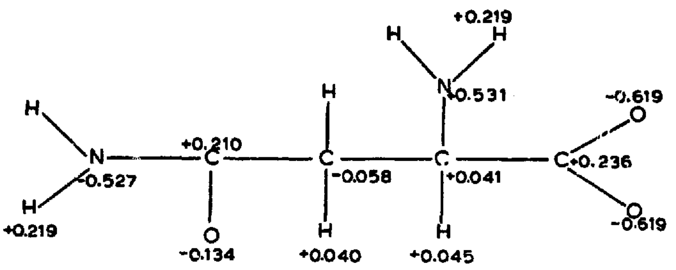


6. Serine

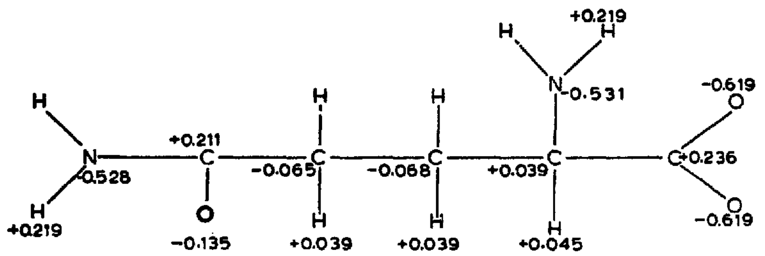
Fig. 4. (1–6)



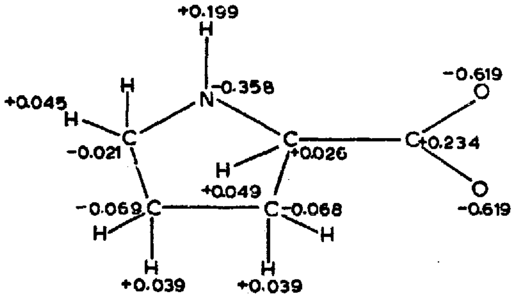
7. Threonine



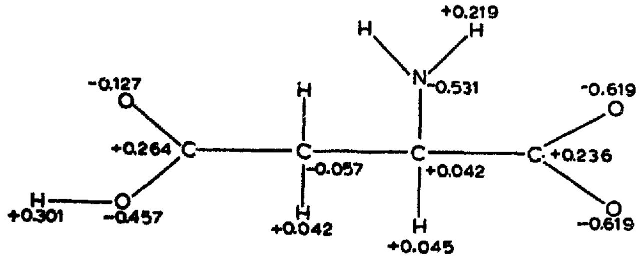
8. Asparagine



9. Glutamine

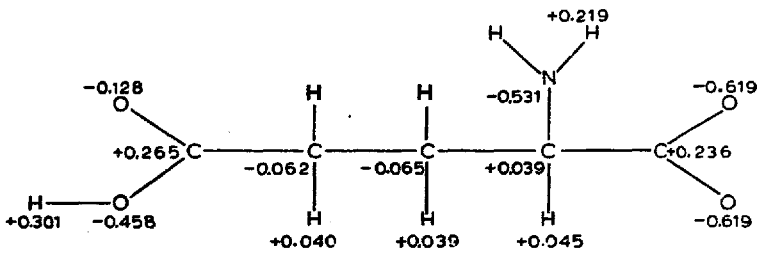


10. Proline

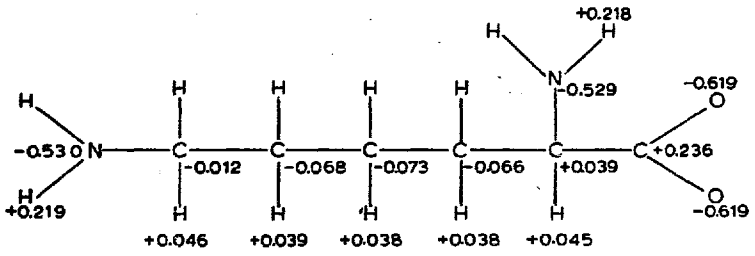


11. Aspartic acid

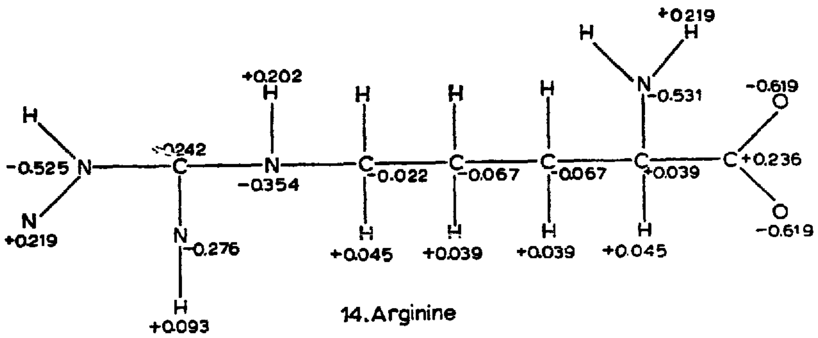
Fig. 4. (7–11)



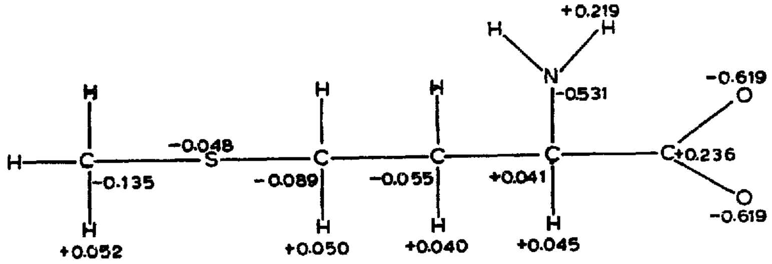
12. Glutamic acid



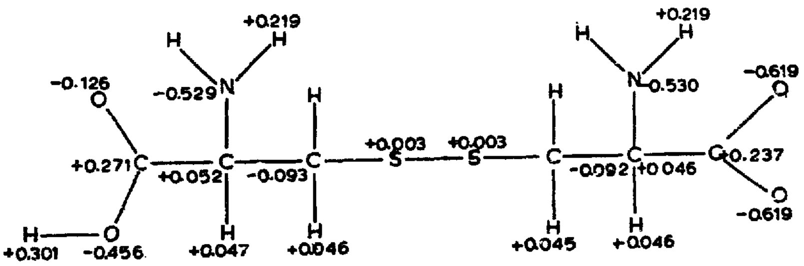
13. Lysine



14. Arginine

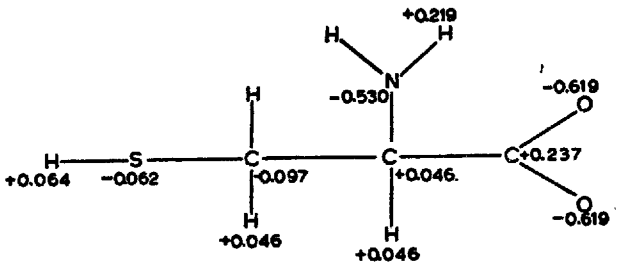


15. Methionine

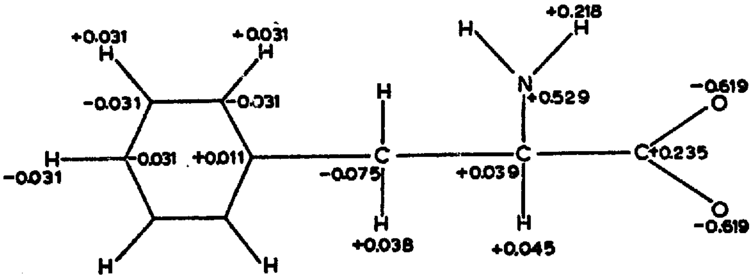


16. Cystine

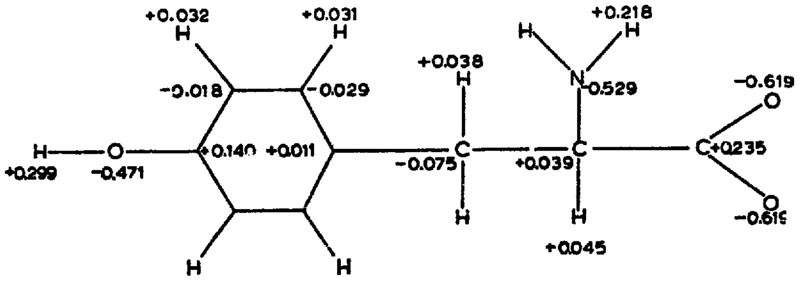
Fig. 4. (12–16)



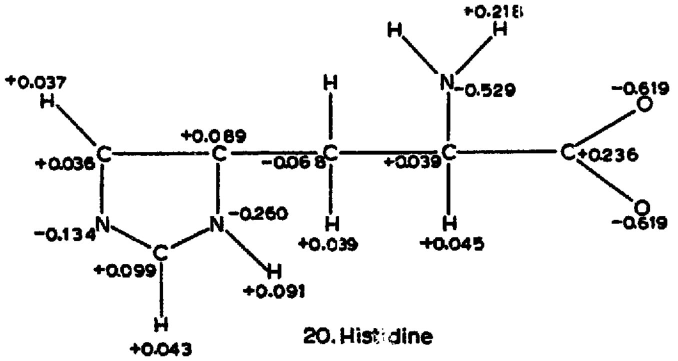
17. Cysteine



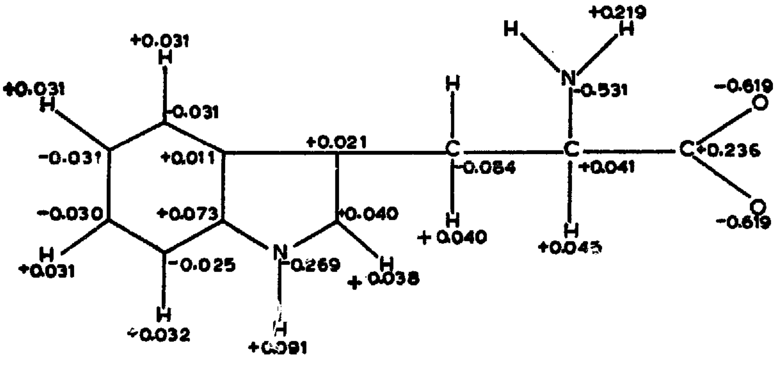
18. Phenylalanine



19. Tyrosine

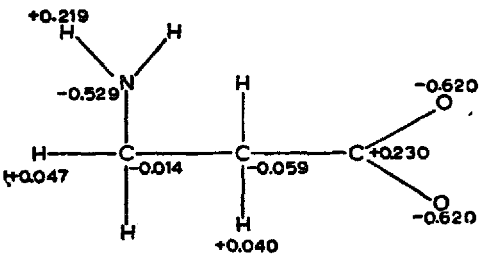


20. Histidine

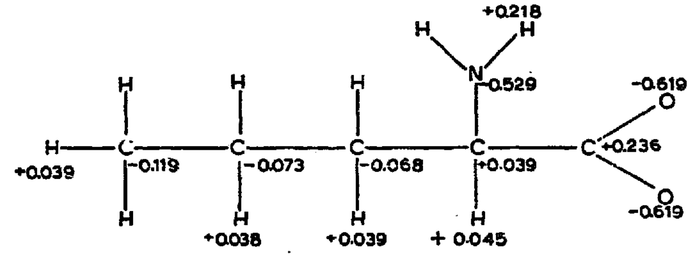


21 Tryptophan

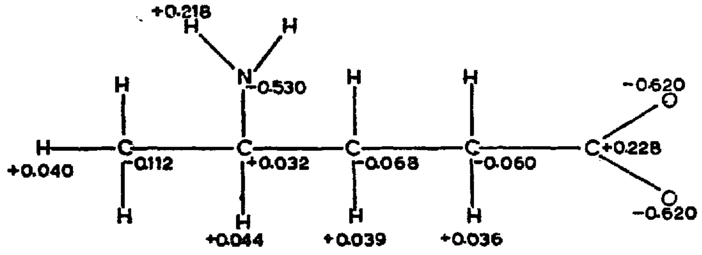
Fig. 4. (17–21)



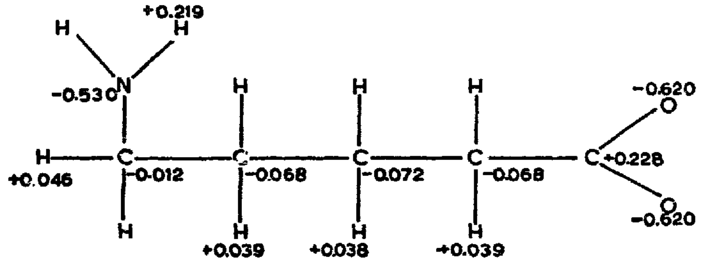
22. β-alanine



23. α-aminovaleric acid



24. γ-aminovaleric acid



25. γ-aminonorvaline

Fig. 4. (22–25)

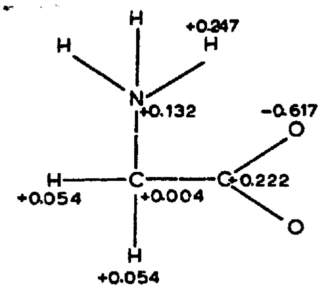
The following general remarks are significant for the subsequent discussion of the correlation between the charges of Figs. 2–5 and the experimental proton chemical shifts.

1. The σ-charges around the different carbon atoms of the main chain decrease in the order $RN^+H_3COOH > RN^+H_3COO^- > RNH_2COOH > RNH_2COO^-$.

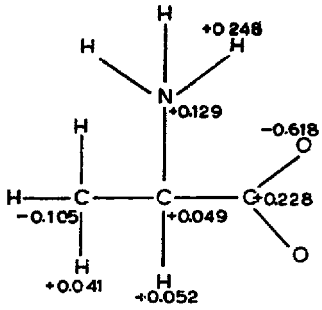
2. The carbon atom whose net charge is most affected by the ionization of the α-amino or of the α-carboxyl group, is, quite naturally, the α-carbon. The other carbon atoms are affected less and less by these changes, as their distance from the α-carbon atom increases.

FIGURE 5

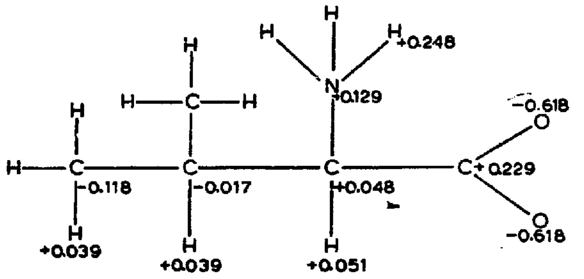
Distribution of electronic charges in α-amino acids (dipolar ions).



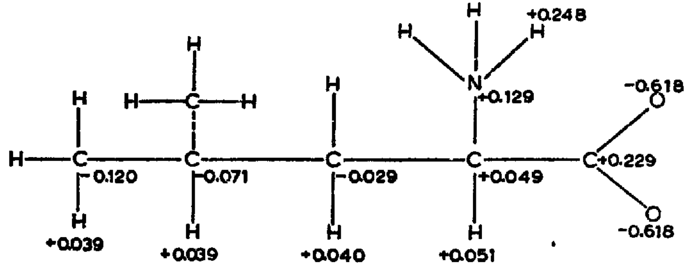
1.Glycine



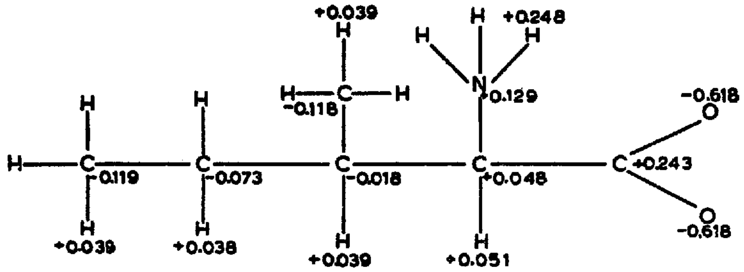
2.Alanine



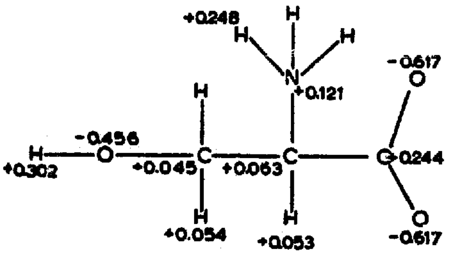
3.Valine



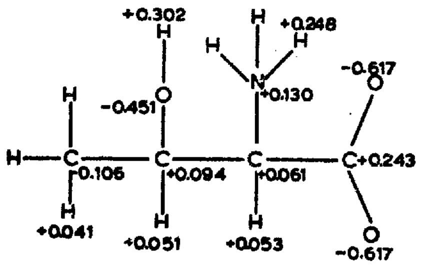
4.Leucine



5. Isoleucine

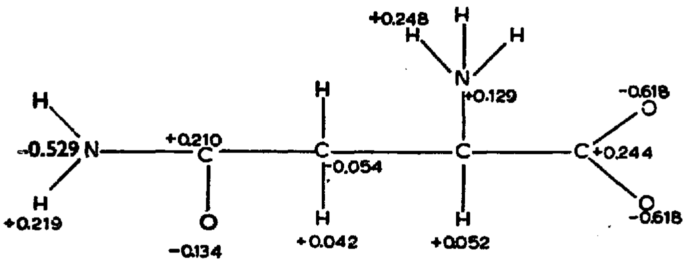


6.Serine

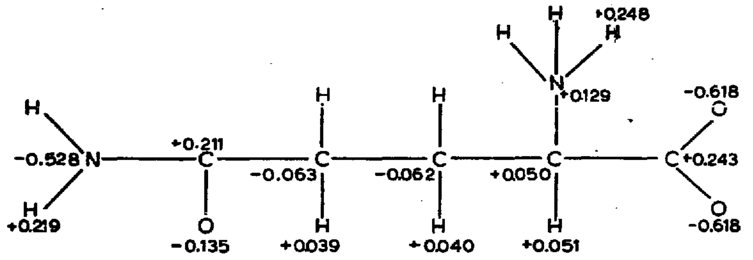


7. Threonine

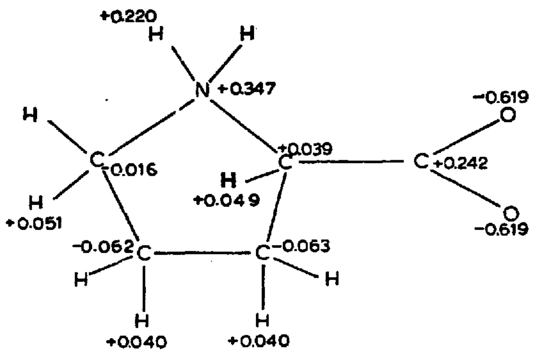
Fig. 5. (1–7)



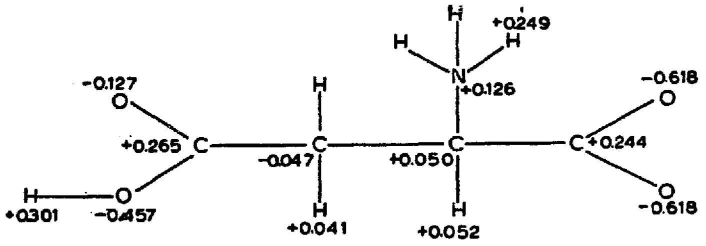
8. Asparagine



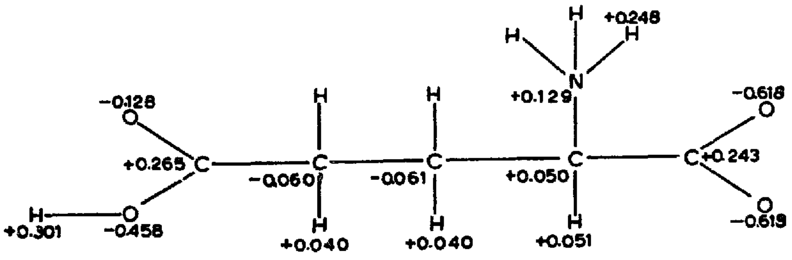
9. Glutamine



10. Proline

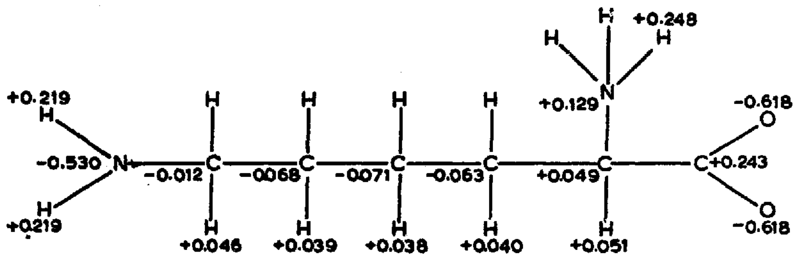


11. Aspartic acid

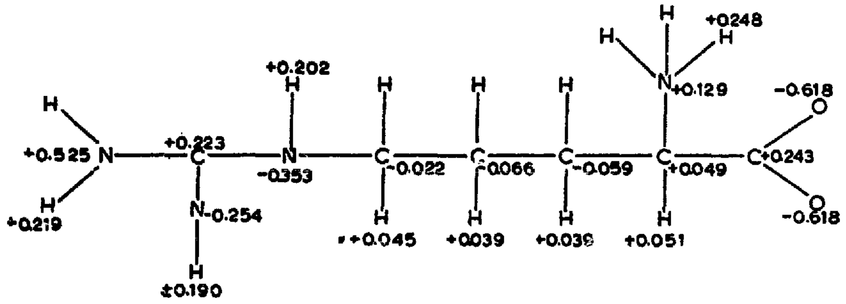


12. Glutamic acid

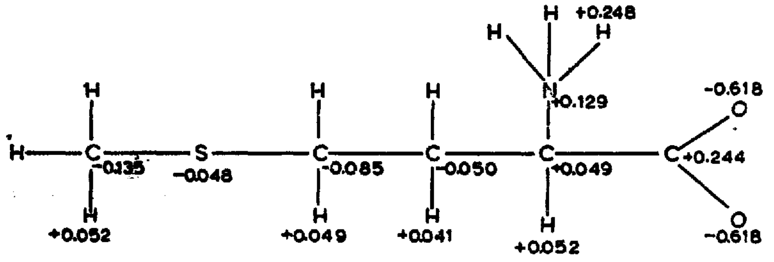
Fig. 5. (8–12)



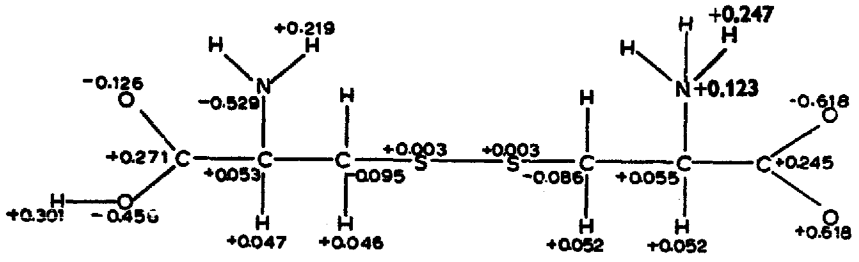
13. Lysine



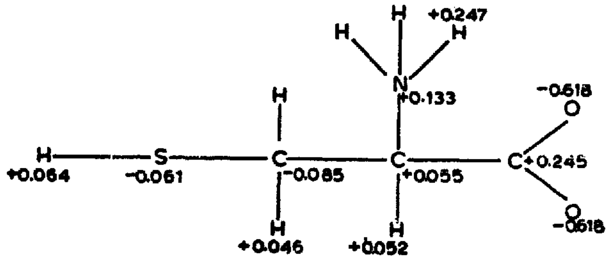
14. Arginine



15. Methionine

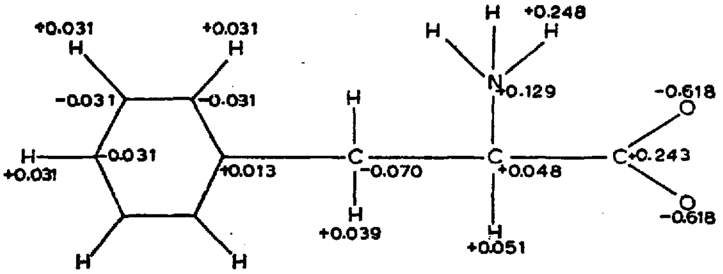


16. Cystine

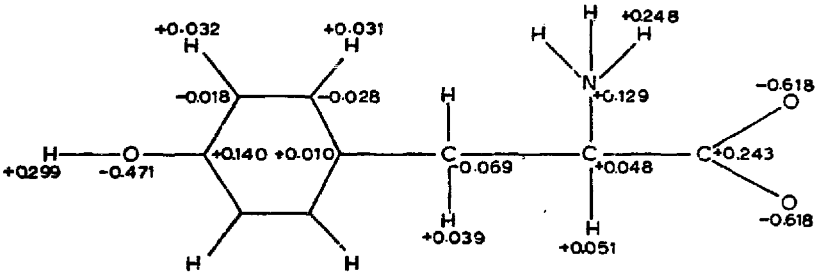


17. Cysteine

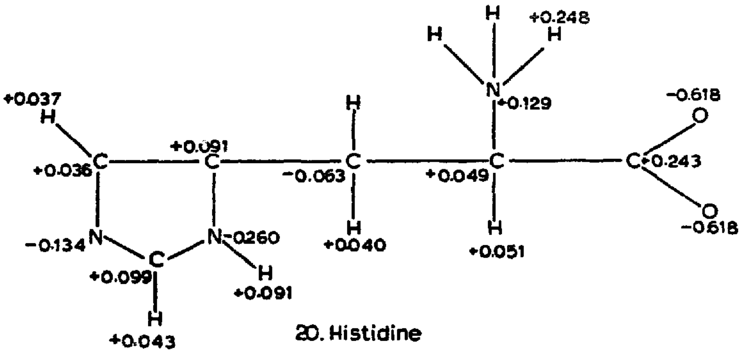
Fig. 5. (13–17)



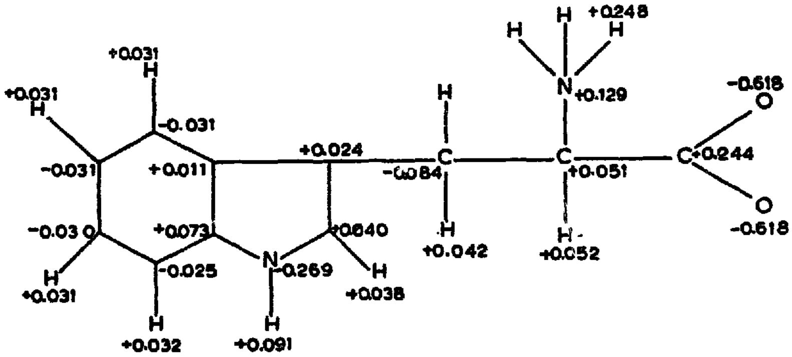
18. Phenylalanine



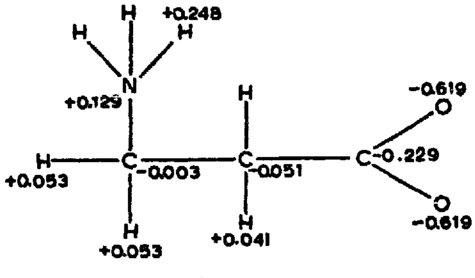
19. Tyrosine



20. Histidine

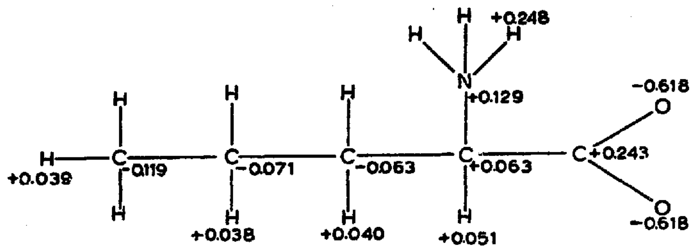


21. Tryptophan

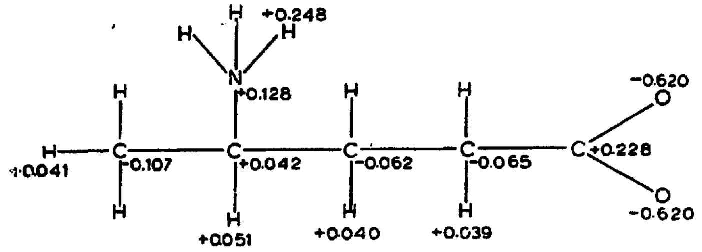


22. β-alanine

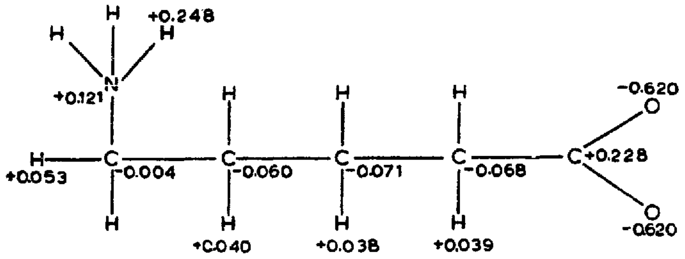
Fig. 5. (18–22)



23. α-aminovaleric acid



24. γ-aminovaleric acid



25. γ-aminonorvaline

Fig. 5. (23–25)

3. The absolute magnitudes of the σ-charges of primary, secondary, and tertiary carbon atoms decrease in the order primary > secondary > tertiary.

4. The net charges of the hydrogen atoms fixed on the main chain decrease in the same order as those of the carbons: $RN^+H_3COOH > RN^+H_3COO^- > RNH_2COOH > RNH_2COO^-$.

PROTON CHEMICAL SHIFT

Experimental determinations of the proton magnetic resonance spectra of the amino acids were carried out by two groups of authors[13–15]. The measurements were made in anhydrous H_2SO_4, in trifluoroacetic acid, and in aqueous and alkaline solutions. A quite remarkable solvent shift is observed.

The interpretation of proton chemical shifts is usually given in terms of the so-called magnetic shielding constants, from which they differ by constant amounts depending on the reference compounds. General theoretical considerations[10] suggest that the magnetic shielding constant of a proton can be treated as resulting from two different contributions, $\sigma_{\text{loc.}}$ and $\Sigma\sigma'$, *i.e.*,

$$\sigma_H = \sigma_{\text{loc.}} + \Sigma\sigma' \tag{9}$$

where $\sigma_{\text{loc.}}$ denotes the local contribution due to the electron cloud around the proton, and $\Sigma\sigma'$ represents the sum of the contributions arising from the anisotropy effects of the neighbouring atoms, and from other influences such as field effects[15].

As is well known[17], the relationship between the local diamagnetic shielding constant and the net charge of the hydrogen atom can be approximated satisfactorily by the formula

$$\sigma_{\text{loc.}} = B'(1 - Q_{\text{H}}) \tag{10}$$

where B' is a positive coefficient. Qualitatively, expression 10 states that the more positive the charge, the smaller the shielding.

As to the term $\Sigma\sigma'$, it seems reasonable to suppose that the essential part of it arises from the contribution $\sigma_{\text{C, adj.}}$ of the magnetic anisotropy of the adjacent carbon atom. This contribution can be written[18]:

$$\sigma_{\text{C,adj.}} = -\frac{1}{3R^3H}\left[\mu_x(1 - 3\cos^2\theta_x) + \mu_y(1 - 3\cos^2\theta_y) + \mu_z(1 - 3\cos^2\theta_z)\right] \tag{11}$$

where μ_x, μ_y, and μ_z denote the components of the magnetic dipole moments induced by the paramagnetic current around a carbon atom, H is the external field, $\cos\theta_x$, $\cos\theta_y$, and $\cos\theta_z$ are the direction cosines of the vector $\vec{R}$ connecting the carbon atom to the hydrogen atom under consideration, and R is the magnitude of this vector.

An LCAO-MO evaluation of μ_x, μ_y, and μ_z has been made by one of us[9], according to the procedure proposed by POPLE[18]. The sign of $\sigma_{\text{C, adj.}}$ has been shown to be negative for charges ranging from -0.15 to $+0.15$, and its absolute value has been found to decrease in the order tertiary > secondary > primary carbon.

A comparison of these results with those indicated in the previous section for the evolution of the σ-charges of these carbon atoms, adds further support to the belief that the approximations which lead to the formula

$$\sigma_{\text{C,adj.}} = -C' - A'Q_{\text{C}} \tag{12}$$

(where A' and C' are positive constants) are reasonable. We shall not discuss these approximations, but will accept Eqn. 12 for our subsequent considerations. It should be noted, by the way, that Eqn. 12 is not to be confused with a similar equation holding between the proton chemical shifts of hydrogen atoms attached to aromatic carbon atoms, and their charges.

Combining Eqns. 9, 10 and 12, and taking into account the fact that the shielding constant σ_{H} differs from the proton chemical shifts, δ, by a constant amount depending on the reference compound, we can write the general equation

$$\delta = -AQ_{\text{C}} - BQ_{\text{H}} + C \tag{13}$$

where A, B, C, are positive constants.

The arguments supporting Eqn. 13 are clearly of a qualitative nature, and the actual dependence of δ on the charges of the protons and of their neighbours may very well be much more complicated. Nevertheless, especially in view of the smallness of the charges in question, we can expect that the linear approximation (Eqn. 13) should be rather satisfactory, if the calculated charges are of the good order of magnitude.

As a preliminary test of Eqns. 10, 12 and 13, the net charges of the hydrogens and those of the adjacent carbon atoms have been plotted against the proton shifts in H_2SO_4 and NaOH solutions (referred to water). In these conditions the amino acids exist essentially as protonated cations and dissociated anions, respectively.

A satisfactory over-all parallelism can be observed between the shifts and Q_H's (Figs. 6 and 7) as well as between the shifts and the Q_C's (Figs. 8 and 9). Moreover, the results are in agreement with the general remarks made about the evolution of the charge distributions in the previous section. Thus Remarks 1 and 4 find their parallels in the fact that the proton chemical shifts measured in NaOH solutions are the largest, and those in H_2SO_4 are the smallest (*e.g.* the proton shifts for the CH_2 of glycine, are +1.55 ppm and +0.55 ppm, respectively).

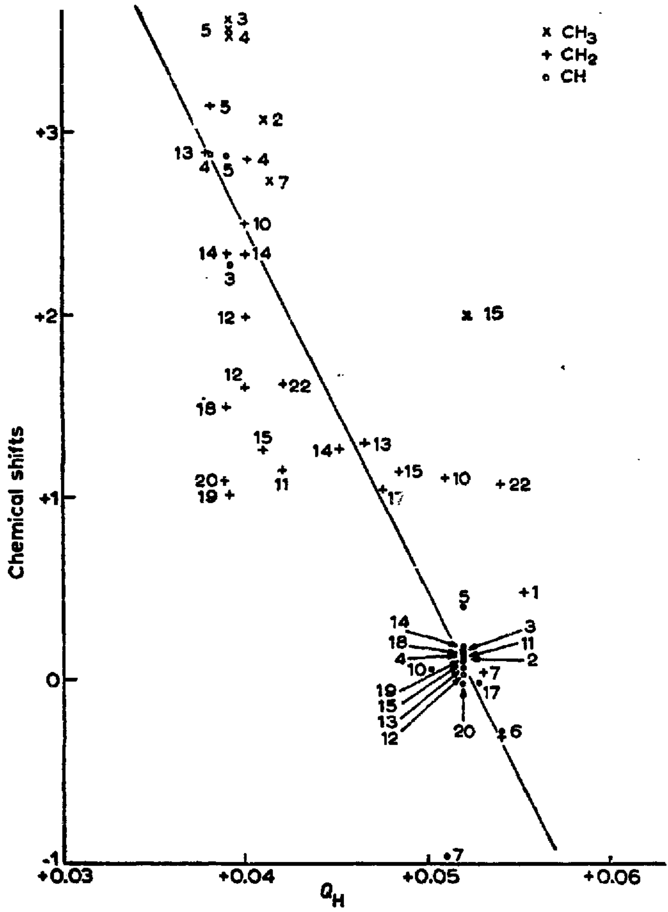


Fig. 6. Relations between the net charges of the hydrogen atoms in $RCHN^+H_3COOH$ and the proton chemical shifts in H_2SO_4 solution.

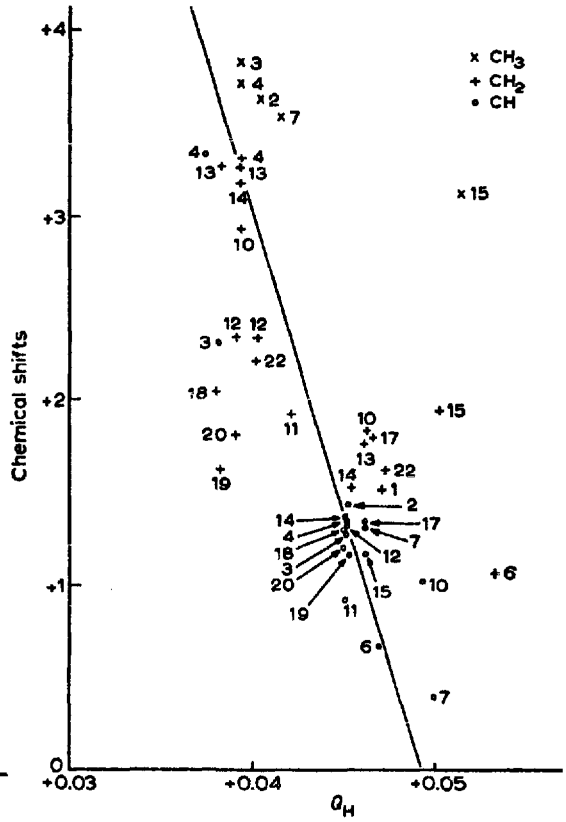


Fig. 7. Relation between the net charges of the hydrogen atoms in $RCHNH_2COO^-$ and the proton chemical shifts in NaOH solution.

The trend indicated in Remark 3 is in agreement with the fact that the proton chemical shifts of the CH_3, CH_2, and CH groups are always in the order $CH_3 > CH_2 > CH$, in three different solvents. Finally, Remark 2 agrees with the experimental fact that the solvent effect is the strongest on the proton attached to the α-carbon atom (it varies by 1.2–1.3 ppm over the entire pH range), and that it is the less affected by a change in pH, the further from the α-carbon atom is the hydrogen atom to which it refers. For instance, the solvent shift of the methylene group of leucine amounts to only 0.4 ppm.

The above discussion indicates the existence of the expected correlation between the proton chemical shifts of the α-amino acids and the charges calculated theoretically at least as an over-all correlation. A further step in the study of this correlation can be made by trying to find the values of the coefficients *A*, *B*, *C* of Eqn. 13 by

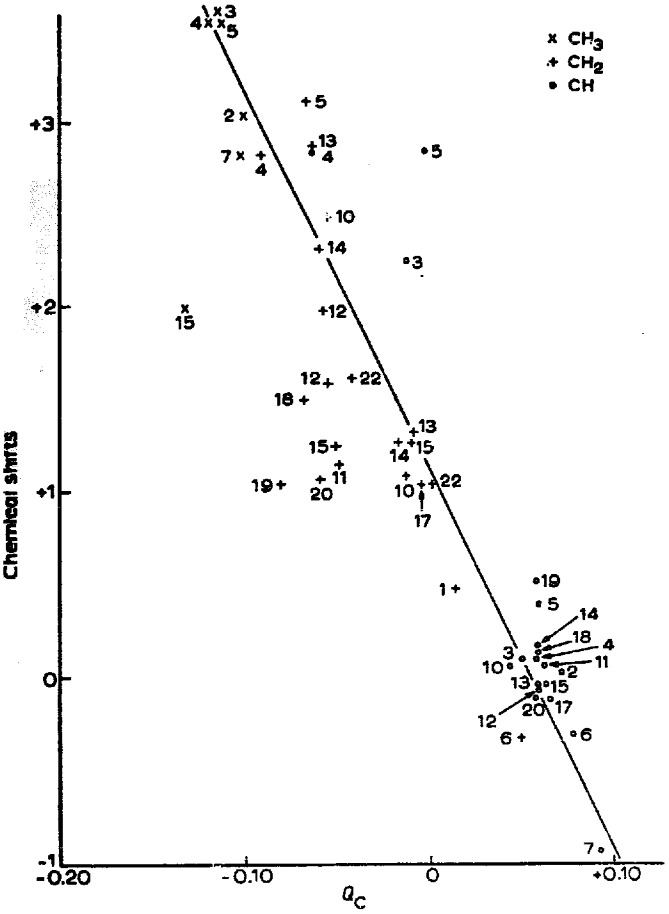


Fig. 8. Q_C of RCHN$^+$H$_3$COOH and proton chemical shifts in H_2SO_4 solution.

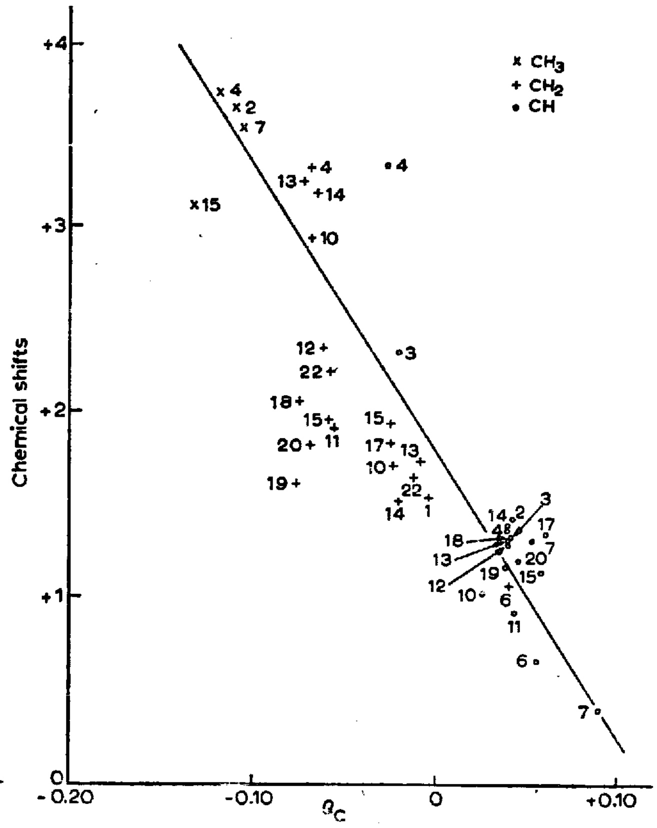


Fig. 9. Q_C of RCHNH$_2$COO$^-$ and proton chemical shifts in alkaline solution.

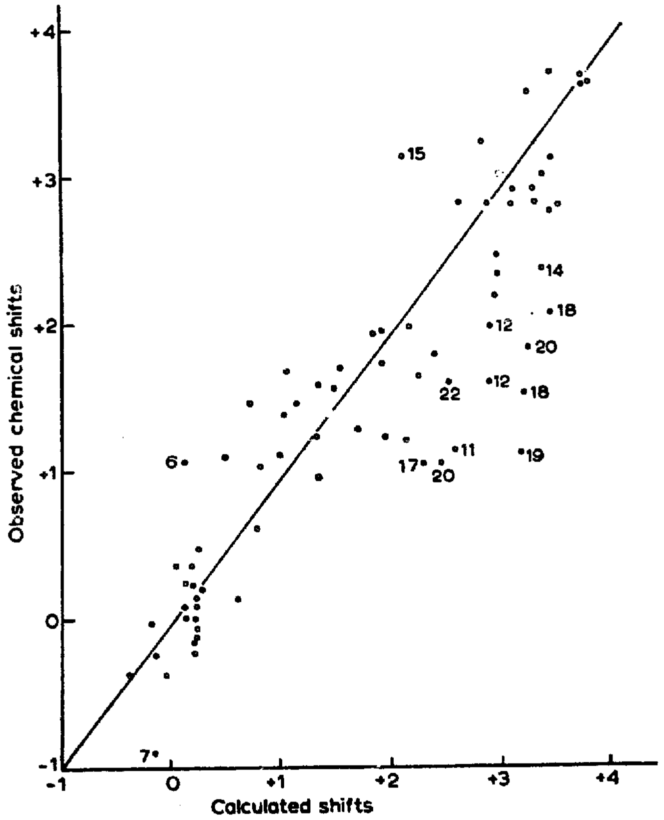


Fig. 10. Observed and calculated proton shifts in NaOH and H_2SO_4 solution.

the method of least squares, using all but the most exceptional data. This we have done for the data relating to alkaline and acidic solutions; the values thus obtained are: $A = 9.92$, $B = 133.93$, $C = 9.67$.

These values are positive, as they should be, and indicate, in agreement with general theory, that the contribution of the carbon charge to the proton chemical shift is much smaller than that of the charge of the proton concerned. A clear idea

TABLE II

SHIFTS IN AQUEOUS SOLUTION AND CALCULATED SHIFTS

Compound	*Group*	$\delta_{obs.}$*	$\delta_{calc.}$ (RNH_2COOH)	$\delta_{calc.}$ $(RN^+H_3COO^-)$
Glycine	CH_2	+1.20	+1.35	+0.39
Alanine	CH	+0.94	+1.04	+0.21
	CH_3	+3.36	+3.41	+3.22
β-Alanine	CH_2	+2.00	+2.71	+2.69
		+1.33	+1.62	+0.59
Cysteine	CH	+0.45	+0.86	+0.16
	CH_2	+1.48	+2.37	+2.35
Threonine	CH	0	+0.08	—0.11
		+0.94	+0.81	—0.04
	CH_3	+3.35	+3.24	+3.23
Valine	CH	+1.14	+1.07	+0.35
		+2.23	+2.79	+2.62
	CH_3	+3.73	+3.63	+3.62
Leucine	CH	+0.96	+1.06	+0.34
		+3.03	+3.11	+2.74
	CH_2	+3.03	+3.13	+2.95
	CH_3	+3.63	+3.64	+3.63
Isoleucine	CH	+1.46	+1.07	+1.35
		+3.43	+2.81	+2.63
	CH_2	+3.43	+3.32	+3.31
	CH_3	+3.61	+3.63	+3.63
Lysine	CH	+0.78	+1.06	+0.34
	CH_2	+3.01	+3.12	+3.12
			+3.20	+3.19
			+3.26	+2.94
		+1.46	+1.63	+1.63
Arginine	CH	+0.71	+1.96	+0.34
	CH_2	+2.74	+3.11	+3.10
				+3.03
		+1.31	+1.86	+1.86
Proline	CH	+0.49	+1.28	+0.71
	CH_2	+2.85	+3.12	+2.93
				+2.94
		+1.39	+1.85	+0.99
Methionine	CH	+0.80	+1.04	+0.22
	CH_2	+1.59	+2.75	+2.68
			+1.85	+1.96
	CH_3	+2.76	+2.05	+2.05

---

* In units of ppm referred to water.

of the deviations of the experimental data from the linear relationship (Eqn. 13) using the above values of the coefficients is given by Fig. 10, where the shifts calculated (in units of ppm) from that relation are plotted against the experimental ones, again for NaOH and H_2SO_4 solutions. In view of the roughness of Eqn. 13, the agreement observed can be considered as very satisfactory, especially as the few important discrepancies are most probably due to the presence, in the corresponding compounds, of particular effects which our simplified theory did not take into account. This is indeed the case for such compounds as glutamic acid, aspartic acid, arginine, serine, cysteine, tyrosine, phenylalanine, and histidine. Thus, in our simplified treatment, not only have we used a linear relationship, but we have included in it, besides the charge of the hydrogen atom itself, only the contribution of the neighbouring carbon atom. Now although this may be assumed to be a sufficiently good approximation in most cases, it is evidently insufficient in some compounds. For instance, in aromatic amino acids, as a result of π-induced currents, the values of C should be reduced by 0.6–1 ppm, when the shifts of protons adjacent to the aromatic rings are calculated. It can be seen that the agreement between the observed chemical shifts of the aromatic amino acids and the calculated ones would indeed be much improved by such a correction. A similar justification could possibly hold for the discrepancies between experimental and calculated shifts for methylene groups adjacent to COOH, NH, SH, or OH groups, since some of these groups may very well give a paramagnetic contribution to the proton shift of the adjacent CH_2 (ref. 20).

Returning to Eqn. 13, we note that the data for aqueous solutions could not be used for the determination of the parameters, because in such conditions the amino acids exist as equilibrium mixtures of their dipolar and uncharged forms. In this case we have therefore preferred to calculate the expected shifts from the above coefficients A, B, C, for both forms, and to compare them with experiment (Table II). With very few exceptions, the experimental shifts lie between the calculated values for the uncharged and dipolar forms. We expect that a further investigation will enable us to give an explicit relationship between the calculated and experimental shifts, on one hand, and the equilibrium constants, on the other.

CONCLUSION

The above discussion clearly shows that, in spite of the many crude approximations of the theoretical treatment the procedure used in the present paper leads to charge distributions which permit a fairly quantitative interpretation of the proton chemical shifts in the amino acids of proteins.

As the charge distributions calculated by the procedure used in this paper also lead to theoretical dipole moments in satisfactory agreement with experiment we feel encouraged to believe that they will be useful for interpreting and predicting other physico-chemical properties of amino acids, such as dissociation constants, chemical reactivities, etc. Investigations of these problems are carried out presently in our laboratory.

ACKNOWLEDGEMENTS

This work was sponsored by grant CY-3073 of the U.S. Public Health Service (National Cancer Institute) and by the Convention No. 61-FR-134 of the Délégation Générale à la Recherche Scientifique et Technique (Comité Cancer et Leucémie).

RÉSUMÉ

La méthode des orbitales moléculaires a été utilisée, dans l'approximation développée par **Del Re**, à l'étude de la distribution des charges électroniques dans les acides α-aminés des protéines et cela dans leurs formes neutres et leurs différentes formes ionisées. Les résultats du calcul sont mis en corrélation avec les déplacements chimiques observés en résonance nucléaire magnétique dans différents solvants.

Un parallélisme général assez satisfaisant est obtenu entre les valeurs des déplacements chimiques (δ_H) et celles des charges électroniques sur les hydrogènes (Q_H) ou celles des carbones porteurs de ces hydrogènes (Q_C). Les résultats peuvent être représentés par l'équation générale $\delta_H = -AQ_C - BQ_H + C$, la méthode des moindres carrés conduisant aux valeurs $A = 9{,}92$, $B = 133{,}93$, $C = 9{,}67$ pour les données relatives aux solutions alkalines et acides. Ce résultat montre qu'en accord avec la théorie générale la contribution de Q_H à la valeur du déplacement chimique est plus importante que celle de Q_C.

REFERENCES

[1] A. Pullman and B. Pullman, *Quantum Biochemistry*, Wiley, New York, 1963.
[2] A. Pullman and B. Pullman, in M. Kasha and B. Pullman, *Horizons in Biochemistry, Albert Szent-Györgyi Dedicatory Volume*, Academic Press, New York, 1962, p. 533.
[3] A. Pullman and B. Pullman, *Nature*, 196 (1962) 1137.
[4] J. Lennard-Jones and G. Hall, *Discussions Faraday Soc.*, 10 (1951) 18.
[5] G. Hall, *Proc. Roy. Soc. London, A*, 205 (1951) 541.
[6] R. D. Brown, *J. Chem. Soc.*, (1953) 2615.
[7] C. Sandorfy, *Can. J. Chem.*, 33 (1955) 1337.
[8] H. Yoshizumi, *Trans. Faraday Soc.*, 53 (1957) 123.
[9] K. Fukui, H. Kato and T. Yonezawa, *Bull. Chem. Soc. Japan*, 34 (1961) 442.
[10] G. Del Re, *J. Chem. Soc.*, (1958) 4031.
[11] J. L. Franklin, *J. Chem. Phys.*, 22 (1954) 1304.
[12] G. W. Wheland and L. Pauling, *J. Am. Chem. Soc.*, 57 (1935) 2086.
[13] O. Jardetzky and C. D. Jardetzky, *J. Biol. Chem.*, 233 (1958) 383.
[14] M. Takeda and O. Jardetzky, *J. Chem. Phys.*, 26 (1957) 1346.
[15] F. A. Bovey and G. V. D. Tiers, *J. Am. Chem. Soc.*, 81 (1959) 2870.
[16] B. Pullman and A. Pullman, *Les Théories Électroniques de la Chimie Organique*, Masson, Paris (1952).
[17] *Cf. e.g.*, J. A. Pople, W. G. Schneider and H. J. Bernstein, *High-Resolution Nuclear Magnetic Resonance*, McGraw-Hill, New York, 1959.
[18] J. A. Pople, *Proc. Roy. Soc. London. A*, 239 (1957) 541, 550.
[19] T. Yonezawa, *Bull. Chem. Soc. Japan*, to be published.
[20] H. Spieseck and W. G. Schneider, *J. Chem. Phys.*, 35 (1961) 722.