XXXVII. THE EFFECT OF FORMALDEHYDE ON THE HYDROGEN ION DISSOCIATION CURVE OF EGG ALBUMIN.

By RALPH AMBROSE KEKWICK¹ AND ROBERT KEITH CANNAN.

From the Department of Chemistry, New York University College of Medicine.

(Received August 15th, 1935.)

It is well known that the cations of many nitrogenous bases which cannot be directly titrated with alkali in aqueous solution submit to titration in the presence of a suitable concentration of formaldehyde. This is true of certain of the cationic groups in the proteins and their hydrolytic products. In the case of the natural amino-acids and their peptides, the various cationic groups which may occur are (a) primary amino-groups, (b) the imino-groups of proline and hydroxyproline, (c) the guanidino-group of arginine and (d) the glyoxaline group of histidine. Only the first two types however react with formaldehyde to an extent sufficient to produce any profound change in their acidic properties. Birch and Harris [1930] first employed the hydrogen electrode to investigate the nature of these changes. More recently, Levy [1933; 1935] has used the same method in an extensive study of the principles involved in the "formol" titration. It appears that the acid-base changes accompanying the addition of formaldehyde may be quantitatively described by an increase in the apparent hydrogen ion dissociation constant of certain cationic groups. Levy has shown that the constants of the cations of primary NH_2 groups vary with the square of the concentration of formaldehyde above certain low limits of concentration. In the cases of proline and hydroxyproline on the other hand the change in the apparent dissociation constant is directly proportional to the formaldehyde concentration.

The present paper is a report on an extension to egg albumin of the electro-metric method of conducting formaldehyde titrations. The electrode behaviour of egg albumin solutions is decidedly less satisfactory in the presence of formaldehyde than in its absence. Nevertheless reproducible titration curves may be traced if occasional irrational observations and the electrodes responsible for them be rejected. The results meet the purposes of the present inquiry, though some features of the reaction remain obscure. It appears that the addition of formaldehyde is accompanied by an immediate readjustment of the acid-base equilibrium. This is followed however by slow secondary changes which persist for many hours and are accompanied by small but persistent changes in hydrogen electrode potential. In a period of one hour however these changes in potential are small in comparison with the large initial change. We have therefore ignored them and have accepted the substantially constant potentials observed during the first hour after the addition of formaldehyde. In doing so we believe that our curves faithfully reproduce the gross acid-base changes which accompany the usual process of formaldehyde titration.

The method of calculation differs from that used in the first paper of this series [Kekwick and Cannan, 1936] only to the extent that the anion of form-

¹ Commonwealth Fund Fellow.

aldehyde contributes to the balance of ions which establishes electrical neutrality in the system. Equation 1 of that paper becomes

$$[H^+] + [Na^+] + g(h - h_0) = [Cl^-] + [OH^-] + [F^-] \quad \dots (1A),$$

where $[F^-]$ is the molal concentration of the anion of formaldehyde. In order to calculate this from the total concentration of formaldehyde ($[F]$) we have used the relation given by Levy [1934]

$$p_H = 12.87 + \log \left[\frac{[F]}{[F^-]} - X \right] - 0.19 [F] \quad \dots (2).$$

The final term in this equation is an empirical correction introduced by Levy to accommodate the effect of the concentration of formaldehyde upon its apparent p_K . The correction is of such magnitude as to imply changes in the activities of the reactants which are much greater than is to be expected from the changing composition of the system. Presumably this dependence of p_K upon $[F]$ results from the presence of secondary equilibria which are dependent upon the primary dissociation of a hydrogen ion. However this may be, Levy's equation permits the calculation of $[F^-]$ in equation (1A) though its approximate character introduces a new source of error into the calculated values of h . This error will become serious at p_H values at which $[F^-]$ attains values comparable with those of $g \cdot h$ and thus places a new upper limit to the p_H to which titrations may be carried with precision. When $[F] = 1$ this limit will be close to p_H 10 and when $[F] = 0.1$ to p_H 11.

In Fig. 1 the hydrogen ion dissociation curve of 1 g.mol. of egg albumin [Kekwick and Cannan, 1936] in dilute aqueous solution is compared with the corresponding curves which were observed in the presence of four concentrations

Fig. 1. H^+ dissociation curves in the presence of formaldehyde. Curves *S*, *A*, *B*, *C* and *D* correspond to 0, 0.1, 0.2, 0.6 and 1.26 *M* formaldehyde respectively. Curves *L*, *M*, *N* and *P* are the comparison curves (see text) for 0.1, 0.2, 0.6 and 1.26 *M* formaldehyde respectively. H = equiv. H^+ per g.mol. protein.

of formaldehyde (0.1, 0.2, 0.6 and 1.26 *M*). At a sufficiently low p_H all the curves coincide. As the p_H rises the curves successively diverge from the control, become roughly parallel and finally converge again to a common limiting value of $-H = 27-28$ equiv. per g.mol. A graphic analysis of these curves is illustrated

by the inset curves on the left of Fig. 1. At successive small increments of p_H we determined the differences in H , (ΔH), between the value in water and in each concentration of formaldehyde. ΔH was then plotted against p_H to give the family of comparison curves shown. Discussion will be facilitated by designating as the F_x groups those groups in the protein whose affinity for H^+ is notably decreased by formaldehyde. The subscript " x " will designate the particular concentration of formaldehyde under discussion. As a first approximation, we may regard these curves as the resultant of the H^+ dissociation curves of the F_x groups and the H^+ association curve of the F_0 groups. The latter will be represented by the common reentrant section above p_H 10 and the former by the sections from p_H 4 to 10. These, it will be seen, vary in position with the concentration of formaldehyde. The four curves attain different maximum values of H indicating an overlapping of the F_0 and F_x curves. Curve P for 1.26 M however exhibits a fairly flat plateau for $-\Delta H=16-17$. We may conclude that at this and higher concentrations of formaldehyde there is little overlapping and the maximum value for $-\Delta H$ corresponds to the stoichiometric number of F groups in the protein. We are inclined to place the latter at a value of 17-18 per g.mol. of egg albumin.

Consider now only the section of each curve which extends from p_H 4 to 10. Each curve exhibits a small segment between p_H 4 and 6 amounting to 1-2 equiv. per g.mol. Since the calculated values of ΔH are probably subject to an error of ± 0.5 equiv. per g.mol. one may not place great emphasis on this small segment. It appears, however, in all the curves and its position on the p_H scale varies in an orderly fashion with $[F]$. We may conclude that 1-2 of the F groups are distinct in type from the remainder.

Confining consideration to the latter the comparison curves suggest that they are uniform in type. We have attempted to indicate this by superimposing on the appropriate section of the $F_{1.26}$ comparison curve a theoretical dissociation curve (broken line) for 16 groups. The comparison curve is manifestly symmetrical. Its departure from the theoretical curve is of the type that might be expected as a result of the spread of p_K values in a polyvalent substance with uniform groups.

In terms of the polypeptide theory of protein structure the F groups to be expected are the $\epsilon-NH_2$ groups of lysine together with a single $\alpha-NH_2$ group at the end of the polypeptide chain. It would be tempting to identify the 1-2 unique F groups indicated above with this $\alpha-NH_2$ group. There would then remain 16-17 lysine groups. This will be a minimum value for the following reason. The p_K values which have been found for $\alpha-NH_2$ groups in simple polypeptides are close to 8. The low p_H at which our F group appears in the presence of formaldehyde is also consistent with a low p_K in its absence. If the latter were at any value below 9 then the total number of F groups betrayed by the maximum on the comparison curve would not include this one unique group because it would have been dissociated completely in the absence of formaldehyde at the p_H of the maximum value of $-\Delta H$. In such case $-\Delta H$ measures only the lysine groups. Until we can establish the position which the unique group occupies in the curve for aqueous solution we must be content to conclude that titration with formaldehyde indicates the presence of 16-18 lysine groups. In so far as the identification of the unique group with $\alpha-NH_2$ is plausible the higher value for lysine is the more probable.

The recent careful analyses of egg albumin by Vickery and Shore [1932] and by Calvery [1932] agree in finding not more than 12 mols. of lysine in 1 mol. of albumin. The formaldehyde titration suggests 16-18. We sought to clarify this

discrepancy by a determination of the number of groups in egg albumin which react with nitrous acid. Choosing conditions of reaction under which the terminal NH_2 group of lysine itself reacts quantitatively and making allowances for secondary reactions, we find that 19–20 groups react with nitrous acid (Fig. 2). One of these will be the $\alpha\text{-NH}_2$ group leaving 18–19 presumed lysine groups. The agreement seems satisfactory when the higher value for the formaldehyde titration is accepted. In any case the important point is that the nitrous acid method

Fig. 2. The NH_2 groups of egg albumin from the nitrous acid reaction showing the effect of reaction time at $21\text{--}23^\circ$ on the results.

and the formaldehyde titration both indicate the presence of an amount of free NH_2 groups much greater than the amount of lysine which has actually been isolated from the hydrolysed protein. The discrepancy is confirmed by determinations of the nitrogen distribution of the products of hydrolysis of egg albumin. Using the method of Van Slyke, we find that the nitrogen of the lysine fraction corresponds to 17–18 mols. of this amino-acid in a molecule of protein. It is doubtful however if much weight can be put on the results of this method of determining lysine.

It is of some interest to attempt an estimate of the relation between $[\text{F}]$ and the p_K values of the F groups. This may be done in a rather empirical manner from the comparison curves by reading off at a given ΔH the p_H values corresponding to the several concentrations of formaldehyde. The following results may be quoted to illustrate the conclusion that $\frac{[\text{H}^+]}{[\text{F}]^2}$ is substantially constant when F is greater than 0.2.

$-\Delta\text{H}$	$[\text{F}]$	p_H	$p_H + 2 \log [\text{F}]$
7.5	0.10	9.13	7.13
7.5	0.20	8.67	7.27
7.5	0.60	7.65	7.30
7.5	1.26	7.10	7.30

This conclusion which presumably defines the variation of the p_K of the lysine groups with $[\text{F}]$ conforms to the behaviour of NH_2 groups in general as established by Levy.

There remains for consideration the section of the comparison curves above $p_H 10$. This is to be interpreted as the hydrogen ion association curve of the lysine groups in the absence of formaldehyde. Were it possible to trace the full course of

this curve we should have a measure of the effective p_K of the lysine groups. Unfortunately the errors of the formaldehyde titration at high p_H do not permit this. The available segment of the curve is however consistent with the conclusion that the dissociation of the lysine groups of the protein in aqueous solution simulate those of an acid of p_K close to 11.

In the light of the data which have been presented we may in conclusion submit certain recommendations respecting the use and interpretation of formaldehyde titrations of protein solutions conducted with the aid of indicators. If it is desired to measure the total F groups other than $\alpha\text{-NH}_2$, the most satisfactory conditions would appear to be (a) adjustment of the solution to an initial p_H 9, (b) addition of neutral formaldehyde to a concentration of 2 *M* and (c) titration of the mixture back to p_H 9. At this p_H the buffering due to the protein alone and to the formaldehyde alone is small and the titration of the F groups is substantially complete. The end-point is fairly sharp. Should steps (b) and (c) lead to an important change of total volume then a correction must be introduced for the H^+ dissociated by the formaldehyde in this increased volume. It is probable though not established that the same conditions should prove the best for the majority of proteins.

If the titration is to include $\alpha\text{-NH}_2$ groups, then the initial adjustment of p_H should be to p_H 6.5. Owing to the rather effective buffering power of the protein in this region, this adjustment will not be very precise. The formaldehyde is then added and the titration conducted to p_H 9. Such a titration includes the H^+ dissociated by the protein between p_H 6.5 and 9 in the absence of formaldehyde. This can be ignored if relative changes in $\alpha\text{-NH}_2$ groups are sought but may be approximately controlled by an independent titration in the absence of formaldehyde. It should be stated in conclusion that salt effects on the ionisation of the protein are so large that serial titrations conducted without control of the ionic strength may be subject to serious errors.

EXPERIMENTAL.

Electrometric titrations.

The technique of titration was the same as that reported in the first paper of the series [Kekwick and Cannan, 1936], except for the addition to each reaction mixture of a measured volume of a standard formaldehyde solution. The latter was prepared and standardised after the manner described by Levy [1933]. The titration curve for 0.6 *M* formaldehyde shown in Fig. 1 was calculated from the results of titrations of two different protein preparations—both of which had been used for titrations described in the first paper. One of these was a Na_2SO_4 preparation which had not been dried. The other was one of the dried preparations. The titration curve for 0.1 *M* formaldehyde refers to the former preparation only and the curves for 0.2 and 1.26 *M* formaldehyde to the latter only.

The nitrous acid reaction.

The amount of nitrogen formed when nitrous acid reacts with egg albumin under the conditions of the Van Slyke manometric method was found to vary with the time of reaction. In Fig. 2 are shown the results obtained when the time was varied from 6 to 96 min. at 21–23°. The results are calculated as equiv. of NH_2 per mol. of egg albumin. It is clear that the rapid initial reaction is followed by a relatively slow increase which is approximately linear in rate. We may presume that the former is the reaction which is due to the presence of

NH₂ groups in the native protein. We believe therefore that the most reliable determination of these is secured by extrapolating to zero time the linear rate of secondary change.

The nitrogen in the basic fraction of hydrolysed egg albumin.

Van Slyke's method [1911] was followed with one modification. The original precipitate of the phosphotungstates was recrystallised once from a volume of washing fluid (2.5% H₂SO₄ containing 0.5% phosphotungstic acid) equal to the volume of the mixture in which it had been formed originally. It was then washed and treated in the manner described by Van Slyke. The recrystallisation was carried out with the hope of reducing the danger of including neutral amino-acids in the final lysine nitrogen.

About 8 g. of albumin were used in each of three determinations. The three experiments agreed within the limits arginine 10.5–11.0, histidine 2.8–3.1 and lysine 17.7–18.4 equiv. of amino-acid per mol. The figures for arginine and histidine agree well with those of Vickery and Shore, though it must be pointed out that our results for histidine are subject to a large error due to the solubility correction being a large fraction of the whole. Our figure for lysine is, as has been discussed, much larger than the values found by authors using isolation methods.

SUMMARY.

1. Hydrogen ion dissociation curves of crystalline egg albumin in 0.1, 0.2, 0.6 and 1.26 *M* formaldehyde are reported.
2. From an analysis of these curves it is concluded that the egg albumin molecule contains 16–18 NH₂ groups other than α -NH₂ groups. These are tentatively identified with the ϵ -NH₂ groups of lysine.
3. The reaction of egg albumin with nitrous acid indicates the presence of 19–20 NH₂ groups per g.mol. The lysine-nitrogen of the phosphotungstate fraction of the hydrolysed protein corresponds with 17–18 molecules of lysine per mol.

REFERENCES.

- Birch and Harris (1930). *Biochem. J.* **24**, 1080.
Calvery (1932). *J. Biol. Chem.* **94**, 613.
Kekwick and Cannan (1936). *Biochem. J.* **30**, 227.
Levy (1933). *J. Biol. Chem.* **99**, 767.
—— (1934). *J. Biol. Chem.* **105**, 157.
—— (1935). *J. Biol. Chem.* **109**, 365.
Van Slyke (1911). *J. Biol. Chem.* **10**, 15.
Vickery and Shore (1932). *Biochem. J.* **26**, 1101.