Stryer, L., and Blout, E. R. (1961), *J. Am. Chem. Soc. 83*, 1411.
Toohey, J. I. (1965), *Proc. Natl. Acad. Sci. U. S. 54*, 934.
Ulmer, D. D., and Vallee, B. L. (1965), *Advan. Enzymol. 27*, 37.
Weissbach, M., and Taylor, R. (1966), *Federation Proc. 25*, 1649.

A Sedimentation Equilibrium Study of the Association of Purine in Aqueous Solutions*

K. E. Van Holde and G. P. Rossetti

ABSTRACT: The association of purine in aqueous solution has been studied by the sedimentation equilibrium technique. The results are in qualitative agreement with the conclusion of Ts'o *et al.* (Ts'o, P. O. P., Melvin, I. S., and Olson, A. C. (1963), *J. Am. Chem. Soc. 85*, 1289) that a simple reversible polymerization is involved, with a constant free-energy increment for the addition of each successive purine molecule. However, for quantitative analysis of the equilibrium, the thermodynamic nonideality of monomer and polymers must be taken into account; when such corrections have been made, the apparent equilibrium constants become invariant over the entire concentration range. Studies at a series of temperatures have provided values for the enthalpy and entropy changes in the process. The results obtained for ΔG° and ΔH° are somewhat larger than those found previously (Ts'o, P. O. P., Melvin, I. S., and Olson, A. C. (1963), *J. Am. Chem. Soc. 85*, 1289; Gill, S. J., Downing, M., and Sheats, G. F. (1967), *Biochemistry 6*, 272).

Several years ago, Ts'o and collaborators demonstrated that purine (and a number of similar compounds as well) was capable of forming "stacked" aggregates in aqueous solution (Ts'o *et al.*, 1963; Chan *et al.*, 1964). These results lent substance to the earlier suggestion of Sturtevant *et al.* (1958) that hydrogen bonding alone probably could not account for the stability of DNA. Further confirmation for the importance of stacking interactions came from the demonstration that a number of synthetic polynucleotides could form single-strand helices in aqueous solution (see, for example, Witz and Luzzati, 1965; Holcomb and Tinoco, 1965; Van Holde *et al.*, 1965; Brahms *et al.*, 1966; Leng and Felsenfeld, 1966; Poland *et al.*, 1966).

The purine association, which may be regarded as the prototype for such processes, was studied by Ts'o *et al.* by vapor pressure osmometry and nuclear magnetic resonance spectrometry. It was concluded that the reaction was probably a simple polymerization, with equal equilibrium constants for successive additions of monomers to the stack. More recently, Gill *et al.* (1967) have used heat of dilution measurements to obtain the enthalpy and entropy changes for the reaction at 25°.

Since these association processes are of such fundamental importance to molecular biology, we have begun a series of physicochemical and optical studies of such systems. As a first step, we felt that a critical reexamination of the purine association, using the more powerful method of sedimentation equilibrium, would be worthwhile. Such measurements should provide a more stringent test of the association mechanism, and also allow nonideality of the solutions, which might be expected to be appreciable at high concentrations, to be taken into account. We present here the results of that study.

Experimental Section

Reagents. The purine used in all experiments was A grade, purchased from Calbiochem. According to the manufacturer's specifications, it was chromatographically homogeneous and contained 46.71% nitrogen (theory 46.66%). All solutions were prepared by weight.

Partial Specific Volume. Apparent partial specific volumes were determined at 25.00°, using pycnometers fabricated by fusing 1-mm i.d. capillaries to 25-ml volumetric flasks. Results at purine concentrations of 5.63 and 9.56 g/l. were 0.702 and 0.701 ml/g, respectively. Since the difference is within anticipated experimental uncertainty, the partial specific volume has been taken to be 0.701 ml/g.

Sedimentation Equilibrium. A Beckman-Spinco Model

* From the Department of Chemistry and Chemical Engineering, University of Illinois, Urbana, Illinois. *Received March 2, 1967.* This research was supported in part by the U. S. Public Health Service Grant GM-13373.

E analytical ultracentrifuge was used. The RTIC calibration was checked twice during the course of the experiments. Except for one series of experiments at different temperatures, all were at 24.9°. Because of the high concentrations employed, phase-plate schlieren optics were used in all experiments. The optical system had been focussed by the method of Gropper (1964) just before the series of experiments was begun. All photographs were measured with a Bausch and Lomb two-way comparator.

In order to determine the initial concentrations in terms of areas on the photographic plates, synthetic boundary experiments were performed with 0.2 M purine solutions. Duplicate experiments gave results which agreed within 0.2%.

In the sedimentation equilibrium experiments, column lengths between 4 and 9 mm were used. In each case, sufficient time was allowed for equilibrium, according to the criterion of Van Holde and Baldwin (1958). Rotor speeds were so chosen as to yield a readily measurable refractive index gradient; nominal values ranged from 35,600 to 56,100 rpm. All were checked by odometer readings over the last several hours of each experiment.

Results

In a precise study of the association of a low molecular weight substance in concentrated solutions, attention must be paid to the effect of solute nonideality on the observed apparent molecular weights. While the introduction of arbitrary activity coefficients for all possible species will make the analysis unmanageably complicated, it is to be hoped that some simple and reasonable correction for nonideality can be accommodated. A test for the success of such a treatment will be the invariability of apparent equilibrium constants over a wide concentration range.

Adams and Williams (1964) have shown that sedimentation equilibrium of a reversibly associating, nonideal solute can be analyzed so as to yield information on both the association equilibrium and the nonideality. If a general system, containing monomer, dimer, trimer... i-mer, etc., is considered, it can be shown that the apparent weight-average molecular weight at a point in the cell where the weight concentration is c given by

$$\frac{1}{M_{wa(c)}} = \frac{1}{M_{w(c)}} + Bc \tag{1}$$

where $M_{w(c)}$ is the true weight-average molecular weight of the mixture at concentration c and B is a virial coefficient. Equation 1 requires the reasonable assumption that the effective virial coefficient for species i is proportional to the molecular weight of species i; that is

$$\ln y_i = iBM_1c \tag{2}$$

where y_i is the activity coefficient. Other assumptions may be made, but this leads to the simplest results, and is consistent with the concept of the virial coefficient as a measure of molar excluded volume. It also may be shown (Adams and Williams, 1964) that if eq 2 is valid, the apparent equilibrium constant for a step in the polymerization will be equal to the thermodynamic equilibrium constant

$$K_i = \frac{a_{i+1}}{a_1 a_i} = \frac{C_{i+1}}{C_1 C_i} \tag{3}$$

where a's are activities and C's molar concentrations.

The quantity $M_{wa(c)}$ in eq 1 may be determined unambiguously from sedimentation equilibrium data

$$M_{wa(c)} = \frac{RT}{\omega^2(1 - \bar{v}\rho(c))} \frac{dc/dr}{rc} \tag{4}$$

or

$$M_{wa(c)} = \frac{RT}{\omega^2(1 - \bar{v}\rho(c))r\left(A_a + \int_a^r Z dr\right)} \frac{Z}{} \tag{5}$$

In these equations ω is the rotor velocity, R the gas constant, T the temperature, $\bar{v}$ the partial specific volume, $\rho(c)$ the solution density at concentration c (grams per milliliter), r the distance from the center of rotation, and Z the phase-boundary image displacement. The quantity A_a is a photographic plate area corresponding to the concentration at the meniscus. It is calculated in the usual fashion, from the area under a synthetic boundary and an integration across the sedimentation equilibrium pattern, utilizing the conservation of mass.

Values of $M_{wa(c)}$ are shown *vs.* C for purine at 24.9° in Figure 1. Data from different experiments are included; it is clear that these all lie upon a single curve, as they should for a reversibly associating system. The curve extrapolates at $C = 0$ to the known monomer weight of purine (120). The maximum error in the molecular weight values are indicated by the vertical bars in Figure 1. Additivity of reasonable uncertainties in ω, r, T, and Z was assumed, but an uncertainty of 1% in $\bar{v}$ (which will affect all points almost equally) was not considered. The error in $\bar{v}$ must not be large, since M_1 was obtained with accuracy.

It can be judged immediately from Figure 1 that: (a) association takes place; (b) that it proceeds, at least in part, past dimerization, and (c) as evidenced from the downward curvature at very high purine concentrations, there is an appreciable positive virial coefficient. A more detailed analysis will be given in the following section.

Given these experimental data, it is also possible to calculate the apparent number-average molecular weight as a function of concentration. According to Adams (1965)

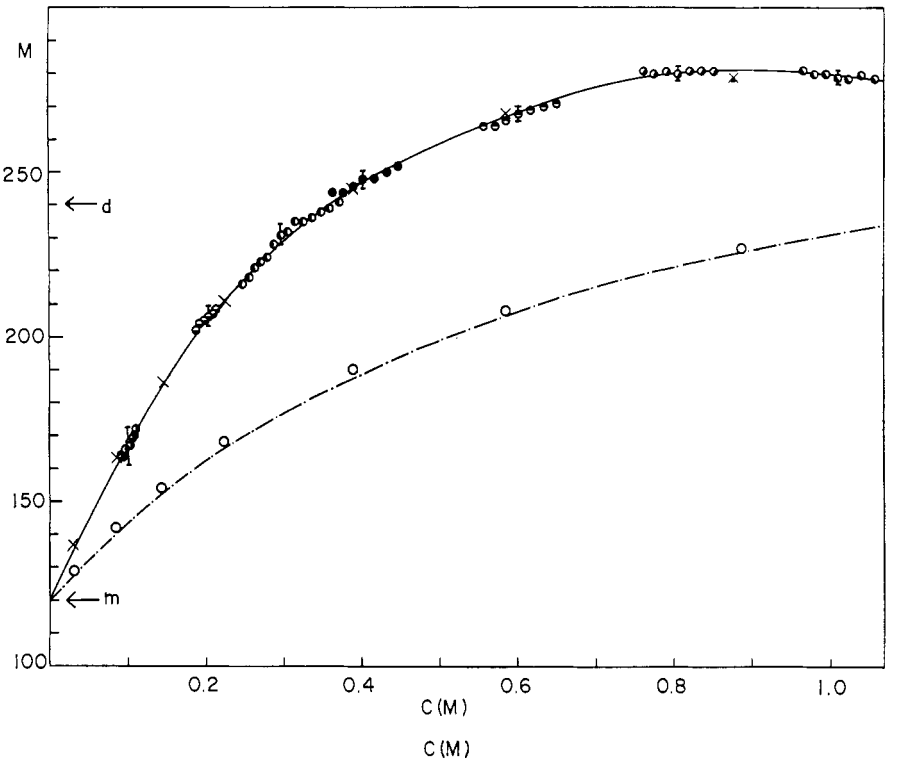

FIGURE 1: Apparent average molecular weights *vs.* molar concentration for purine in aqueous solution at 24.9°. Weight-average molecular weights ($M_{wa(c)}$) from seven experiments at different initial concentrations are shown: (◑) 0.1, (◒) 0.2, (◐) 0.3, (●) 0.4, (◓) 0.6, (◕) 0.8, and (◔) 1.0 M. The vertical bars show estimated maximum error at each concentration (see text). The crosses (X) are $M_{wa(c)}$ values calculated from the deduced equilibrium constant (k) and virial coefficient (B). The broken line shows $M_{na(c)}$, as calculated from the smoothed curve (solid line) through the $M_{wa(c)}$ data. The open circles (O) are $M_{na(c)}$ values calculated from k and B. The arrows show theoretical molecular weights for monomer and dimer, respectively.

$$M_{na(c)} = \frac{c}{\int_0^c \frac{dc}{M_{wa(c)}}} \tag{6}$$

The smoothed curve passed through the $M_{wa(c)}$ points in Figure 1 has been used to carry out the integration needed for eq 6. The apparent number-average weight obtained (broken line in Figure 1) does not exhibit as much curvature as $M_{wa(c)}$. This is to be expected, since

$$\frac{1}{M_{na(c)}} = \frac{1}{M_{n(c)}} + \frac{B}{2}c \tag{7}$$

In fact, the $M_{na(c)}$ curve shows, as it should, curvature comparable to that of the osmotic coefficient data of Ts'o *et al.* (1963). In the following section, we shall use the weight- and number-average molecular weights to test possible mechanisms of association for purine.

Discussion

Analysis of the Data at 24.9°. Given the curves of $M_{wa(c)}$ and $M_{na(c)}$ *vs.* c, it remains to seek a kind of equilibrium reaction which will predict these data. As has been pointed out above, a simple monomer–dimer reaction cannot possibly account for the results. Likewise, attempts to fit the data to a simple monomer–n-mer scheme ($n > 2$) were wholly unsuccessful. In particular, use of eq 19 of Adams (1965) led to unreasonable (negative) values of the virial coefficient, which varied enormously with concentration.

Since polymerization beyond the dimer stage is possible, we may hypothesize, with Ts'o *et al.*, that this is a stacking process in which an essentially constant increment of free energy accompanies the addition of each monomer. That is, we assume the set of reactions

$$\begin{aligned} 2P_1 &\rightleftharpoons P_2 & K_1 &= C_2/C_1^2 \\ P_2 + P_1 &\rightleftharpoons P_3 & K_2 &= C_3/C_2C_1 \\ &\vdots & &\vdots \\ P_i + P_1 &\rightleftharpoons P_{i+1} & K_i &= C_{i+1}/C_iC_1 \\ &\vdots & &\vdots \end{aligned} \tag{8}$$

with the special assumption that $K_1 = K_2 = \ldots = K_i = \ldots = K$. Since we will use weight concentrations

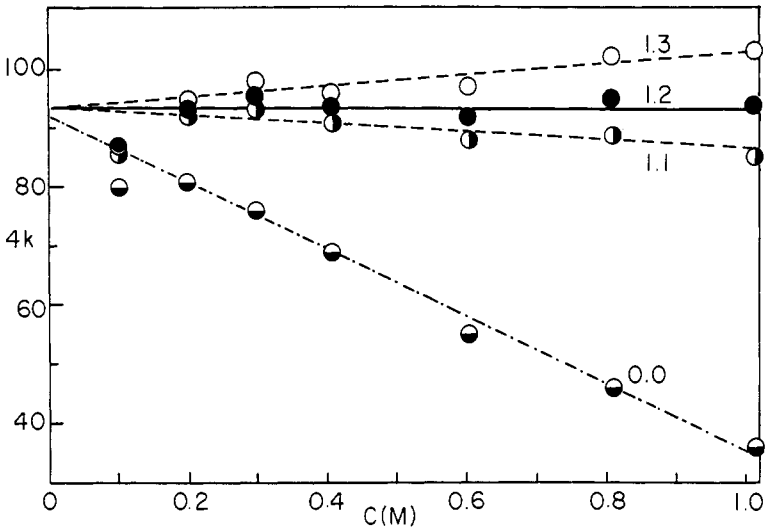

FIGURE 2: Determination of BM_1. Values of the right side of eq 12 are graphed *vs.* concentration, for the following assumed values of BM_1 (◒) 0.0, (◑) 1.1, (●) 1.2, and (○) 1.3.

in defining the average molecular weights, we convert to an equilibrium constant on the grams-per-milliliter scale; $k = 1000K/M_1$. If we define a quantity $R = M_{w(c)}/M_1$, it is easy to show that

$$R^2 - 1 = 4kc \tag{9}$$

if the above reaction scheme holds. Recalling that we must work with apparent weight-average molecular weights, we define $R_a = M_{wa(c)}/M_1$, and note from eq 1 that

$$R = \frac{R_a}{1 - BM_1R_ac} \tag{10}$$

which with eq 9 yields

$$\frac{R_a^2}{(1 - BM_1R_ac)^2} - 1 = 4kc \tag{11}$$

or

$$4k = \left\{\frac{R_a^2}{(1 - BM_1R_ac)^2} - 1\right\}\Big/c \tag{12}$$

Thus, if the proper value of B is chosen, the quantity on the right of eq 12 should be a constant, independent of concentration over the entire concentration range. The procedure for analysis is then straightforward; one may calculate the right side of eq 12 from the data, assuming various values of B, and seek the value which yields a constant k. Such calculations are shown in Figure 2, for data at 24.9°. From each experiment one point (the one marked with a vertical bar in Figure 1) has been used. Since the precision within an experiment is better than that between experiments, inclusion of more points would not really yield more information. It is evident that the value of $BM_1 = 1.2$ does quite nicely, and that the results are fairly sensitive to the value of BM_1. The point at lowest concentration, which is clearly the least accurate, has been neglected in arriving at the best value. The assumption employed by previous workers (that $B = 0$) leads to the highly inconstant values of $4k$ shown by the lowest curve. Of course, at infinite dilution, where ideality is approached, any nearly correct guess for B leads to values tending toward the true value of $4k$. But data at low concentrations are certain to be the least reliable.

This value of BM_1 is, in fact, quite reasonable. On the more common g/100-ml concentration scale, B will have the small value 1.0×10^{-4}. The nonideality effects are appreciable only because the solutions are so very concentrated.

Given values of B and k, the $M_{na(c)}$ data can be used as a semiindependent test. It should be recalled that the calculation of $M_{na(c)}$ from $M_{wa(c)}$ involved no assumption about the equilibrium constant, or even the type of equilibrium involved. Thus, if the fit of the $M_{wa(c)}$ data to the proposed model is fortuitous, values of $M_{na(c)}$ calculated from k and B should not agree with the observed results. It is easy to carry out this calculation, for the model chosen. One may readily demonstrate that

$$\frac{1}{M_{na(c)}} = \frac{1 - kc_1}{M_1} + \frac{B}{2}c \tag{13}$$

and

$$c = \frac{c_1}{(1 - kc_1)^2} \tag{14}$$

where c_1 is the concentration of monomer. The easiest

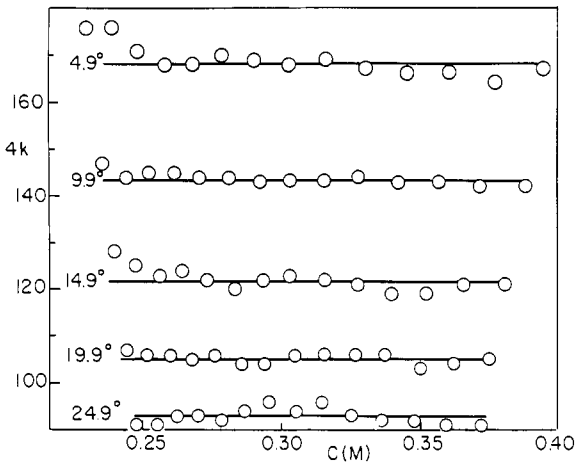

FIGURE 3: Apparent equilibrium constants (according to eq 12, with $BM_1 = 1.2$) graphed as a function of concentration at different temperatures (numbers given). All data from experiments at initial concentration of 0.3 M.

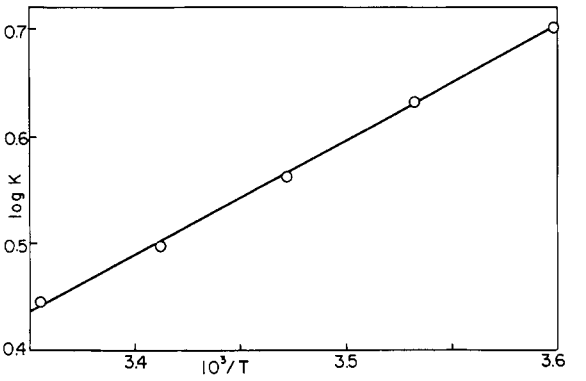

FIGURE 4: A graph of the logarithms of the equilibrium constants deduced from Figure 3 *vs.* $1/T$.

way to use these equations is to pick an arbitrary series of c_1 values, calculate c (knowing k), and then calculate $M_{na(c)}$ from eq 13. The results are shown as the open circles in Figure 1; within experimental error, they are in agreement with the observed values. A similar (though not nearly so critical) check is provided by back calculating $M_{wa(c)}$ values from k and BM_1. These results are shown as crosses in Figure 1. It should be noted that because of the weaker dependence of $M_{na(c)}$ upon concentration (compare eq 1 and 7) the effect of nonideality would not be very evident if osmotic pressure measurements alone were used.

The value of the equilibrium constant obtained at 24.9° is somewhat higher than that found by Ts'o *et al.* (1963), and essentially the same as found by Gill *et al.* (1967) by an independent procedure. Converted to a molar scale, the value of $k = 23.3 \pm 0.5$ $(g/ml)^{-1}$ yields $K = 2.80 \pm 0.06$ $(moles/l.)^{-1}$, whereas Ts'o *et al.* found an average value of about 2.1 m^{-1}. The difference between the molar and molal scales is not significant here; apparent equilibrium constants on the molal scale would show a slightly greater concentration dependence. An alternative procedure to obtain K, by extrapolation of apparent K values to $C = 0$ without explicit consideration of the nonideality effects, leads to essentially the same result, as shown in Figure 2. While the difference is not qualitatively important, one should point out that our data indicate an even *stronger* tendency to self-association of purine than had been hitherto presumed.

The question must be raised as to whether such analysis uniquely fixes the details of the polymerization. We believe that the answer must be *no*. While such alternatives as a monomer–*n*-mer equilibrium are excluded, it is quite likely that an arbitrary set of equilibrium constants (plus a nonideality coefficient) for, say, a monomer–dimer–trimer–tetramer equilibrium could be found that would fit the data quite well. Such a system would involve four arbitrary parameters, allowing considerable latitude in curve fitting. However, it should not be presumed that the necessity for a finite B value could be avoided by some other reaction scheme. The fact that $M_{wa(c)}$ values indisputably level off and then tend downward at high concentration requires the inclusion of such a term. Furthermore, the precision with which both the $M_{wa(c)}$ and $M_{na(c)}$ data are predicted from k and B over a wide concentration range argues that no *very* different set of equilibrium constants would serve. Since there is no *a priori* reason to assume a variation of K_i values with degree of polymerization, and the model employed here is capable of fitting both the $M_{wa(c)}$ data and the $M_{na(c)}$ data within experimental error, no compelling reason to seek a more complex model is evident.

Data at Other Temperatures. Thermodynamics of the Reaction. In order to determine the enthalpy and entropy changes associated with the reaction, a series of sedimentation equilibrium experiments were performed at a given initial concentration (0.3 M) and different temperatures. Actually, to minimize errors, this series of experiments employed a single sample; the temperature of the rotor was simply changed and time was allowed for reequilibration at each new temperature to be established. Values of the equilibrium constant obtained over the range of concentrations in the ultracentrifuge cell are shown at the several temperatures in Figure 3. It should be noted that the *same* value of BM_1 is employed in each of these experiments as had already been deduced from the more extensive data at 24.9°. Since B should not be expected to vary markedly with T, the fact that this one value suffices for data over a range of concentration *and* temperature argues for the essential validity of the analysis. There seems to be a very slight downward trend of k with increasing c at low T, even if values near the meniscus (generally found to be the least precise) are discounted. This could be removed by postulating a slight increase of

BM_1 (to about 1.3 at 4.9°) with decreasing temperature. However, the effect is so small that it hardly seems justified to make this correction.

In Figure 4 is shown a van't Hoff graph of the logarithm of the equilibrium constant *vs.* $1/T$. From the slope of the straight line an average $\Delta H°$ can be calculated; using the molar equilibrium constant to obtain $\Delta G°$, we have, at 25° $\Delta G° = -0.61$ kcal/mole, $\Delta H° = -4.9$ kcal/mole, and $\Delta S° = -14$ cal/°mole. According to the model chosen, we interpret these values as corresponding to the addition of 1 mole of purine to previously existing stacks of any size.

The average value for $\Delta H°$ is somewhat larger than the values obtained by Gill *et al.* (1967) from heat of dilution measurements at 24.9°. This apparent difference may or may not be real. There is no *a priori* reason why $\Delta H°$ should be independent of temperature, and there is some slight indication of curvature in Figure 4. While this is close to the experimental uncertainty, if taken literally it yields a value of $\Delta H° \cong -4$ kcal/mole at 25°, in good agreement with the results of Gill *et al.* (1967).

Both the enthalpy and entropy changes are considerably smaller than the values obtained by Van Holde *et al.* (1965) for the stacking of bases in the nucleotide dimer, ApA. A more meaningful comparison, however, should come from results of studies of adenosine 5′-phosphate now in progress. This compound also shows a strong tendency to associate in aqueous solution (Rossetti and Van Holde, 1967).

References

Adams, E. T., Jr. (1965), *Biochemistry 4*, 1646.

Adams, E. T., Jr., and Williams, J. W. (1964), *J. Am. Chem. Soc. 86*, 3454.

Brahms, J., Michelson, A. M., and Van Holde, K. E. (1966), *J. Mol. Biol. 15*, 467.

Chan, S. I., Schweizer, M. P., Ts'o, P. O. P., and Helmkamp, G. K. (1964), *J. Am. Chem. Soc. 86*, 4182.

Gill, S. J., Downing, M., and Sheats, G. F. (1967), *Biochemistry 6*, 272.

Gropper, L. (1964), *Anal. Biochem. 7*, 401.

Holcomb, D. N., and Tinoco, I., Jr. (1965), *Biopolymers 3*, 121.

Leng, M., and Felsenfeld, G. (1966), *J. Mol. Biol. 15*, 455.

Poland, D., Vournakis, J. N., and Scheraga, H. A. (1966), *Biopolymers 4*, 223.

Rossetti, G., and Van Holde, K. E. (1967), *Biochem. Biophys. Res. Commun. 26*, 717.

Sturtevant, J. M., Rice, S. A., and Geiduschek, P. (1958), *Discussion Faraday Soc. 25*, 1317.

Ts'o, P. O. P., Melvin, I. S., and Olson, A. C. (1963), *J. Am. Chem. Soc. 85*, 1289.

Van Holde, K. E., and Baldwin, R. L. (1958), *J. Phys. Chem. 62*, 734.

Van Holde, K. E., Brahms, J., and Michelson, A. M. (1965), *J. Mol. Biol. 12*, 726.

Witz, J., and Luzzati, V. (1965), *J. Mol. Biol. 11*, 620.