CLXXX. THE POTENTIOMETRIC DETERMINA-
TION OF POLYPEPTIDES AND AMINO-ACIDS.
II. THE FORMALDEHYDE TITRATION.

By ERNEST WILLIAM BALSON AND ALEXANDER LAWSON.

From University College, Southampton.

IN a previous paper [Balson *et al.* 1935] a potentiometric examination was made of the Willstätter-Waldschmidt-Leitz titration. This titration was found to be of limited application to the study of proteins and their hydrolysis products owing to the low solubility of such substances in 90 % aqueous alcohol. We have accordingly made an examination of the formaldehyde titration of Sørensen, which as far as solubility is concerned, appears to be better suited to the examination of protein hydrolysates.

Previous work on the formaldehyde titration falls into two main divisions, firstly the isolation of compounds between formaldehyde and amino-acids, and secondly, the titration of amino-acids in the presence of formaldehyde with indicators or by electrometric methods.

Krause [1918], in a study of the reaction between glycine and formaldehyde, postulated the formation of a compound formed from two molecules of glycine and three of formaldehyde having the constitution
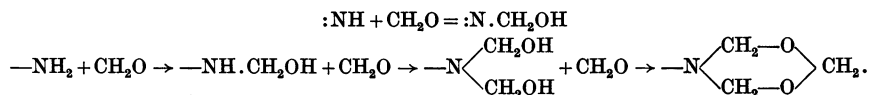
$$CHOH(CH_2.NH.CH_2.COOH)_2.$$
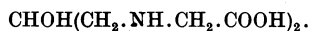
Bergmann *et al.* [1923] obtained compounds containing three molecules of formaldehyde with one amino-acid residue, and later [Bergmann & Ensslin, 1925] the triformylglycine obtained in this reaction was shown to be identical with the product of Krause. Bergmann's compounds were of the type $C_3H_6O_2.R$, one molecule of water being eliminated in the reaction, and they were readily converted into the corresponding methyleneimino-compounds of Sørensen [1908] and Schiff [1902] with loss of formaldehyde.

Harris [1924; 1929] and Birch & Harris [1930] have examined, both by colorimetric and potentiometric methods, the titration curves of amino-acids in formaldehyde solution. The treatment was not quantitative, but from the results Harris concluded that the action of the formaldehyde was to give the methyleneimino-acid of Sørensen, having a dissociation constant a thousand times as great as the parent acid. In the last paper of the series the effect of formaldehyde was used to prove the zwitterionic structure of the amino-acids. Richardson [1934] has made a comprehensive survey of the methods available for the determination of amino-N and -COOH and has discussed the significance of the results obtained in the light of the zwitterion hypothesis. The only quantitative work on the effect of formaldehyde on amino-acids is due to Levy [1933] and Tomiyama [1935]. Levy made a physico-chemical study of the equilibria involved and was able, assuming that formaldehyde reacts with the amino-acid anion, to give a quantitative interpretation of the reactions. With proline the reaction was found to produce only the binary compound, whereas with amino-acids both the mono- and di-formyl compounds were obtained. Tomiyama, restricting himself to the examination of the effect of small concen-

trations of formaldehyde (less than $0.2\,M$), concluded that only one molecule reacts with the amino-acids glycine, alanine and proline. With such small concentrations, however, it is not surprising that the reaction does not proceed beyond the first stage. This restriction in formaldehyde concentration was applied to obviate possible errors in pH determination due to electrode failure and/or solvent change, but a consideration of the data presented in this paper shows that it is unlikely that any significant errors of such a nature are present. The equations developed describe quite adequately the data for formaldehyde concentrations up to about $5M$. In the few cases where electrode failure was experienced the results were quite at variance with theoretical requirements, and under normal conditions the values for the reaction constants could be satisfactorily reproduced.

This work has been extended to include other amino-acids and imino-acids, dimethylglycine and a di- and tri-peptide. It has been found in agreement with Levy, that only the monoformyl compound is formed with the imino-acids proline and sarcosine, and also that both mono- and di-formyl compounds are formed with the amino-acids, but in addition, evidence is brought forward to show that a third molecule also enters into combination in the second case, and this behaviour has been found generally true for all the amino-acids examined. This is in agreement with the results of Bergmann. A di- and tri-peptide were also examined and found to give qualitatively the same results as the amino-acids. Dimethylglycine was found to be very slightly affected by formaldehyde, the effect being qualitatively the same as is obtained with boric acid. In view of this, and also because the results do not fit any mechanism involving combination with formaldehyde, it is concluded that no reaction takes place. The changes which do occur are probably in the main due to the weakly acidic function of the formaldehyde and in a lesser degree to solvent changes.

In view of these results, and taking into account the high reaction rate, it seems probable that the reaction takes place in stages. We therefore suggest that compound formation is dependent on the number of hydrogen atoms attached to the nitrogen. This will account for the addition of one molecule of formaldehyde to the imino-group. With the amino-group two molecules can be directly attached and the third molecule will most probably form a ring of the type of trioxymethylene. The following mechanism is suggested:

$$:NH + CH_2O = :N.CH_2OH$$

$$-NH_2 + CH_2O \rightarrow -NH.CH_2OH + CH_2O \rightarrow -N\begin{smallmatrix}CH_2OH\\\\CH_2OH\end{smallmatrix} + CH_2O \rightarrow -N\begin{smallmatrix}CH_2-O\\\\CH_2-O\end{smallmatrix}CH_2.$$
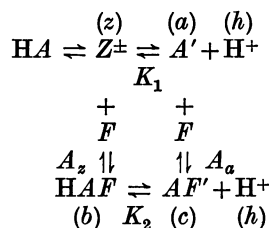
Tomiyama postulates the formation of a co-ordination compound between the formaldehyde and the amino-acid. Such a mechanism would imply that the deciding factor for combination is the presence of the lone pair of electrons on the nitrogen atom. On this basis one would expect only quantitative differences in behaviour between glycine, sarcosine and dimethylglycine. This, however, is not in agreement with our observations.

Birch and Harris have used the effect of formaldehyde on the titration curves of amino-acids to prove that they possess, in preponderating amount, the zwitterionic structure. It can be shown, however, that all the observations connected with this effect are not necessarily due to the presence of the zwitterion. Let us consider the titration of glycine with hydrochloric acid in the presence of formaldehyde (Fig. 3). Here we have a small but definite shift in the titration curve corresponding to a lowering of the apparent basicity caused presumably by reaction with formaldehyde. If the formaldehyde reacts with the $-NH_2$

group the effect on the apparent basic constant must of necessity be small since the apparent pK_b is 12; with a primary amine, on the other hand, the apparent basic constant is about 10^{-5} and by addition of formaldehyde this can be lowered to reach in the limit a value of 10^{-14}. Thus in the second case there can be a large possible effect on the titration curve with hydrochloric acid as Harris has shown.

The apparent acid constant of an amino-acid is about 10^{-10}, and the effect of formaldehyde is to increase this, so that a marked shift in the titration curve with sodium hydroxide towards a lower limit corresponding to the dissociation constant of a carboxylic acid can take place. The change in the acid constant can be quantitatively accounted for by assuming that the $-NH_2$ group reacts with formaldehyde. As is shown later, and as Levy has pointed out in this case, it is again immaterial whether the zwitterionic form, the normal form or the acid anion reacts. It will be seen therefore that the only effect of formaldehyde is to combine with the $-NH_2$ group to reduce its basic power, the $-NH_2$ group being either present as such or arising from the zwitterion.
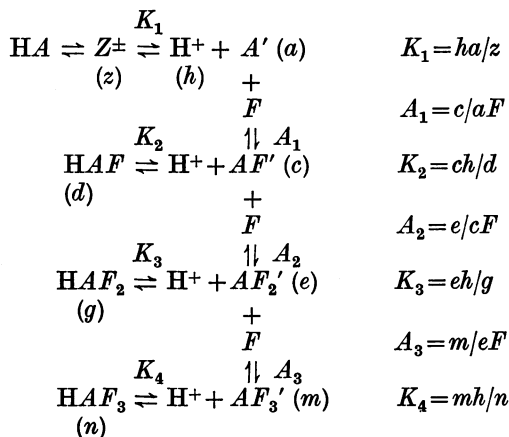
It can be shown that it is immaterial as far as the results from the sodium hydroxide titration are concerned whether the formaldehyde attacks the anion, or the zwitterion or both. Thus if we consider the following equilibria:

$$
\begin{array}{ccc}
(z) & (a) & (h) \\
HA \rightleftharpoons Z^{\pm} \rightleftharpoons & A' + H^+ \\
& K_1 & \\
+ & + & \\
F & F & \\
A_z \updownarrow & \updownarrow A_a & \\
HAF \rightleftharpoons & AF' + H^+ \\
(b) \quad K_2 \quad (c) & (h) &
\end{array}
$$

Where z, h, a, F, b, c are the molar concentrations, K_1 and K_2 the dissociation constants, and A_z and A_a the association constants for the zwitterion and the anion respectively. We then have $K_1 = ha/z$, $K_2 = hc/b$, $A_z = b/zF$, $A_a = c/aF$,

so that $$K_1/K_2 = A_z/A_a \qquad \ldots\ldots(1).$$

Let us assume that the anions are attacked and that new acids are formed with dissociation constants different from the parent acids. If it is assumed that compounds containing up to three molecules of formaldehyde are formed, then the following equilibria are involved:

$$
\begin{array}{lll}
\quad\quad K_1 & & \\
HA \rightleftharpoons Z^{\pm} \rightleftharpoons H^+ + A' \ (a) & \quad K_1 = ha/z \\
\quad (z) \quad\ (h) \quad + & & \\
\quad\quad\quad\quad F & \quad A_1 = c/aF \\
\quad\quad K_2 \quad\ \updownarrow A_1 & & \\
HAF \rightleftharpoons H^+ + AF' \ (c) & \quad K_2 = ch/d \\
(d) \quad\quad\quad + & & \\
\quad\quad\quad\quad F & \quad A_2 = e/cF \\
\quad\quad K_3 \quad\ \updownarrow A_2 & & \\
HAF_2 \rightleftharpoons H^+ + AF_2' \ (e) & \quad K_3 = eh/g \\
\ (g) \quad\quad\quad + & & \\
\quad\quad\quad\quad F & \quad A_3 = m/eF \\
\quad\quad K_4 \quad\ \updownarrow A_3 & & \\
HAF_3 \rightleftharpoons H^+ + AF_3' \ (m) & \quad K_4 = mh/n \\
\ (n) & &
\end{array}
$$

Since the Henderson-Hasselbach equation describes the titration curves in the presence of formaldehyde we know that the apparent dissociation constant is given by:
$$K_0 = H^+ \text{ (sum of all anions)/(all undissociated acids)},$$

i.e.
$$K_0 = h \, (a+c+e+m)/(z+d+g+n).$$

This equation together with the above relations leads to the expression:

$$(K_0/K_1-1) + A_1F \, (K_0/K_2-1) + A_1A_2F \, (K_0/K_3-1) + A_1A_2A_3F \, (K_0/K_4-1) = 0$$
$$\dots\dots(2).$$

Now if K_0/K_2, K_0/K_3 etc. are small compared with 1 this reduces to

$$K_0/K_1 - 1 = A_1F + A_1A_2F^2 + A_1A_2A_3F^3 \qquad \dots\dots(3),$$

which, apart from the last term, is Levy's expression. This expression is found to hold in practice and so the approximation is justified.

The value of K_0/K_1 can readily be obtained from the observed E.M.F. values as follows:

The amino-acid is approximately half neutralized with sodium hydroxide and titrated with formaldehyde. Before addition of formaldehyde is made the pK_a is given by:

$$pK_1 = pH + \log \, (M-Na^+)/Na^+ \qquad (M = \text{molar conc. of amino-acid}).$$

This will also be true for the apparent pK_a values in the presence of formaldehyde.

Thus if $E_1 = $ E.M.F. of hydrogen electrode before CH_2O addition,

$\qquad\quad E_0 = $ E.M.F. of hydrogen electrode at any CH_2O addition,

then at 25°: $\qquad E_1/0\cdot0591 = pH_1 \qquad$ and $\qquad pH_1 + C = pK_1,$

where $C = \log \, (M-Na^+)/Na^+$ which is constant and independent of volume, and

$$E_0/0\cdot0591 = pH_0 \qquad \text{and} \qquad pH_0 + C = pK_0.$$

Therefore $\qquad (E_1 - E_0)/0\cdot0591 = pH_1 - pH_0 = pK_1 - pK_0 = \log K_0/K_1.$

This does not involve the E.M.F. of the half cell.

The constants in equation (3) were obtained by graphical methods as follows:

Equation (3) is $\qquad K_0/K_1 - 1 = A_1F + A_1A_2F^2 + A_1A_2A_3F^3 = R.$

Dividing through by F we obtain:

$$(K_0/K_1-1)/F = R/F = A_1 + A_1A_2F + A_1A_2A_3F^2 \qquad \dots\dots(4).$$

The plot of R/F against F yields a curve whose intercept is A_1 (see Fig. 1).

Taking the term A_1 to the left-hand side and dividing through by F we obtain:

$$(R/F - A_1)/F = A_1A_2 + A_1A_2A_3F \qquad \dots\dots(5).$$

This is a straight line intercept A_1A_2 slope $A_1A_2A_3$ (see Fig. 2).

The amino-acids give results which fit equation (3) and the various constants have been evaluated (Table I). Proline and sarcosine combining with only one molecule of formaldehyde give straight lines when equation (4) is applied, the slope being zero and the intercept A_1 (see Fig. 1). By means of equation (1), equation (3) may be converted into the corresponding equation which describes the attack on the zwitterion. Since however equation (1) does not involve the

Table I.

Amino-acid	α	β	γ
Glycine	160	500	69
Alanine	24	66	36·5
α-Aminobutyric acid	21	28	7·5
β-Phenylalanine	22	24·5	6·7
Valine	23	9·5	2·7
Leucine	17	32	2·7
Leucylglycine	22	35	9·4
Leucylglycylglycine	25	38	13·8
Proline	126	—	—
Sarcosine	320	—	—
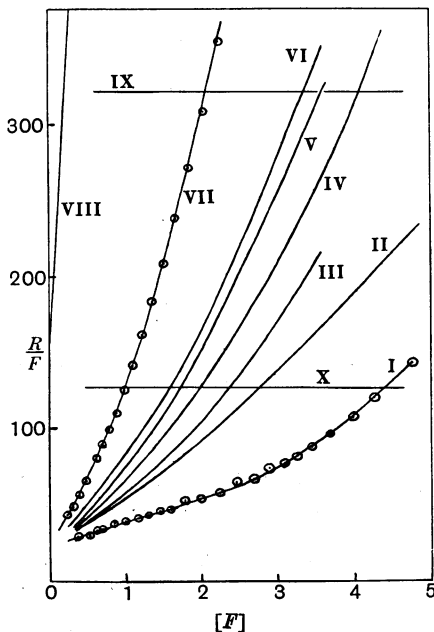
Fig. 1. Plot of CH_2O concentration against R/F (equation 4). Curve I, Valine. II, Leucine. III, α-Aminobutyric acid. IV, Leucylglycine. V, Leucylglycylglycine. VI, Glycine (on reduced scale R/F divided by 100). VII, Alanine. VIII, Glycine. IX, Sarcosine. X, Proline.

concentration of formaldehyde, then the new equation will be of the same order with respect to F as is (3), both being of the form:

$$K_0/K_1 - 1 = \alpha F + \beta F^2 + \gamma F^3.$$

The constants α, β, γ, may be determined by the methods of plotting already described. From the form of the curve no decision can be made as to whether the anion, the zwitterion or both react. These constants α, β, γ are thus composite association constants and from the data it is impossible to determine with which equilibria they are associated.

It would appear from the table of results that the constant for the addition of the first molecule of formaldehyde is practically unaltered with increase in chain length with the exception of glycine, whereas the second and third constants are markedly lowered. This is most noticeable in the cases of valine and leucine. It

is interesting to note that chain increase caused by peptide formation results in the reverse effect in the case examined, the second and third constants showing an increase, as in the example leucine, leucylglycine and leucylglycylglycine.
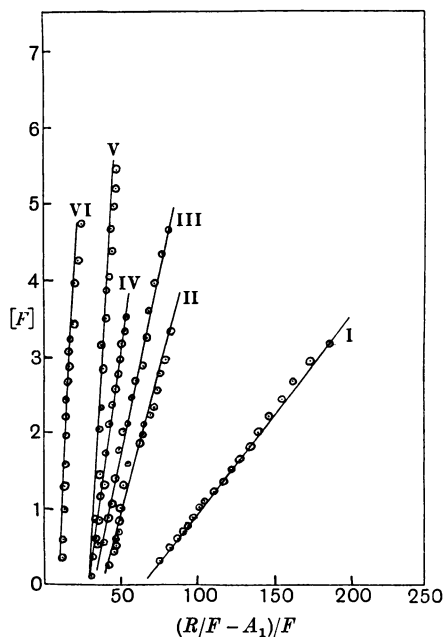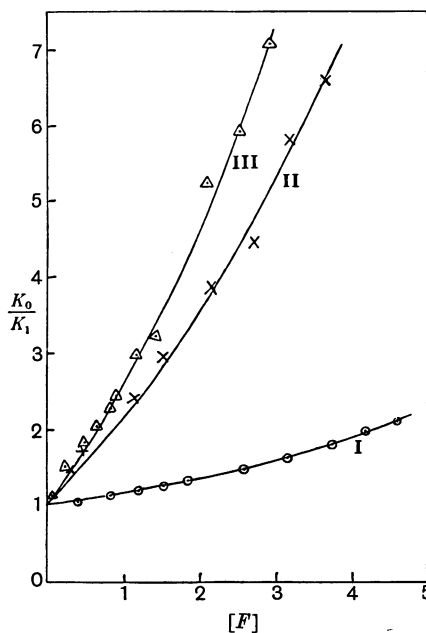

Fig. 2.

Fig. 3.

Fig. 2. Plot of CH_2O concentration against $(R/F - A_1)/F$ (equation 5). Curve I, Alanine. II, Leucylglycylglycine. III, Leucylglycine. IV, α-Aminobutyric acid. V, Leucine. VI, Valine.

Fig. 3. Plot of CH_2O concentration against K_0/K_1, where K_0 is apparent basic dissociation constant at zero CH_2O conc., and K_1 the value at conc. F. Curve I, Effect of CH_2O on glycine half-titrated with HCl. II, Effect of CH_2O on dimethylglycine half-titrated with NaOH. III, Effect of CH_2O on boric acid half-titrated with NaOH.

EXPERIMENTAL.

Glycine (A.R.) was used as standard, the other amino-acids being checked against this by electrometric titration. The formaldehyde was of A.R. purity and immediately before use was distilled over calcium carbonate, giving a distillate (pH 4·8) containing practically no formic acid. The distillate was filtered and immediately transferred to a sealed microburette, saturated with hydrogen and maintained in an atmosphere of hydrogen. Its strength (10–13 M) was accurately determined by the method of Romijn [1897]. All titrations were carried out in a thermostat at 25° using the titration vessel and hydrogen electrodes previously described.

SUMMARY.

1. The effect of formaldehyde on the titration curves of simple amino-acids and polypeptides has been quantitatively examined.

2. It is shown that mechanisms involving reaction between the formaldehyde and the amino-acid either in the form of the zwitterion, or the anion

or as both simultaneously yield results which fit the experimental data, so that it is impossible to decide in which form the amino-acid reacts.

3. All the primary amino-acids and polypeptides react with up to three molecules of formaldehyde, the secondary with one only, the tertiary being without action. A reaction mechanism is suggested.

4. The reaction constants for a series of amino-acids and two polypeptides have been evaluated.

REFERENCES.

Balson, Earwicker & Lawson (1935). *Biochem. J.* **29**, 2700.
Bergmann, Jacobsohn & Schotte (1923). *Hoppe-Seyl. Z.* **131**, 18.
—— & Ensslin (1925). *Hoppe-Seyl. Z.* **145**, 194.
Birch & Harris (1930). *Biochem. J.* **24**, 1080.
Harris (1924). *Proc. roy. Soc.* B **95**, 500.
—— (1929). *Proc. roy. Soc.* B **104**, 412.
Krause (1918). *Ber. dtsch. chem. Ges.* **51**, 136.
Levy (1933). *J. biol. Chem.* **99**, 767.
Richardson (1934). *Proc. roy. Soc.* B **115**, 121.
Romijn (1897). *Z. anal. Chem.* **36**, 18.
Schiff (1902). *Liebigs Ann.* **325**, 348.
Sørensen (1908). *Biochem. Z.* **7**, 45, 407.
Tomiyama (1935). *J. biol. Chem.* **111**, 51.