

Hydrophobicity of Long Chain *n*-Alkyl Carboxylic Acids, as Measured by Their Distribution Between Heptane and Aqueous Solutions

(fatty acids/lipids/hydrogen bonds)

ROSS SMITH AND CHARLES TANFORD

Department of Biochemistry, Duke University Medical Center, Durham, North Carolina 27710

Contributed by Charles Tanford, November 17, 1972

ABSTRACT Previous data for the distribution of the *n*-alkyl carboxylic acids between aqueous buffers and heptane have suggested that the free energy of transfer between the two solvents does not increase with the length of the alkyl chain beyond palmitic acid, indicating a limit to the hydrophobicity of the alkyl chain at a length of about 15 carbon atoms. It is shown in this paper that previous measurements were influenced by unexpectedly strong pH-dependent association of the long-chain acids in the aqueous phase. Under conditions where such association becomes negligible, the free energy of transfer becomes a strictly linear function of alkyl chain length for all acids up to behenic acid (C₂₁H₄₃COOH).

This work was undertaken to examine the suggestion of Mukerjee (1, 2) that the hydrophobicity of *n*-alkyl carboxylic acids does not increase indefinitely with the length of the alkyl chain, but reaches a limit at a chain length of about 15 carbon atoms. Mukerjee's suggestion was based on the studies of Goodman (3) of the distribution of fatty acids between aqueous solutions and liquid heptane, and in particular on anomalous results obtained for C₁₅COOH and C₁₇COOH. In this paper the measurements of Goodman have been repeated and extended, and their interpretation has been reinvestigated. An additional acid, behenic acid (C₂₁COOH), has been included in the study.

THERMODYNAMIC FORMULATION

The unitary free energy of transfer of an acid HA between water and a hydrocarbon such as heptane may be taken as a measure of its hydrophobicity; this quantity, in turn, is proportional to the logarithm of the distribution coefficient $K_p = [HA]_{HC}/[HA]_W$. The extremely high values of K_p for the long chain acids in fact preclude direct measurement of $[HA]_W$ under conditions where $[HA]_{HC}$ is sufficiently small to permit approach to ideal solution behavior. Goodman (3) circumvented this difficulty by use of a pH far above the expected pK_a for the acid. Under these conditions HA in the aqueous phase is in equilibrium with a large excess of the anionic form A⁻, and, assuming otherwise ideal behavior in both solvents, the observed distribution ratio R of total acid may be written as

$$R_0 = [HA]_{HC}/[A^-]_W = K_p \alpha_H + K_a \quad [1]$$

Abbreviations: HA represents any acid in the series, and A⁻ is the corresponding anion. C_{*n*}COOH is used to designate a particular acid, e.g., C₁₅COOH is palmitic acid. Square brackets represent molar concentrations, and the subscripts W and HC refer to the aqueous and hydrocarbon solvents, respectively.

where R_0 is used to indicate that the equation applies only under limiting ideal conditions.

Goodman (3) observed that experimental values of R depend on total acid concentration. He attributed this dependence to dimer formation in the organic phase, $2HA \rightleftharpoons H_2A_2$, for which there is abundant evidence (4-6). For C₇COOH, C₉COOH, C₁₁COOH, and C₁₃COOH, the concentration dependence could indeed be accounted for in terms of reasonable values for the equilibrium constant for the dimerization process. For these acids, therefore, measurements of R at sufficiently low concentration, where R becomes independent of concentration, can confidently be used to evaluate K_p by Eq. 1. Values of $\log K_p$ obtained in this way are linear functions of chain length, and both absolute values and the dependence on chain length are reasonable in the light of other data, as has been discussed elsewhere (7).

The observed concentration dependence of R for C₁₅COOH and C₁₇COOH is, however, less than expected on the basis of dimerization in the organic phase. Mukerjee (1, 2) suggested that this anomaly could be explained by assuming a dimerization reaction in the aqueous phase also, with the formation of A₂⁻. Since this process, too, should become unimportant at low acid concentrations, it remains valid to calculate K_p by Eq. 1 from limiting values of R . Values of $\log K_p$ obtained in this way by Mukerjee do not lie on an extension of the linear relation between $\log K_p$ and chain length for the shorter acids. The values for C₁₅COOH and C₁₇COOH are, in fact, almost the same (1, 2, 7). This result was the basis for Mukerjee's suggestion that the alkyl chains of the longer fatty acids fold upon themselves, so that no increase in hydrocarbon-water interface occurs when the alkyl chain length exceeds about 15 carbon atoms.

There is considerable evidence that dimerization of moderately long-chain fatty acids in water occurs (2, 8), but it appears probable (9) that the principal dimer is not A₂⁻ but HA₂⁻, and that a hydrogen bond between a protonated and an unprotonated carboxyl group plays an important role in the association. Evidence for the formation of remarkably strong hydrogen bonds of this type also comes from pK_a measurements of dicarboxylic acids containing hydrocarbon groups (10-13). Moreover, the work of Eagland and Franks (9) has shown that formation of HA₂⁻ alone cannot account for anomalous titration behavior of C₁₁COOH, C₁₃COOH, and C₁₅COOH, and that higher aggregates such as H₂A₃⁻, H₂A₄⁻, etc., must be invoked to explain the results. Since the formation of all such associated species depends on pH as well as total concentration, these findings suggest that pH may have

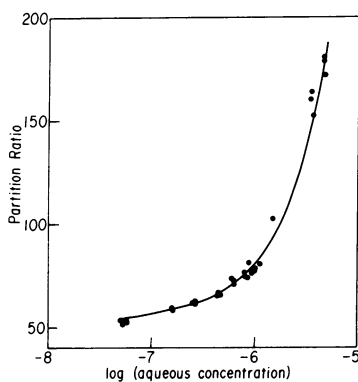


FIG. 1. Results for $C_{13}COOH$ at pH 7.40, ionic strength 0.10, 25° . The solid line is calculated from Eq. 2, with $R_0 = 53.5$, $K_d = 4700$, and the denominator equal to unity.

an important influence on the distribution measurements, more complex than its effect on the equilibrium between HA and A^- , which is incorporated in Eq. 1. The pH of 7.45 used by Goodman may not be sufficiently high to prevent interference by HA_2^- and/or higher aggregates with the interpretation of the data.

The procedure followed in the present study is the same as that of Goodman (3), except that we have investigated the effect of pH, as well as the effect of concentration, on the results. If the formation of H_2A_2 in the organic phase and of HA_2^- and higher aggregates in the aqueous phase is taken into account, the effects of these variables may be described in terms of the relation

$$R = R_0 \frac{1 + 2K_d K_p [HA]_w}{1 + 2K_b [HA]_w + \text{higher terms}} \quad [2]$$

where R_0 is the ideal expression for R as given by Eq. 1, K_d is the equilibrium constant for dimerization in the organic phase, and K_b is the equilibrium constant for the reaction $A^- + HA \rightleftharpoons HA_2^-$ in the aqueous phase.

Eq. 2 shows that R will approach R_0 with increase in pH (i.e., decrease in the proportion of HA in the aqueous solution), as well as with decrease in total concentration. To use Eq. 1 for the calculation of K_p , two criteria must be satisfied. One is the criterion already used by Goodman and Mukerjee, namely that R (at constant pH) must approach a value inde-

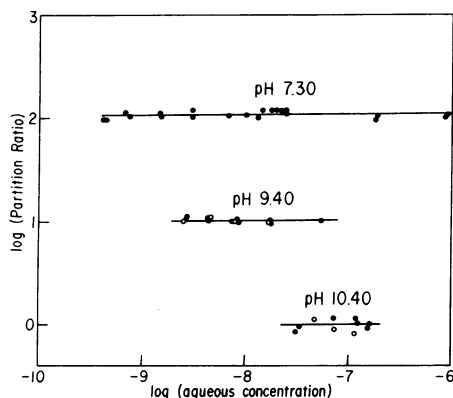


FIG. 2. Results for $C_{15}COOH$ at three pH values. Ionic strength 0.10 or 0.15, 25° , except that open circles are at ionic strength 0.01.

pendent of acid concentration. The second criterion is that the pH must be sufficiently high for R/a_{H^+} to approach a value independent of pH. Goodman showed that this second condition was satisfied in his study for $C_{11}COOH$ at pH 7.45, but the pH-dependence was not investigated for the longer acids. The results of the present work will show that a pH-dependence of R/a_{H^+} in fact persists for $C_{15}COOH$ and the longer acids to pH values well above the pH at which Goodman's measurements were made, and that this fact alone accounts for the anomalous values of K_p obtained from his data.

Eq. 2 can also be used to obtain information on the importance of higher aggregates in the aqueous solution. If the higher terms in the denominator of Eq. 2 can be neglected, the direction of change in R with concentration is determined solely by the relative magnitudes of $K_d K_p$ and of K_b . If $K_d K_p > K_b$, R will increase with concentration at any pH, and if $K_d K_p < K_b$, the reverse result will be obtained, again regardless of pH. Furthermore, K_d , which represents the tendency for dimerization in the organic phase, should be little influenced by alkyl chain length, and the values of K_d for C_7COOH , C_9COOH , and $C_{11}COOH$ obtained by Goodman are, in fact, nearly the same. On the other hand, K_p and K_b both involve the hydrophobicity of the acid, and they should, therefore, increase with alkyl chain length. Simple considerations, such as those discussed by Mukerjee (2), suggest that $\log K_p$ and $\log K_b$ should vary about equally with chain length. As a first approximation, therefore, the ratio $K_d K_p / K_b$ should be independent of alkyl chain length. Thus, a criterion for the neglect of higher aggregates is that the direction of change in R with concentration be independent of pH for a given acid, and that it should not vary dramatically as the alkyl chain length is increased.

If higher terms in the denominator of Eq. 2 are important, however, the direction of change of R with increasing concentration will depend both on pH and alkyl chain length. Thus, if $H_2A_4^-$ contributes significantly to the species in the aqueous phase, the denominator of Eq. 2 would become

$$1 + 2K_b [HA]_w (1 + 2K_c K_b [HA]_w [A^-]_w)$$

where K_c is the equilibrium constant for the reaction $2HA_2^- \rightleftharpoons H_2A_4^-$. $\log K_c$ would be expected to increase more rapidly with chain length than $\log K_b$ or $\log K_p$, and the net effect on Eq. 2 would be that the dominant effect of concentration could shift from the numerator to the denominator both as pH is decreased (i.e., as $[HA]_w$ increases) and as the alkyl chain is lengthened.

In calculating K_p from the limiting values of R by use of Eq. 1, Goodman assumed that K_d could be taken as indepen-

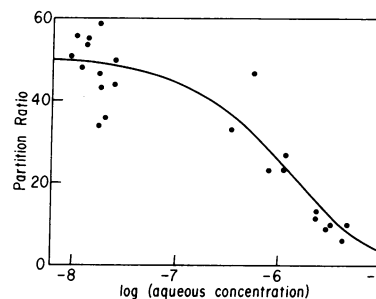


FIG. 3. Results for $C_{17}COOH$ at pH 9.84, ionic strength 0.15, 25° . R_0 was taken to be 50 ± 15 .

dent of alkyl chain length. We have made the same assumption, as it is supported by all available data. Measurements in aqueous solutions (14) show that pK_a values for the series C_2COOH to C_7COOH are identical to better than ± 0.1 . Measurements in 2-ethoxyethanol (15) show no effect of chain length up to $C_{17}COOH$. *n*-Alkyl ammonium ions show only slight scattered variations in pK_a (in water) as the chain length varies from methyl to decyl (14). In calculating K_p from experimental values of R_0 by Eq. 1, we have set $pK_a = 4.76$ —somewhat smaller than reported values for zero ionic strength—because most of the experiments were made at ionic strength 0.10 or 0.15. A small error in the choice of the value for pK_a would not affect the conclusions, because it is the variation in K_p with alkyl chain length, and not its absolute value, that is of primary interest. Erroneous conclusions would be drawn if pK_a were to vary significantly with chain length for the longer acids, contrary to expectation, but the experimentally observed anomalies cannot in fact be explained by invoking this possibility, as the results will show.

EXPERIMENTAL PROCEDURE

[^{14}C]myristic, [^{14}C]palmitic, and [^{14}C]stearic acids, and [^{14}C]behenic acid, as well as the corresponding unlabeled compounds, were purchased from Applied Science Laboratories (State College, Pa.). They were judged pure within experimental error by thin-layer and gas-liquid chromatography. Radiochemical purity of the labeled compounds was 98% or better. Other reagents were the best available commercial products. Water was glass distilled.

Ammonium phosphate, sodium carbonate, or tris-(hydroxymethyl)aminomethane buffers were used, with addition of NaCl to an ionic strength of 0.10 or 0.15. The nature of the buffer used did not affect the results. The pH was determined at the conclusion of each set of experiments with a Radiometer Type B glass electrode. Corrections for sodium ion error were applied.

Partition experiments were performed with equal volumes (4–8 ml) of heptane and aqueous buffer, in glass tubes with Teflon-sealed stoppers. Fatty acids were initially dissolved in the heptane phase, or they were initially dissolved in the aqueous phase after neutralization with NaOH. In this way equilibrium was approached in both directions in each series of experiments. The extent of enrichment of the unlabeled acid with the radioactive compound was modified as needed to obtain reasonable counts above background. The radioactive acids were stored in benzene solution and the unlabeled acids were sometimes added as concentrated solutions in ethanol. In no experiment did the amount of benzene or

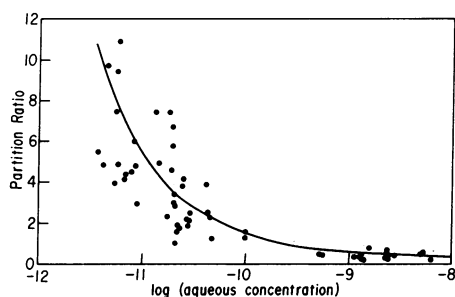


FIG. 4. Results for $C_{21}COOH$ at pH 11.87, ionic strength 0.15, 25° . A limiting value for R cannot be obtained from the data.

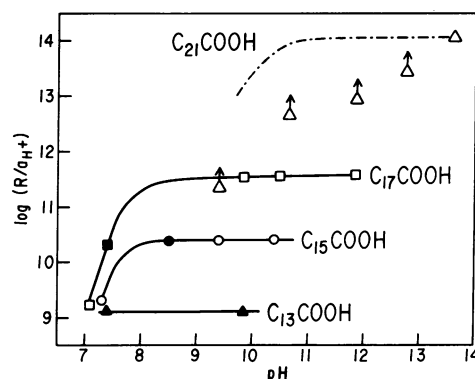


FIG. 5. Limiting values for R/a_{H^+} as a function of pH. For $C_{21}COOH$, except at the highest pH, the largest experimental values of R/a_{H^+} are shown, and the arrows indicate the direction of change with decreasing concentration. The filled square is from the data of Goodman (3), the filled circle from the data of Arvidsson *et al.* (16).

ethanol introduced in this way exceed 0.05% of the total volume.

The tubes were shaken in a water bath at $25.00 \pm 0.05^\circ$. Equilibrium was attained in 3–5 days, and additional shaking for a week or more had no effect. Calibrated glass pipets were used to transfer samples from each phase into the scintillation fluid. The pipets were washed to remove adhering or adsorbed acid and the washings were also counted. The scintillation fluid contained 2.67 g of 2,5-diphenyloxazole, 67 mg of 1,4-bis-[2-(5-phenyloxazolyl)]benzene and 333 ml of Triton X-100, made up to 1 liter with toluene. 15 ml of this solution was used for each sample to be counted. Solutions were acidified where necessary to keep the counting solutions acidic. Counting was done in a Packard model 3320 Tri-Carb Liquid Scintillation Spectrometer, at $79 \pm 3\%$ efficiency. Addition of a large quantity of unlabeled fatty acid to the scintillation vial before addition of samples did not alter the counting rates, indicating that adsorption onto the walls of the vials was insignificant.

The partition ratio (R) is defined as the ratio of total counts per unit volume of the heptane phase to total counts per unit volume of the aqueous phase, corrected for background counts. Actual concentrations were determined from the known specific activity of each labeled acid and the proportion of labeled to unlabeled acid in the solution.

RESULTS

Partition ratios were measured as a function of concentration at several pH values. Typical data are shown in Figs. 1–4. Each set of results includes experiments in which equilibrium was approached in opposite directions, i.e., with all the solute initially in a different phase. Identical results within experimental error were always obtained. Limiting values of R/a_{H^+} were obtained from the concentration-independent values of R approached at low solute concentrations; they are plotted as a function of pH in Fig. 5. It should be noted that experiments at the very lowest concentrations used tend to become inaccurate, as the counting rate approaches background noise.

The results for $C_{13}COOH$ at pH 7.4 (Fig. 1) agree closely with those of Goodman at nearly the same pH. The concentration dependence shows that $K_d K_p > K_b$ and can, in fact, be

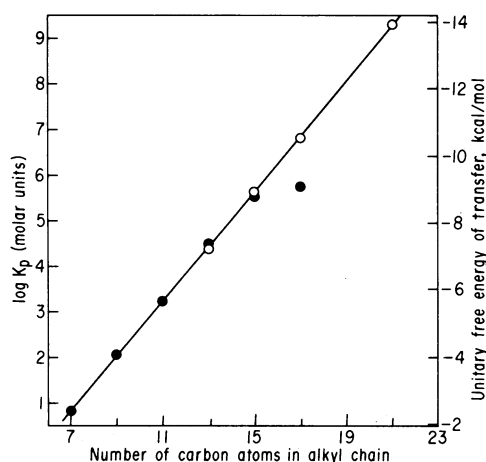


FIG. 6. Logarithm of the true distribution constant (K_p) for $C_n\text{COOH}$ as a function of alkyl chain length. Open circles represent the present results, filled circles are data of Goodman (3). The right-hand ordinate shows unitary free energies of transfer from water to heptane, calculated as discussed earlier (7).

described by Eq. 2 without inclusion of any association terms in the denominator. The value of K_d obtained in this way is 4700 liter/mol, somewhat smaller than the value of 9300 liter/mol reported by Goodman on the same basis. Measurements were also made at pH 9.82, yielding the same limiting value of R/a_{H^+} as at pH 7.4. The conditions for applicability of Eq. 1 are, therefore, satisfied. The value of $\log K_p$ (molar units) is 4.37 ± 0.03 .

The partition ratios obtained for $C_{15}\text{COOH}$ at pH 7.3 (Fig. 2) are considerably lower than those obtained by Goodman at pH 7.45. Limiting values of R/a_{H^+} obtained from both determinations are lower than those obtained from Fig. 2 at pH 9.4 and 10.4 or from the data of Arvidsson *et al.* (16) for the same acid at pH 8.5. For this acid the conditions for applicability of Eq. 1 are evidently not met at neutral pH, but Fig. 5 shows that they are satisfied at pH 8 and above. The value of $\log K_p$ obtained from the data at higher pH is 5.64 ± 0.05 . A striking aspect of all of the studies with $C_{15}\text{COOH}$, at all pH values, is that R is virtually independent of concentration, although concentrations in the organic phase are as high in some of the experiments as those at which marked association occurs for $C_{13}\text{COOH}$. Since K_d is expected to be independent of chain length, this result means that the association-dependent terms in the denominator of Eq. 2 must under the conditions of the experiments be of comparable magnitude to the term $2K_dK_p[HA]_w$ in the numerator.

For $C_{17}\text{COOH}$, R is essentially independent of concentration at pH 11.82, but at pH 10.5 and 9.8 (Fig. 3) there is marked concentration-dependence, in the opposite direction from that observed for $C_{13}\text{COOH}$. This effect of pH on the concentration dependence of R can be explained in terms of Eq. 2 only if association in the aqueous phase involves higher aggregates than HA_2^- ; the dramatic reversal in direction in going from $C_{13}\text{COOH}$ to $C_{17}\text{COOH}$ is consistent with this conclusion. At pH 7.1 we again observed R to be nearly independent of concentration, and a similar result was obtained by Goodman at pH 7.45. The limiting values of R/a_{H^+} shown in Fig. 5, are all self-consistent and indicate that for this acid the conditions for applicability of Eq. 1 are not

satisfied except at pH 9 and above.* The value of $\log K_p$ obtained from the results above pH 9 is 6.8 ± 0.15 .

For $C_{21}\text{COOH}$, all measurements were made above pH 9. The concentration dependence was always in the same direction as for $C_{17}\text{COOH}$ in Fig. 3, except at the highest pH used (pH 13.60), where R became independent of concentration. In the experiments where concentration dependence was observed, it was impossible to extend measurements to a sufficiently low concentration to obtain a limiting value of R , as is illustrated by Fig. 4. At pH 12.8, however, the values of R/a_{H^+} observed at the lowest attainable concentration approach sufficiently close to the value at pH 13.6 that it is reasonable to assume that conditions for the validity of Eq. 1 are being approached. The value of $\log K_p$ obtained from the data at pH 13.6 is 9.30 ± 0.15 .

The values of $\log K_p$ calculated from the values of R/a_{H^+} by Eq. 1, when only data obtained at pH values where the applicability of this equation can be established are used, are plotted as a function of alkyl chain length in Fig. 6. It is seen that the results of this paper, together with Goodman's results for the shorter acids, are strictly linear to the highest chain length tested. There is no limit to the hydrophobicity of the alkyl chain, such as the earlier data obtained at lower pH had suggested.

It may be noted that the results obtained strongly suggest the correctness of the assumed constancy of pK_a for all of the acids in this series. If the anomalous K_p values derived from Goodman's data were to be ascribed entirely to anomalous pK_a values for the longer acids, the limiting values of R/a_{H^+} in Fig. 5 would have been independent of pH for each individual acid. In fact they are not, and the anomaly in the K_p values disappears if the pH-dependence of R/a_{H^+} is taken into consideration. This result does not prove the absence of an anomaly in pK_a , but introduction of such an anomaly would lead to uninterpretable variation in K_p with chain length.

DISCUSSION

It is evident from Fig. 6 that the free energy of transfer of $C_n\text{COOH}$ between heptane and aqueous solutions increases linearly with the length of the alkyl chain, at least as far as $n = 21$, corresponding to the longest acid used in this work. The apparent limit to the hydrophobicity of an alkyl chain at a length of 15 carbon atoms, deduced from earlier data, has been shown to result from the use of Eq. 1 for the calculation of K_p under conditions where this equation is, in fact, invalid. Our result is consistent with the observation to Evans (17) that the free energy of micellization of alkyl sulfates changes linearly with chain length to at least 18 carbon atoms.

Failure of the results to obey Eq. 1 arises from association between solute molecules in either phase. Our results show unequivocally that dimeric species cannot account for the observed deviations from Eq. 1, and that higher aggregates must be formed in the aqueous phase. We do not have sufficient detailed data at relatively low pH to be able to determine the stoichiometry or formation constants of the dominant aggregated species. The very high pH that has to be used before aggregation can be neglected indicates that the aggregated species must retain one or more protons with remark-

* One set of results we obtained at pH 8.3 has been omitted. It was inconsistent with all other data, both in the limiting value of R/a_{H^+} and the concentration dependence.

able tenacity. A corollary is that hydrogen bonds between COO^- and COOH are extraordinarily stable when carboxyl groups are attached to long alkyl chains, a conclusion that is supported by other evidence (9-13).

All our experimental results were obtained at concentrations in the aqueous phase below the critical micelle concentrations of the acid anions. However, mixed micelles containing both HA and A^- would be stabilized by hydrogen bonds between the charged and uncharged head groups, and the formation of such micelles at very low concentrations is possible. The strength of the hydrogen bond between COO^- in a hydrophobic environment also suggests that it may not be entirely safe to assume the absence of aggregates containing acid anions (and, presumably, accompanying counterions) in the organic phase. Further study of the system at lower pH is needed to provide convincing evidence on these questions.

This work was supported by National Science Foundation Grant GB 14844 and by U.S. Public Health Service Grant AM-04576. C. T. is a Research Career Awardee, National Institutes of Health, U.S. Public Health Service.

1. Mukerjee, P. (1965) *J. Phys. Chem.* **69**, 2821-2828.
2. Mukerjee, P. (1967) *Advan. Colloid Interface Sci.* **1**, 241-275.
3. Goodman, D. S. (1958) *J. Amer. Chem. Soc.* **80**, 3887-3892.
4. Pohl, H. A., Hobbs, M. E. & Gross, P. M. (1951) *J. Chem. Phys.* **9**, 408-414.
5. Maryott, A. A., Hobbs, M. E. & Gross, P. M. (1949) *J. Amer. Chem. Soc.* **71**, 1671-1674.
6. Pauling, L. (1960) *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, N.Y.), 3rd ed., p. 477.
7. Tanford, C. (1972) *J. Mol. Biol.* **67**, 59-74.
8. Danielson, I. & Stenius, P. (1971) *J. Colloid Interface Sci.* **37**, 264-280.
9. Eagland, D. & Franks, F. (1965) *Trans. Faraday Soc.* **61**, 2468-2477.
10. Westheimer, F. H. & Benfey, O. T. (1956) *J. Amer. Chem. Soc.* **78**, 5309-5311.
11. Tanford, C. (1957) *J. Amer. Chem. Soc.* **79**, 5348-5352.
12. Ebersson, L. (1959) *Acta Chem. Scand.* **13**, 211-223.
13. Ebersson, L. & Wadsö, I. (1963) *Acta Chem. Scand.* **17**, 1552-1562.
14. Sober, H. A., ed. (1970) *Handbook of Biochemistry* (Chemical Rubber Co., Cleveland, Ohio), 2nd ed., pp. J191-J205.
15. Garvin, J. E. & Karnovsky, M. L. (1956) *J. Biol. Chem.* **221**, 211-222.
16. Arvidsson, E. O., Green, F. A. & Laurell, S. (1971) *J. Biol. Chem.* **246**, 5373-5379.
17. Evans, H. C. (1956) *J. Chem. Soc.* 579-589.