BIOPHYSICAL CHEMISTRY 1 (1973) 1–10. NORTH-HOLLAND PUBLISHING COMPANY

THEORY OF ZONE SEDIMENTATION FOR NON-COOPERATIVE LIGAND-MEDIATED INTERACTIONS*

John R. CANN

Department of Biophysics and Genetics, University of Colorado Medical Center, Denver, Colorado 80220, USA

Received 15 May 1973

Theoretical zone sedimentation patterns have been computed for a variety of non-cooperative ligand-mediated, macromolecular interactions. The generalization is drawn that rapidly equilibrating ligand-mediated interactions in general have the potentiality for showing bimodal zones irrespective of reaction mechanism subject to the provisos that overall ligand-binding is sufficiently strong to generate large gradients of unbound ligand along the sedimentation column but that binding to the reactant itself is not so strong that in effect the system behaves like the analogous nonmediated interaction. A physical explanation of the resolution of bimodal reaction zones is given. Qualitatively similar results have been obtained for molecular-sieve chromatography.

1. Introduction

Previously (chapters 4 and 5 in ref. [1], [2, 3]), we have described the theory of sedimentation of reversible ligand-mediated interactions of the type (reaction I)

$$mM + nX \rightleftharpoons M_m X_n \,, \qquad \text{(I)}$$

in which a macromolecule, M associates into an m-mer, with the mediation of a small ligand molecule or ion, X, of which a fixed number, n, are bound into the complex. The theoretical sedimentation patterns reveal that such interactions can given rise to a well-resolved bimodal reaction boundary or zone despite instantaneous establishment of equilibrium. Resolution of the two peaks depends upon the production and maintenance of concentration gradients of unbound ligand along the centrifuge cell by re-equilibration during differential transport of macromonomer and polymer; and the peaks correspond to different equilibrium mixtures of monomer and polymer. This is so even for ligand-mediated dimerization (i.e., $m = 2$ and $n = 1$ or 6) which is in contradistinction to simple dimerization (reaction II)

$$2M \rightleftharpoons M_2 \,, \qquad \text{(II)}$$

which always gives sedimentation patterns showing a single peak when re-equilibration is rapid [4–7]. Of course, at sufficiently high ligand concentration (or low centrifugal field) ligand-mediated dimerization can also give patterns that show a single peak. In the limit where for any reason the concentration of unbound ligand along the centrifuge cell is not significantly perturbed by the reaction during differential sedimentation of the macromolecular species, the sedimentation behavior of ligand-mediated association effectively approaches that of the simple association reactions considered by Gilbert [4, 5] in the case of analytical sedimentation and by Bethune and Kegeles [8] for zone sedimentation.

Subsequently, the prediction of bimodal reaction boundaries found experimental verification in studies on the reversible dimerization of New England lobster hemocyanin mediated by the binding of 4–6 Ca^{2+} and 2–4 H^+ which occurs when the pH is lowered from above 9.6 to below 9.2 [9, 10] and on the dimerization of tubulin through the mediation of 1 vinblastine molecule [11]. One notes, however, that

* Supported in part by Research Grant 5 RO1 HL13909-21 from the National Heart and Lung Institute, National Institutes of Health, United States Public Health Service. This publication is No. 544 from the Department of Biophysics and Genetics, University of Colorado Medical Center, Denver, Colorado 80220.

except for the case in which $n = 1$, reaction I is cooperative with respect to ligand; and the question remains whether non-cooperative interactions in general can give bimodal sedimentation patterns. Thus, for example, the author was asked recently: "Will the zone sedimentation pattern of a macromolecule undergoing the set of sequential and simultaneous interactions (reaction set III)

$$\begin{aligned} &M + X \rightleftharpoons MX \quad (K_1), \\ &MX + M \rightleftharpoons M_2X \quad (K_2), \\ &2MX \rightleftharpoons M_2X_2 \quad (K_3), \end{aligned} \tag{III}$$

show two peaks?" In order to answer such questions a theoretical investigation of the sedimentation behavior of some representative, non-cooperative ligand-mediated interactions has been made. In addition to reaction set III, the following interactions have been examined: the sequential monomer-tetramer reaction (reaction set IV)

$$\begin{aligned} &M + X \rightleftharpoons MX \quad (K_1), \\ &4MX \rightleftharpoons M_4X_4 \quad (K_2), \end{aligned} \tag{IV}$$

sequential and progressive tetramerization

$$\begin{aligned} &M + X \rightleftharpoons MX \quad (K_1), \\ &2MX \rightleftharpoons M_2X_2 \quad (K_2), \\ &M_2X_2 + MX \rightleftharpoons M_3X_3 \quad (K_3), \\ &M_3X_3 + MX \rightleftharpoons M_4X_4 \quad (K_4), \\ &K_2 = K_3 = K_4 \equiv K \end{aligned} \tag{V}$$

dissociation of a dimer driven by binding of ligand to the monomer

$$\begin{aligned} &M_2 \rightleftharpoons 2M \quad (K_1), \\ &M + X \rightleftharpoons MX \quad (K_2), \end{aligned} \tag{VI}$$

and sequential isomerization

$$\begin{aligned} &M + X \rightleftharpoons MX \quad (K_1), \\ &MX \rightleftharpoons DX \quad (K_2), \end{aligned} \tag{VII}$$

in which the conformational isomer, DX, has different diffusion and sedimentation coefficients than does isomer M, either uncomplexed or complexed with X. The results of the calculations predict that under appropriate conditions each of these rapidly equilibrating interactions can give rise to zone patterns showing two peaks.

2. Theory

The theory of zonal transport to be described is for rectilinear coordinates and constant driven velocities of the various species and is thus valid for zone sedimentation through a preformed linear density gradient in the preparative ultracentrifuge. This is so because macromolecules sediment with constant velocities under the conditions described by Martin and Ames [12].

Theoretical sedimentation patterns have been computed by numerical solution of the appropriate set of transport equations for constituent macromolecule and constituent ligand. The computations are for the limiting case of rates of reaction so fast that, in effect, there is local equilibrium among the interacting species. The salient features of the calculation are as follows: Suppose the sedimentation column to be divided into a number of discrete segments. Given the initial equilibrium concentrations of the several species in each segment, we calculate the change in the distribution of material during a short interval of time, Δt, due to transport. It is assumed that there is no re-equilibration by interaction during Δt, each species migrating independently from its distribution at the beginning of the interval. After the concentrations have been advanced, equilibrium is recalculated. That is, for known constituent concentrations of macromolecule and ligand computed from the concentrations of the several species as changed by transport, new equilibrium concentrations are calculated by applying the law of mass action. We then compute the change in this new distribution of material due to transport over the next Δt; recalculate the equilibrium; and so on, constructing the entire evolution of the distribution of material in the sedimentation column from the initial condition.

To illustrate the numerical procedures let us consider the dimerization reaction set III. In this case there

are five sedimenting species including unbound ligand; and the simplifying assumption is made that both monomeric species, M and MX, have the same diffusion and sedimentation coefficients and that both dimeric species, M_2X and M_2X_2, likewise have the same transport coefficients. The molar concentrations of the macromolecular species are designated as $C_{i,k}$ where $i = 1$ or 2 designates monomer or dimer and $k = 0$, 1 or 2 designates the number of bound ligand molecules (e.g., $C_{1,0}$ is the concentration of M; $C_{2,1}$ the concentration of M_2X; etc.). Their transport coefficients bear the subscript i. The molar concentration, diffusion coefficient and sedimentation coefficient of unbound ligand are designated as C_3, D_3 and s_3 respectively.

The changes in concentration of the four macromolecular species with time, t, and position in the sedimentation column, x, during independent transport are described by the conservation equations

$$\frac{\partial C_{i,k}}{\partial t} = D_i \frac{\partial^2 C_{i,k}}{\partial x^2} - v_i \frac{\partial C_{i,k}}{\partial x} \qquad (i = 1, 2; k = 0, 1, 2), \qquad (1a)$$

where D_i is the diffusion coefficient; and v_i is the driven velocity. $v_i = s_i \omega^2 \bar{x}$ in which s_i is the sedimentation coefficient; $\omega^2 \bar{x}$, the constant field strength; ω, angular velocity and $\bar{x}$, an average position. In our calculations, $\omega^2 \bar{x} = 2.6056 \times 10^8$ cm sec^{-2} which corresponds to 60,000 rpm and $\bar{x} = 6.6$ cm. The conservation equation for independent transport of unbound ligand is

$$\frac{\partial C_3}{\partial t} = D_3 \frac{\partial^2 C_3}{\partial x^2} - v_3 \frac{\partial C_3}{\partial x}. \qquad (1b)$$

These equations have been solved in a frame of reference moving with the average macromolecular velocity $\bar{v} = \frac{1}{2}(v_1 + v_2)$, which introduces the new position variable $x' = x - \bar{v}t$. Eqs. (1a) and (1b) are transformed into the moving coordinate system simply by making the replacements $x' \to x$, $v_i - \bar{v} \to v_i$ and $v_3 - \bar{v} \to v_3$.

We now introduce discrete time and position variables

$$t_n = n\Delta t \qquad (n = 0, 1, 2, \ldots),$$

$$x'_l = l\Delta x' \qquad (l = 0, 1, 2, \ldots L)$$

and replace the continuous variables $C_{i,k}(t, x')$ and $C_3(t, x')$ by the discrete variables $C_{i,k}(n\Delta t, l\Delta x') \equiv C_{i,k}(t_n, x'_l)$ and $C_3(n\Delta t, l\Delta x') \equiv C_3(t_n, x'_l)$. Proceeding as in a previous investigation [13] the transformed eqs. (1a) are approximated by the finite difference equations

$$C_{i,k}(t_{n+1}, x'_l) = C_{i,k}(t_n, x'_l) + \frac{D_i \Delta t}{(\Delta x')^2} \{C_{i,k}(t_n, x'_{l+1}) - 2C_{i,k}(t_n, x'_l) + C_{i,k}(t_n, x'_{l-1})\} - \frac{(v_i - \bar{v})\Delta t}{\Delta x'} [\delta C_{i,k}]_{\text{for}} \qquad (i = 1, 2; k = 0, 1, 2), \qquad (2a)$$

in which

$$[\delta C_{i,k}]_{\text{for}} = C_{i,k}(t_n, x'_l) - C_{i,k}(t_n, x'_{l-1}) \quad \text{if } (v_i - \bar{v}) > 0,$$

$$= C_{i,k}(t_n, x'_{l+1}) - C_{i,k}(t_n, x'_l) \quad \text{if } (v_i - \bar{v}) < 0.$$

Similarly, for eq. (1b). Given values of $C_{i,k}$ and C_3 at any time t, we can calculate their values at $t + \Delta t$ as changed by transport using these equations.

At each time, t_{n+1}, new values of the constituent concentrations of macromolecular, $\bar{C}_M$, and ligand, $\bar{C}_X$, are computed from $C_{i,k}$ and C_3 as changed by transport

$$\bar{C}_M = C_{1,0} + C_{1,1} + 2(C_{2,1} + C_{2,2}),$$

$$\bar{C}_X = C_3 + C_{1,1} + C_{2,1} + 2C_{2,2}$$

and then re-equilibration is imposed by simultaneous solution of the mass action expressions

$$\bar{C}_M = C_{1,0} + K_1 C_{1,0} C_3 + 2K_1 K_2 C_{1,0}^2 C_3 + 2K_1^2 K_3 C_{1,0}^2 C_3^2, \qquad (3a)$$

$$\bar{C}_X = C_3 + K_1 C_{1,0} C_3 + K_1 K_2 C_{1,0}^2 C_3 + 2K_1^2 K_3 C_{1,0}^2 C_3^2 \qquad (3b)$$

for the equilibrium values of $C_{1,0}$ and C_3 which, in turn, are used to calculate the equilibrium values of

$$C_{1,1} = K_1 C_3 C_{1,0},$$

$$C_{2,1} = K_1 K_2 C_3 C_{1,0}^2,$$

$$C_{2,2} = \tfrac{1}{2}(\bar{C}_M - C_{1,0} - C_{1,1}) - C_{2,1}.$$

To solve eqs. (3a) and (3b) we apply the Newton–Raphson method [14], using the old equilibrium values of $C_{1,0}$ and C_3 as starting approximations except when the old value of $C_{1,0}$ is zero in which case 0.5 $\bar{C}_M$ is used. The new values of the equilibrium concentrations serve as the starting distribution of material for the next time cycle of transport followed by re-equilibration; and so on. Given initial conditions and boundary values, this recursive calculation allows one to follow the evolution of the sedimentation pattern.

The initial conditions are

$$C_{1,0}(0, x_l') = C_{10}/(1 + R_{10}),$$

$$C_{1,1}(0, x_l') = C_{10} - C_{1,0}(0, x_l'),$$

$$C_{2,1}(0, x_l') = C_{20}/(1 + R_{20}),$$

$$C_{2,2}(0, x_l') = C_{20} - C_{2,1}(0, x_l'),$$

$$C_3(0, x_l') = C_{30} (l = L/2 - 5, \ldots L/2 + 5),$$

$$C_{1,0}(0, x_l') = C_{1,1}(0, x_l') = C_{2,1}(0, x_l') = C_{2,2}(0, x_l') = 0,$$

$$C_3(0, x_l') = C_{30} (l \neq L/2 - 5, \ldots L/2 + 5).$$

The quantities C_{10}, C_{20} and C_{30} are the initial equilibrium concentrations of total monomer, total dimer and unbound ligand, respectively; and R_{10} and R_{20} are the initial equilibrium ratios

$$R_{10} = C_{1,1}(0, x_l)/C_{1,0}(0, x_l), \tag{4a}$$

$$R_{20} = C_{2,2}(0, x_l)/C_{2,1}(0, x_l). \tag{4b}$$

C_{10}, C_{20}, C_{30}, R_{10} and R_{20} constitute computer input data, and the values of the equilibrium constants are calculated therefrom:

$$K_1 = R_{10}/C_{30}, \tag{5a}$$

$$K_2 = \frac{C_{20}}{C_{10}^2} \cdot \frac{(1 + R_{10})^2}{R_{10}(1 + R_{20})}, \tag{5b}$$

$$K_3 = \frac{C_{20}}{C_{10}^2} \cdot \frac{R_{20}(1 + R_{10})^2}{R_{10}^2(1 + R_{20})}. \tag{5c}$$

The boundary values are

$$C_{1,0}(t_n, 0) = C_{1,1}(t_n, 0) = C_{2,1}(t_n, 0) = C_{2,2}(t_n, 0) = 0,$$

$$C_{1,0}(t_n, x_L') = C_{1,1}(t_n, x_L')$$

$$= C_{2,1}(t_n, x_L') = C_{2,2}(t_n, x_L') = 0,$$

$$C_3(t_n, 0) = C_3(t_n, x_L') = C_{30}.$$

As stated mathematically by the initial conditions given above, the calculations are for unbound ligand initially distributed uniformly throughout the centrifuge column.

Computations were made on the University of Colorado's CDC 6400 electronic computer. The values of $\Delta t = 10$ sec and $\Delta x' = 0.02$ cm used in these calculations satisfy the stability criterion employed previously [13]. The computed sedimentation patterns are displayed as plots of constituent concentration of macromolecule, $(C_1 + 2C_2) \equiv \bar{C}_M$, against position, x', where C_1 and C_2 are the concentrations of total monomer and total dimer, respectively. Two vertical arrows indicate where the peaks in the patterns would have been located had sedimentation been carried out on a mixture of noninteracting macromolecules having the same transport parameters as monomer and dimer.

The foregoing mathematical formulation has been applied in principle to the other interactions which we have examined. In each case it is assumed that M and MX have the same transport parameters. The notation used is as follows:

Reaction IV. C_1, C_4 and C_5 desigante molar concentrations of total monomer both uncomplexed and complexed with ligand; tetramer; and unbound ligand. Subscripts 10, 40 and 50 designate initial concentrations.

Diffusion and sedimentation coefficients bear the same subscripts as the concentrations of the corresponding species. R_{10} is defined as in eq. (4a).

Reaction V. C_1, C_2, C_3, C_4 and C_5 designate molar concentrations of total monomer; dimer; trimer; tetramer; and unbound ligand. Initial concentrations and transport parameters are subscripted in a manner analogous to that indicated above; and R_{10} is defined as in eq. (4a).

Reaction VI. C_1, C_2 and C_3 designate molar concentrations of total monomer; dimer; and unbound ligand. R_{10} is defined as in eq. (4a).

Reaction VII. C_1, C_2 and C_3 designate molar concentrations of total isomer M both uncomplexed and complexed with X; isomer DX; and unbound ligand. Initial concentrations and transport parameters are subscripted in an analogous manner as above; and R_{10} is defined as in eq. (4a).

In the case of reactions IV and V we chose to increase $\Delta x'$ to 0.03 cm in order to insure stability while maintaining $\Delta t = 10$ sec to minimize the amount of calculation, rather than decreasing Δt at $\Delta x' = 0.02$ cm. This compromise was made at the expense of increased truncation error which expresses itself as excess spreading ("truncation diffusion") of the peaks in the sedimentation pattern [15] but has no effect on their positioning. The error is not serious and has no bearing on the conclusions. Thus, the half-widths of the peaks shown by a noninteracting mixture of monomer and tetramer after 1.5×10^4 sec of sedimentation were only about 15% greater than expected from translational diffusion of the macromolecules.

The theory of molecular-sieve chromatography is for small zones [16] and was formulated for reaction set IV in the same way as zone sedimentation except that in most cases the ligand was taken to be initially present only in the starting zone. Calculations were made for the total column frame of reference [17] with system parameters as given by Zimmerman et al. [18] for monomer (17,000 daltons) and tetramer on G-200R. The "truncation diffusion", $|\upsilon_i - \bar{\upsilon}|\Delta x'/2$, was subtracted from the axial dispersions of monomer and tetramer in order to eliminate this source of error (correction $\leqslant 10\%$). The concentration profiles are displayed as plots of total constituent concentration of macromolecule [$C'_T = C'_1 + 4C'_4$ where, in the notation of Zimmerman and Ackers [17], the prime designates the total column frame of reference] vs x'.

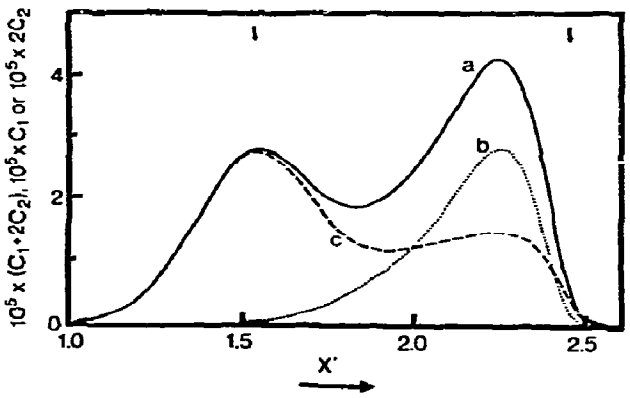


Fig. 1. Theoretical zone sedimentation pattern computed for the ligand-mediated dimerization reaction set III, 50% dimerization with $R_{10} = 0.10$ and $R_{20} = 1.00$; a – plot of constituent molar concentration of macromolecule ($C_1 + 2C_2$) vs. position, x'; b – twice the molar concentration of dimer ($2C_2$) vs x'; c – molar concentration of monomer (C_1) vs x'. The following values of the macromolecule concentrations and parameters were chosen to approximate the sedimentation of a protein of molecular weight 60,000: $C_{10} = 7 \times 10^{-5}$ M, $C_{20} = 3.5 \times 10^{-5}$ M, $s_1 = 4S$, $s_2 = 6.35S$, $D_1 = 6 \times 10^{-7}$ cm^2 sec^{-1}, $D_2 = 4.76 \times 10^{-7}$ cm^2 sec^{-1}. For the small ligand molecule: $C_{30} = 5 \times 10^{-7}$ M, $s_3 = 0.1S$, $D_3 = 10^{-5}$ cm^2 sec^{-1}. Time of sedimentation, 1.5×10^4 sec.

3. Results and discussion

Let us first consider the dimerization reaction set III according to which the binding of a single ligand molecule to the macromolecule mediates formation of two kinds of dimer—one containing a single ligand molecule and the other containing two ligand molecules. As for mechanism, it could be that the monomer conformation that favors formation of dimer is stabilized by the binding of a ligand molecule as might obtain, for example, in an allosteric interaction. As illustrated by curve a in fig. 1 for 50% dimerization such an interaction can give a well-resolved bimodal reaction zone even though re-equilibration is rapid. The individual distributions of dimer and monomer are shown by curves b and c, respectively. In contrast to dimer, the monomer is bimodally distributed – a fact which would have to be borne in mind if the pattern of biological activity were to be measured through the zone. The slower peak in the zone is composed largely of monomer, while the faster one is a mixture of dimer with an appreciable amount of monomer. Consequently, the slower peak migrates only slightly faster than monomer, while the faster

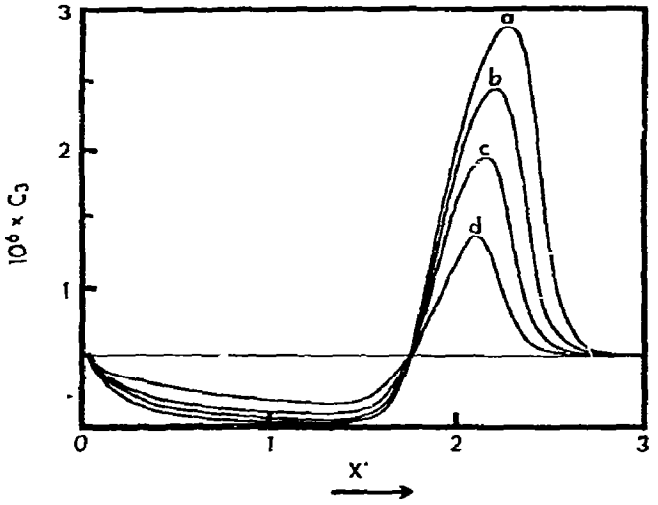


Fig. 2. Plots of molar concentration of unbound ligand (C_3) vs position, x': a, b, c and d pair with the correspondingly designated macromolecular patterns displayed in fig. 3C; pattern a also pairs with the macromolecule pattern presented in fig. 1. Horizontal line represents the initially uniform distribution of unbound ligand throughout the sedimentation column.

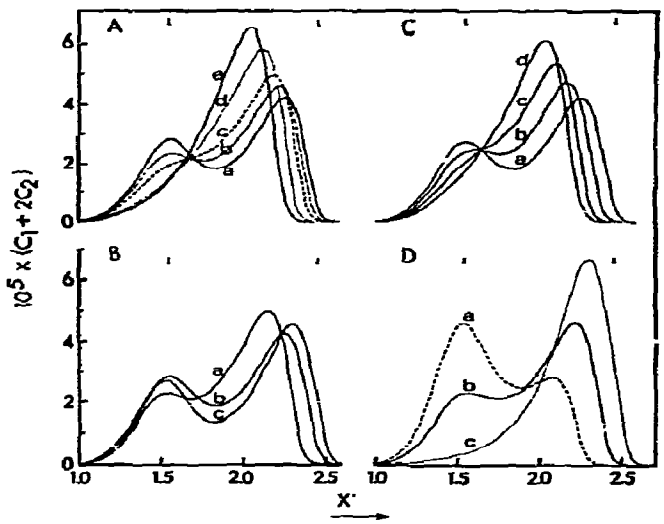


Fig. 3. Factors governing the shape of the theoretical zone sedimentation pattern for the dimerization reaction set III. A—Dependence of shape on C_{30} for 50% dimerization with $R_{10} = 0.10$ and $R_{20} = 1.00$: pattern a – $C_{30} = 5 \times 10^{-7}$ M; b – 1×10^{-6}; c – 1.7×10^{-6}; d – 5×10^{-6}; and e – 2×10^{-5}. B – Dependence on R_{20} for 50% dimerization with $R_{10} = 0.10$ and $C_{30} = 5 \times 10^{-7}$ M: pattern a – $R_{20} = 0.10$; b – 1.00 and c – 10.00. C – Dependence on R_{10} for 50% dimerization with $R_{20} = 1.00$ and $C_{30} = 5 \times 10^{-7}$ M: pattern a – $R_{10} = 0.10$; b – 0.25; c – 0.50; and d – 1.00. D – Dependence on percent of dimerization at fixed $K_1 = 1 \times 10^5\ M^{-1}$, $K_2 = 4.32 \times 10^4\ M^{-1}$ and $K_3 = 4.32 \times 10^5 M^{-1}$: pattern a – 25% dimerization, total ligand concentration 2.49×10^{-5} M; b – 50%, 5.99×10^{-5}; and c – 75%, 10.80×10^{-5}. $C_{10} + 2C_{20} = 14 \times 10^{-5}$ M; other parameters as in fig. 1.

one migrates rather slower than dimer. The macromolecule contained in the faster peak is about 60% dimerized even though the overall extent of dimerization throughout the spreading zone has decreased to about 35% due to mass action.

The distribution of unbound ligand is given by curve a in fig. 2. We see that as the zone migrates down the sedimentation column ligand is swept out of the upper part of the column and concentrated in the region corresponding to the position of the faster peak in the pattern of macromolecule. Comparison of the macromolecule and ligand distributions (figs. 1 and 2) with the aid of the mass action expression, $C_{1,1}/C_{1,0} = K_3 C_3$, shows that, whereas less than 3% of the monomer contained in the slower peak is complexed with ligand, more than 30% is complexed in the faster peak. It is immediately apparent that resolution of the bimodal reaction zone depends upon the production and maintenance of concentration gradients of unbound ligand along the centrifuge column by re-equilibration during differential transport of macromonomer and dimer; and that the peaks correspond to different equilibrium mixtures.

The mechanism of resolution is as follows: Because of the transport of ligand bound into dimer the back half of the sedimenting zone is almost depleted of ligand. Consequently, the remaining monomer in this region is deprived of the dimerization-mediating small molecule and thus lags behind to form a second peak. At the same time the fast peak is enriched in ligand to an extent which more than compensates for dilution of the macromolecule in the spreading zone. As a result, the macromolecule in this peak is actually more highly dimerized than in the starting zone*. The two peaks never completely resolve, however, because the association reactions are reversible. With only minor

* The redistribution of ligand during sedimentation of macromolecule is particularly striking. From the thermodynamic point of view the decrease in entropy accompanying these processes is at the expense of centrifugal work. Mechanistically, the sedimentation velocity of the ligand is increased by almost two orders of magnitude as a result of binding into the dimer.

modification of details this argument is expected to apply to any ligand-mediated interaction which leads to a change in sedimentation coefficient of the macromolecule provided that the overall binding of ligand is sufficiently strong (i.e., ratio of bound to free concentration sufficiently large) to generate large gradients of unbound ligand along the sedimentation column; but binding to the reactant itself must not be so strong that the sedimentation behavior of the system approaches that of the analogous nonmediated interaction. A survey of the factors governing the shape of reaction zones follows.

These factors are delineated in fig. 3 for the dimerization reaction set III. The families of patterns displayed in figs. 3A, 3B and 3C for 50% dimerization at fixed macromolecule concentration show how the shape of the zone depends upon the strength of ligand-binding as adjusted by systematic and independent variation of each of the parameters [C_{30}, R_{10} and R_{20}; see eqs. (4a) and (4b)] which determine the values of the three equilibrium constants [see eqs. (5a)–(5c)]. The family of patterns in fig. 3A was computed by varying C_{30} at fixed values of $R_{10} = 0.10$ and $R_{20} = 1.00$. At the lowest value of C_{30} the bimodal zone is well resolved. Increasing C_{30} (i.e., decreasing the strength of overall ligand-binding by decreasing K_1 at fixed K_2 and K_3) causes progressive loss of resolution until sufficiently high C_{30} the two peaks have coalesced to form a skewed unimodal zone whose apex migrates with a velocity only slightly greater than the weight average velocity of monomer and dimer. This is so because redistribution of unbound ligand along the sedimentation column becomes progressively weaker as the limit is approached in which the initial uniform distribution cannot be significantly perturbed. In this limit the sedimentation behavior approaches that of the simple dimerization reaction II.

The strength of ligand-binding can be increased at a fixed low value of C_{30} by increasing either R_{10} or R_{20} with differing effects on resolution because here the controlling factor is the change in proportions of ligand bound to the different macromolecular species rather than the increased strength of overall binding. Thus, increasing R_{20} (i.e., increasing the proportion of ligand bound into the dimer M_2X_2) enhances resolution (fig. 3B) because in the limit of very large R_{20} and small R_{10} the reaction set III collapses to the cooperative interaction

$$2M + 2X \rightleftharpoons M_2X_2 \ , \qquad \text{(VIII)}$$

whose equilibrium position is more sensitive to the concentration of unbound ligand than in the case of the noncooperative interaction. Consequently, the macromolecule in the faster peak is more highly dimerized and thus migrates more rapidly. In contrast, increasing R_{10} (i.e., increasing the proportion of ligand bound to monomer) causes progressive loss of resolution (fig. 3C) * because the redistribution of ligand along the sedimentation column becomes progressively weaker (fig. 2) as the fraction of monomer initially complexed with ligand increases.

The set of patterns presented in fig. 3D illustrates how the shape of the reaction zone depends upon the initial percent of dimerization as determined by total ligand concentration at constant macromolecule concentration and fixed equilibrium constants. The zones for 25% and 50% dimerization are bimodal, and the faster peak grows at the expense of the slower one as the percent of dimerization increases. However, the distribution of macromolecule between the two peaks does not faithfully reflect the initial composition, the amount in the faster peak being disproportionally large. The disparity increases progressively with increasing percent of dimerization until resolution is completely lost at 75% dimerization. This behavior is due to progressive increase in both the fraction of monomer initially complexed with ligand and the initial concentration of unbound ligand. As we have already seen, either factor when taken alone leads to eventual loss of resolution. Finally, the velocity of the faster peak in the bimodal zone increases with increasing percent dimerization; as dimerization is

* It is interesting that both families of patterns displayed in figs. 3B and 3C exhibit a stationary point reminiscent of an isosbestic point in absorption spectra except that the constituent macromolecule concentration at the stationary point decreases with time of sedimentation. On the other hand the point moves at a constant velocity different from the local velocity which indicates that it is a property of the equilibrium equations rather than the transport equations *per se*. The analogy to an isosbestic point holds only insofar as the amount of monomer contained in the zone is the same for each pattern in a given family. This is so within the error of the calculation; e.g., only a 4% drift over a 10-fold range of R_{10}-values for 7.5×10^3 sec of sedimentation and 8% for 1.5×10^4 sec which is consistent with the fact that the truncation error is propagated in time.

driven to completion, the velocity of the apex of the unimodal zone approaches the velocity of the dimer.

Although the foregoing results are for zone sedimentation, the same behavior can be expected of analytical sedimentaiton in which case bimodal reaction boundaries are predicted for essentially the same conditions.

The several factors delineated above for reaction set III also apply to other ligand-mediated interactions as illustrated, for example, in fig. 4 for the tetramerization reaction sets IV and V. Even progressive tetramerization (fig. 4B) can give bimodal zones (e.g., pattern a) in which the slow peak is composed largely of monomer and the faster one is an equilibrium mixture of monomer, dimer, trimer and tetramer. The leading edge of the zone is hypersharp because of the strong dependence of reversible higher-order polymerization reactions on macromolecule concentration; as higher-order polymers are differentially transported into the fresh solution immediately ahead of the advancing zone they dissociate into slower sedimenting species which lag behind. Increasing either the initial concentration of unbound ligand (C_{50}) or R_{10} leads to loss of resolution. In the first instance this is due to decreased strength of overall ligand-binding, while in the second case the limit is approached in which all of the monomer is complexed with ligand so that the system in effect behaves like the nonmediated progressive tetramerization reaction

$$2M \rightleftharpoons M_2 \quad (K_1),$$

$$M_2 + M \rightleftharpoons M_3 \quad (K_2),$$

$$M_3 + M \rightleftharpoons M_4 \quad (K_3),$$

$$K_1 = K_2 = K_3. \qquad \text{(IX)}$$

Ligand-mediated dissociation and isomerizaiton can also give well-resolved bimodal reaction zones. In the case of the dissociation reaction set VI resolution is enhanced by increasing the proportion of monomer complexed with ligand (i.e., increasing R_{10}) because in the limit of very large R_{10} the system collapses to the cooperative interaction

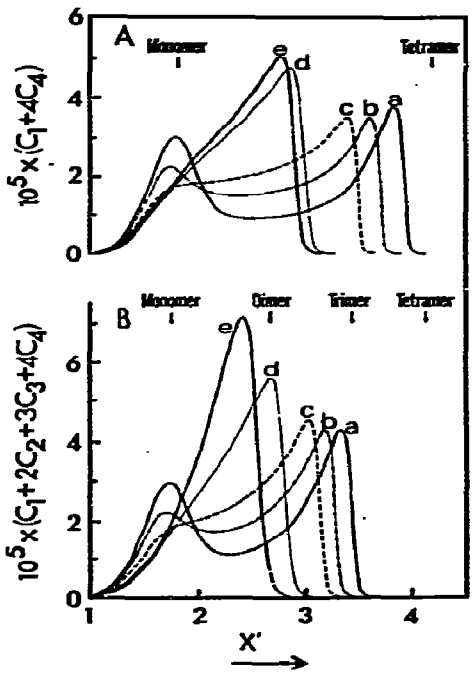


Fig. 4. Factors governing the shape of the theoretical zone sedimentation pattern for ligand-mediated tetramerization. A – Reaction set IV, 50% tetramerization with $R_{10} = 0.10$ for patterns a, b, c and d and $R_{10} = 10.00$ for pattern e: pattern a – $C_{50} = 5 \times 10^{-7}$ M; b – 2×10^{-6}; c – 5×10^{-6}; d – 7×10^{-5}; and e – 5×10^{-7}. B – Reaction set V, 50% association with $R_{10} = 0.10$ and $K = 1.72 \times 10^5$ M^{-1} for patterns a, b, c and d and $R_{10} = 10.00$ and $K = 5.10 \times 10^3$ M^{-1} for pattern e: pattern a – $C_{50} = 5 \times 10^{-7}$ M; b – 2×10^{-6}; c – 5×10^{-6}; d – 5×10^{-5}; and e – 5×10^{-5}. Initial constituent concentration of macromolecule equals 14×10^{-5} M, $s_1 = 4S$, $s_2 = 6.35S$, $s_3 = 8.32S$, $s_4 = 10.08S$, $D_1 = 6 \times 10^{-7}$ cm^2 sec^{-1}, $D_2 = 4.76 \times 10^{-7}$ cm^2 sec^{-1}, $D_3 = 4.16 \times 10^{-7}$ cm^2 sec^{-1}, $D_4 = 3.78 \times 10^{-7}$ cm^2 sec^{-1}. Other parameters as in fig. 1.

During the course of these calculations it was noted that a good approximation to zone sedimentation could be made by (1) transporting total macromonomer and total dimer without distinguishing between species containing different numbers of ligand molecules; (2) assuming that the ratios $R_1 = C_{1,1}/C_{1,0}$ and $R_2 = C_{2,2}/C_{2,1}$ do not change significantly during transport over a small time interval, Δt; (3) calculating at time, t_{n+1}, the concentrations $C_{i,k}$ from $C_{1,0} + C_{1,1}$ and $C_{2,1} + C_{2,2}$ as changed by transport using the old values of R_1 and R_2; and (4) calculating new values of R_1 and R_2 after re-equilibration. For ease of calculation this approximation has been used to compute analytical sedimentation patterns employing a modification of the computer program developed by Goad (Chapter 5 in ref. [1]) for reaction I. The major modification in the program is a new Subroutine Equil. The computed sedimentation behavior is strikingly similar to that shown in fig. 3 for zone sedimentation.

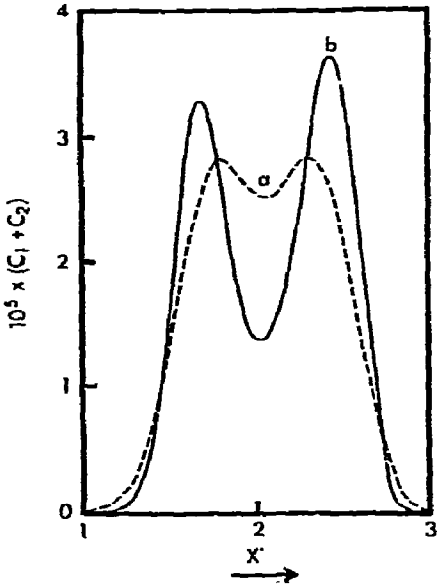


Fig. 5. Theoretical zone sedimentation patterns computed for the ligand-mediated isomerization reaction set VII for 50% isomerization. Pattern a – $s_1 = 5S$, $s_2 = 4S$, $D_1 = 7.5 \times 10^{-7}$ cm^2 sec^{-1}, $D_2 = 6 \times 10^{-7}$ cm^2 sec^{-1}, and 3×10^4 sec of sedimentation. Pattern b – $s_1 = 6.35S$, $s_2 = 4S$, $D_1 = 4.76 \times 10^{-7}$ cm^2 sec^{-1}, $D_2 = 3 \times 10^{-7}$ cm^2 sec^{-1}, and 1.5×10^4 sec. $C_{10} + C_{20} = 14 \times 10^{-5}$ M; $C_{30} = 5 \times 10^{-7}$ M and $R_{10} = 0.10$; other parameters as in fig. 1.

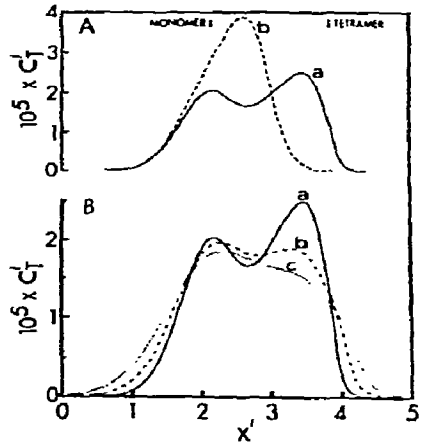


Fig. 6. Factors governing the shape of the molecular-sieve chromatographic profile computed for reaction set IV on Sephadex G-200R, 50% tetramerization. A – Dependence of shape on R_{10} for flow rate $F = 1.2$ ml/hr: a – $R_{10} = 0.10$; b – 10.00. B – Dependence of shape on F for $R_{10} = 0.10$: a – 1.2 ml/hr; b – 5 ml/hr; c – 9.6 ml/hr. Ligand initially present only in starting zone; where tested (conditions in A) virtually the same profiles were obtained when unbound ligand was initially distributed throughout the chromatographic column. The times are such that the volume passed ($V = Ft$) is 5 ml for all curves. $C_{10}' + 4C_{40}' = 14 \times 10^{-5}$ M; $C_{50}' = 5 \times 10^{-7}$ M.

$$M_2 + 2X \rightleftharpoons 2MX \qquad (X)$$

The patterns for 50% dissociation with $C_{30} = 5 \times 10^{-7}$ M and $R_{10} = 1.00$, 2.50 or 5.00 are similar to although better resolved than patterns c, b and a in fig. 3C, respectively. Not only does the shape of the zone show a dependence upon R_{10} which is opposite to that for dimerization; but the redistribution of ligand along the sedimentation column is also opposite in sense, i.e. ligand is concentrated in the upper portion of the column and depleted in the region of the faster peak of the pattern of macromolecule. Finally, bimodal reaction zones computed for the isomerization reaction set VII are displayed in fig. 5.

Thus, the forementioned generalization seems justified that rapidly equilibrating ligand-mediated interactions in general have the potentiality for showing bimodal sedimenting zones irrespective of reaction mechanism provided that overall ligand-binding is strong but that the product of the interaction binds ligand more strongly than does the reactant. We recognize the second condition as characteristic of allosteric interactions. Nor, as shown previously [2], need the ligand be initially distributed throughout the sedimentation column for resolution to occur; virtually the same bimodal pattern was obtained for cooperative ligand-mediated dimerization (reaction I with $m = 2$ and $n = 6$) when ligand was initially present only in the starting zone. Since the generation of bimodal zones is not unique for any one reaction mechanism, it is imperative that appeal be made to the combined application of zone sedimentation with one or more other physical methods such as sedimentation equilibrium and light scattering in order to elucidate the mechanism of reaction.

The above results for zone sedimentation also apply to zone electrophoresis when the mobility of the associated macromolecule happens to be greater than the monomer, and as illustrated in fig. 6, a qualitatively similar behavior is predicted for molecular-sieve chromatography on Sephadex or other gel-permeation supports. It is particularly interesting that increasing the flow rate (fig. 6B) causes progressive coalescence of the two peaks in the broadening reaction zone. This is due to progressive increase in the axial dispersions

of the several species especially the increase in the axial dispersion of the ligand which tends to smooth its concentration gradients through the zone.* Although the chromatographic patterns were computed for the total column frame of reference and are thus directly related to the column-scanning mode of data acquisition [16], the elution profile will show the same features. Accordingly, the same precaution that was emphasized previously for electrophoresis and sedimentation (chapter 6 in ref. [1]; [2, 19]) also applies to gel filtration; namely, fractions must be rechromatographed to see if they run true. Otherwise bimodality due to interactions could be misinterpreted in terms of inherent heterogeneity.

Finally, it is anticipated that the concepts elaborated in these investigations will find application to a variety of biochemical reactions such as the interaction of enzymes with cofactors and allosteric affectors. A recently reported practical application of the mass transport of interacting systems is the determination of the equilibrium constants for the binding of complementary tetranucleotide to *t*-RNA and proflavin to chymotrypsin from acrylamide gel electrophoresis [20].

References

[1] J.R. Cann, Interacting macromolecules. The theory and practice of their electrophoresis, ultracentrifugation, and chromatography (Academic Press, New York, 1970).
[2] J.R. Cann and W.B. Goad, Science 170 (1970) 441.
[3] J.R. Cann and W.B. Goad, Arch. Biochem. Biophys. 153 (1972) 603.
[4] G.A. Gilbert, Dis. Faraday Soc. 20 (1955) 68.
[5] G.A. Gilbert, Proc. Roy. Soc. A250 (1959) 377.
[6] E.O. Field and J.R.P. O'Brien, Biochem. J. 60 (1955) 656.
[7] E.O. Field and A.G. Ogston, Biochem. J. 60 (1955) 661.
[8] J.L. Bethune and G. Kegeles, J. Phys. Chem. 65 (1961) 433.
[9] K. Morimoto and G. Kegeles, Arch. Biochem. Biophys. 142 (1971) 247.
[10] M. Tai and G. Kegeles, Arch. Biochem. Biophys. 142 (1971) 258.
[11] R.C. Weisenberg and S.N. Timasheff, Biochemistry 9 (1970) 4110.
[12] R.G. Martin and B.N. Ames, J. Biol. Chem. 236 (1961) 1372.
[13] J.R. Cann and W.B. Goad, J. Biol. Chem. 240 (1965) 148.
[14] B. Carnahen, H.A. Luther and J.O. Wilkes, Applied numerical methods (Wiley, New York, 1969) p. 319.
[15] W.B. Goad and J.R. Cann, Ann. N.Y. Acad. Sci. 164 (1969) 172.
[16] G.K. Ackers, Adv. Protein Chem. 24 (1969) 343.
[17] J.K. Zimmerman and G.K. Ackers, J. Biol. Chem. 246 (1971) 1078.
[18] J.K. Zimmerman, D.J. Cox and G.K. Ackers, J. Biol. Chem. 246 (1971) 4242.
[19] J.R. Cann and D.C. Oates, Biochemistry 12 (1973) 1112.
[20] J. Eisinger and W. Blumberg, Fed. Proc. 32 (1973) 662 Abs.

* Decreasing the centrifugal field has an analogous effect on the analytical sedimentation patterns for ligand-mediated dimerization (chapter 4 in ref. [1]).

BIOPHYSICAL CHEMISTRY 1 (1973) 11–17. NORTH-HOLLAND PUBLISHING COMPANY

RELATIONSHIP BETWEEN HILL PLOTS WITH VARIABLE EXPONENTS AND DETERMINATION OF AVERAGE FREE ENERGY OF INTERACTION PER SITE*

Magar E. MAGAR and Paul W. CHUN**
Department of Biochemistry, College of Arts and Sciences, University of Florida, Gainesville, Florida 32601, USA

Received 10 January 1973

The Hill plot has been used by many investigators in protein ligand equilibria. As a general rule, it is used with a single exponent. In this paper, we propose the use of the Hill equation with an exponent that varies with the ligand concentration. Since Wyman [1,2] has shown this exponent is related to free energy of interaction of the site, computing the exponent as a function of ligand concentration provides a history of the conformational changes that occur with increasing ligand concentration.

1. Introduction

It has long been known that cooperative effects [3–6] in the binding of oxygen by hemoglobin follow a sigmoid saturation curve which differs from the typical Michaelian saturation character [7].

It is clear, however, from a multitude of recent studies, that many enzymes display sigmoid saturation dependence, and such behavior may be affected by cooperactive interaction [8–17]. Work by Perutz et al. [18, 19] on the structure of hemoglobin suggests that the interaction between hemes or binding sites is indirect, resulting from conformational changes when the molecule binds oxygen.

The recent emphasis on such cooperative conformational effects suggests the view that all allosteric enzymes are composed of subunits and are thought to possess distinct stereospecific binding sites. The binding of ligand to its site produces a specific conformational change in the protein (an allosteric transition).

Since the concept of allosteric protein was introduced by Monod et al. [20], four representative models for allosteric phenomena have been proposed in recent years: (1) the two-state model [21]; (2) the induced-fit or square model [22]; (3) the probability Ising model [23]; and (4) the model of progressive change in ligand interaction [24]. These models, which are equivalent, all pay attention to the mechanism of binding and may be derived using Adair's expression, as described by Magar and Steiner [17]. Wyman [1, 2] has shown that Hill's coefficient (n) [4] may be used as an index of the interaction that occurs in the cooperative binding process.

The interaction parameter, formally defined by Wyman [2, 25] as $n \equiv \mathrm{d}\ln(\bar{y}/1-\bar{y})/\mathrm{d}\ln x$, where $\bar{y}$ is the degree of saturation and x is the ligand concentration, is related to the average free energy of interaction of the sites designated as ΔF_{xx}. By convention, $\Delta F_x = RT \ln x$. $\mathrm{d}\Delta F_x$ measures the effect of a gradual change by $\mathrm{d}\bar{y}$ in the fractional saturation on the free energy. Then, $\mathrm{d}\Delta F_x/\mathrm{d}\bar{y} = RT\,\mathrm{d}\ln x/\mathrm{d}\bar{y}$. When no interaction occurs in the system, $\mathrm{d}\Delta F_x/\mathrm{d}\bar{y} = 0$ and when there is interaction, $\mathrm{d}\Delta F_x/\mathrm{d}\bar{y} \neq 0$. Thus, at any given value of $\mathrm{d}\bar{y}$ with fractional saturation $\bar{y}$, ΔF_{xx} is defined as the free energy of interaction at a given saturation. When $n > 1$, the interaction is a stabilizing one, where ΔF_{xx} gives the stabilizing energy per site for the system as a whole at saturation [2]. Wyman derived the following equation for ΔF_{xx}:

$$\Delta F_{xx} = RT \int_{x=0}^{\infty} (n-1)\,\mathrm{d}\ln x \qquad (1)$$

for fixed exponents of n.

* This work was supported by National Science Foundation Grant GB 28223-A*-1 and in part by the Department of Health, Education, and Welfare Grant HL 05784-07(T01).
** Reprint requests and all correspondence to: Dr. Paul W. Chun.

The Hill plot has been extensively used by Wyman [2], as well as by Antonini et al. [26], Brunori et al. [27], and Banerjee and Cossaoly [28] in studies on the oxygenation of hemoglobin to obtain the average free energy of interaction per site. This procedure has been disputed by Saroff and Minton [29], who claim that Hill plots can only yield the difference between the free energies of interaction between the first and last sites. Saroff and Yap [30] further conclude that the relationship between the value of n and the interaction energy is meaningless without the application of a model. Whatever the precise interpretation of the quantities yielded by the Hill plots, it is clear that the analysis to date indicates the value of n is not, in fact, constant throughout the binding curve, but is a function of ligand concentration. Therefore, a single, constant n value is not sufficient to describe the cooperative binding process which occurs.

In this communication, we propose the use of the Hill plots with a variable exponent which varies as a function of ligand concentration. This new procedure permits a more detailed evaluation of the thermodynamic parameters of protein-ligand interaction.

2. Theoretical considerations

A.V. Hill [4] proposed an empirical equation which describes the oxygenation of hemoglobin:

$$\frac{\bar{y}}{1-\bar{y}} = kx^n. \tag{2}$$

The linear transform of this equation yields the parameter n, that is, $\ln(\bar{y}/1-\bar{y})$ versus $\ln x$. When the sites are all identical and independent of each other, then this plot gives a straight line with $n = 1$. If cooperative interaction exists, i.e., $n > 1$, then the Hill plot will no longer give a straight line.

Wyman [1] derived the stabilizing free energy of binding per site at saturation x, given by:

$$\Delta F_{xx} = \frac{RT}{(1-\bar{y})\bar{y}}\left(1-\frac{1}{n}\right). \tag{3}$$

This equation (3) can also be expressed as:

$$\Delta F_{xx} = RT \int_0^{\infty} (n-1)\,\mathrm{d}\ln x.$$

Mounting studies in the literature indicate that a constant value of n is not suitable to describe the binding process. In work on hemoglobin, Saroff [31] has asserted that a fixed exponent n characterizes only a limited region of the binding curve, neglecting regions at high and low saturations of oxygen.

In order to assess the full range of cooperative efforts, n must be considered as a function of ligand concentration, and therefore variable.

In this paper, Hill's exponent will be designated $\psi(x)$ to distinguish it from n defined as the number of binding sites, as it often appears in the literature. Thus, Wyman's expression would be written as:

$$\Delta F_{xx} = RT \int_0^{x_H} [\psi(x) - 1]\,\mathrm{d}\ln x. \tag{4}$$

In eq. (4), x_H is the highest ligand concentration at which $\psi(x) \neq 1$. Since, when $\psi(x) = 1$, then $\psi(x) - 1 = 0$, and there is no contribution to the average stabilizing free energy.

In order to compute $\psi(x)$ as a function of ligand concentration $\psi(x)$ can be represented as a polynomial in x, thus:

$$\psi(x) = 1 + \sum_{i=1}^{m} \alpha_i x^i, \tag{5}$$

where α_i are constants to be determined. Using Hill's equation at a given ligand concentration x_j, we obtain:

$$\bar{y}_j = kx_j^{\psi(x_j)}/(1 + kx_j^{\psi(x_j)}) + \delta_j, \tag{6}$$

where δ_j is the experimental error in making a measurement. Substituting $\psi(x)$ into eq. (6) gives:

$$\frac{\bar{y}_j}{1-\bar{y}_j} = [kx_j^{(1+\Sigma_{i=1}^{m}\alpha_i x_j^i)}] + \delta_j. \tag{7}$$

The experimental error δ_j can be written as:

$$\delta_j = \left[\left(\frac{\bar{y}_j}{1-\bar{y}_j}\right) - kx_j^{(1+\Sigma_{i=1}^{m}\alpha_i x_j^i)}\right],$$

whose sum can be experessed as a function of the parameters k and the α_i (written collectively as $\boldsymbol{\alpha}$) as follows:

$$\phi(k,\boldsymbol{\alpha}) = \left[\sum_{j=1}^{k} \frac{\bar{y}_j}{1-\bar{y}_j} - k x_j^{1+\sum_{i=1}^{m}\alpha_i x_j^i}\right]^2. \tag{8}$$

We note from eq. (8) that we are trying to determine k and $\boldsymbol{\alpha}$, so as to minimize the sum of the squares of the experimental error or the function,

$$\phi(k,\boldsymbol{\alpha}) = \sum_{j=1}^{k} \delta_j^2. \tag{9}$$

Once we find the values of k and $\boldsymbol{\alpha}$ that determine that minimum, we can compute $\psi(x)$ and, using eq. (4), we can compute ΔF_{xx} as a function of concentration. Lyster's results on horse hemoglobin quoted in ref. [32] and the results of Antonini et al. [33] on *Spirographis chlorocruorin* hemoglobin were used to make the computations.

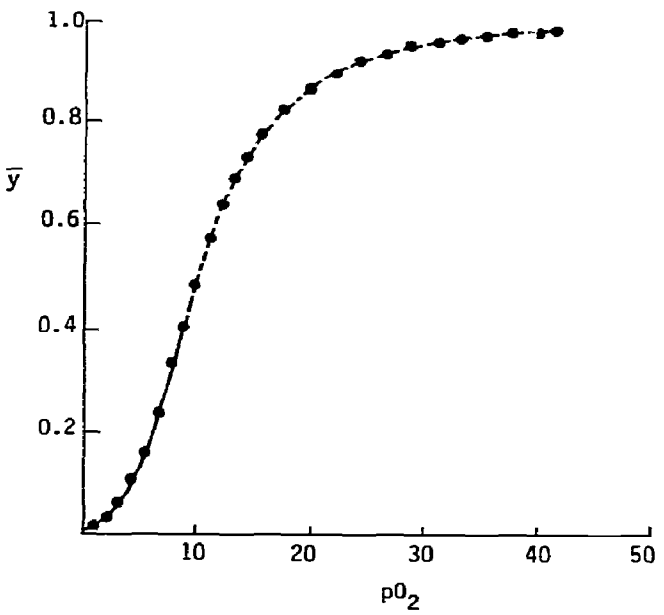


Fig. 1. Nonlinear least square analysis of the oxygenation of horse hemoglobin. The black circles (•) represent horse-hemoglobin data (Lyster, quoted in ref. [32]); the dotted line is based on eq. (8) where $k = 0.085$; and the solid line on eq. (10) where $k = 0.033$.

3. Computational procedure

There are a number of methods for obtaining the minimum of $\phi(k,\boldsymbol{\alpha})$ [34]. In this manuscript we used the non-linear least square procedure, BMD-OIR (UCLA Biomed program of Dixon [35]), which employs a modified Gauss–Newton method. This is a widely used method available at many computing centers. For our particular application we recommend that all initial estimates of the constants of k and all the α's be set equal to zero. Failure to do so results in an overflow, since the magnitude of the first derivatives is quite large. Having put in the initial estimates, the computer program makes continuous changes in the parameters, with each iteration resulting in successively smaller mean square deviations of the observed $\bar{y}$ values with respect to x for the curve generated at that stage. The procedure is continued until the mean square deviation is negligible and provides the best fit to the data. For our purposes, we fit the curves in sections, since it is clear that if

$$\psi(x) = 1 + \sum_{i=1}^{m} \alpha_i x^i,$$

then data at lower ligand concentrations must be eliminated from the fit. At such low concentrations, a plot of $\ln(\bar{y}/1-\bar{y})$ versus $\ln x$ is a straight line with a slope equal to unity. Such a straight line cannot be fitted with a polynomial function unless $\alpha_i = 0$ for all $i \geqslant 1$. In this region of low concentration, the contribution to the average free energy of interaction per site is negligible. A fit at low concentration was made using the expression of eq. (10):

$$\bar{y}/(1-\bar{y}) = kx, \tag{10}$$

where the exponent of x is unity. At higher concentrations we made fits with variable exponents.

4. Results

In fig. 1, we see the effects of such a fit for data on horse hemoglobin (Lyster, quoted in ref. [32]). The parameter $k = 0.033$ describes the region $x = 0$ to $x = 12.5$. From $x = 12.5$ to $x = 45$, k was found to be 0.085 and

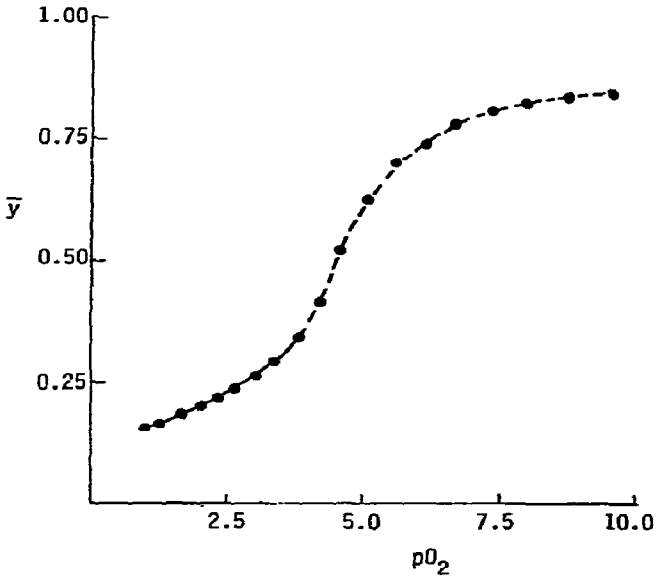


Fig. 2. Regeneration of the curve for $\bar{y}$ as a function of pO_2 from a Hill plot of $\log \bar{y}/(1-\bar{y})$ versus $\log P$. Data on *Spirographis chlorocruorin* taken from the results of Antonini et al. [33] as plotted by Wyman [2]. Here again the solid line is a fit of eq. (10) and the dotted line is a fit using eq. (8).

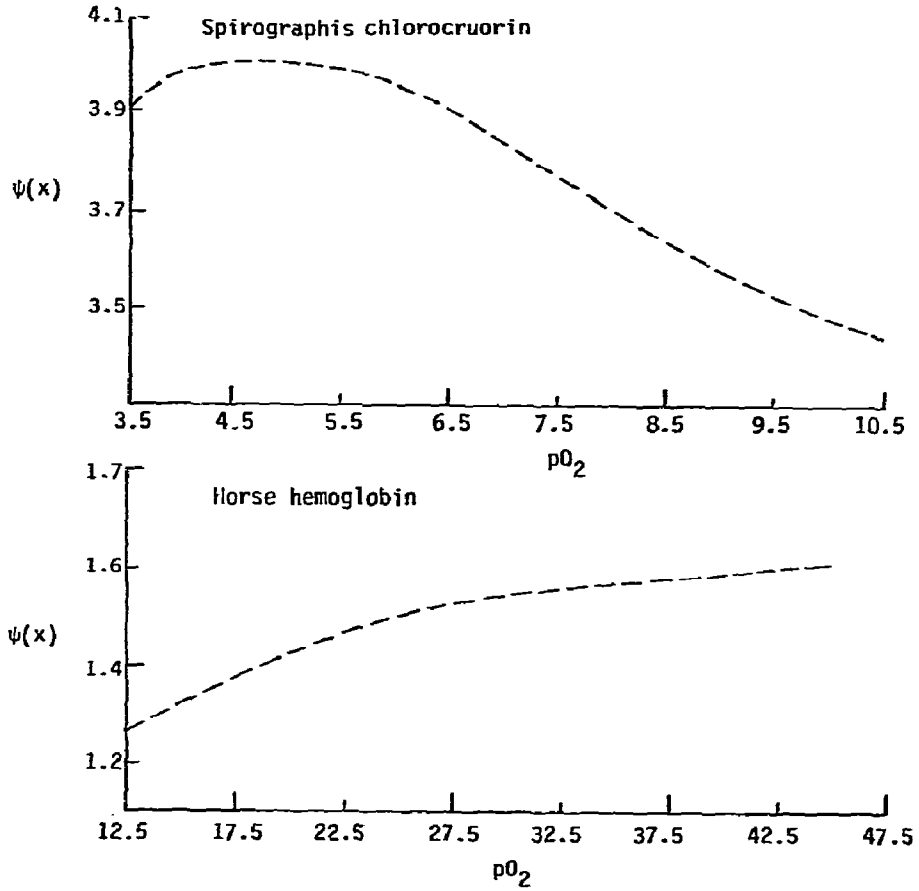


Fig. 3. A plot of $\psi(x)$ versus pO_2 for *Spirographis chlorocruorin* and horse hemoglobin. Data on *Spirographis* taken from the results of Antonini et al. [33] as plotted by Wyman [2]. Lyster's data on horse hemoglobin as quoted in Monod et al. [32].

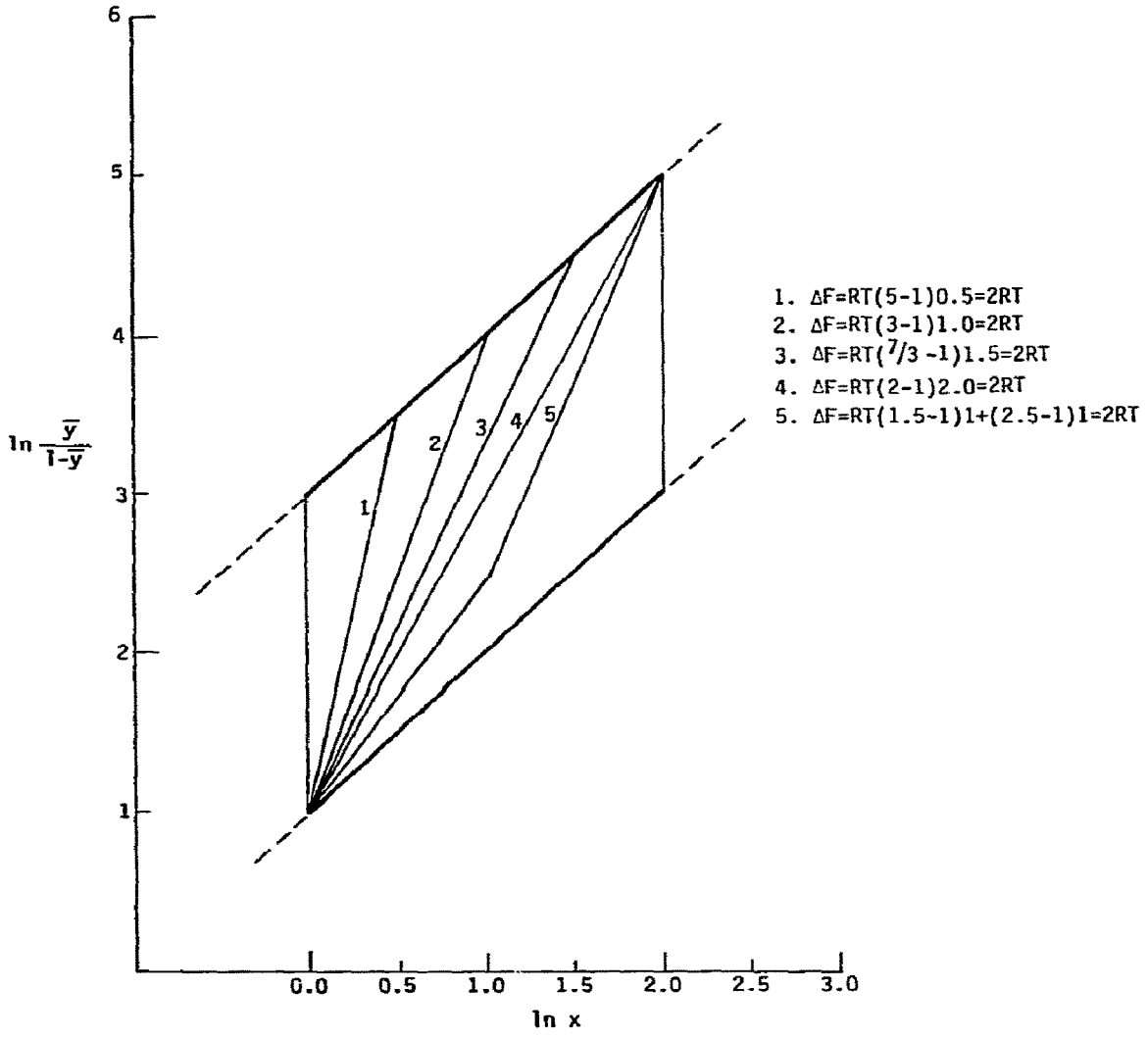


Fig. 4. Demonstration of the fact that, irrespective of the slope of the Hill plot, the Wyman equation $RT\int(n-1)\mathrm{d}\ln x$ always gives the same free energy for the same distance between the asymptotes. Curve 5 shows that even if we have several different slopes on the same curve, we also have the same value for free energy between the same asymptotes. It also can be proven that the same energy is obtained if a curve is broken up into an infinite number of slopes.

$$\psi(x) = 1 + 0.0113x + 0.00145x^2 - (6.11 \cdot 10^{-5})x^3 + (6.66 \cdot 10^{-7})x^4$$

for horse hemoglobin.

In fig. 2, on *Spirographis chlorocruorin*, the parameter $k = 0.3362$ describes the low-concentration region, $x = 0$ to $x = 3.6$. From $x = 3.6$ to $x = 10.6$, $k = 0.00235$ and

$$\psi(x) = 1 + 1.65x - 0.307x^2 + (2.27 \cdot 10^{-2})x^3 - (6.01 \cdot 10^{-4})x^4$$

for *chlorocruorin*. A plot of $\psi(x)$ versus x for horse hemoglobin and *Spirographis chlorocruorin* is shown in fig. 3.

5. Discussion

The relationship of Hill's parameter n to the free energy of interaction makes it a good indication of cooperativity. The use of $\psi(x)$ as a variable exponent rather than a single value of n permits continuous characterization of the cooperativity at all ligand concentrations, compiling a more complete "history" of the ligand binding.

It is important to note that the ΔF_{xx} for the entire ligand range can be computed by Wyman's equation (1) by taking the distance between the asymptotes of the Hill plots, irrespective of the value of n and even if n is not constant. This is demonstrated geometrically in fig. 4. However, if one requires a "history" of ΔF_{xx} as a function of x, then $\psi(x)$ must be used.

As defined in eq. (4):

$$\Delta F_{xx} = RT \int_0^{xH} [\psi(x) - 1] \, \mathrm{d} \ln x,$$

a plot of $[\psi(x) - 1]$ versus $\ln x$ yields the average stabilizing free energy whenever $[\psi(x) - 1] > 0$. Such a plot provides an idea of the free energy of interaction with changing ligand concentration. The area under the $[\psi(x) - 1]$ versus $\ln x$ between the two ligand concentrations thus provides a sort of "history" of the conformational changes as the ligand concentration increases.

Saroff [36], in his work on hemoglobin, has emphasized that n is not constant, although he stops short of advocating a solution for that problem. Furthermore, a variety of single exponents can be shown to fit the data. For example, he has shown that a given set of data can be fitted with a series of k and n values, with n varying from 2.6 to 3.2 for a given portion of the Hill plot.

His results serve to emphasize the point that unless the whole range of the $\ln(\bar{y}/1 - \bar{y})$ versus $\ln x$ curve is fitted, many single values of n are possible, depending on which portion of the curve is fitted. We feel that the plot of $\psi(x)$ versus x is a more faithful reproduction of the cooperative binding which is taking place.

Our plot of $\psi(x)$ versus ligand concentration (fig. 3) for horse hemoglobin approaches a maximum value of 1.6 for the region $x = 12.5$ to $x = 45$ in contrast to a value reported by Wyman [2] of approximately $n = 3$ for one portion of the Hill plot. Because our results were considerably lower than the slope of the Hill plot reported by other researchers, we attempt enumerable fits, always obtaining a result close to 1.6.

Thus, we must conclude that the maximum value of $\psi(x) = 1.6$ for this particular region. This conclusion reinforces the point that any two or three points on the Hill plot will yield many different slopes, particularly when the data is rising so steeply. Linear least square analysis, however, yields the best fit for all points on the Hill plot (fig. 1).

Also to be noted in fig. 3 is the fact that the curve of $\psi(x)$ versus x for horse hemoglobin does not reach a maximum and start to decline again as would be expected because $n = 1$ at both extremes of the Hill plot. That $\psi(x)$ does not decline in the case of horse hemoglobin can be attributed to the data being analyzed. It is a truism that data analysis methods are only as good as the data they seek to analyze. As recorded by Wyman [2], Lyster's unpublished data show that at high ligand concentration, the Hill plot had yet to approach a slope of unity. Once a slope of unity is approached at high concentration, a fit using equation (10) may be entertained, as there is no contribution to the average free energy of interaction of the site in this region.

Because of the empirical nature of the plot, the computation of the average free energy of interaction per site, it should be emphasized that plots of $\psi(x)$ versus x are not to be extrapolated beyond the ligand concentration for which they were computed.

The use of a polynomial to express the variability of the Hill exponent is not the only device to achieve that end. If an investigator feels that other functions are appropriate or even more suitable to obtain good fits, they could be employed. In particular, since we are fitting the saturation curves piecewise, spline-function interpolation may be used [37].

Since the occurrence of negative cooperativity in L-threonine deaminase has been reported by Changeux [38] and in digested hemoglobin by Brunori et al. [27], as well as in other systems (Monod et al., [32]), we feel that an additional advantage of our procedure is that, in those rare instances where the free energy of interaction at the site may fluctuate from positive to negative at a range of ligand concentration, use of $\psi(x)$ offers additional information over earlier procedures. Careful analysis using the variable $\psi(x)$ may reveal the magnitude of cooperative interaction operating in these systems, and the nature of the delicate regulatory mechanism which controls it.

Acknowledgement

We are grateful for the assistance provided by the University of Florida Computing Center. Our thanks

are also extended to Mrs. Jean Holtzer for her time in making computations for this work.

Appendix

For the convenience of investigators who wish to employ function optimization techniques which use the first derivative, we have the following derivations:

$$F(k,\alpha) = kx^{(1+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3+\alpha_4 x^4)}$$

$$\frac{\partial F(k,\alpha)}{\partial k} = [x^{(1+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3+\alpha_4 x^4)}]$$

$$\frac{\partial F(k,\alpha)}{\partial \alpha_1} = [kx^{(1+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3+\alpha_4 x^4)}]\ \ln x\, x$$

$$\frac{\partial F(k,\alpha)}{\partial \alpha_2} = [kx^{(1+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3+\alpha_4 x^4)}]\ \ln x\, x^2$$

and in general

$$\frac{\partial F(k,\alpha)}{\partial \alpha_i} = [kx^{(1+\alpha_1 x+\alpha_2 x^2+\alpha_3 x^3+\alpha_4 x^4)}]\ \ln x\, x^i$$

or

$$\frac{\partial F(k,\alpha)}{\partial \alpha_i} = \left[\frac{F(k,\alpha)}{\alpha_{i-1}}\right] x\,.$$

References

[1] J. Wyman, Jr., Adv. Proc. Chem. 4 (1948) 407.
[2] J. Wyman, Jr., Adv. Proc. Chem. 19 (1964) 233.
[3] C. Bohr, Zentr. Physiol. 17 (1903) 682.
[4] A.V. Hill, J. Physiol. 40 (1910) 190.
[5] G.S. Adair, J. Biol. Chem. 63 (1925) 529.
[6] L. Pauling, Proc. Natl. Acad. Sci. USA 21 (1935) 186.
[7] L. Michaelis and M. Menten, Biochem. 8 (1913) 49, 333.
[8] R.A. Alberty, J. Am. Chem. Soc. 75 (1953) 1925.
[9] W.W. Cleland, Biochim. Biophys. Acta 67 (1963) 173.
[10] A. Rossi-Fanelli, E. Antonini and A. Caputo, Adv. Prot. Chem. 19 (1964) 73.
[11] D.E. Atkinson, Ann. Rev. Biochem. 35 (1966) 85.
[12] D.E. Koshland, Jr. and K.E. Jeet, Ann. Rev. Biochem. 37 (1968) 359.
[13] C. Frieden, J. Biol. Chem. 242 (1967) 4045.
[14] C. Frieden, J. Biol. Chem. 245 (1970) 5788.
[15] E. Whitehead, Biochemistry 9 (1970) 1440.
[16] M.E. Magar, R.F. Heiner and J. Fletcher, J. Theor. Biol. 32 (1971) 59.
[17] M.E. Magar and R.F. Steiner, J. Theor. Biol. 32 (1971) 495.
[18] M.F. Perutz, W. Bolton, R. Diamond, H. Muirhead and H.C. Watson, Nature 203 (1964) 687.
[19] M.F. Perutz, M.G. Rossmann, A.F. Cullis, H. Muirhead, G. Hill and A.C.T. North, Nature 185 (1960) 416.
[20] J. Monod, J.P. Changeux and F. Jacob, J. Mol. Biol. 6 (1963) 306.
[21] J. Monod, J. Wyman, Jr. and J.P. Changeux, J. Mol. Biol. 12 (1965) 88.
[22] D.E. Koshland, Jr., G. NeMethy and D. Filmer, Biochemistry 5 (1966) 365.
[23] C.J. Thompson, Biopolymers 6 (1968) 1106.
[24] D.E. Atkinson, J.A. Hathaway and E.C. Smith, J. Biol. Chem. 240 (1965) 2682.
[25] J. Wyman, Jr., Cold Spring Harb. Symp. Quant. Biol. 28 (1963) 483.
[26] E. Antonini, J. Wyman, M. Brunori, J.F. Taylor, A. Rossi-Fanelli and A. Caputo, J. Biol. Chem. 239 (1964) 907.
[27] M. Brunori, E. Antonini, J. Wyman, Jr., R. Zito, J.F. Taylor and A. Rossi-Fanelli, J. Biol. Chem. 239 (1964) 2340.
[28] R. Banerjee and R. Cassoly, J. Mol. Biol. 42 (1969) 351.
[29] H.A. Saroff and A.P. Minton, Science 175 (1972) 1253.
[30] H.A. Saroff and W.T. Yap, Biopolymers 11 (1972) 957.
[31] H.A. Saroff, J. Phys. Chem. 76 (1972) 1597.
[32] J. Monod, J. Wyman, Jr. and J.P. Changeux, J. Mol. Biol. 12 (1965) 88.
[33] E. Antonini, A. Rossi-Fanelli and A. Caputo, Arch. Biochem. Biophys. 97 (1962) 336.
[34] M.E. Magar, Data analysis in molecular biology (Ac. Press, New York, 1972).
[35] W.J. Dixon, BMD, Biomedical computer programs, nr. 2 (Univ. of Calif. PRess, Berkeley, 1969).
[36] H.A. Saroff, J. Phys. Chem. 76 (1972) 1597.
[37] J.R. Rice, The approximation of function, Part II (Addison-Wesley, Reading, Mass., 1968).
[38] J.P. Changeux, Cold Spring Harb. Symp. Quant. Biol. 28 (1963) 497.

BIOPHYSICAL CHEMISTRY 1 (1973) 18–23. NORTH-HOLLAND PUBLISHING COMPANY

COMPARISON OF METHODS FOR ENUMERATION OF INTERACTING SPECIES IN A BIOLOGICAL SYSTEM

Magar E. MAGAR and Paul W. CHUN*

Department of Biochemistry, College of Medicine, J. Hillis Miller Health Center, University of Florida, Gainesville, Florida 32601, USA

Received 15 October 1972
Revised manuscript received 27 January 1973

Matrix rank analysis, a method extensively used to enumerate interacting species or components in various biological systems, has been compared with the more readily available technique of factor analysis. Data previously analyzed by matrix rank analysis has been subjected to factor analysis. The results show that there are some agreements and disagreements between the two methods. The sources and nature of the disagreements are discussed. Our work indicates that factor analysis may be preferable to matrix rank analysis.

1. Introduction

Matrix rank analysis, a technique to determine the number of linearly independent components, has had many applications in biochemistry and molecular biology. It can be applied to determine the number of components in a mixture of absorbing species [1, 2] to analyze the optical rotatory dispersion of tobacco mosaic virus [3], and to determine the number of interacting species in fluorescent spectroscopy [4, 5]. It has also been applied to the kinetics of the oxidation of hemoglobin derivatives [6], to gel filtration data [7], and to the determination of interacting species from a diffusion experiment in the ultracentrifuge [8].

Factor analysis and component analysis, two long-recognized techniques, have been suggested as alternatives to the method of matrix rank analysis [9].

Fundamentally, the question asked of matrix rank analysis and these two methods is: given a matrix array of numbers resulting from one or the other of the analytical techniques mentioned above, how many linearly independent components or interacting species produce this array? One has to ask this question because from a matrix array of essentially experimental numbers one cannot determine the rank of the matrix by simply determining the smallest nonzero minor. It is manifest that since the matrix elements are experimental quantities (and hence subject to error), such a procedure will not do. In fact, because of experimental error, it is probably that the entire data matrix under consideration is nonsingular. In what follows, we briefly describe both methods of determining the number of components, apply component and factor analysis to examples in the literature in which matrix rank analysis has been used, and, lastly, compare results.

2. Comparison of methods

Consider a matrix array whose elements x_{ij} are the result of a set of measurements. To fix the idea and show the applications we pick the analysis of absorption spectroscopy as an example. Here, the index i would run over wavelength and the index j would identify the solution under various conditions of pH, ionic strength, etc., or conditions that purportedly alter the contributions of the various components. Thus, in the case of Wallace [1], a solution of indicator was measured at different wavelengths and at different pH. McMullen et al. [3] measured the ORD of a solution of tobacco mosaic virus RNA at various conditions of temperature and ionic strength. Ainsworth

* To whom to address correspondence.

and Bingham [6] measured the initial velocity of the combination of various concentrations of partially oxidized hemoglobin with carbon monoxide at different pH's, and so on.

To determine the rank of the matrix, Wallace [1] first suggested that the values and the probable error of all the determinant minors of the data matrix be computed and compared. There is some merit to this procedure and it has been extensively applied to biological systems. Its merits and applications will be detailed in the discussion. Unfortunately, it was felt to be too tedious and was abandoned by Wallace and Katz [2] in favor of another method. This method took the original data matrix and applied a series of elementary transformations (which do not change the rank of the matrix) to produce an upper triangular matrix. These transformations were performed in such a way that the largest element in the rows of the transformed matrix was on the principal diagonal. Schematically, this may be represented as:

$$\begin{bmatrix} x_{11} \cdots x_{1n} \\ x_{21} \cdots x_{2n} \\ \vdots \quad \vdots \\ x_{m1} \cdots x_{mn} \end{bmatrix} \xrightarrow{\text{elementary transformation}} \begin{bmatrix} x'_{11} & x'_{12} \cdots x'_{1n} \\ & x''_{22} \cdots x''_{2n} \\ & \ddots \quad \vdots \\ & \qquad x^{(m-1)}_{mn} \end{bmatrix}$$

Original matrix — Reduced data matrix

with the $(m-1)$ elements x'_{11} and x''_{22} being larger than all other elements on their corresponding rows. In a similar fashion, another matrix whose elements are the standard errors of the corresponding elements of the original matrix is subjected to a series of elementary transformations to a reduced error matrix [2]. The reduced error matrix is also upper triangular.

$$\begin{bmatrix} \text{Elements of this matrix are standard errors of data matrix} \end{bmatrix} \xrightarrow{\text{elementary transformations}} \begin{bmatrix} s'_{11} & \cdots & s'_{1n} \\ & s''_{22} & \cdots & s''_{2n} \\ & & s'''_{33} \\ & & \ddots \\ & & & s^{(m-1)}_{mn} \end{bmatrix}$$

Original error matrix — Reduced error matrix

To determine the number of components or interacting species, one compares the $(m-1)$ elements of the reduced data matrix with those of the reduced error matrix and comes to a conclusion as to how many components there are. This last step is best illustrated for the reader by showing two actual examples as to how it was done. In a case analyzed by Wallace and Katz [2], the diagonal elements for the reduced data matrix were:

1.015 0.414 0.060 −0.016
0.006 0.005 0.006 0.004,

and for the reduced error matrix they were:

0.003 0.003 0.004 0.005
0.006 0.010 0.017 0.021.

From these data Wallace and Katz [2] came to the conclusion, and we quote, "The results show the definite existence of three colored components and perhaps a fourth, since the element 4, 4 i.e., −0.016 in the reduced matrix is three times its estimated error, i.e., 0.005". In the case analyzed by Ainsworth and Bingham [6], the elements for the reduced data matrix were:

2.22 0.3968 −0.2150 0.0808,

and the elements of the reduced error matrix were:

0.120 0.0900 0.157 0.1682.

The authors then came to the following conclusions regarding the number of components: "Comparing the principal diagonals of the two reduced matrices

2.22 0.369 0.2150 0.0807
0.120 0.090 0.157 0.1682

indicates the existence of at least two variables. The possibility that a third independent variable might exist receives only slender support; thus, the ratio of element 3, 3 to its estimated error is 1.4".

The exact statements of the authors are inserted to give the reader a feeling for what has been involved in estimating the number of components. Ideally, there should be a specific statistical test giving a certain probability level that one can use in order to

determine the number of interacting species, and we presume that Wallace and Katz [2] and Ainsworth and Bingham [6] were using one. However, no such probability or confidence statements were given by those authors, a matter that is taken up later.

With respect to factor analysis, the method can be outlined as follows:

Factor analysis can analyze any system which can be written in the form

$$x_{ji} = \sum_{k=1}^{m} \ell_{jk}\zeta_k + e_j \qquad (j = 1, 2 \ldots p).$$

The various x_j's are samples thought to be made up of m components denoted by ζ_k $(k = 1, 2 \ldots m)$; ℓ_{jk} is the quantitative contribution of the kth factor to the jth sample; and e_j stands for the errors. In previous papers we advocated the use of maximum likelihood estimates to determine the ℓ_{jk} and the σ_j^2 (which are the variances of the set of e_j). Once those values are obtained, a statistical test proposed by Bartlett [10] was suggested to determine the number of components. In this paper we apply a much simpler criterion to determine the number of components, using the data of Wallace and Katz [2] and Ainsworth and Bingham [6].

There are several reasons for using this simpler method. First and foremost, as far as we know, most computer centers have only programmed packages using the simpler criteria, which are therefore readily accessible. The extensively distributed UCLA Biomed package (which we used) and the University of Wisconsin factor analysis package do not use the test based on maximum likelihood estimates, but rather a simpler one. Second, and we discuss this in greater detail, the criteria adopted for determining the number of factors or interacting species in those programs, as a practical matter, have proven very useful.

This is not to say that the maximum likelihood methods are not to be preferred. On the contrary, the fact that they emanate from a sound statistical basis would, in the eyes of many, make them preferable. However, since our objective is to simplify matters for the working chemist, biochemist and molecular biologist, and since questions and canned programs associated with maximum likelihood solutions are still in a state of flux [11, 12], we will use the simpler criteria.

For this method, we proceed as follows. The data matrix, call it X, is transformed into its standard form, given by the matrix Z as follows. Define the sample variance of x_j by

$$s_j^2 = \sum_{i=1}^{n} (x_{ji}^2/n).$$

The elements of Z are given by

$$Z_{ji} = x_{ji}/s_j.$$

From the matrix Z, the matrix of observed correlations, which is given by

$$R = ZZ'/n\ ,$$

is obtained. In the above equations, Z' is the transpose of Z. The eigenvalues of R are then found, call them $\theta_1, \theta_2, \ldots \theta_n$, and there are as many significant factors as there are eigenvalues greater than unity. For the data of Wallace and Katz [2] the correlation matrix was found to be as in table 1 and the eight eigenvalues of this matrix were:

6.36461	1.63065	0.00296	0.00142
0.00022	0.00013	0.00001	0.00000,

accounting (respectively) for the following proportion of the total variance:

0.79558	0.99941	0.99978	0.99995
0.99998	1.00000	1.00000	1.00000.

In the case analyzed by Ainsworth and Bingham [6] the correlation matrix was found to be as in table 2. The eigenvalues of this correlation matrix were:

3.68502	1.49854	0.49854
0.32751	0.00000	0.00000,

which in turn accounted for the following cumulative proportion of the total variance:

0.61417	0.86393	0.94541	1.0000	1.0000.

Table 1
Correlation matrix of Wallace & Katz (1964)

1	2	3	4	5	6	7	8
1.00000							
0.99858	1.00000						
0.99739	0.99826	1.00000					
0.99713	0.99495	0.99464	1.00000				
0.94498	0.93222	0.93595	0.96346	1.00000			
−0.19808	−0.23567	−0.22189	−0.13772	0.13173	1.00000		
−0.74682	−0.77277	−0.76196	−0.70572	−0.49077	0.79851	1.00000	
−0.85601	−0.87563	−0.86709	−0.82305	−0.64159	0.67553	0.98305	1.00000

Using the criteria of eigenvalues of one or greater as representing the number of interacting species, we find that in the case analyzed by Wallace and Katz [2], according to this analysis only two components were present; and in the case analyzed by Ainsworth and Bingham [6], our analysis yields two components, which is roughly the same conclusion that Ainsworth and Bingham came to.

3. Discussion

The test used here to determine the number of interacting species is a simple rule of thumb, and as such commends itself to investigators involved in biological problems such as those mentioned in the introduction. There is, however, more to justify it than mere simplicity. There are a number of techniques used over the years by factor analysts to determine the number of factors [13, 14]. Statisticians, as a rule, would be inclined to favor tests of significance, and in this respect it appears that the method of using maximum likelihood estimations [10, 15, 16] in statistical tests would be best. On the basis of a wide use of factor analysis and after considering statistical significance, algebraic necessity conditions and other criteria, Kaiser [17, 18] came to the conclusion that eigenvalue one criterion is one of the "best" methods available. This criterion has gained wide popularity among factor analysts [14]. It is not infrequent that empirical criteria sometimes yield results which are just as useful as more formal criteria. Speaking of this problem in the domain of factor analysis, Harman [13] has stated, "It has been found by a number of workers that empirical tests of significance used by factor analysts frequently lead to about the same results as the more proper statistical tests".

Perhaps the biggest drawback of this simple criterion is the problem posed when one has eigenvalues close to unity. For example, it does not appear altogether proper to retain an eigenvalue of 1.03 and drop a subsequent one of 0.97, or to consider that an eigenvalue of 0.989, for example, does not belong to a factor contributing significantly to the result. In such instances, other empirical criteria or perhaps rigorous statistical tests may be useful to supplement the analysis. The examples analyzed here do not pose this problem.

Table 2
Correlation matrix of Ainsworth and Bingham (1968)

1	2	3	4	5	6
1.00000					
0.57669	1.00000				
0.59778	0.71901	1.00000			
0.65424	0.56354	0.90744	1.00000		
0.60740	0.22901	0.35830	0.70074	1.00000	
0.27492	0.02390	0.16171	0.54816	0.93066	1.00000

Comparing the results obtained by factor analysis and matrix rank analysis, we find that we obtained only two components for the absorption spectra data of Wallace and Katz [2]. This is in contrast to the authors' statement of at least three or four components. In the case of Ainsworth and Bingham's data [6], these authors felt that there were two factors, possibly three. We found two. In view of those differences, it is pertinent to look at the procedure of matrix rank analysis.

The current procedure of rank analysis due to Wallace and Katz [2] was suggested in preference to an earlier procedure by Wallace [1]. In the earlier procedure, Wallace decided to use an elementary statistical criterion to determine rank. His argument runs as follows: The probable error of a function ψ of several variables $x_1, x_2, \ldots x_n$, written as $\psi(x_1, x_2, \ldots x_n)$, is given by:

$$P_{x_1}^2 \left(\frac{\partial \psi}{\partial x_1}\right)_{\bar{x}_1}^2 + P_{x_2}^2 \left(\frac{\partial \psi}{\partial x_2}\right)_{\bar{x}_2}^2 + \ldots P_{x_n}^2 \left(\frac{\partial \psi}{\partial x_n}\right)_{\bar{x}_n}^2,$$

where $P_{x_1}, P_{x_2}, \ldots P_{x_n}$ are the probable errors of $x_1, x_2, \ldots x_n$, respectively, where the derivatives are evaluated at the mean value of the measured variable. Since a determinant of a square matrix is a function of all its elements, the standard deviation of a determinant is:

$$\left[\sum_i \sum_j (\text{probable error of } a_{ij})^2 \times \left(\frac{\partial A}{\partial a_{ij}}\right)^2\right]^{\frac{1}{2}}.$$

Since the derivative of a square matrix with respect to any of its elements is its cofactor, then the standard deviation of the determinant $\sigma_{|A|}$ is

$$\sigma_{|A|} = \left[\sum_i \sum_j (\text{probable error of } a_{ij})^2 \times (\text{cofactor } a_{ij})^2\right]^{\frac{1}{2}}.$$

If the a_{ij} or the standard deviations of the individual elements are known, one can compute $\sigma_{|A|}$ and the probability that a matrix is singular can be determined from the ratio $|A|/\sigma_{|A|}$ by consulting statistical tables. In his first paper Wallace applied this test to all the minors of increasing order and concluded that there were only two components in his system. Later, analyzing the same system with the new method, Wallace and Katz came to the conclusion noted at the beginning, and they believed that the more recent analysis was better, partly because Wallace had to restrict the size of the matrices he had analyzed the first time, and partly because Wallace and Katz felt that they were confirming the result of Reilley and Smith [19], who found three components by another method.

A word of caution is in order with respect to the remark sometimes made by matrix rank analysts to the effect "that other *experimental evidence* shows that there are more components in the system". A biological system could well contain other interacting species detectable by *experimental* (as opposed to mathematical) means. But if matrix rank analysis and factor analysis are used, then these methods determine in principle only *linearly independent components or species. All linearly dependent components appear as one component.*

Wallace's older method is based on a fairly rigorous statistical test, and despite the large number of computations that one may have to make, the test is conceptually easy to understand. It is possible to devise more sophisticated tests for Wallace's older method.

On the other hand, there are a few things about the newer method of Wallace and Katz which need clarification. It is not clear that a series of elementary transformations will yield a unique upper triangular matrix, but it is true that in the absence of experimental error any series of elementary transformations which yields an upper triangular matrix will have the same number of nonzero rows as the rank of the matrix. Here we have experimental error, and the question of uniqueness is of some import. If we do not have uniqueness, then there would be several possible diagonal elements that must be compared and no clear choice as to what should be done.

Another point that must be noted is that workers using matrix rank analysis seem to have decided that comparison of the elements on the principal diagonal of the reduced data matrix and the reduced error matrix is all that is required to come to a conclusion regarding the rank of the matrix. This state of affairs seems to have originated with Wallace and Katz. Although these authors stated quite explicitly that the

entire row of the reduced data matrix has to be zero, they only compared the elements on the principal diagonals of the pertinent matrices. Ainsworth and Bingham followed that procedure. Since it is clear that we have to compare all elements of a given row, the question arises as to what would happen if, having selected a certain criterion for termination, we find that some elements of the rows satisfy this criterion, while others do not.

As indicated earlier, the authors have not given the probability level by which we are to accept termination. The statements (already noted) by Wallace and Katz and by Ainsworth and Bingham, regarding termination criteria, are rather loose; and only Godschalk states clearly that whenever a matrix element (in the data matrix) after a reduction step falls below its error value (in the error matrix) is it considered zero.

Wallace's first paper presumably uses a Student *t* test to give a probability level, and we suppose that it was the intent of Wallace and Katz and of Ainsworth and Bingham to do the same when they compared an element with its error. Whether or not we can actually do that using the newer method of Wallace and Katz depends on a clarification of the precise transformations one uses in their methods. Factor analysis as used here does not use a statistical test, but once procedures for making maximum likelihood estimates become readily available, a statistical test will be used. The criterion used here for the number of factors is the Kaiser rule of eigenvalue unity. This criterion has undergone extensive comparisons with statistical tests and other criteria with favorable results. Accordingly, because of the coincidences between results using this criterion and statistical testing, it has emerged as one of the most popular criteria of factor analysis.

Our preference for factor analysis insofar as we wish to determine the number of factors stems from the fact that its foundation and methodology have undergone extensive scrutiny over the last sixty-five years [11, 13, 14]. This, however, is not to say that we will be unable to achieve the same results in enumerating interacting species with matrix rank analysis as with factor analysis. It is probable that the method will be made more rigorous and statistical tests will be found so that several quantities are simultaneously compared.

Acknowledgement

This work was supported by National Science Foundation Grant GB28223-A*-1, and in part by the Membrane Biophysics Program, University of Florida.

References

[1] R.M. Wallace, J. Phys. Chem. 64 (1960) 899.
[2] R.M. Wallace and S.M. Katz, J. Phys. Chem. 68 (1964) 3890.
[3] D.W. McMullen, S.R. Jaskunas and I. Tinoco, Biopolymers 5 (1967) 589.
[4] S. Ainsworth, J. Phys. Chem. 65 (1961) 1968.
[5] G. Weber, Nature (London) 190 (1961) 27.
[6] S. Ainsworth and W.S.W. Bingham, Biochim. Biophys. Acta 160 (1968) 10.
[7] G.K. Ackers, J. Biol. Chem. 243 (1968) 2056.
[8] W. Godschalk, Biochemistry 10 (1971) 3284.
[9] M.E. Magar, Biopolymers (1972) in press.
[10] M.S. Bartlett, J. Roy. Stat. Soc. B16 (1954) 296.
[11] D.N. Lawley and A.E. Maxwell, In: Factor analysis as a statistical method (American Elsevier, New York, 1971).
[12] M.G. Kendall and A. Stuart, In: The advanced theory of statistics, Vol. 3 (Hafner, London, 1968).
[13] H.A. Harman, In: Modern factor analysis (University of Chicago Press, Chicago, 1967).
[14] R.J. Rummel, In: Applied factor analysis (Northw. Univ. Press, Chicago, 1970).
[15] D.N. Lawley, Proc. Roy. Soc., Edin. 60 (1940) 64.
[16] C.R. Rao, Psych. 20 (1955) 93.
[17] H.F. Kaiser, Ed. Psych. Measurement 20 (1960) 141.
[18] H.F. Kaiser, J. Stam. 12 (1964) 238.
[19] C.N. Reilley and E.M. Smith, Anal. Chem. 32 (1960) 1233.

BIOPHYSICAL CHEMISTRY 1 (1973) 24–34. NORTH-HOLLAND PUBLISHING COMPANY

HELIX-COIL EQUILIBRIA OF POLY (rA) · POLY (rU)

R.D. BLAKE

*Department of Biochemical Sciences, Frick Chemical Laboratory, Princeton University, Princeton, New Jersey 08540, USA**

Received 15 June 1973

Absorbance melting curves of the double-stranded (rA) · (rU) helix, made with fractionated homopolynucleotides of matched length, have been obtained over a 15-fold range of $[Na^+]$ and 30° range of temperature. An excellent fit of the observed profiles was obtained with theoretical curves calculated on the basis of the simplest interpretation for the occurrence of particular equilibria [1–3]; the complete molecular partition function being evaluated by the power series method developed by Applequist [4–6]. The stability constant was evaluated from literature values for the calorimetric enthalpy. The loop closure exponent was best represented by 2.22 ± 0.04 for the mismatching loop mode of melting and 1.22 for the matching mode and was independent of $[Na^+]$ and temperature. Assuming the applicability of the nonintersecting random walk value of 1.9 ± 0.1, these results would suggest a slight bias toward matched loop formation during melting of homopolynucleotides that might be expected to form only mismatched loops. The value of the stacking parameter at 60°C was only ~6% higher than that at 30°C, 0.0221 (0.0184 for the matching case). Calculated melting curves indicate the occurrence of a fifth-order phase transition when the mean helix length is only ~13 base-pairs, or about one full turn of the helix.

1. Introduction

In order to gain a more detailed appreciation of the molecular dynamics of DNA, it is desirable to know the equilibrium concentrations of individual base-pairs in the various conformational states as a function of temperature. Such information may provide some insight into the molecular mechanisms by which information encoded in DNA or double-stranded viral RNA is expressed at temperatures far below their transition temperature.

Difficulties associated with fulfilling assumptions of the statistical thermodynamical analysis in melting studies on DNA or viral RNA's are compounded by the presence of two different base-pairs: A · T(U) and G · C. Not only do they have different H-bounding stabilities, but they probably exhibit differences in stacking energy with each of the 14 possible nearest neighbors. To avoid some of these complexities, we have made a detailed study of the melting behavior of poly (rA) · poly (rU) ((rA) · (rU)) composed of two simple homopolymers.

* Present address: Department of Biochemistry, Hitchner Hall, University of Maine, Orono, Maine 04473, USA.

The average equilibrium concentration of a particular intermediate in the course of polynucleotide melting is determined in terms of base-pair 'stability' and loop 'interruption' constants [1–7] for each of the N residues in a chain. The sum over their statistical weights for all states in the chain then specifies the complete molecular partition function Z_N. Various methods have been developed to evaluate the molecular partition function; that found most convenient for extracting the various parameters from our experimental data is due to Applequist's application [4–6] of the Lifson–Zimm power series method [8, 9]. This method states that in the limit of very long chains $Z_N^{-1/N}$ approaches ρ, the radius of convergence of the sum $\Sigma_{N=1}^{\infty} Z_N \rho^N$; which occurs in the same interval as the sum $\Sigma_{i=1}^{\infty} [sG(\rho)]^i$; where for the case in which loops that occur in complementary chains are restricted to equal numbers of residues ('matched'):

$$G(\rho) \equiv \rho + \sum_{j=1}^{\infty} b_1 [j+1]^{-a_1} \rho^{j+1} \qquad (1)$$

$$= 1/s = \rho(1 - b_1) + b_1 \, \mathrm{Li}_{a_1}(\rho) \text{ (matching case).}$$

$\mathrm{Li}_{a_1}(\rho)$ is the polylogarithm of order a_1 [6, 10], b_1 is the "stacking parameter," j is the number of non-bonded residue pairs in a disordered loop segment, s is the stability constant, and a is the 'loop closure exponent'. Z_N is related to an experimentally measurable quantity f, the fraction of bases that are paired through the relationship (e.g., refs. [5, 8]):

$$f=(1/N)(\partial \ln Z_N/\partial \ln s)_{b_1} = -(\partial \ln \rho/\partial \ln s)_{b_1}. \quad (2)$$

Eqs. (1) and (2) then lead to the following expression for the melting of the theoretical model in the matching mode [6]:

$$f=[s\rho(1-b_1)+b_1 s\,\mathrm{Li}_{a_1-1}(\rho)]^{-1} \text{ (matching case)}. \quad (3)$$

The melting behavior of polynucleotide helices is thus associated with the properties of a polylogarithm function. In the calculation of theoretical transition curves, ρ is eliminated as an explicit variable through eq. (1). Similarly, for melting with mismatched loops:

$$f=\{s\rho(1-2^{-a_2}b_2) + 1/2\, b_2 s\,[\mathrm{Li}_{a_2-2}(\rho^{1/2})-\mathrm{Li}_{a_2-1}(\rho^{1/2})]\}^{-1} \quad (4)$$

(mismatching case).

Variables subscribed by 1 or 2 specify the variable as being applicable to the matching or mismatching mode of loop formation, respectively. Expressions (3) and (4) serve in the fit to the experimental melting profiles for (rA) · (rU).

A tractable analysis of helix-coil equilibria requires the assumption that the stability constant be independent of the length of the helical segment. Consequently, forces that extend beyond the nearest neighbor boundaries of a given base-pair are ignored. Significant long-range electrostatic forces may exist, however [11, 12]; such a possibility would be assessable from an examination of the melting profile over a wide range of concentrations of counter-ions that screen the coulombic forces between neighboring phosphates. Consequently, the melting properties of (rA) · (rU) were investigated over a 15-fold $[Na^+]$, the maximum practical range.

Due to the potential for total binding or sliding degeneracy, homopolynucleotides such as (rA) · (rU) would be expected to form 'mismatched loops', that is, loops of equal probability with regard to size in either chain. Applequist has suggested [6] that long-range electrostatic forces of repulsion between the negative charges on opposing chains might lead to a bias toward the formation of matched loops during the melting of homopolynecleotide helices. A bias toward matched loops could result in an intermediate situation for which a theoretical values for the loop closure exponent [13–16] would not be applicable. For this reason the exponent shall be regarded as an adjustable parameter. Since it is to be determined over a wide range of $[Na^+]$, the effects of charge screening on a possible bias toward matched loops can also be evaluated.

2. Materials and methods

2.1. Polynucleotides

Poly (rA) and poly (rU) were synthesized with polynucleotide phorphorylase [17] by J.R. Fresco and purified and fractionated as previously described [18]. The poly (rU) had an $S_{20,w}$ of 7.6 corresponding to a residue length $N \approx 1000$ (mol. wt. ≈ 340,000) [19, 20]. The UV sedimentation boundary of the sample was extremely sharp; the distribution indicated that >90% of the polymer chains fell within 5% of the average length, but was skewed to the long side. The poly (rA) had a sedimentation coefficient indicating a chain length the same as that for the poly (rU) [21, 22].

2.2. Preparation of the poly (rA) · (rU) complex

The (rA) · (rU) complex was prepared according to Blake and Fresco [18]. There was no evidence of any three-stranded (rA) · 2 (rU); it is estimated that as little as 3% (rA) · 2 (rU) could be detected. In the range of ionic strengths of the experiments below (rA) · 2 (rU) would always melt at a lower temperature than (rA) · (rU). Nucleotide residue concentrations were determined from the molar extinction coefficients [23].

2.3. Solvents

The solvent contained 0.0010 M cacodylate (Na^+),

pH 7.1, plus NaCl to give the desired $[Na^+]$. The $[Na^+]$ was determined within ± 0.5% of the nominal value by conductivity.

2.4. Melting curves

The melting of (rA) · (rU) at 6.5×10^{-5} moles total nucleotide concentration was monitored by the loss of hypochromicity at 259 mμ. The fraction of base-pairs remaining at temperature T, contiguous to $\langle \ell_B \rangle$ stacked nearest neighbors, is calculated from the expression:

$$f^*(T) = [A_u - A(T)]/[A_u - A_\ell] \ , \qquad (5)$$

where A_u is the upper and A_ℓ the lower limiting absorbancies, linearly extrapolated from short (~3°) temperature regions immediately above and below the transition region respectively, as illustrated in the example of fig. 1. It is assumed that the fractional loss of hypochromicity originates only from nearest neighbor interactions and is therefore related to the true fraction of base-pairs remaining $f(T)$ according to the expression (e.g., ref. [5]):

$$f = f^* \langle \ell_B \rangle / (\langle \ell_B \rangle - 1) \ . \qquad (6)$$

On the basis of results below, it was found that the difference between f and f^* never becomes truly significant, that is $\gtrsim 1\%$, before melting is finished.

Melting was conducted in a Gilford Model 2000 spectrophotometer equipped with an automatic sample changer and nulling device. Temperature was increased linearly at a rate of 3.2°/hr. with a synchronous motor attached to the thermoregulator. This rate is ~17°/hr. less than that at which a deviation between thermal and chemical equilibrium was found to become significant. The temperature is estimated to be within ± 0.2° of absolute.

Absorption data were recorded at 0.02° intervals, smoothed (not averaged) over 0.1° intervals and corrected for thermal volume change using Kell's empirical expression for the temperature dependence of the

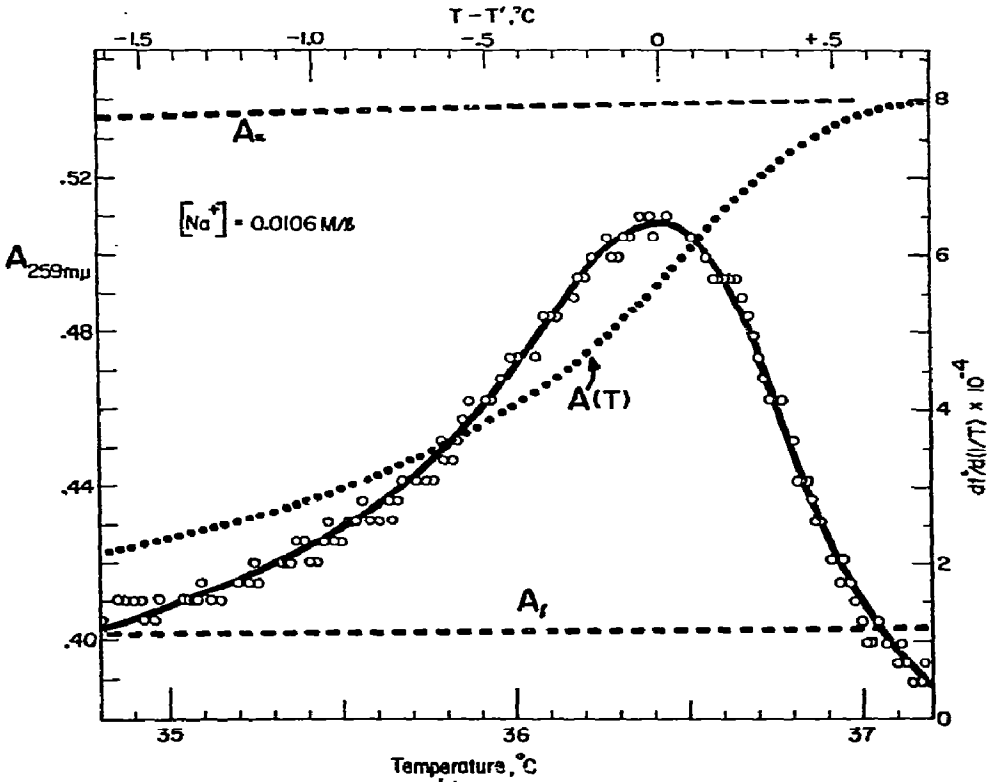


Fig. 1. Melting curve for (rA) · (rU) obtained in 0.10 M-Na^+/l. The sigmoidal (dotted) curve illustrates the analog output of the spectrophotometer. The line through open circles represents the derivative of the fraction of base-pairs remaining with the absolute temperature, $\delta f^*/\delta(1/T)$. The upper ordinate is defined by the temperature at which the line bisecting the derivative profile intersects with it at $[\delta f^*/\delta(1/T)]'$.

specific volume of water [24]. Processing was on a Hewlett-Packard 2116B mini-computer.

2.5. Theoretical transitions

The calculation of theoretical melting transitions based on eqs. (3) and/or (4) were carried out on an IBM 360/91. The program that evaluates polylogarithm, zeta and gamma functions was kindly supplied by J. Applequist, whose 1969 paper [6] gives full details of all computations.

3. Results

3.1. Experimental melting curves

The sigmoidal curve in fig. 1 represents a direct tracing of the melting transition. Results are corrected for volume change and sloping boundaries in the computation of the fraction f^* of base-pairs remaining according to eq. (5). The slope of the upper boundary is caused by the continued break-up of some residual stacking in the single-stranded (rA) chain after being released from helix. The lower plateau is virtually flat, i.e., $5(\pm 1) \times 10^{-4} A_{259}/^\circ C$.

It was found more convenient to work with the derivative profile $[\delta f^*/\delta(1/T)]$, for several reasons. A much greater sensitivity to subtle changes in melting behavior is exhibited; and a convenient reference temperature T' is presented, corresponding to a maximum in $[\delta f^*/\delta(1/T)]$ or inflection in f^*. Though a higher order derivative would be more sensitive, the demands on the quality of the data became unusually severe.

The derivative in f^* is taken over $(1/T)$, $^\circ K^{-1}$ so that curves obtained at different temperatures can be compared directly. Operationally T' represents the temperature at which a line bisecting the derivative curve intersects with it at $[\delta f^*/\delta(1/T)]'$. (The primed variable everywhere specifies its value at the inflection point in the melting curve.)

Some of the energy needed to dissociate the helix preexists in the mutual repulsion of anionic charges on the backbone phosphates. The addition of Na^+ counterions reduces this energy, resulting in a net increase in thermal energy needed to dissociate the helix. Extensive investigation has shown that T' (Celsius scale) increases linearly with the logarithm of the $[Na^+]$; for (rA) · (rU) [23, 25–28]):

$$T' = 19.8 \log [Na^+] + 76.9^\circ C . \quad (7)$$

For these studies (rA) · (rU) was melted in the presence of various $[Na^+]$ within the range 0.008 – 0.12 M-Na^+, corresponding to $T' \sim 30$–$60^\circ C$. These boundaries represent practical limits of investigation. Above 0.15 M-Na^+ ($T' = 61^\circ C$) (rA) · (rU) does not melt directly, but rearranges to the (rA) · 2(rU) helix [23, 26]; while below 0.005 M-Na^+ the transition becomes particularly sensitive to low concentrations of contaminating multivalent cations whose effect on T' is very much greater. (Melting was always conducted in 10^{-3} M-NaEDTA.)

With T' as a temperature of reference, an internal $(T-T')$ scale was established for each experiment. That for the derivative profile in fig. 1 is given along the upper horizontal scale. It was found that the fraction of base-pairs remaining at the temperature of inflection f', showed *no* dependence on $[Na^+]$ or temperature. Thus, from a least squares plot, $f' = 0.25 \pm 0.02$ at $T' = 30^\circ C$ (5 mM-Na^+) and the same at $60^\circ C$ (0.15 M-Na^+) (cf. table 1). Since T' occurs at constant f', the $(T-T')$ scale serves as a convenient basis for comparing profiles of different experiments obtained at different $[Na^+]$ and temperature. Fig. 2 illustrates how three values for $[\delta f^*/\delta(1/T)]$ at the same $(T-T')$ temperature are related by $[Na^+]$ and temperature. Results of a point-by-point comparison at 0.10° intervals across the entire profile are summarized in fig. 3, where the open circles represent values taken near 30°C from the best straight line through the full range of data such as that in fig. 2. The diameter of each open circle reflects the standard deviation at different temperatures from least squares lines such as those in fig. 2.

A derivative melting curve thus averaged at $T' = 30^\circ C$ (fig. 3) clearly reflects the characteristic asymmetric melting profile with a value for $[\delta f^*/\delta(1/T)]'$ of $6.29 (\pm 0.11) \times 10^4$, and that ~90% of the transition occurs within two degrees. The shape of the profile is the same, whether obtained by an increasing or decreasing temperature through the transition region.

3.2. Evaluation of the stability constant, s

Having the properties of an equilibrium constant, s is determined from the standard thermodynamic relationship:

$$s = \exp\left[\frac{\Delta H}{R}\left(\frac{1}{T_s} - \frac{1}{T}\right)\right], \quad (8)$$

Table 1
Thermodynamic parameters for the helix-coil equilibria of (rA) · (rU)

							Matching		Mismatching	
T'(°C)	$f^{*'}$	$[\delta f^*/\delta(1/T)]'/10^4$	T_t(°C)	$\langle \ell_B \rangle_t$ (base pairs)	ΔH[a] (cal/mol)	ΔS[b] (e.u.)	a_1	b_1*	a_2	b_2*
30.00	0.254 ± 0.18	6.29 ± 0.11	30.84	13.1	−6607	−21.79	1.22	0.0184	2.22	0.0221
60.00	0.262 ± 0.20	7.43 ± 0.15	60.84	12.2	−8397	−25.22	1.22	0.0197	2.22	0.0236

a) $\Delta H = -59.8\ T'$ (°C) − 4813 cal/mol
b) $\Delta S = \Delta H/[T'(°K) - 1.32]$

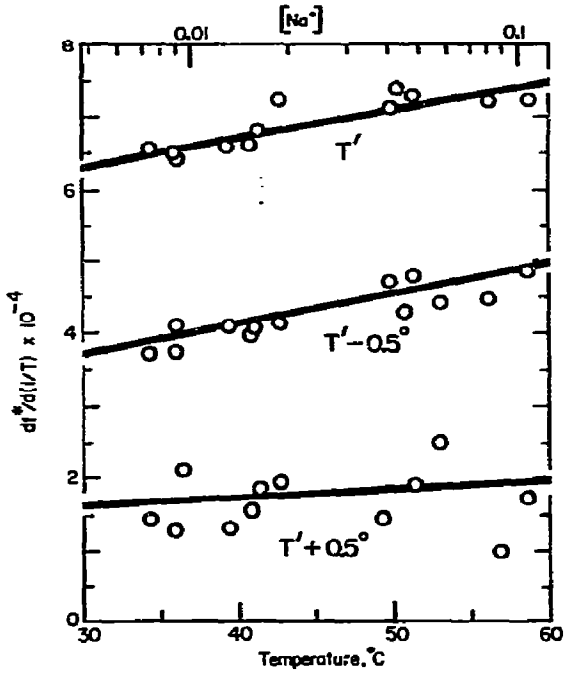


Fig. 2. Representative variations in the value observed for $\delta f^*/\delta(1/T)$ at constant $T'-T$ with temperature. The upper scale shows the relationship to the secondary variable, $[Na^+]$, as given by eq. (7).

where ΔH is the enthalpy for formation of a mole of base-pairs and where T_s, the temperature at which $s = 1$, is given by the relationship: $T_s = T'$ (°K) − 1.32°. This slight difference between T' and T_s is due to the effect of loops on the overall equilibrium between bonded and non-bonded bases, as found below. Calorimetric determinations of ΔH are available from the work of Rawitscher et al. [29], Neumann and Ackermann [30] and Krakauer and Sturtevant [28]. Those from the latter two groups were obtained from measurements of the dependence on temperature of the release of excess heat capacity on passing through the region of the helix-coil transition. The profiles that result resemble the derivative melting curves of figs. 1 and 3; though from the one illustrated example, the melting of (rA) · (rU) in the calorimeter occurred over 2 1/2 times the breadth in temperature of that in fig. 1, achieving the same change in state. (The likelihood is that this greater breadth is due to chain length heterogeneity.) In any case, the temperature of maximum release of excess heat will not vary greatly from T' of the optical results above. Fig. 4 shows a plot of the integrated calorimetric enthalpies obtained as a function of T'. As with the optical melting curves herein, this latter dependence was achieved by varying $[Na^+]$ over the range indicated by the upper horizontal scale in fig. 4 (cf. fig. 2). A linear relationship exists between T' and $[Na^+]$, as indicated by eq. (7), as well as with the electrostatic repulsive component of the free energy per base-pair [9, 27]. Therefore, the dependence of ΔH for base-pair *formation* on T' must likewise be linear, and is defined on the Celsius scale by the relationship:

$$\Delta H = -59.8\ T'\ (°C) - 4813 \text{ calories} \qquad (9)$$

from a least squares analysis of the results in fig. 4.

Since $\Delta G = 0$ at $T' - 1.32°$, eq. 9 serves in the determination of the unitary conformational entropy per (rA) · (rU) base-pair according to the relationship:

$$\Delta S = \Delta H/[T'\ (°K) - 1.32]. \qquad (10)$$

The dependence of ΔS on T' is indicated in fig. 4 by

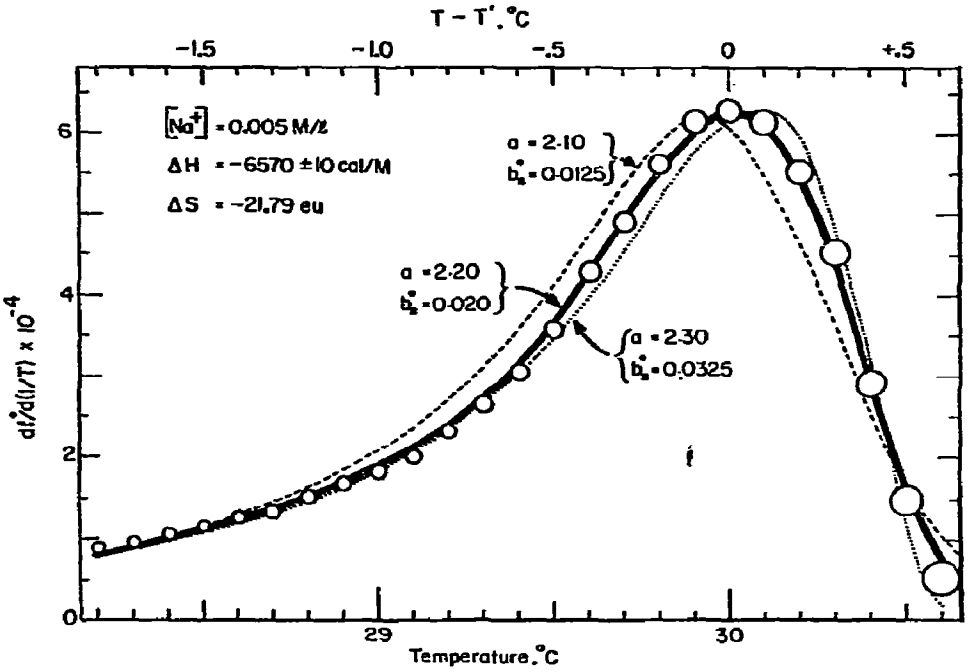


Fig. 3. Observed and theoretical derivative melting curves of (rA) · (rU). The variable diameter of the open circles represents the standard deviation from the indicated value at 0.005 M-Na^+/l and T' = 30°C for $\delta f^*/\delta(1/T)$ taken from linear least squares lines through experimentally observed results obtained over a 30° range of temperature, as illustrated in fig. 3. The lines represent the variation of theoretical $df^*/d(1/T)$ with temperature with three trial values for the loop closure exponent and stacking parameter for which $[df^*/d(1/T)']/10^4 = 6.29$, $\Delta H = -6570 \pm 10$ cal/mol, and $\Delta S = -21.79$ e.u.

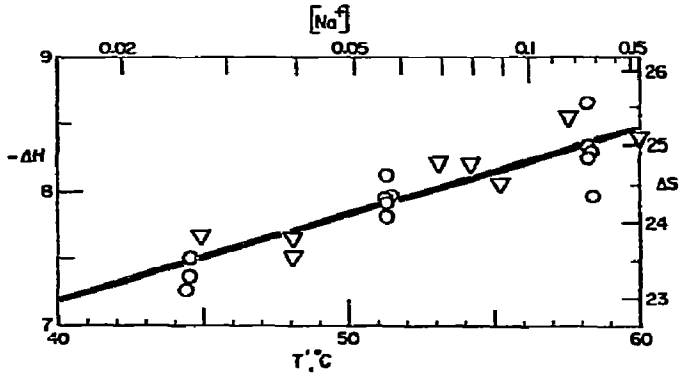


Fig. 4. Variation of calorimetric enthalpy (kcals) and entropy (e.u.) on association of one mole of (rA) · (rU) base-pairs with temperature and $[Na^+]$. Data from Krakauer and Sturtevant (○); from Neumann and Ackerman (▽). The least squares line is given by eq. (9).

the right-hand-side vertical scale. For the most part, this variation reflects the release of an increasingly less structured poly (rA) component with increasing T', as does much of the change in ΔH with T' (e.g., ref. [27]). Additional contributions to an apparent change in heat capacity of −60 cal per mole-degree (eq. 9) are due to the increase in interphosphate charge screening with increasing $[Na^+]$, the decrease in bulk dielectric constant with temperature (both effects anticipated in eq. 7), 'solvophobic effects', etc.

3.3. Evaluation of the loop closure exponent a and stacking parameter b

The variation in f^* with temperature for the theoretical model is calculated from eqs. (1), (8) and (4) [or (3), depending on the loop mode] where [5, 6]:

$$f^* = fs\rho = 1/G'(\rho)\,. \tag{11}$$

The derivative of f^* is taken over s:

$$(\partial f^*/\partial \ln s)_{b^*} = f^{*3}\,\rho\,G''(\rho)\,, \tag{12}$$

where $b^* = bs$. From the van 't Hoff equation:

$$\partial f^*/\partial(1/T) = -(\partial f^*/\partial \ln s)\,\Delta H/R\,, \tag{13}$$

which is required for direct comparison of observed

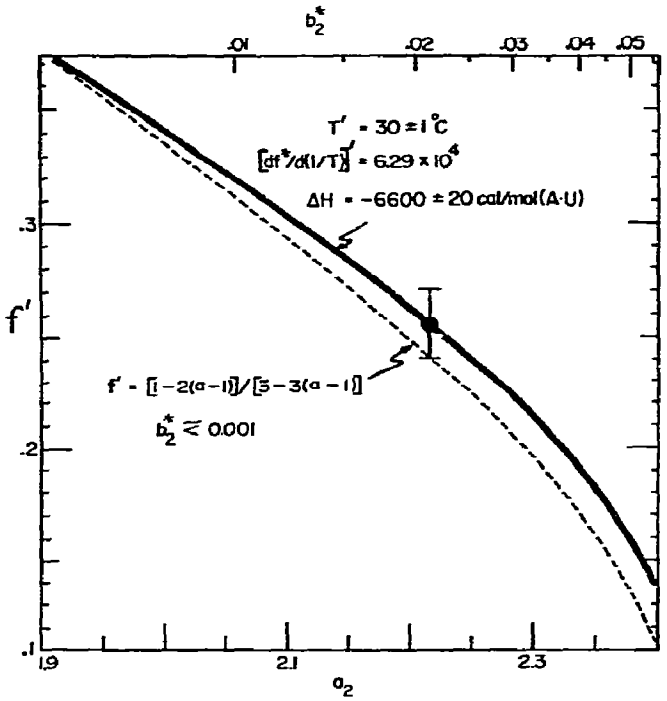


Fig. 5. Variation in the theoretical fraction of base-pairs remaining at the temperature of the melting curve inflection point with loop closure exponent. The dashed line represents a plot of eq. (14), and is independent of b_2^* only when the latter is $\lesssim 0.005$. For $b_2^* > 0.005$, the dependence of a_2 on f' was determined from eq. (11) where $[df^*/d(1/T)]'/10^4 = 6.29$, $\Delta H = -6600 \pm 20$ cal/mol, $T' = 30 \pm 1$°C, and the value of b_2^* given by the upper scale. The point represents the experimental value for $f^{*\prime}$ between 30–60°C for (rA) · (rU).

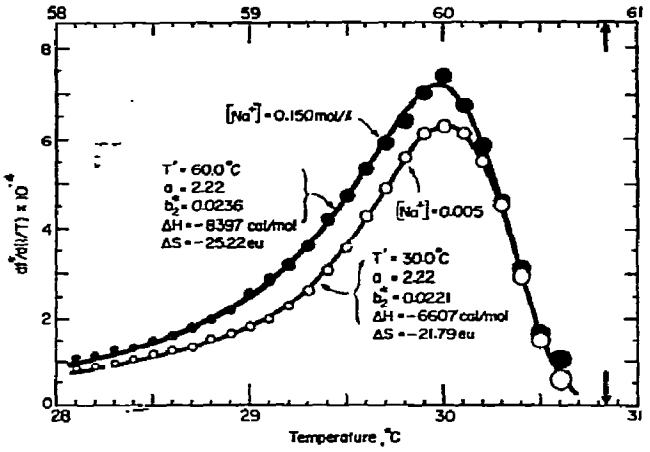


Fig. 6. The best fit of theoretical to observed derivative melting curves of (rA) · (rU). Results at 0.005 M-Na^+/l and $T' = 30$°C for $\delta f^*/\delta(1/T)$ taken from the least squares lines through data at constant $T'-T$ over a range of temperature (cf. fig. 3) (○); results at 0.150 M-Na^+/l and $T' = 60°$ similarly obtained (●). Circle size defines the standard deviation. The lines represent theoretical results calculated with values for the required parameters as given in the fig. Heavy arrows indicate phase transition temperatures.

and theoretical curves. The remaining adjustable parameters a and b, are evaluated by fitting to the experimental curves.

Applequist [6] has shown that when b is sufficiently small, its variation is accompanied by virtually no change in the value of a. On this basis he proposed a convenient means for approximating the latter. The sole requirement is an experimental value for the fraction of base-pairs remaining at the temperature of inflection:

$$a_2 = 2 + [(1 - 3f')/(2 - 3f')]\,. \qquad (14)$$

A plot of this function (14) is shown in fig. 5, represented by the dashed line; and indicates a value of $a_2 \approx 2.2$ when $f' \approx f^{*\prime} = 0.25$ (cf. table 1). With this approximate value for a_2, some trial derivative melting curves were calculated for which $[\partial f^*/\partial(1/T)]'/10^4 = 6.29$; shown in fig. 3. That for $a_2 = 2.2$, represented by the solid line, clearly exhibits the closest fit to the observed results over most of the transition. An identical family of curves is obtained for the matching case when:

$$a_1 = a_2 - 1 \qquad (15)$$

and

$$b_1^* = b_2^*[\zeta(a_2 - 1) - \zeta(a_2) - 2^{-a_2}]/[\zeta(a_2 - 1) - 1]\,, \qquad (16)$$

where $\zeta(x)$ is the Riemann zeta function.

Expression (14) seems to hold well only for $b_2 \lesssim 0.005$. The value for a_2 was refined from the dependence of $f'(\approx f^{*\prime})$ on a_2, given by the solid line in fig. 5. This line was computed from eq. (11) for different values of $b_2^* > 0.005$ (upper horizontal scale), when $[\partial f^*/\partial(1/T)]'/10^4 = 6.29$. The placement of the experimentally observed $f^{*\prime}$ on the line indicates a value for $a_2 = 2.22$ and $b_2^* = 0.022$; which actually led to the best fit of theoretical to observed transition curves (cf. table 1), illustrated in fig. 6. The upper half of fig. 7 then shows the dependence on temperature of the fraction of base-pair remaining through the transition region. Both observed and theoretical curves are represented by the solid line, since they are virtually indistinguishable. For reference, the dependence on temperature of the stability constant is

also given (upper horizontal scale). Note that the temperature at which $s = 1$, T_s is lower than T' by 1.32°; while the temperature at which $f^* = 0.5$, (T_m) is lower than T' by 0.37°.

By a similar routine a best fit was obtained at $T' = 60$°C, shown in fig. 6. Values for the various parameters are again given in table 1; the stacking parameter is increased by only 6% of that at 30°, but otherwise there is *no* change in a.

Thus, a region of the helix-coil transition most sensitive to subtle changes in melting parameters has been examined over a moderately wide range of $[Na^+]$ and temperature, with no significant effect of these variables. It could be that the mean helix length $\langle \ell_B \rangle$ does not become sufficiently short over this particular region of the transition to reveal any depency on length of the stability constant, as might be anticipated from the effects of electrostatic forces. Similarly, the mean loop size $\langle \ell_A \rangle$ may be so large that matched-loop biasing forces no longer exert a very significant influence.

3.4. Evaluation of mean sequence lengths

The mean helical segment length is given by [5–7]:

$$\langle \ell_B \rangle = 1/(1 - s\rho) . \qquad (17)$$

The dependence of $\langle \ell_B \rangle$ on temperature, calculated with parameters given in table 1 (irrespective of the loop mode), is plotted in the lower half of fig. 7 (on a logarithmic scale for convenience). It can be seen that by the temperature at the inflection point, $\langle \ell_B \rangle$ has dropped to an average of 20 contiguous base-pairs per helical segment, which is still a considerable length. The mean loop size, calculated as a function of temperature from [6]:

$$\langle \ell_A \rangle = \langle \ell_B \rangle (1 - f)/f , \qquad (18)$$

is also plotted in fig. 7. $\langle \ell_A \rangle$ reaches 50 residue pairs by T'. This value seems rather large to experience a significant coulombic bias and is parhaps the reason why an ionic strength effect was not observed.

Fig. 7 shows very different rates of change of loop and helix sizes. While the average helical segment length is dropping from 50 to 17 base-pairs between 27° and 30.5°, the average loop size is rising from 5 to 800 residue pairs. Applequist [6] has found this behavior to characterize melting when $2.1 \lesssim a_2 \lesssim 3$.

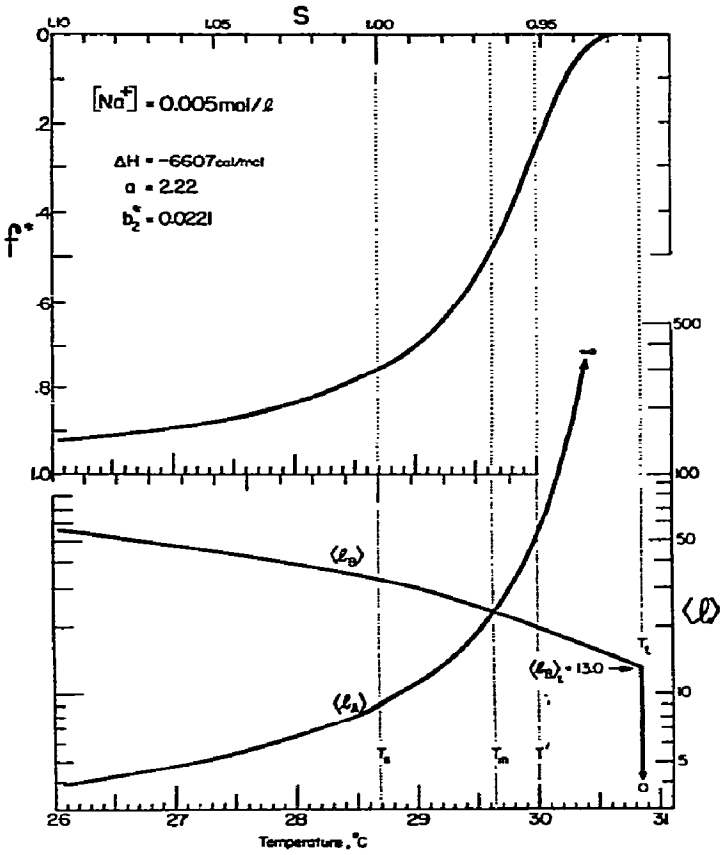


Fig. 7. The curve in the upper half represents the variation in both observed and theoretical fraction of (rA) · (rU) base-pairs remaining with temperature; the latter being calculated with the same values for parameters as used in fig. 6, viz., $a_2 = 2.22$, $b_2^* = 0.0221$, $\Delta H = -6607$ cal/mol and $\Delta S = -21.79$ e.u. Curves in the lower half represent the theoretical variation in mean helix length, $\langle \ell_B \rangle$ (base pairs), and mean loop size, $\langle \ell_A \rangle$ (residue pairs), with temperature, calculated from eqs. (19) and (20) respectively. ($\langle \ell \rangle$ is plotted on a logarithmic scale only for convenience.)

Since loops are increasing in size at a much greater rate than helix lengths are decreasing, they must be doing so at the expense of *numbers* of helical segments, implying that helical segments have some critical minimum length.

3.5. Phase transition

The requirement for a true phase transition is a discontinuity in a derivative of the molecular partition function with temperature. The unusual structural feature of nucleic acid helices that leads to this condition is the occurrence of long-range interactions, par-

ticularly within the disordered loop regions of the chain. Large closed loops, as appear late in melting, constitute weak, long-range entropy effects. It was shown by Poland and Scheraga [31] and by Applequist [4] that in the limit of infinite chain length such interactions give rise to a discontinuity in the function $(\partial^m \ln \rho / \partial \ln s^m)$ (cf. eq. (2)), where m, the order of the phase transition, is a function of the loop closure exponent a.

Applequist [6] has demonstrated that if $a_1 > 1$, a solution for eq. (1) exists only for $s > s_t$, where:

$$s_t = 1/[1 - b_1 + b_1 \zeta(a_1)] \text{ (matching case)} \qquad (19)$$

fulfilling the condition for a phase transition. The order of the transition is then determined from the precise increase of a_1 above unity:

$$1 + \frac{1}{m} < a_1 < 1 + \frac{1}{m-1}.$$

The temperature at which the phase transition occurs, T_t, is obtained by setting $s = s_t$ in eq. (8) and solving for T_t. The fraction of base-pairs remaining at T_t is determined from eq. (3) by setting $\rho = 1$ and $s = s_t$.

Likewise, if $a_2 > 2$, then:

$$s_t = 1/(1 + b_2[\zeta(a_2 - 1) - \zeta(a_2) - 2^{-a_2}]) \qquad (20)$$

(mismatching case).

The order for the mismatching model is determined from:

$$2 + \frac{1}{m} < a_2 < 2 + \frac{1}{m-1}.$$

On the basis of these conditions (rA) · (rU) would appear to undergo a *fifth*-order transition (irrespective of loop mode) at $T_t = 30.84°C$ and $60.84°C$ (table 1) for the lower and higher temperature curves in fig. 6, respectively (indicated by the heavy arrows). The mean helix length at the phase transition temperature $\langle \ell_B \rangle_t$, would be ~13 base-pairs (cf. table 1 and fig. 7) from eq. (17). It is to be noted, however, that such an analysis does not provide a direct experimental determination of the order of the transition.

4. Discussion

4.1. Applicability of the theoretical model for melting

The obviously good fit of theoretical to experimental melting curves over a large part of the transition would seem to give considerable support to Zimm's *simplex munditiis* model [1]. Many qualifications to this model have been proposed, but their relevance appears to be restricted to regions of the melting transition outside the practical temperature boundaries of the analysis herein. Of course, a fit of one semiempirical and two adjustable theoretical parameters to experimental data is a necessary but not sufficient criterion to test the applicability of the basic model.

There is a hierarchy in the precision with which various parameters are determined. Assuming its invariance with helix length, s is presumed known to approximately ± 2% from calorimetric measurements of the enthalpy. The foregoing assumption has, however, been challenged recently on the grounds that at low ionic strength 'long-range' electrostatic forces make a significant contribution to the net energy per base-pair. Such forces are likely responsible for the greater relative stability of the shorter oligomer helices [3, 11]. Such helices have also been demonstrated to exhibit an increase in the dependence of $(\partial T_m / \partial [M^+])_{C,N}$ with increasing helix length, asymptotically approaching that for the polynucleotide. On the basis of such results Elson et al. [11] proposed an 'effective' stability constant:

$$s^*(N) = s \cdot \exp[\Delta H - \Delta H^*(N)]/RT, \qquad (21)$$

where $[\Delta H - \Delta H^*(N)]$ represents the difference in total energy between the infinite polymer and oligomer helix, respectively, averaged over the chain length. The magnitude of $[\Delta H - \Delta H^*(N)]$ increases with decreasing oligomer length due to the truncation of disruptive 'long' range coulombic forces. (By a similar argument, the penultimate base-pairs of the finite *poly*nucleotide helix would be expected to be more stable.) However, the quantitative relevance of an 'effective' stability constant computed on the basis of results from short helices to the short helical segments of a partially melted polynucleotide is questionable. The proximity of charge density from disordered loop segments in the partially melted polynucleotide cannot be totally ignored in the summation of net electrostatic energy for a given base-pair. For example, in helices with extrahelical non-complementary residues there is a contrasting *decrease* in $(\partial T_m / \partial [M^+])$ with increasing mean

helical segment length between loops of noncomplementary residues asymptotically approaching that for the homo · homopolynucleotide [32]. Since the coordinates for all pairwise coulombic interactions between phosphates in discretely localized bonded and statistically localized nonbonded segments of the chain cannot be defined in any simple geometric fashion, a more pertinent $s^*(N)$ cannot be readily computed. In any case, it does not seem necessary, since the average helical length in the critical inflection region of the melting curve is > 20 base-pairs, and certainly not less than 13 (the critical minimum length). The former helix length would seem to be beyond the range of *significant* length-dependent sensitivity to intramolecular coulombic forces. Moreover, the analysis above was quite insensitive to a 15-fold increase in $[Na^+]$ that presumably provides a greater screening of the coulombic forces of repulsion between PO_2^- groups on opposing backbone chains. It is concluded, therefore, that a value for s independent of helical segment length serves as a reasonable approximation for analyses of helix-coil equilibria of polynucleotides.

4.2. Loop formation

The interpretation of the value measured for the loop closure exponent depends upon the precise mode of loop formation in these homopolynucleotides, which is unknown. On the assumption that loops on opposing chains occur in mismatched fashion, the value observed, 2.2, is significantly greater than that calculated, 1.9 ± 0.1, from theoretical random walk studies for non-intersecting chains [13–16]. This would suggest the occurrence of forces leading to a bias toward matched loop formation. Alternatively, it could mean the inapplicability of the random walk calculation. Recent theoretical and experimental results suggest that the weighting function for loop closure somewhat underestimates the true value for small loops, from which the latter derive a greater relative stability [3, 33, 34]. A dependence of the loop weighting function on loop size was not detected in this analysis due to the intensitivity of theoretical melting curves to rather widely different values of a at $T' - T \gtrsim 1.5°$ (cf. fig. 3), that is, when $\langle \ell_A \rangle \lesssim 10$ residue pairs. Perhaps for the same reason, a dependence of a upon ionic strength was not observed, since a coulombic bias probably influences a predominance of matched loops early in melting. The screening of coulombic forces between opposing phosphates is not complete within the $[Na^+]$ range examined $[Na^+] \leq 0.12$ M, so that the later growth of larger matched loops may also be unable to form in either chain with an equal probability with regard to size. This would explain the slightly high value observed for the loop closure exponent.

The value found herein for the exponent, $a_2 = 2.22$, falls in the range of empirical values from other laboratories. For (rA-H$^+$) · (rA-H$^+$): $2.5 > a_2 > 2.3$ (ref. [6]), and $a_2 = 1.8$ (Hennage and Crothers, quoted in ref. [35]); for (dA) · (dT): $1.9 > a_2 > 1.5$ and for (dA-T) · (dA-T): $1.7 > a > 1.5$ (ref. [36]). The reason for such a range of values for a is not clear, that is, it may not simply represent experimental uncertainty. It could, for example, be a reflection of the effects of differing finite chain lengths. The theoretical melting curves are calculated on the assumption that end effects are nonexistent; that the polynucleotide chains are of effectively infinite length. The extent to which experimental melting curves reflect the effects of finite length is not known at present.

Acknowledgement

I am indebted to Jon Applequist and Jacques Fresco for encouragement and many invaluable discussions, and to the Princeton University Computer Center for time on their IBM 360/91.

This investigation was supported by grants to J.R. Fresco from the National Institutes of Health (GM-07654) and the National Science Foundation (GB-18865).

References

[1] B.H. Zimm, J. Chem. Phys. 33 (1960) 1349.
[2] D.M. Crothers, Accts. Chem. Res. 2 (1969) 225.
[3] R.L. Baldwin, Accts. Chem. Res. 4 (1971) 265.
[4] J. Applequist, J. Chem. Phys. 45 (1966) 3459.
[5] J. Applequist, in: Conformation of biopolymers, Vol. 1, ed. G.M. Ramachandran (Academic Press, New York, 1967) p. 403.
[6] J. Applequist, J. Chem. Phys. 50 (1969) 600.

[7] D.M. Crothers and B.H. Zimm, J. Mol. Biol. 9 (1964) 1.
[8] S. Lifson and B.H. Zimm, Biopolymers 1 (1963) 15.
[9] S. Lifson, J. Chem. Phys. 40 (1964) 3705.
[10] L. Lewin, Dilogarithms and associated functions (MacDonald and Co., London, 1958) p. 169.
[11] E.L. Elson, I.E. Scheffler and R.L. Baldwin, J. Mol. Biol. 54 (1970) 401.
[12] C.L. Schildkraut and S. Lifson, Biopolymers 3 (1965) 195.
[13] F.T. Wall, L.A. Hiller Jr. and W.F. Atchison, J. Chem. Phys. 23 (1955) 2314.
[14] B.J. Hiley and M.F. Sykes, J. Chem. Phys. 34 (1961) 1531.
[15] C. Domb, J. Gillis and G. Wilmers, Proc. Phys. Soc. 85 (1965) 625.
[16] M.E. Fisher, J. Chem. Phys. 45 (1966) 1469.
[17] M. Grunberg-Manago, P.J. Ortiz and S. Ochoa, Science 122 (1955) 907.
[18] R.D. Blake and J.R. Fresco, J. Mol. Biol. 19 (1966) 145.
[19] E.G. Richards, C.P. Flessel and J.R. Fresco, Biopolymers 1 (1963) 431.
[20] L.D. Inners and G. Felsenfeld, J. Mol. Biol. 50 (1970) 373.
[21] J.R. Fresco and P. Doty, J. Amer. Chem. Soc. 79 (1957) 3928.
[22] H. Eisenberg and G. Felsenfeld, J. Mol. Biol. 30 (1967) 17.
[23] R.D. Blake, J. Massoulié and J.R. Fresco, J. Mol. Biol. 30 (1967) 291.
[24] G.S. Kell, J. Chem. Eng. Data 12 (1967) 66.
[25] J.R. Fresco, in: Informational macromolecules, eds. H.J. Vogel, V. Bryson and J.D. Lampen (Academic Press, N.Y., New York, 1963) p. 121.
[26] C.L. Stevens and G. Felsenfeld, Biopolymers 2 (1964) 293.
[27] J. Massoulié, Europ. J. Biochem. 3 (1968) 428.
[28] H. Krakauer and J.M. Sturtevant, Biopolymers 6 (1968) 491.
[29] M.A. Rawitscher and P.D. Sturtevant, J.Amer. Chem. Soc. 85 (1963) 1915.
[30] E. Neumann and T. Ackermann, J. Phys. Chem. 71 (1967) 2377; J. Phys. Chem. 93 (1969) 2170.
[31] D. Poland and H.A. Scheraga, J. Chem. Phys. 45 (1966) 1464.
[32] A. Lomant and J.R. Fresco, Biopolymers 12 (1973), to be published.
[33] P.J. Flory and J.A. Semlyen, J. Amer. Chem. Soc. 88 (1966) 3209.
[34] I.E. Scheffler, E.L. Elson and R.L. Baldwin, J. Mol. Biol. 58 (1970) 145.
[35] D.M. Crothers, Biopolymers 10 (1971) 2147.
[36] R.M. Wartell, Biopolymers 11 (1972) 745.

BIOPHYSICAL CHEMISTRY 1(1973) 35–45. NORTH-HOLLAND PUBLISHING COMPANY

SEDIMENTATION COEFFICIENTS OF SELF-ASSOCIATING SPECIES I. BASIC THEORY

Charlotte A. WEIRICH*, E.T. ADAMS, Jr.** and Grant H. BARLOW
Chemistry Department, Texas A & M University, College Station, Texas 77843, USA
and Molecular Biology Department, Abbott Laboratories, Inc., North Chicago, Illinois 60064, USA

Received 27 April 1973

Under the same solution conditions, the apparent weight average sedimentation coefficient, s_{wa}, and some quantities obtained from it can be combined with the equilibrium constant or constants, K_i, and the monomer concentration, c_1, obtained from sedimentation equilibrium, light scattering or osmotic pressure experiments on the same self-associating solute, so that the individual sedimentation coefficients, s_i, of the self-associating species, and also the hydrodynamic concentration dependence parameter, g or g, can be evaluated. Using two different models for the hydrodynamic concentration parameter, four different methods are presented for the evaluation of the s_i's. Methods for evaluating g or g, once the s_i's are known, are presented. A method for obtaining the number average sedimentation coefficient, s_N, and its application to self-associations is presented. Methods are shown for the evaluation of z average properties, x_{zc}, as well as number average properties, x_{Nc}, of a self-associating solute from its weight average properties, x_{wc}.

1. Introduction

Self-associations are reactions of the type

$$nP_1 \rightleftarrows P_n, \quad n = 2, 3, \ldots, \qquad (1)$$

and related self-associations. Here P represents the self-associating solute. These reactions are quite widely encountered, and they can be studied by a variety of techniques [1], including sedimentation equilibrium [2–8] and sedimentation velocity experiments [1, 9–13].

While there is great interest in self-associating solutes, it is only in the last twenty years that real progress has been made in the analysis of self-associations. Steiner [14–16] was the first to apply a thermodynamic analysis that allowed the evaluation of the equilibrium constant or constants for a variety of self-associations. His methods were restricted to ideal, dilute solutions and were developed individually for light scattering and osmotic pressure experiments. Rao and Kegeles [17] extended the Steiner analysis to the Archibald experiment. They developed some novel methods for the analysis of a monomer-n-mer self-association. The first attempt to analyze a non-ideal self-association was done by Adams and Fujita [2], whose procedure was restricted to a monomer-dimer self-association. This restriction was removed by Adams [4], who also established the interrelation between average molecular weights [4, 7]. Since then much progress has been made in the analysis of self-associations, and newer and much simpler methods have been developed for the analysis of various self-associations [5, 8].

Transport methods have also been used extensively for the analysis of self-associations [1]. Much of the pioneering work in this area was done by Gilbert and his associates [9–12]. While the Gilbert method did

* This paper is based in part on the thesis submitted by Charlotte A. Weirich to the Graduate College of Texas A & M University in partial fulfillment of the requirements for the degree of Master of Science (August 1972).

** Please address all correspondence regarding this paper to Dr. E.T. Adams, Jr., Chemistry Department, Texas A & M University, College Station, Texas 77843, U.S.A. Part of this material was presented at the 16th Annual Meeting of the Biophysical Society, Toronto, Canada, Feb. 24–27, 1972.

have some limitations, the most serious being that the continuity equation when self-associations exist could not be solved if diffusion were included, it nevertheless did stimulate much interest in the analysis of self-associations. More recently Goad and Cann [18], using sophisticated computer methods, have been able to include the effects of diffusion and solve the continuity equation when self-associations are present. Cox [19] has developed some elegant methods for simulating sedimentation velocity experiments on self-associating solutes; he has been able to show how variation in the values of several parameters (the equilibrium constant or constants, the sedimentation coefficients, the hydrodynamic concentration dependence parameter) affects the concentration gradient (schlieren) patterns. Godshalk [20] has developed a method for the evaluation of diffusion coefficients from a series of sedimentation velocity experiments at different speeds; he has proposed a matrix rank analysis method to enumerate the number of associating species present, if this be unknown beforehand. Very recently Kegeles [21] has developed a method for obtaining the z average sedimentation coefficient, s_z, for self-associating solutes; he has proposed some ways to estimate the sedimentation coefficients in a monomer-n-mer self-association.

Generally these developments in sedimentation equilibrium and sedimentation velocity experiments on self-associating solutes were done separately, and no attempt was made to combine the two experiments. The first attempt to combine sedimentation velocity data with sedimentation equilibrium data on the same self-associating solute, so that the sedimentation coefficients of the associating species and also the hydrodynamic concentration dependence parameter could be evaluated, was done by Kakiuchi and Williams [22]. This procedure was also used by Iso and Williams [23]. Unfortunately their method involved extrapolation from a region of high solute concentration for the evaluation of the desired quantities.

The purpose of this paper is to show some new ways of combining sedimentation velocity and sedimentation equilibrium data to extract as much information as possible about the self-associating system under investigation. Two models for the hydrodynamic concentration dependence parameter will be proposed; we can generate several independent equations for the evaluation of the sedimentation coefficients of the associating species. We will also show how the hydrodynamic concentration dependence parameter can be obtained from the data with either model. In addition we will show how s_{zc} can be applied to this problem. We will also show that if any weight average property of the system, x_{wc}, is available experimentally for a self-association, then it is a simple matter to obtain the z average property; here x could represent a sedimentation coefficient, a partition coefficient, a diffusion coefficient, etc. We will also show how to obtain the number average sedimentation coefficient, s_{Nc}. The methods used here can be extended to obtain other number average properties, x_{Nc}, of the associating system. Knowledge of the sedimentation coefficients of the associating species may lead to insights into the shape of the associating species, as well as how the associating species are joined into n-mers. Values of the sedimentation coefficients and the hydrodynamic concentration dependence parameter may be useful for testing various methods for simulating sedimentation velocity schlieren patterns for self-associating solutes; they could also be used to help test various theories regarding self-associations and sedimentation velocity experiments. The sedimentation coefficients could also be used for the evaluation of translational diffusion coefficients of the associating species

2. Evaluation of the weight average (s_{wc}) of apparent weight (s_{wa}) average sedimentation coefficient

In order to proceed further it is necessary to make the following assumptions.

(1) The natural logarithm of the activity coefficient, y_i, for each self-associating species can be represented by the expression

$$\ln y_i = iBM_1c, \qquad i = 1, 2, \ldots . \tag{2}$$

Here c is the total solute concentration on the gram per liter (or other weight over volume concentration) scale; for a monomer-n-mer self-association c is given by

$$c = c_1 + K_n c_1^n \tag{3}$$

as a consequence of eq. (2). The quantity BM_1 in eq. (2) is known as the second virial coefficient.

(2) Sedimentation equilibrium, light scattering or osmotic pressure experiments have been performed on the self-associating solute under the same solution conditions so that c_1, the monomer concentration, and K_n, the association equilibrium constant, are known.

(3) The third assumption is that interacting flows are absent; this is an assumption that has been made by all the workers in this field so far, [9–12, 19–23].

When a self-association is present, the weight average sedimentation coefficient, s_{wc}, is defined [13] by

$$s_{wc} = \frac{c_1 s_1 + K_n c_1^n s_n}{c_1 + K_n c_1^n} = \frac{s_1 + K_n c_1^{n-1} s_n}{1 + K_n c_1^{n-1}} \tag{4}$$

for a monomer-n-mer self-association; analogous equations can be written for other types of self-associations. The subscript c in eq. (4) indicates that a self-association is present; note that as $c \to 0$, $s_{wc} \to s_1$, whereas as $c \to \infty$, $s_{wc} \to s_n$. Studies on nonassociating solutes have established the existence of a hydrodynamic concentration dependence parameter, g; this was the reason for the use of the term sedimentation coefficient in place of sedimentation constant. Thus with nonassociating solutes one encountered the following empirical expressions to describe the observed concentration dependence of s:

$$s = s^0(1 - gc) \tag{5}$$

or

$$s = s^0/(1 + gc). \tag{6}$$

Here s^0 is the value of s at infinite dilution of the sedimenting solute.

When hydrodynamic concentration dependence is present, and this is usually the case, we will use the term s_{wa}, the apparent value of s_{wc}, in place of s_{wc}. Two models for defining s_{wa} will be used here. One model which we have proposed defines s_{wa} by the relation

$$\frac{1}{s_{wa}} = \frac{1}{s_{wc}} + \underline{g}c \qquad \text{(Model I)}. \tag{7}$$

In this model, which we will call Model I, the hydrodynamic concentration dependence parameter is designated by $\underline{g}$; generally this model is simpler to use mathematically in the treatment that follows. The other model for s_{wa}, which we will call Model II, is given by

$$s_{wa} = s_{wc}/(1 + gc) \qquad \text{(Model II)}. \tag{8}$$

Model II was first used by Gilbert [9–12] and is the model for s_{wa} that was used by Kakiuchi and Williams [22] and by Iso and Williams [23]. Here the hydrodynamic concentration dependence parameter is designated by g.

There are two ways that s_{wa} can be evaluated. With refractometric optics (Rayleigh and schlieren optics), s_{wa} can be evaluated from the rate of movement of the square root of the second moment of the boundary gradient curve, $\bar{r}$, since

$$\frac{\mathrm{d} \ln \bar{r}}{\mathrm{d}t} = \omega^2 s_{wa}. \tag{9}$$

Here ω is the angular velocity of the rotor and is given by $\omega = 2\pi(\mathrm{RPM})/60$. The quantity $\bar{r}$ is defined by the general equation [24, 25]

$$\bar{r}^2 = \left(c_m r_m^2 + \int_{r_m}^{r_p} r^2 \frac{\mathrm{d}c}{\mathrm{d}r} \, \mathrm{d}r \right) / c_p. \tag{10}$$

Here c_m is the value of the concentration of solute at the air-solution meniscus, r_m, and c_p is the concentration of solute in the plateau region, the region of constant concentration. The quantity r_p is the radial position from the center of rotation at which the plateau region begins. In order to calculate $\bar{r}$ it is necessary to know c_m and c_p. Kegeles and Rao [26] showed that c_m can be evaluated for self-associations from the equation

$$c_m = c_0 - (1/r_m^2) \int_{r_m}^{r_p} r^2 \frac{\mathrm{d}c}{\mathrm{d}r} \, \mathrm{d}r. \tag{11}$$

The concentration at the plateau, c_p, can be obtained from

$$c_p = c_m + \int_{r_m}^{r_p} \frac{\mathrm{d}c}{\mathrm{d}r} \, \mathrm{d}r. \tag{12}$$

When $c_m = 0$, then eq. (10) reverts to the form obtained by Goldberg [27].

With absorption optics one can use the continuity equation to evaluate s_{wa}. In the plateau region it is easy to show that

$$\left(\frac{\partial c}{\partial t}\right)_r = -\omega^2 c_p s_{wa}. \tag{13}$$

Thus s_{wa} can be evaluated from the slope of a plot of $\ln c_p$ vs. t.

3. Elimination of the hydrodynamic concentration dependence parameter

The key to the evaluation of the sedimentation coefficients of the associating species lies in the elimination of the concentration dependence parameter. When this is done one obtains an equation containing only sedimentation coefficients, s_i, the monomer concentration, c_1, and the equilibrium constant or constants, K_n. Note that the s_i, $i = 1, 2, \ldots$, are the true values of the sedimentation coefficients of the associating species. We will illustrate the procedure for eliminating the hydrodynamic concentration dependence parameter with a monomer-n-mer self-association, and then the arguments will be extended to other self-associations. With nonassociating solutes one can easily obtain g, the hydrodynamic concentration dependence parameter, by simply plotting s or $1/s$ vs. c. For instance, from eq. (6) one notes that

$$\frac{1}{s} = \frac{1}{s_0} + \frac{gc}{s_0}. \tag{14}$$

Thus one notes that $\lim_{c\to 0}(1/s) = (1/s_0)$ and the slope of a plot of $1/s$ vs. c is g/s_0. With self-associations the situation is more complicated. With model I one notes that

$$\lim_{c\to 0} \frac{ds_{wa}}{dc} = -\underline{g}s_1^2 + \left(\frac{ds_{wc}}{dc}\right)_{c=0}, \tag{15}$$

whereas with model II we note that

$$\lim_{c\to 0} \frac{ds_{wa}}{dc} = -gs_1 + \left(\frac{ds_{wc}}{dc}\right)_{c=0} \tag{16}$$

If one used Model I and made a plot of $1/s_{wa}$ vs. c, the intercept would be s_1, but the limiting slope would be $\underline{g} - s_1^{-2}(ds_{wc}/dc)_{c=0}$. Thus we cannot obtain g or $\underline{g}$ from plots of $1/s_{wa}$ vs. c, and we must look for alternative methods when dealing with self-associations. Fortunately it is easy to obtain these alternatives, and four different methods will be used to eliminate g or $\underline{g}$. Model I will be used first.

Model I: $$\frac{1}{s_{wa}} = \frac{1}{s_{wc}} + \underline{g}c.$$

In all of the four procedures to be described below, a monomer-n-mer self-association will be used as an example. Since it is assumed that c_1 and K_n are available from light scattering, sedimentation equilibrium, or osmotic pressure experiments on the self-associating solute, and since

$$\lim_{c\to 0} s_{wa} = s_1, \tag{17}$$

the elimination of $\underline{g}$, the hydrodynamic concentration dependence parameter for model I, leads to an equation in one unknown, s_n. One can then use successive approximations to obtain the value of s_n.

Method 1:

Here one uses $1/s_{wa}$ and $d(1/s_{wa})/dc$ to eliminate $\underline{g}$, since

$$\frac{1}{s_{wa}} - c\frac{d}{dc}\left(\frac{1}{s_{wa}}\right) = \frac{1}{s_{wc}} - c\frac{d}{dc}\left(\frac{1}{s_{wc}}\right). \tag{18}$$

For a monomer-n-mer association

$$\frac{1}{s_{wc}} = \frac{1 + K_n c_1^{n-1}}{s_1 + K_n c_1^{n-1} s_n} \tag{19}$$

and

$$\frac{d}{dc}\left(\frac{1}{s_{wc}}\right) = \frac{(n-1)K_n c_1^{n-2}(s_1 - s_n)}{(s_1 + K_n c_1^{n-1} s_n)^2 (1 + nK_n c_1^{n-1})}. \tag{20}$$

Thus,

$$\frac{1}{s_{wa}} - c\frac{d}{dc}\left(\frac{1}{s_{wa}}\right) = \frac{1 + K_n c_1^{n-1}}{s_1 + K_n c_1^{n-1} s_n} + \frac{c(n-1)K_n c_1^{n-2}(s_n - s_1)}{(s_1 + K_n c_1^{n-1} s_n)^2 (1 + nK_n c_1^{n-1})}. \quad (21)$$

Method 2:

Suppose the plot of $1/s_{wa}$ vs. c has a minimum at $c = c_*$, then

$$\frac{d}{dc}\left(\frac{1}{s_{wa}}\right)_{c=c*} = 0, \quad (22)$$

$$g = -\frac{d}{dc}\left(\frac{1}{s_{wc}}\right)_{c=c*}, \quad (23)$$

and

$$\frac{1}{s_{wa}} = \frac{1}{s_{wc}} - c\frac{d}{dc}\left(\frac{1}{s_{wc}}\right)_{c=c*} \quad (24)$$

The insertion of eq. (19) and eq. (20), evaluated at $c = c_*$, leads to an equation with one unknown, s_n.

Method 3:

Here we note that

$$\frac{1}{s_{wa}} - \frac{2}{c}\int_0^c \frac{dc}{s_{wa}} = \frac{1}{s_{wc}} - \frac{2}{c}\int_0^c \frac{dc}{s_{wc}}. \quad (25)$$

For a monomer-dimer association one can show that

$$\frac{1}{s_{wa}} - \frac{2}{c}\int_0^c \frac{dc}{s_{wa}} = \frac{1 + K_2 c_1}{s_1 + K_2 c_1 s_2} - \frac{2}{c}\left\{\frac{1}{K_2 s_2}\left(1 - \frac{3s_1}{s_2} + \frac{2s_1^2}{s_2^2}\right)\ln\left(\frac{s_1 + K_2 c_1 s_2}{s_1}\right) + \frac{c_1}{s_2}\left(3 + K_2 c_1 - \frac{2s_1}{s_2}\right)\right\}. \quad (26)$$

The only unknown here is s_2, and it is obtained by successive approximations. The integrals necessary for the evaluation of $\int_0^c dc/s_{wc}$ for $n = 2$ to 4 are found in tables of integrals; when $n \geq 5$ one must use a reduction formula to evaluate the necessary integrals [28].

Method 4:

Kegeles [21] recently pointed out that one could obtain the z average sedimentation coefficient, s_{zc}, from the equation

$$s_{zc} = \frac{d(cs_{wc})}{dc} = \sum ic_i s_i / \sum ic_i. \quad (27)$$

For a monomer-n-mer self-association

$$s_{zc} = \frac{s_1 + nK_n c_1^{n-1} s_n}{1 + nK_n c_1^{n-1}}. \quad (28)$$

Now note that with model I s_{za} is given by

$$s_{za} = \frac{d(cs_{wa})}{dc} = s_{za}/(1 + gcs_{wc})^2. \quad (29)$$

From the definition of s_{wc} given by model I (see eq. 7), one can obtain

$$1 + gcs_{wc} = \frac{s_{wc}}{s_{wa}}. \quad (30)$$

With model I, s_{wa} is written as

$$s_{wa} = \frac{s_{wc}}{(1 + gcs_{wa})}.$$

The insertion of eq. (30) into eq. (29) followed by a rearrangement leads to

$$s_{za}/s_{wa}^2 = s_{zc}/s_{wc}^2. \quad (31)$$

For a monomer-n-mer self-association, eq. (31) can be written as

$$\frac{s_{za}}{s_{wa}^2} = \frac{c^2(c_1 s_1 + nK_n c_1^n s_n)}{(c_1 + nK_n c_1^n)(c_1 s_1 + K_n c_1^n s_n)^2}. \quad (32)$$

Eq. (32) has only one unknown, s_n.

Model II: $s_{wa} = s_{wc}/1 + gc$.

Again a monomer-n-mer association will be used as an example. It will be assumed here also that s_1, c_1

and K_n are already known so that the elimination of g leads to an equation in one unknown, s_n, which we can solve for by successive approximations. The methods described in the previous section for model I can also be tried with model II, but in this case the mathematical relations are somewhat more complicated.

Method 1:

Here we note that

$$\frac{1}{s_{wa}} = \frac{1}{s_{wc}}(1 + gc). \tag{33}$$

When model II is used eq. (33) can be used to obtain an expression for $d(1/s_{wa})/dc$, and this can be used to obtain an expression for g. Doing this we obtain

$$g = \frac{s_{wc}\left[\dfrac{d(1/s_{wa})}{dc} + \dfrac{1}{s_{wc}^2}\dfrac{ds_{wc}}{dc}\right]}{\left(1 - \dfrac{c}{s_{wc}}\dfrac{ds_{wc}}{dc}\right)}. \tag{34}$$

The substitution of eq. (34) into eq. (33) leads to

$$\frac{1}{s_{wa}} = \frac{1}{s_{wc}} + \frac{c\left(\dfrac{d(1/s_{wa})}{dc} + \dfrac{1}{s_{wc}^2}\dfrac{ds_{wc}}{dc}\right)}{\left(1 - \dfrac{c}{s_{wc}}\left[\dfrac{ds_{wc}}{dc}\right]\right)}. \tag{35}$$

When eqs. (19 and 20) are substituted into eq. (35), one obtains an equation in one unknown, s_n.

Method 2:

If s_{wa} has a minimum at $c = c_*$ or if $1/s_{wc}$ has a maximum at $c = c_*$, then one can show that

$$g = \frac{ds_{wc}/dc}{s_{wc} - c\left(\dfrac{ds_{wc}}{dc}\right)_{c=c_*}} \tag{36}$$

and

$$s_{wa} = \frac{s_{wc}}{1 + c\left[\dfrac{ds_{wc}/dc}{s_{wc} - c(ds_{wc}/dc)}\right]_{c=c_*}}. \tag{37}$$

When a monomer-n-mer association is present the substitution of eqs. (19 and 20) into eq. (37) leads to an equation containing only one unknown, s_n.

Method 3:

This method can be used in a different form; here we note that

$$\int_0^c \frac{dc}{s_{wa}} = \int_0^c \frac{1}{s_{wc}}(1 + gc)dc. \tag{38}$$

Solving for g one obtains

$$g = \frac{\displaystyle\int_0^c \frac{dc}{s_{wa}} - \int_0^c \frac{dc}{s_{wc}}}{\displaystyle\int_0^c \frac{cdc}{s_{wc}}}. \tag{39}$$

One can insert eq. (39) into eq. (33), and then use eq. (4) to obtain an equation in the unknown, s_n.

Method 4:

When model II applies, Kegeles [21] showed that

$$s_{za} = \frac{d(cs_{wa})}{dc} = \frac{s_{zc} - gs_{wa}}{1 + gc}. \tag{40}$$

With a monomer-n-mer self-association one can use eq. (40) to obtain an equation in onw unknown, s_n, using the following steps:

$$s_{wc}/s_{wa} = 1 + gc, \tag{41}$$

$$g = \left(\frac{s_{wc}}{s_{wa}} - 1\right)\Big/ c, \tag{42}$$

and

$$\frac{s_{za}}{s_{wa}} = \frac{s_{zc}}{s_{wc}} - \frac{(s_{wc} - s_{wa})}{cs_{wc}}. \tag{43}$$

The substitution of eqs. (4 and 27) leads to an equation in one unknown, s_n.

4. Evaluation of g or $\underline{g}$

Since we have shown that one cannot obtain g or $\underline{g}$ from the slope of a plot of $1/s_{wa}$ vs. c (see eqs. 15 or 16), we must use another approach. The simplest procedure is to calculate s_{wc} first; this can be done once the sedimentation coefficients of the self-associating species have been obtained (see the preceeding section). Then with model I $\underline{g}$ can be obtained from the following equation:

$$\frac{1}{s_{wa}} - \frac{1}{s_{wc}} = \underline{g}c \,. \tag{44}$$

A plot of $(1/s_{wa}) - (1/s_{wc})$ vs. c will have a slope of $\underline{g}$. If method 2 applies one could also use eq. (23) to evaluate $\underline{g}$, or one could also use eq. (30) for this purpose.

When model II applies and when s_{wc} is known, eq. (33) can be rearranged to give

$$\frac{1}{s_{wa}} - \frac{1}{s_{wc}} = g(c/s_{wc}). \tag{45}$$

Thus a plot of $(1/s_{wa}) - (1/s_{wc})$ will have a slope of g. Alternatively g could be evaluated from eq. (41), since g is the slope of a plot of s_{wc}/s_{wa} vs. c. Eqs. (34 or 36) could also be used for obtaining g.

5. Extension of these methods to other self-associations

5.1. Monomer-dimer-trimer association

For this association the expression for s_{wc} becomes

$$s_{wc} = \frac{s_1 + K_2 c_1 s_2 + K_3 c_1^2 s_3}{1 + K_2 c_1 + K_3 c_1^2} \,. \tag{46}$$

Since s_1, K_2, K_3 and c_1 are known, we are left with only two unknowns to solve for, s_2 and s_3. All of the previous methods can be used to eliminate g or $\underline{g}$, but we are faced with the necessity of evaluating the two unknowns. Note that method 3 can be used here since the integrals necessary for this method can be found in a good table of integrals. However, it may be more difficult to apply method three to other associations, such as a monomer-dimer-tetramer self-association.

With a monomer-dimer-trimer self-association we would have enough equations available to solve for s_2 and s_3, but admittedly we would have messy algebra to do this. One way to obtain s_2 and s_3 would be to use an iterative procedure, such as the Gauss–Seidel method or a curve crawling procedure [29] or we could use Monte Carlo methods [30]. These procedures could be carried out on a computer or on some programmable calculators. The only restriction here would be that $s_3 > s_2 > s_1$. Another but more restrictive way to attack this problem would be to assume that the three sedimentation coefficients are related by an equation of the type

$$s_i = K(iM_1)^{\alpha}, \qquad i = 1, 2, 3. \tag{47}$$

Here K is a constant whose value depends on the temperature and the combination of solvent and solute used. If eq. (47) were valid, then $s_2/s_1 = 2^{\alpha}$ and $s_3/s_1 = 3^{\alpha}$, so that eq. (46) can be written as

$$s_{wc} = \frac{s_1(1 + 2^{\alpha}K_2 c_1 + 3^{\alpha}K_3 c_1^2)}{1 + K_2 c_1 + K_3 c_1^2} \,. \tag{48}$$

The quantity α is the only unknown in eq. (48). So we could use eq. (48) along with other required quantities in the methods previously described for both models for s_{wa}, but here α instead of s_2 or s_3 would be evaluated by successive approximations. The disadvantage of this procedure is that all of the $s_i (i = 1, 2, 3)$ must obey eq. (47), whereas the other iterative method mentioned in this section does not have this restriction. In principle one could evaluate s_2 and s_3 first without assuming that they are related through eq. (47). Then we could see if they are interrelated, since if this were true then

$$\alpha = \frac{\log(s_2/s_1)}{\log 2} = \frac{\log(s_3/s_1)}{\log 3} \,. \tag{49}$$

If all of the self-associating species were spheres, then $\alpha = 2/3$; and if all the self-associating species were random coils, then $\alpha = 1/2$.

5.2. Indefinite self-association

Here the self-association reaction can be written as

$$nP_1 \rightleftarrows qP_2 + mP_3 + \dots .$$

It appears that the only way to treat this case is to assume that eq. (47) applies. When one does this, then one notes that $s_j/s_i = j$, where $j = 2, 3, \dots$. The expression for s_{wc} becomes

$$s_{wc} = \frac{s_1(1 + 2^{\alpha+1}kC_1 + 3^{\alpha+1}k^2C_1^2 + \dots)}{(1 + 2kC_1 + 3k^2C_1^2 + \dots)^2}$$
$$= s_1(1 + 2^{\alpha+1}kC_1 + 3^{\alpha+1}k^2C_1^2 + \dots)(1 - kC_1)^2, \quad (50)$$

if $kC_1 < 1$. Here $k = 1000K/M_1$ is the intrinsic equilibrium constant. In arriving at eq. (49) we have used the relationship

$$1 + 2kC_1 + 3k^2C_1^2 + \dots = 1/(1 - kC_1)^2,$$

when $kC_1 < 1$. Note that C_1 is the monomer concentration in g/ml. At present the sum of the series in the numerator of eq. (49) is not known. Since the sum of the series in the denominator of eq. (50) is $(1 - kC_1)^{-2}$ for $kC_1 < 1$, we might be able to truncate the series in the numerator after the fifth or sixth term. Just where to truncate it would depend on the magnitude of kC_1 and on the expected experimental error in the values of s_{wc} or s_{wa}. Assuming that the truncation is justifiable, one could use eq. (50) and the other necessary quantities in the four methods previously proposed for the elimination of g or $\underline{g}$ and the estimation of the s_i from either model for s_{wa}, except here one would obtain α by successive approximations. Note that the series in the numerator of eq. (50) has the sum of

$$(1 + kC_1)/(1 - kC_1)^3 \text{ for } \alpha = 1,$$

and

$$(1 + 4kC_1 + k^2C_1^2)/(1 - kC_1)^4 \text{ for } \alpha = 2.$$

Thus one could see whether α is less than one or if its value lies between one and two.

6. Discussion

How do our methods differ and what is their advantage over the Kakiuchi and Williams [22, 30] method? Consider a monomer-dimer-trimer self-association. In the Kakiuchi and Williams method model II s_{wa} was used; they made an attempt to evaluate s_2 and g from a plot of $1/s_{wa}$ vs. c. The first point to note is that Kakiuchi and Williams used the wrong equation for s_{wa}. Instead of using eq. (4) for s_{wc}, they used (sic)

$$s_{wc} = \frac{s_1 + K_2cs_2}{1 + K_2c}. \quad (51)$$

The error here is the use of c, the total solute concentration, instead of c_1, the monomer concentration. The correct expression for s_{wc} is given by eq. (4). Kakiuchi and Williams then proceeded to write s_{wa} as (sic)

$$s_{wa} = \frac{s_1 + K_2s_2c}{(1 + K_2c)(1 + gc)}, \quad (52)$$

when model II applies. The correct expression for s_{wa}, when model II applies, is

$$s_{wa} = \frac{s_1 + K_2c_1s_2}{(1 + K_2c_1)(1 + gc)}. \quad (53)$$

The Kakiuchi and Williams equation (eq. 45) can be rewritten as

$$\frac{1}{s_{wa}} = \frac{1 + gc}{s_2\left[1 + \frac{s_1 - s_2}{s_2(1 + K_2c)}\right]}. \quad (54)$$

If this equation were correct then as c becomes large, eq. (47) approaches the form

$$\frac{1}{s_{wa}} \approx \frac{1 + gc}{s_2} \quad (55)$$

Eq. (55) is linear in c, and if the tangent to a plot of $1/s_{wa}$ vs. c is taken at large values of c, one obtains g/s_2 for the slope and $1/s_2$ for the intercept at $c = 0$. One can justify this procedure by using eq. (53), which is the correct equation. Here one obtains

$$\frac{1}{s_{wa}} = \frac{1 + gc}{s_2\left[1 + \left(\frac{s_1 - s_2}{1 + K_2 c_1}\right)\right]}. \quad (56)$$

If $K_2 c_1$ gets large as c gets large, then eq. (56) can be reduced to eq. (55).

The disadvantage of the Kakiuchi and Williams method [22, 31] is that the experiment must be carried out to large values of c, the plot of $1/s_{wa}$ vs. c should be linear at these high values of c, and the results depend on how the tangent to the plot of $1/s_{wa}$ vs. c is drawn. At very high concentrations other hydrodynamic concentration terms might have to be included. The main advantage of this method is its simplicity.

Our methods, while somewhat more complex, do offer a possibility of evaluating s_2 (or s_n) and g or $\underline{g}$ at several different concentrations. Thus we can obtain an average value for these quantities, whereas the Kakiuchi and Williams procedure is a one shot method. Also we have available several ways to evaluate s_2 (or s_n) and g or $\underline{g}$, whereas they do not have this versatility.

We can extend the methods described here to any other weight average property, x_{wc}, of an associating system; here x could be a translational diffusion coefficient, a partition coefficient, a partial specific volume, etc. As long as a physical method gives x_{wc}, where x_{wc} is defined by

$$x_{wc} = \sum_i c_i x_i / c, \quad i = 1, 2, \ldots, \quad (57)$$

then it is possible to obtain x_z, the z average value of x, and also x_N, the number average value of x. If these physical properties exhibit some type of concentration dependence, then one can use models, analogous to models I or II for s_{wa}, to define x_{wa}, the apparent x_{wc}, and the individual x_i for the associating species can be evaluated from x_{wa} using methods analogous to those described in this paper. Note that x_{zc} is defined by

$$x_{zc} = \frac{\sum_i c_i M_i x_i}{\sum_i c_i M_i} = \frac{\sum_i i c_i x_i}{\sum_i i c_i}, \quad (58)$$

since $M_i = iM_1$, $i = 1, 2, \ldots$. We can obtain x_{zc} from

$$x_{zc} = \mathrm{d}(c x_{wc})/\mathrm{d}c, \quad (59)$$

since

$$\frac{\mathrm{d}(c x_{wc})}{\mathrm{d}c} = \sum_i x_i \frac{\mathrm{d}c_i}{\mathrm{d}c} = \sum_i x_i \frac{\mathrm{d}c_i/\mathrm{d}c_1}{\mathrm{d}c/\mathrm{d}c_1}, \quad (60)$$

$$\sum_i x_i \frac{\mathrm{d}c_i}{\mathrm{d}c} = x_1 + 2x_2 K_2 c_1 + 3x_3 K_3 c_1^2 + \ldots$$

$$= (1/c_1) \sum_i i c_i x_i, \quad (61)$$

and

$$\mathrm{d}c/\mathrm{d}c_1 = 1 + 2K_2 c_1 + 3K_3 c_1^2 + \ldots = (1/c_1) \sum_i i c_i. \quad (62)$$

How do we obtain the x_{wc}? Ackers [32] has shown that the weight average partition coefficient, σ_{wc}, can be obtained from analytical gel chromatography experiments with self-associating solutes. It should be noted that Chun et al. [33] have shown how to obtain the weight fraction of monomer from σ_{wc} by a different procedure. The apparent weight average diffusion coefficient, D_{wa}, can be obtained from free diffusion experiments [34]. The weight average partial specific volume, $\bar{v}_{wc}$, could be obtained from careful density measurements.

The number average property, x_{Nc} is defined by

$$s_{Nc} = \sum_i n_i s_i / \sum_i n_i = \sum_i (c_i s_i / M_i) / \sum_i (c_i / M_i)$$

$$= \sum_i (c_i s_i / i) / \sum_i (c_i / i), \quad i = 1, 2, \ldots. \quad (63)$$

Now note that

$$\int_0^c c s_{wc} \mathrm{d}c_1 / c_1 = \int_0^{c_1} (s_1 + K_2 c_1 s_2 + \ldots)\mathrm{d}c_1$$

$$= c_1 s_1 + \frac{K_2 c_1 s_2}{2} + \ldots = \sum_i \frac{s_i c_i}{i}. \quad (64)$$

But it has been shown by Adams [35] that

$$\frac{\mathrm{d}c_1}{c_1} = \frac{M_1 \mathrm{d}c}{c_1 M_{wc}} \quad (65)$$

for self-associating systems. Thus it follows that

$$\int_0^c s_{wc}(M_1/M_{wc})\mathrm{d}c = s_{Nc}(cM_1/M_{nc}) \qquad (66)$$

or

$$s_{Nc} = \frac{M_{nc}}{cM_1} \int_0^c s_{wc}(M_1/M_{wc})\mathrm{d}c\,. \qquad (67)$$

Under conditions where the hydrodynamic concentration dependence parameter is present, one can show using model II that

$$s_{Nc} = s_{Na} - g\left(\frac{M_{nc}}{cM_1}\right) \int_0^c s_{wa}(M_1/M_{wc})\mathrm{d}c. \qquad (68)$$

Here s_{Na} is the apparent values of s_{Nc}; s_{Na} is defined by

$$s_{Na} = 1/(cM_1/M_{nc}) \int_0^c s_{wa}(M_1/M_{wc})\mathrm{d}c. \qquad (69)$$

Eq. (68) can be rearranged to give

$$g = \frac{(s_{Nc} - s_{Na})(cM_1/M_{nc})}{\int_0^c c(M_1/M_{wa})s_{wa}\mathrm{d}c}\,, \qquad (70)$$

which gives us still another way to evaluate g. The evaluation of s_{Na} from model I is more difficult and will not be discussed here. One can generalize the treatment above to show that the number average property can be obtained from

$$x_N = 1/(cM_1/M_{nc}) \int_0^c x_{wc}(M_1/M_{wc})\mathrm{d}c. \qquad (71)$$

It is clear from this discussion that if some weight average property, s_{wc}, of a self-associating solute is available from transport experiments, then one can combine transport data with thermodynamic data to extract the individual x_i of the self-associating species. Here we have emphasized a combination of sedimentation velocity with sedimentation equilibrium experiments. Some test of these procedures will be reported in the next paper in this series.

Acknowledgement

This work was supported in part by a grant (to ETA, Jr.), GB 32242, from the National Science Foundation.

We are very grateful to Dr. David Cox, Dr. Paul Chun, Dr. C.N. Pace, Dr. Russel Larsen, Mr. Peter Wan and Mr. Charles Smith for their interest and comments on this work.

References

[1] L.W. Nichol, J.L. Bethune, G. Kegeles and E.L. Hess, in: The proteins, H. Neurath, ed., 2nd ed. (Academic Press, New York, 1964).
[2] E.T. Adams, Jr., and H. Fujita, in: Ultracentrifugal analysis in theory and experiment, ed. J.W. Williams (Academic Press, New York, 1963) p. 119.
[3] E.T. Adams, Jr. and J.W. Williams, J. Am. Chem. Soc. 86 (1964) 3454.
[4] E.T. Adams, Jr., Biochemistry 4 (1965) 1646.
[5] E.T. Adams, Jr., Biochemistry 6 (1967) 1864.
[6] K.E. Van Holde, G.P. Rossetti and R.D. Dyson, Ann. N.Y. Acad. Sci. 164 (1969) 279.
[7] Dennis E. Roark and David A. Yphantis, Ann. N.Y. Acad. Sci. 164 (1969) 245.
[8] P.W. Chun, S.J. Kim, J.D. Williams, W.T. Cope, L.H. Tang and E.T. Adams, Jr., Biopolymers 11 (1972) 197.
[9] G.A. Gilbert, Disc. Faraday Soc. 20 (1955) 68.
[10] G.A. Gilbert, Proc. Royal Soc. (London) A 250 (1959) 377.
[11] G.A. Gilbert, R.C. Ll. Jenkins, Proc. Royal Soc. (London) A 253 (1959) 420.
[12] L.M. Gilbert and G.A. Gilbert, Nature 194 (1962) 1173.
[13] H. Fujita, Mathematical theory of sedimentation analysis (Academic Press, New York, 1962) ch. 4.
[14] R.F. Steiner, Arch. Biochem. Biophys. 39 (1952) 333.
[15] R.F. Steiner, Arch. Biochem. Biophys. 44 (1953) 120.
[16] R.F. Steiner, Arch. Biochem. Biophys. 49 (1954) 400.
[17] M.S.N. Rao and G. Kegeles, J. Am. Chem. Soc. 80 (1958) 5724.
[18] W. Goad and J.R. Cann, Ann. N.Y. Acad. Sci. 164 (1969) 172.
[19] D.J. Cox, Arch. Biochem. Biophys. 142 (1971) 514.
[20] W. Godschalk, Biochemistry 10 (1971) 3284.
[21] G. Kegeles, Proc. Nat. Acad. Sci. USA, 69 (1972) 2577.
[22] K. Kaiuchi and J.W. Williams, J. Biol. Chem. 241 (1966) 2781.
[23] N. Iso and J.W. Williams, J. Biol. Chem. 241 (1966) 2787.
[24] R.L. Baldwin, Biochem. J. 55 (1953) 644.
[25] R.V. Webber, J. Am. Chem. Soc. 78 (1956) 536.
[26] G. Kegeles and M.S.N. Rao, J. Am. Chem. Soc. 80 (1958) 5721.
[27] R.J. Goldberg, J. Phys. Chem. 57 (1953) 194.
[28] I.S. Gradshteyn and I.M. Ryzhik, Table of integrals, series and products, 4th ed., prepared by Yu.V. Geroniumus and M.Yu. Tseytlin (Academic Press, New York, 1965) pp. 58–60.
[29] Forman S. Acton, Numerical methods that work (Harper and Row, New York, 1970).

[30] W.J. Hammerle, Statistical computations on a digital computer (Blaisdell Publishing Co., Waltham, Mass., 1967).
[31] J.W. Williams, Ultracentrifugation of macromolecules: modern topics (Academic Press, New York, 1972).
[32] G.K. Ackers, Advan. Protein Chem. 24 (1970) 343.
[33] P.W. Chun, S.J. Kim, C.A. Stanley and G.K. Ackers, Biochemistry 8 (1969) 1625.
[34] L.J. Gosting, Advan. Protein Chem. 11 (1956) 429.
[35] E.T. Adams, Jr., Fractions No. 3, 1967.

BIOPHYSICAL CHEMISTRY 1 (1973) 46–50. NORTH-HOLLAND PUBLISHING COMPANY

RATE CONSTANTS FOR THE HEXAMER–DODECAMER REACTION OF LOBSTER HEMOCYANIN

Gerson KEGELES and Mei-Sheng TAI
Section of Biochemistry and Biophysics, University of Connecticut, Storrs, Connecticut 06268, USA

Received 15 May 1973

90° stopped-flow light scattering experiments have been used to determine the forward and reverse rate constants of the reversible hexamer–dodecamer reaction of lobster hemocyanin in glycine buffer at pH 9.6. Results were obtained at 20.5° and 25° at two different levels of calcium ion, and equilibrium constants derived from the data are in good agreement with those measured earlier with the ultracentrifuge.

1. Introduction

A logical method for monitoring macromolecular reactions is to use light scattering whenever a change in molecular weight is involved. In a classical paper on the dimerization of bovine plasma mercaptalbumin [1], such an application of light scattering was made to study a slow reaction. One of us, in a review entitled "Interacting Protein Systems" suggested [2] the application of light scattering methods of observation to relaxation techniques [3]. Stopped–flow measurements of macromolecular interactions were similarly performed by turbidity or scattering at 90° in several laboratories since 1969 [4–7]. A pressure-jump light scattering technique is being described elsewhere [8].

The "concentration-jump" method, in which a sudden dilution of a protein system with buffer produces a relaxation to a new equilibrium, was developed by Fisher and Bard [5]. This procedure has the distinct advantage over many other types of stopped–flow experiments, in that the rate constants for the protein reaction steps, which may in fact be coupled to many steps of ligand binding, are actually determined in a prescribed and constant solvent medium. Such rate constants for single steps are then directly related to equilibrium constants measured by other methods, in the same buffer.

2. Theory

In ion–mediated protein interactions, such as that considered here, in which both protons and calcium ion play a role [9], the detailed expressions for the dependence of the protein reaction relaxation time can be developed, if at all [3], only when a mechanism is known for the ion-binding steps. Otherwise, since the "equilibrium constant" for the protein reaction depends strongly on the free ion concentrations, one or both of the "rate constants" for the protein interaction will similarly depend on these ion concentrations. Thus, for the results of an experiment to be interpretable at all, it is necessary for these "rate constants" to be determined in a well-defined ionic environment which does not change appreciably with time, i.e. in a well-buffered system. This makes the "concentration-jump" method particularly attractive compared to other types of stopped–flow experiments, since the dialyzed protein can be diluted by a small amount (in our case to 5/7 of its total original concentration) by a buffer with which it was already at dialysis equilibrium.

When concentrations are expressed in molar units, the relaxation time for a simple monomer–dimer reaction step is given by [3]

$$\tau^{-1} = 4k'_{12}n_1 + k'_{21} , \qquad (1)$$

where k'_{12} and k'_{21} are the molar rate constants for dimerization and dissociation, and n_1 is the equilibrium molar concentration of monomer, all referred to the final (in this case the diluted) state. This relaxation equation assumes that the system is never far from its

final equilibrium condition, which we have assured by dilution of five volumes of protein with only two volumes of buffer, and appropriate adjustment of calcium ion and pH according to the prior equilibrium data [9]. If concentrations in grams per unit volume are used in the original rate equations to derive the expression for the relaxation time, it is found that in place of eq. (1), as cited in previous publications [5, 10], the correct expression is

$$\tau^{-1} = 2k_{12}c_M + k_{21} , \qquad (2)$$

where k_{12} and k_{21} are the weight concentration rate constants and c_M is the weight concentration of monomer. Since the dissociation constant has the dimensions of reciprocal time, $k_{21} = k'_{21}$, and the relationship between the equilibrium constants is, as required,

$$k_{12}/k_{21} = (2/M_1)(k'_{12}/k'_{21}) . \qquad (3)$$

Here M_1 is the molecular weight of monomer. In our case, the "monomer" is a hexameric unit, and the "dimer" is the dodecamer.

It has been pointed out [3] that a plot of reciprocal relaxation time against equilibrium concentration of monomer gives a straight line, providing the dimerization constant from the slope and the dissociation constant from the intercept. Although we have had the good fortune to have available prior equilibrium data in some cases [9] from which to calculate the monomer concentration, we have nevertheless found this type of plot less than satisfactory for the following reason: when the formation constant is too high because of experimental error, the slope and intercept of the plot from eq. (2) are so distorted that the ratio of the retrieved rate constants gives back a formation constant which is still higher. Conversely, when the formation constant is too low to begin with, the retrieved value from use of eq. (2) is still lower. In other words, successive approximations based on direct use of eq. (2) are uniformly divergent.

This unfortunate circumstance can be circumvented, and rate constants for the monomer–dimer interaction can be determined which are wholly independent of the use of prior equilibrium data, in the following manner. The equilibrium constant is given by

$$(c - c_M)/c_M^2 \approx k_{12}/k_{21} , \qquad (4)$$

where the total weight concentration is denoted by c.

When eq. (2) is squared, and the value of c_M^2 from eq. (4) is substituted therein, the result is

$$\tau^{-2} = 4k_{12}k_{21}c + k_{21}^2 . \qquad (5)$$

According to eq. (5), the square of the reciprocal relaxation time is a linear function of the total weight concentration of protein. The intercept provides the square of the dissociation constant, and the slope and intercept together then provide a value for the rate constant for reassociation. Hence no prior value of the equilibrium constant is required, and the ratio k_{12}/k_{21} provides a completely independent equilibrium constant directly from the kinetic data.

In the case of non-ideal solutions, as shown previously [10], the measured relaxation time τ should be multiplied by the factor $(1 + c\,\partial \ln y/\partial c)$ before performing the plots of eq. (2), or in this case eq. (5). The non-ideality factors were measured for non-reacting dodecameric hemocyanin [9].

3. Experimental

The dilution, or "concentration-jump" experiments [5] were performed in a Durrum–Gibson stopped-flow apparatus [11] equipped with a new cell and photomultiplier assembly [6] for high response to scattering at 90°. For optimal long-time stability, the tungsten iodide light source originally part of the equipment was found to be far superior to either Xenon or Xenon–Mercury sources. The apparatus was also furnished with a set of "ratio-syringes", which permitted the dilution of five volumes of protein with two volumes of buffer. Experiments were performed at either 20.5°C or 25°C by circulating water from a carefully regulated water bath. It had been found earlier that temperature control is quite essential to good scattering experiments. Protein solutions were measured at 5–6 different protein concentrations, leading to final diluted concentrations in the range between 0.18% and 1.14%, in 0.1 ionic strength glycine buffers at pH 9.6 containing either 0.0031 M or 0.0036 M free calcium ion [9]. Each solution was dialyzed against the corresponding buffer with which it was mixed in the stopped-flow experiment. Solutions were kept in closed vessels, and were prepared from

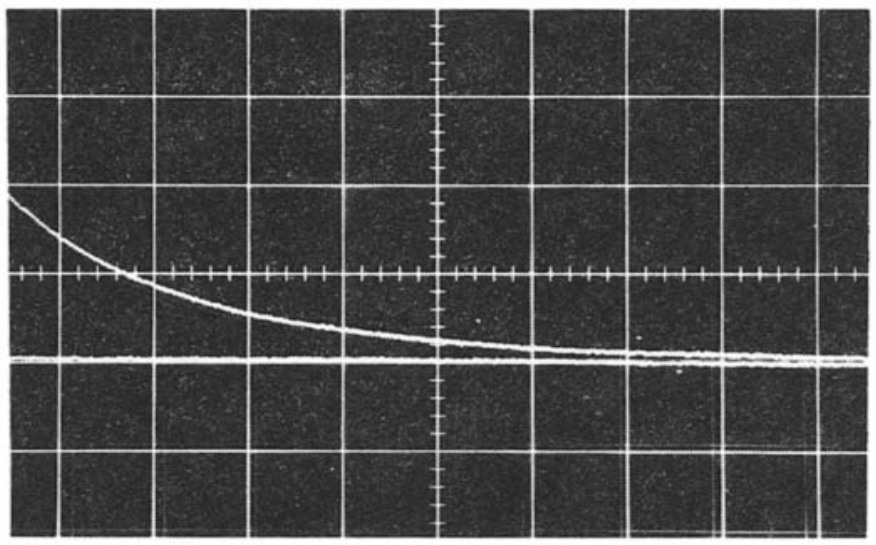


Fig. 1. Light scattering stopped flow "concentration–jump" relaxation curve for lobster hemocyanin. Five volumes of 1.203% hemocyanin dialyzed against buffer were mixed with two volumes of 0.1 ionic strength, pH 9.6 glycine – NaOH buffer containing 0.0031 M Ca^{2+} at 25°C. Scanning speed was 14.43 sec/div. The tungsten iodide light source was used, with the monochromator set at 436 nm.

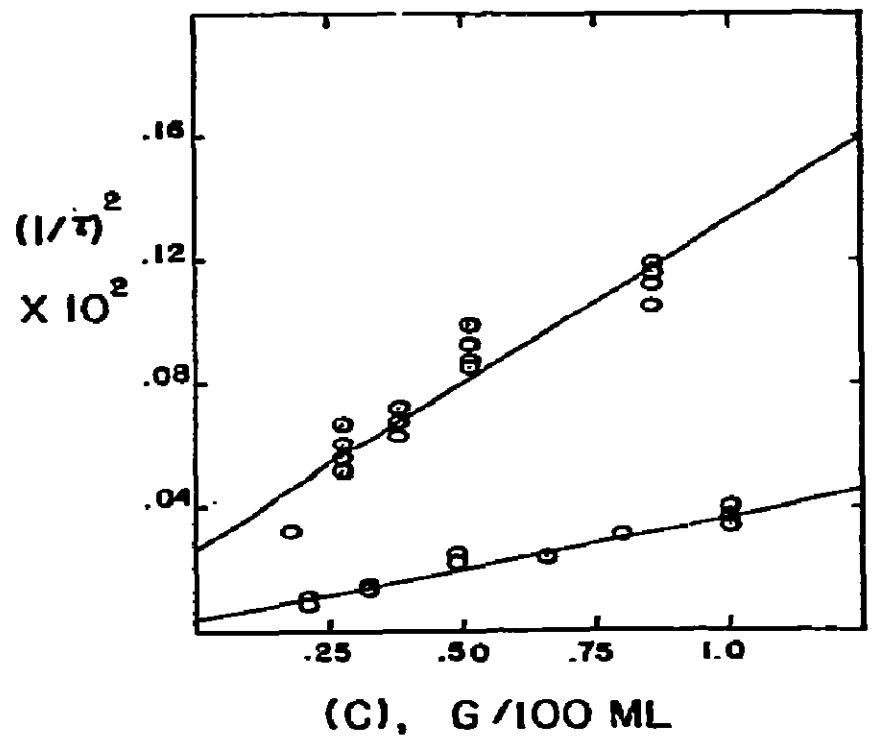


Fig. 2. Reciprocal of square of non-ideality-corrected relaxation time vs. total protein concentration. Points are experimental, and lines are least-squares fit (see text). Upper curve: 25°C. Lower curve: 20.5°C. The buffer was 0.1 ionic strength pH 9.6 glycine-NaOH containing 0.0031 M Ca^{2+}. Upper curve was based on 18 experiments, and lower curve on 20 experiments; where points are missing they overlapped too closely with those shown, to be seen.

freshly boiled distilled water to minimize air bubbles. All hemocyanin preparations were processed by removing the clot, centrifuging briefly, and either freezing the whole serum for storage, or diluting and dialyzing directly. Small amounts of aggregated material were removed from stored thawed samples by centrifugation before dilution and dialysis.

4. Results and data evaluation

For a considerable period we had believed that the only scattering relaxation in lobster hemocyanin was to be found in the time range too fast for stopped–flow experiments, since a small relaxation was found [10] in the time range below 1 millisecond, and repeated efforts to detect any slow relaxation with the earlier version of the stopped–flow scattering cell had failed, as had hand mixing experiments [9]. Once the newer scattering cell and photomultiplier arrangement with greatly improved aperture [6] was available to us, a crucial set of experiments, in which reacting hemocyanin was either mixed with an excess of calcium ion or with a dilute solution of ethylenediaminetetraacetic acid, showed a strong scattering change in the range of 15–25 seconds. These effects lie entirely outside the range of the temperature–jump experiment, which could not detect even an amplitude shift, since the entire cell had cooled back to original temperature before the protein could react.

The use of the concentration–jump technique, as described above, leads to relaxation scattering curves such as the oscilloscope trace shown in fig. 1, where a solution of original concentration of 1.203% was diluted to 0.859% in a buffer containing 0.0031 M Ca^{2+} at 25°C. The oscilloscope traces were mounted and measured directly with reflected illumination in a two-coordinate Mann microcomparator, some 30 points being read (the baseline in fig. 1 being a second scan following the first after a gap of two sweep scale divisions). Between two and five experiments were performed at each concentration.

The raw microcomparator data were least–squared with an IBM 1620 computer as a single exponential decay, in the form of a log–log relation, the uncorrected relaxation time and the zero–time amplitude were printed out, and the computations were used to produce automatic correlation plots between experi-

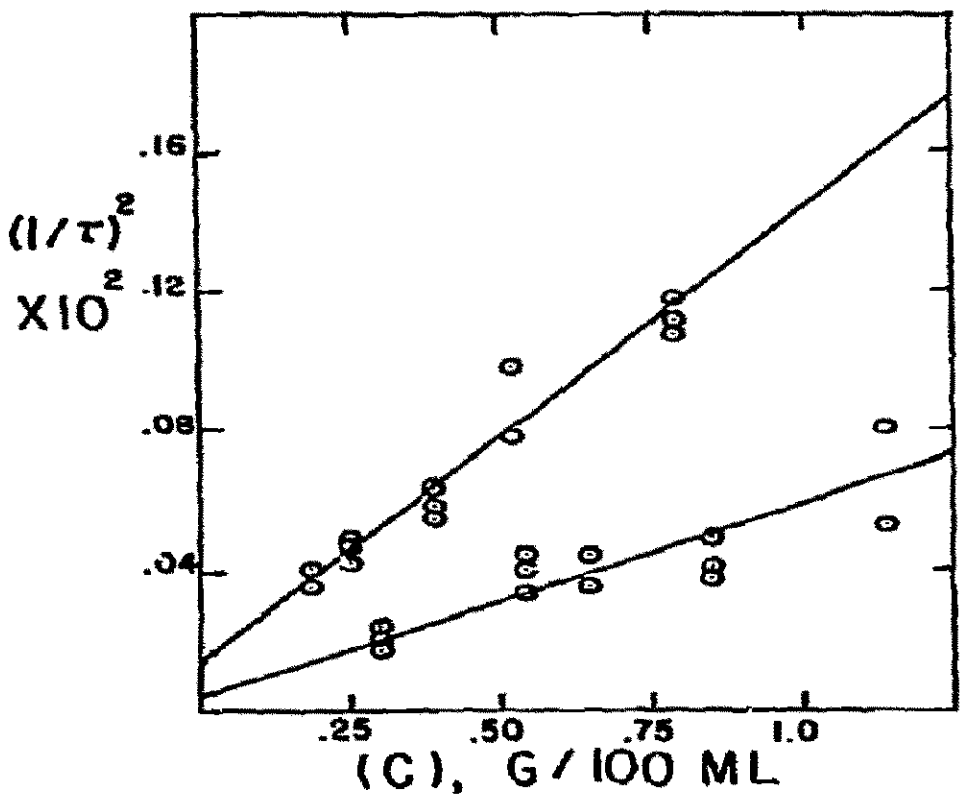


Fig. 3. Reciprocal of square of non-ideality-corrected relaxation time vs. total protein concentration. Points are experimental, and lines are least squares fit (see text). Upper curve: 25°C based on 15 experiments. Lower curve: 20.5°C based on 16 experiments. The buffer was 0.1 ionic strength, pH 9.6 glycine–NaOH containing 0.0036 M Ca^{2+}.

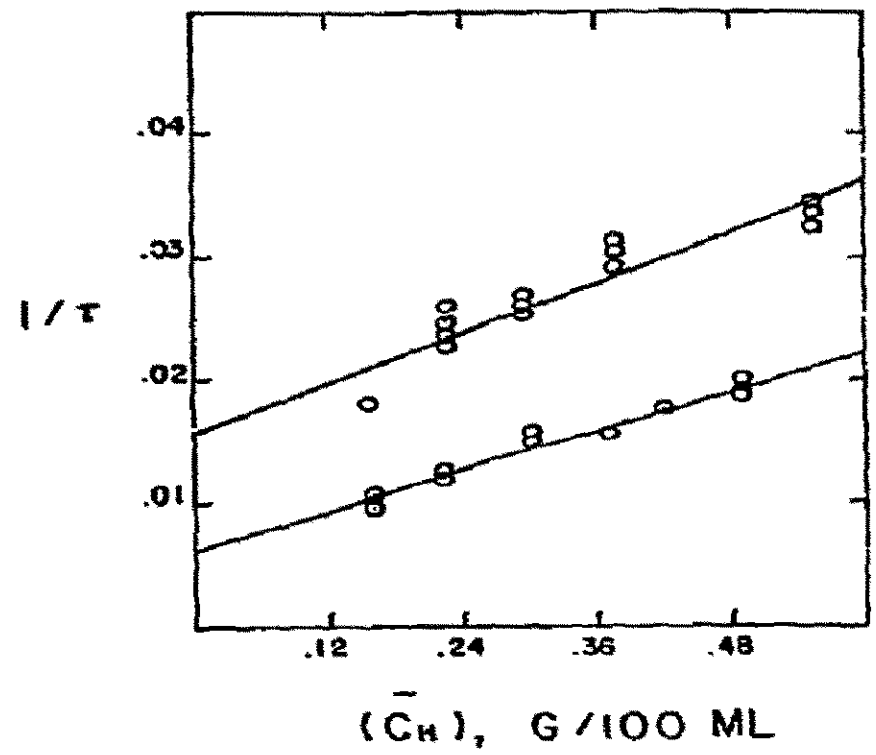


Fig. 4. Reciprocal of non-ideality-corrected relaxation time vs. hexamer concentration calculated from equilibrium constants derived from fig. 2 (see text). Upper curve: 25°C based on 18 experiments. Lower curve: 20.5°C based on 20 experiments. Buffer contained 0.0031 M Ca^{2+}.

mental data and the least square line, with an IBM 1627 plotter. Erroneous or inconsistent data points were then rechecked, and if necessary discarded, and the process repeated.

Once relaxation times were obtained, they were multiplied by the non-ideality factor $(1 + c\,\partial \ln y/\partial c)$, and the square of the corrected reciprocal was least squared as a function of the total (diluted) protein concentration, according to eq. (5) at a given temperature and calcium ion concentration. The ratio of the derived forward and reverse rate constants was obtained, from which the concentration of monomer was obtained at equilibrium, eq. (4). Having this, the reciprocal of the non-ideality-corrected relaxation times were least-squared as a function of monomer concentration, according to eq. (2). These computations were all performed with the IBM 1620 computer, and the punched card output was used to plot the experimental points versus the least square lines, eq. (2) and (5).

Fig. 2 shows the results of the plot according to eq. (5) at 0.0031 M free Ca^{2+}, and fig. 3 shows the plot of the results at 0.0036 M Ca^{2+}. Each figure contains data at both 20.5° and 25°C. Fig. 4 shows the more conventional (corrected) reciprocal relaxation time

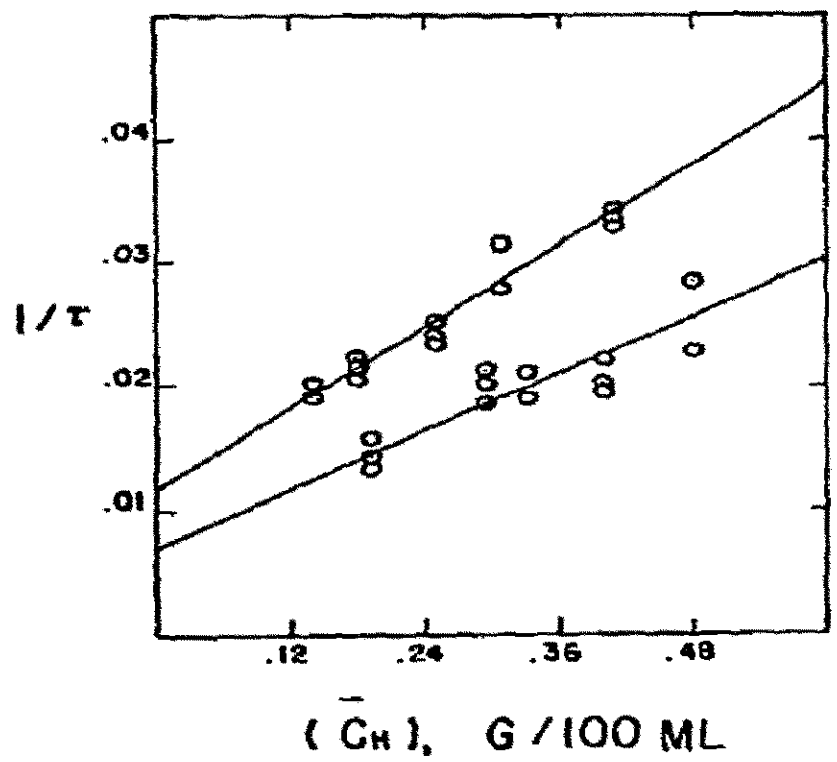


Fig. 5. Reciprocal of non-ideality-corrected relaxation time vs. hexamer concentration calculated from equilibrium constants derived from fig. 3 (see text). Upper curve: 25°C based on 15 experiments. Lower curve: 20.5°C based on 16 experiments. Buffer contained 0.0036 M Ca^{2+}.

Table 1

T(°C)	Free Ca^{2+} (M)	k_{12}	k_{21}	$K_{eq.}$
20.5	0.0031	0.01318	0.00624	2.11
	0.0036	0.0198	0.0069	2.86
25.0	0.0031	0.0161	0.0164	0.982 a)
	0.0036	0.0270	0.0120	2.26 b)

a) Comparison ultracentrifuge value [9]: 1.0
b) Comparison ultracentrifuge value [9]: 2.89.

plots vs. monomer concentration, eq. (2) at 0.0031 M free Ca^{2+}, and fig. 5 shows the corresponding plots at 0.0036 M free Ca^{2+}, each figure containing data at both 20.5° and 25°.

5. Discussion and summary

Eqs. (2) and (5) assume that the reaction involved in producing a change in scattering is entirely a single step, bimolecular in the forward direction and unimolecular in the reverse direction. One additional test of this assumption, beyond figs. 2, 3, 4 and 5 is the consistency of direct oscilloscope data in the log–log plots with a single straight line. Usually the data followed this requirement very well – in cases where this was not so, either uncovery of errors in reading or a repeated experiment usually substantiated the assumption. The biggest experimental defect has been drifting or shifting of the scattering baseline, caused usually by minute leaks through valves or around drive syringe plungers. Although the initial amplitude of the experimental scattering curve is not required, we have found no way to eliminate the requirement of knowing the final scattering baseline.

A summary of the rate constants on the g/dl concentration scale taken from use of eq. (5) and the computer results represented by figs. 2 and 3, is given in table 1. It is noted that the formation constant, $K_{eq} = k_{12}/k_{21}$, is in good agreement with prior ultracentrifuge results [9] determined at 25°C. Again, it is emphasized that k_{12} and k_{21} contain effects of ion binding in an unknown way, and are constants only in the specified solvents.

Thus, it has been found that the major part of the scattering change resulting from the reversible hexamer–dodecamer interaction of lobster hemocyanin in glycine buffer at pH 9.6 lies in the time range of about 15–25 seconds. The previously reported relatively small scattering amplitude [10], earlier ascribed to the hexamer–dodecamer interaction, may, in fact, be due to a minor conformational change coupled to oxygenation. Repeated efforts to investigate the phenomenon have been disappointing and this remains to be further elucidated.

It is rather intriguing, table 1, that the dissociation rate constant is relatively independent of calcium ion concentration, while the association constant bears most of the calcium ion dependence previously observed [9] in the equilibrium constant.

Acknowledgement

This study was made possible under support of National Science Foundation Grant GB 29092X entitled "Kinetics of macromolecular interactions" and a University of Connecticut Research Grant 35–342 of the same title, which permitted purchase of the original stopped-flow unit. Computations were performed at the University of Connecticut IBM 1620 computer center.

References

[1] H. Edelhoch, E. Katchalski, R.H. Maybury, W.L. Hughes, Jr. and J.T. Edsall, J. Am. Chem. Soc. 75 (1953) 5058.
[2] L.W. Nichol, J.L. Bethune, G. Kegeles and E.L. Hess, in: The proteins, Vol. 2, ed. H. Neurath (Ac. Press, New York, 1964) p. 329.
[3] M. Eigen and L. De Maeyer, in: Techniques of organic chemistry, Vol. 8, ed. A. Weissberger (Interscience, New York, 1963), part 2, p. 895.
[4] C. Huang and C. Frieden, Fed. Proc. 28 (1969) 536.
[5] H.F. Fisher and J.R. Bard, Biochim. Biophys. Acta 188 (1969) 168.
[6] Bulletin 63, Durrum Instrument Corp., 3950 Fabian Way, Palo Alto, California, 1971.
[7] D. Riesner and H. Buenemann, Proc. Natl. Acad. Sci. U.S. 70 (1973) 890.
[8] G. Kegeles, V.P. Saxena and R. Kikas, unpublished.
[9] K. Morimoto and G. Kegeles, Arch. Biochem. Biophys. 142 (1971) 247.
[10] M.-S. Tai and G. Kegeles, Arch. Biochem. Biophys. 142 (1971) 258.
[11] Q.H. Gibson and L. Milnes, Biochem. J. 91 (1964) 161.

BIOPHYSICAL CHEMISTRY 1 (1973) 51–53. NORTH-HOLLAND PUBLISHING COMPANY

INTERACTION OF 1,N^6-ETHENOADENOSINE 3′,5′ MONOHYDRATE (E-C-AMP) WITH SOLVENTS AND BIOLOGICAL RECEPTORS*

C. Jayne BIASELLE and David B. MILLAR
Laboratory of Physical Biochemistry, Environmental Biosciences Department, Naval Medical Research Institute, National Naval Medical Center, Bethesda, Maryland 20014, USA

Received 15 May 1973

1. Introduction

The profound biological effects of cyclic-AMP are now widely recognized [1, 2]. It is expected that investigators interested in developmental biology, membrane biology, and cell transformation will find it of increasing importance. Elucidation of the reaction of cyclic-AMP with receptors will be facilitated by the development of fluorescent analogues such as that of Secrist et al. [3, 4]. In this connection, while examining the luminescent properties of the 1,N^6-ethenoadenosine 3′,5′ monohydrate (E-C-AMP) we observed: 1) an unusual solvolysis reaction between *p*-dioxane and E-C-AMP resulting in a complex in which luminescent lifetime varies in an intricate manner with *p*-dioxane concentration, and 2) an interaction between E-C-AMP and a wide range of proteins not presently known to be biologically significant C-AMP binding sites.

2. Methods

Cyclic-AMP, purchased from Calbiochem, was converted to cyclic E-AMP via the reaction procedure of Barrio et al. [4]. The lyophilized product was tested for homogeneity by paper chromatography [5]. Proteins and enzymes were purchased from Calbiochem or Worthington and were the highest purity available. The absorption studies of E-C-AMP were performed at room temperature in a Cary-14 spectrophotometer. Fluorescence spectra were obtained with an Aminco–Bowman spectrophotofluorometer. Lifetimes (25°C) were determined using a TRW Model 31A spectral source and decay time computer using Corning filters #9863 and #3385 to isolate the excitation and emission bands resepectively. All chemicals were of reagent grade or better quality.

3. Results

Only one UV absorbing spot was noted in a paper chromatographic analysis of E-C-AMP employing 5% sodium hydrogen phosphate saturated by isoamyl alcohol. The ultraviolet absorption spectrum was identical with that reported by Secrist et al. for E-C-AMP [3]. As it was anticipated that future studies of receptor proteins would involve a variety of solvents, the fluorescent lifetimes shown in table 1 were measured. In neutral solvents, the value of 19 nanoseconds agrees with that reported by Secrist et al. [3] of "about 20 nanoseconds". There is only a slight decrease in the strong protein denaturing solvents, guanidine and urea. However, in 50% *p*-dioxane the

* From the Bureau of Medicine and Surgery, Navy Department, search Task MRO41.06.01.0004BOEX.

Table 1
Fluorescent lifetimes of E-C-AMP

Solvent	Lifetime [a] (nsec) ± S.D.
0.01 M sodium phosphate buffer, pH 7	18.9 ± 1
0.1 M sodium phosphate buffer, pH 7	16.3 ± 0.5
0.01 M sodium phosphate buffer + 0.2 M NaCl, pH 7	19.0 ± 0.9
0.1 M sodium phosphate buffer + 0.2 M NaCl, pH 7	18.2 ± 0.7
5 M Guanidine and 0.001 M Clelands Reagent and 0.05 M sodium phosphate buffer, pH 7	16.6 ± 0.9
8 M Urea and 0.001 M Clelands Reagent and 0.05 M sodium phosphate buffer, pH 7	17.7 ± 0.7
50% DMSO in 0.05 M sodium phosphate buffer, pH 7	16.3 ± 0.7
50% Dioxane in 0.05 M sodium phosphate buffer, pH 7	11.3 ± 0.6
50% Ethylene Glycol in 0.5 M sodium phosphate buffer, pH 7	19.2 ± 1.2

a) Mean of twenty or more determinations

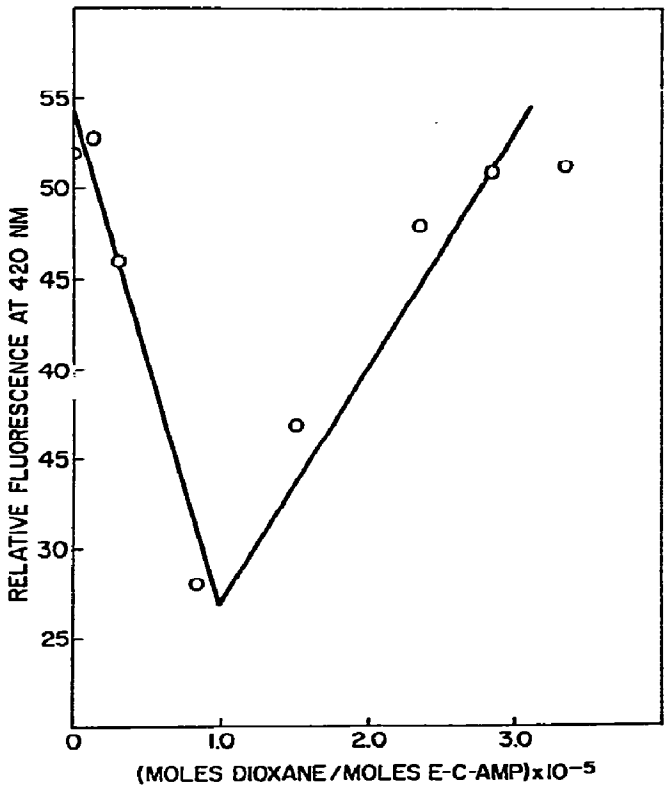
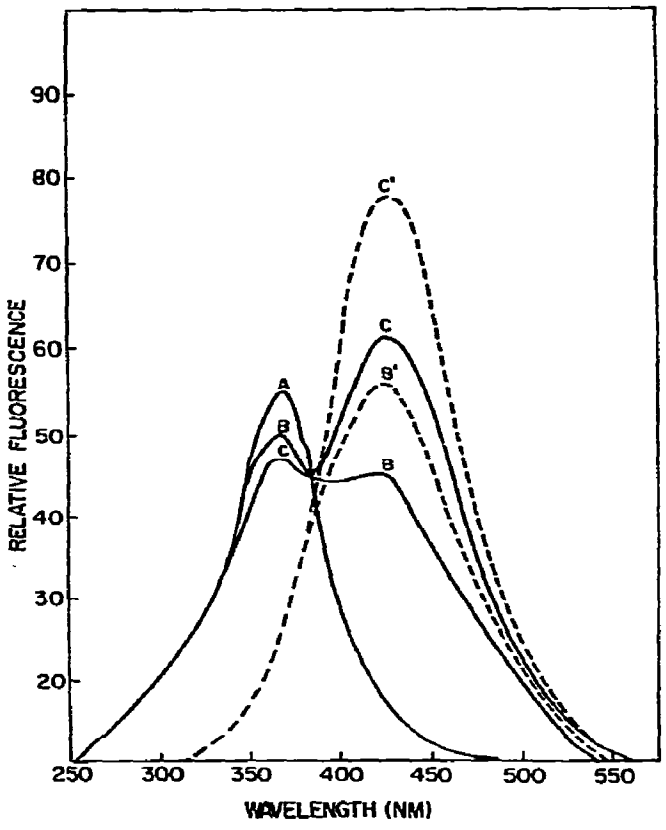


Fig. 1. Dependence of E-C-AMP fluorescence on *p*-dioxane concentration. [E-C-AMP] = 3.88×10^{-4} M; Excitation wavelength = 305 nm.

lifetime falls almost to half. Further examination showed that the fluorescence emission in *p*-dioxane depends on solvent composition in an unusual manner. Fig. 1 shows these results for an E-C-AMP concentration of about 4×10^{-4} M. The drop and subsequent gain in fluorescence presumably results from aqueous solvent displacement on the chromophore by the organic solvent. In 100% *p*-dioxane, the characteristics of the excited state must be fairly similar to that of the aqueous one since the emission spectra in the two cases are virtually identical. At the inflection point, the ratio of *p*-dioxane/E-C-AMP is about 1×10^5. Similar results were obtained with varying concentrations of E-C-AMP. This result suggests that highly organized structures of *p*-dioxane hydrogen bonded to water may exist and that care should be taken in attributing lack of quenching to lack of binding for the very high and very low ratios of *p*-dioxane/E-C-AMP. Neither urea nor guanidine shows this behavior; only simple quenching occurs with these solutes. Quenching of E-C-AMP also occurs in dimethyl sulfoxide, dimethyl formamide and more extensively in methanol, ethanol and butanol. This involved behavior certainly indicates that receptor E-C-AMP studies be performed with the possibility in mind that the receptor may not be a single macromolecule but a noncovalently stabilized network of smaller molecules.

As the adenylate moiety of C-AMP is a component of NADH, we measured the interaction between E-C-AMP and NADH dependent dehydrogenases. With excitation at 280 nanometers quenching of both enzyme and E-C-AMP fluorescence (20% to 40%) occurred for lactate dehydrogenase, phosphorylase A and B, 3-glycerol dehydrogenase and glutamic dehydrogenase. The quenching for 3-glycerol phosphate dehydrogenase is shown in fig. 2, which is typical of the titration spectra for the other proteins. This apparent interaction might be attributed to the known involvement of the adenine moiety in the coenzymes vital to these enzymes. No such explanation can be advanced for the quenching of E-C-AMP by egg albumen, B-lactoglobulin or trypsin. This quenching may represent binding at similar hydrophobic sites



on proteins of widely disimilar function. The reason for the existence of those sites is still uncertain.

References

[1] G.A. Robinson, E.W. Sutherland and R.W. Butcher, Cyclic AMP (Academic Press, New York, 1971).
[2] W.T. Rall, M. Rodbell and P. Condliffe (eds.), The Role of Adenyl Cyclase and Cyclic 3',5'–AMP in Biological Systems (Fogarty International Center Proceedings No. 4, NIH, 1969).
[3] J.A. Secrist, J.R. Barrio and N.J. Leonard, Biochemistry 11 (1972) 3499.
[4] J.A. Secrist, J.R. Barrio, N.J. Leonard, C. Villar-Palasi and A.G. Gilman, Science 177 (1972) 279.
[5] R.F. Steiner, FEBS Letters 23 (1972) 2.

Fig. 2. Titration of E-C-AMP into 3-glycerol phosphate dehydrogenase (0.1 mg/ml). A – enzyme in 0.01 M tris buffer, pH 7.2; B – 3×10^{-5} M E-C-AMP + enzyme; B' – 3×10^{-5} M E-C-AMP in buffer; C – 5×10^{-5} M E-C-AMP + enzyme; C' – 5×10^{-5} M E-C-AMP in buffer. Excitation wavelength = 280 nm.

BIOPHYSICAL CHEMISTRY 1 (1973) 55–61. © NORTH-HOLLAND PUBLISHING COMPANY

NMR AND DENSITY STUDY OF Co^{2+} SITE BINDING BY POLYELECTROLYTES

P. SPEGT, C. TONDRE, G. WEILL and R. ZANA
Centre de Recherches sur les Macromolécules CNRS, Strasbourg, CEDEX 67083, France

Received 16 August 1973

The changes of density and chemical shifts of the water proton upon addition of $CoCl_2$ to aqueous solutions of tetramethylammonium salts of seven polyelectrolytes (polyphosphate, maleic acid-methylvinylether alternated copolymer, polyacrylic acid, carboxymethylcellulose of substitution degree 0.98, 1.3, 2.1, 2.65) have been measured. Assuming a negligible contribution of pseudo-contact interaction to the water proton chemical shift and a constant hyperfine constant upon displacement of water molecules by other ligands, has permitted the calculation of (i) the number of water molecules released by Co^{2+} ions upon binding and the approximate fraction of Co^{2+} ions bound with loss of water, and (ii) the total volume change upon binding and the individual contributions of counterions and polyelectrolyte charged sites to this volume change. Our results are generally in agreement with those obtained using other methods.

1. Introduction

The theory of counterion condensation recently reported by Manning [1] has clarified the picture of polyion–counterion interaction in polyelectrolyte solutions. In summary this theory states that condensed counterions can be considered as bound counterions while interactions between uncondensed counterions and polyions can be described with a Debye–Hückel potential. In its present form, however, Manning's theory does not give the fraction of condensed counterions associated with discrete sites on polyions although a large body of experimental evidence for "specific" or "site" binding has been obtained by means of refractometry [2], dilatometry [3, 4], density [5] and ultrasonic absorption [6–8] measurements. All of these techniques are sensitive to the total volume change occurring upon binding, owing to the release of electrostricted water molecules from the hydration shells of counterions and polyion charged sites. However, both the interpretation of this volume change in terms of numbers of released water molecules and the evaluation of the contributions of counterion and of polyion site to the total volume change are difficult. Another interesting and yet unanswered question is that of the distribution of condensed counterions between those bound with and without dehydration.

Clearly, such problems can be dealt with only by means of techniques such as density, NMR and ultrasonic absorption which are sensitive to the local environment of the ions, i.e., to their state of hydration and/or to the exchange of counterions between the different states of hydration. The use of density for such studies has been already reported [3–5]. On the other hand, NMR measurements using a paramagnetic ion like Co^{2+} provide a powerful way of looking at the state of hydration of counterions with and without added polyions [10–12]. Indeed in Co^{2+} solutions, the water proton signal appears as a single narrow line due to the fast exchange between free and coordinated water and to the fast electronic relaxation time. This line is shifted relative to pure water proportionally to the molality of the solution and to the number of water molecules present in the first hydration sphere of the ion. Replacing some of these water molecules by other ligands will cause a decrease of the shift. Some authors [11, 12] have taken this change in shift as directly proportional to the number of displaced water molecules. Shift measurements can therefore be used to evaluate the loss of electrostricted water molecules from the counterion hydration shell alone in the binding process.

The purpose of this paper is to report the results of measurements of density changes and water proton chemical shift changes upon additions of $CoCl_2$ to the

tetramethylammonium (TMA) salts of a series of polyelectrolytes. It will be shown that, in favorable cases, the combination of density and NMR data can be used to obtain the individual contributions of counterions and polyion sites to the total volume change upon binding and the number of Co^{2+} ions bound with loss of exchangeable water molecules. The usefulness of these last data for the quantitative interpretation of the excess ultrasonic absorption due to site binding of cobalt ions in cobalt-polyphosphate solutions is emphasized in a following paper [14] where values of the fractions of site bound Co^{2+} in different states of hydration are reported.

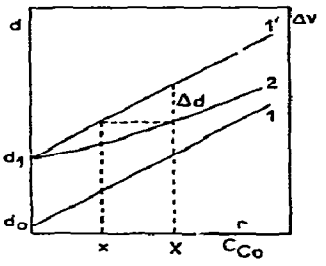


Fig. 1. Calculation of ΔV and $\bar{q}$ from the variation of density (left scale) and chemical shift (right scale) as a function of the concentration of added Co^{2+}.

2. Experimental

2.1. Materials

Measurements have been performed on polyphosphate (PP), polyacrylic acid (PAA), carboxymethylcellulose (CMC) with substitution degrees (SD) 0.98, 1.3, 2.1 and 2.65 and an alternated copolymer of maleic acid and methylvinylether (MA-MVE). The origin of these samples, the purification procedures and the preparation of the stock solutions have been previously reported [5–9].

In most of the measurements, the starting materials were the TMA salts of the above polyacids. All TMA–polysalt solutions had concentrations between 0.05 and 0.1 equivalent/ℓ. TMA was used as a reference counterion as its interaction with polyions gives rise to a negligible volume change [6–9] owing to its large ionic radius and weak hydrophobic and electrostrictive effects on surrounding water molecules [15, 16].

For each polysalt, a series of solutions was prepared by adding both water and aliquots of a standardized $CoCl_2$ stock solution to the polysalt solution. In each case, the ratio of water to $CoCl_2$ stock solution was chosen to keep the final polyelectrolyte concentration c constant and to obtain the desired value of the ratio $r = C_{Co}/c$ where C_{Co} is the concentration of Co^{2+} ions in the polyelectrolyte solution in equivalent/ℓ.

2.2. Density measurements

The densities were measured at 25°C by means of a digital precision densitometer DMA 02 [5, 17]. The total volume change ΔV upon binding of Co^{2+} ions by polyions was obtained as follows. In fig. 1, curve 1 gives the density of $CoCl_2$ solution as a function of C_{Co}. (Curve 1 has been found to be linear in the concentration range investigated; its slope yields for $CoCl_2$ an apparent molal volume of 11 cm^3/mole, as compared with the value 11.6 cm^3/mole listed by Millero [18].) Curve 2 represents the variation of the density d of the TMA-polysalt solution under study upon additions of $CoCl_2$ at constant c. Curve 1′ is obtained from curve 1 by a translation along the d-axis of a quantity $d_1 - d_0$ equal to the difference between the densities of the TMA-polysalt solution and of pure water.

At any value of r the difference Δd between curves 1′ and 2 has been attributed to the binding of Co^{2+} by the polyions. The volume change upon binding per equivalent of added Co^{2+} is then given by

$$\Delta V = 10^3 \times \Delta d/crd, \qquad (1)$$

which assumes that conformational changes undergone by polyions upon $CoCl_2$ additions result in negligible volume changes. This assumption is supported by the results of dilatometric studies [3]. Eq. (1) also assumes that the increase of ionic strength upon $CoCl_2$ additions has a negligible effect on ΔV. Indeed it has been shown [19] that the variation of the volume change of ionic reactions with ionic strength is negligible at low ionic strength (< 0.1 N) such as used in this work. The results given in section 3.1 of this paper also support this assumption.

2.3. NMR measurements

Spectra were recorded on a Varian HA 100 spectrometer at a temperature of about 30°C, i.e., slightly above that at which density measurements were performed. This difference should not matter as it is known [18, 20] that hydration numbers as obtained from NMR and apparent molal volumes show negligible changes in such a narrow temperature range. A precision of about 0.4 Hz of the resonance frequency measurements has been obtained by the use of a frequency synthesizer to sweep the audio frequency.

The chemical shift of water was measured using 2,2-dimethyl-2-silapentane-5-sodium sulfonate (DSS) as an internal reference at a concentration of about 1%, small enough not to disturb appreciably the ionic concentrations but insufficient for locking the magnetic field. A capillary of benzene was used for this purpose. The frequency difference between the benzene lock and the DSS reference signal, which is a direct measure of the magnetic susceptibility of the solution, allows to check that Co^{2+} suffers no change of its apparent magnetic moment upon binding.

Assuming the same proportionality between the shift of water protons $\Delta\nu$ and the number q of water molecules in the first hydration shell, irrespective of the state of binding, a mean hydration number $\bar{q}$ can be calculated from the plot of $\Delta\nu$ as a function of the Co^{2+} concentration in the presence and absence of polyelectrolyte by a procedure very similar to the above calculation of ΔV. Curves 1′ and 2 of fig. 1 now represent the shifts $\Delta\nu$ respectively in absence and presence of polyelectrolyte and the value of $\bar{q}$ at $r = X/c$ is related to the difference between the concentration X and x corresponding to a same shift respectively on curves 2 and 1′. Knowing that the coordination number of Co^{2+} is 6 [20] one has

$$(6 - \bar{q})/6 = (X - x)/X. \qquad (2)$$

At any r, $\bar{q}$ is clearly related to the equilibrium between the different binding states of the Co^{2+} counterions in the same way as ΔV. However, divalent ions are known to bind very strongly on polyions. Therefore the value of $\bar{q}$ at low r represents the value associated with site bound ions and the initial slope of the plot $6(X-x)/c$ versus r is equal to the number n of water molecules released from the inner hydration shell upon binding ($n \leqslant 6$).

This interpretation rests strongly on the assumption of constant proportionality between the shift and the hydration number which has already been shown to hold in some specific cases of replacement of part of the water by another ligand in the first hydration shell of octahedrally coordinated Co^{2+} [11]. The proportionality constant contains two terms arising from (i) the Fermi contact due to the density of unpaired spin at the resonating nucleus, (ii) the pseudo-contact shift resulting from the dipole interaction between the electronic spin with an anisotropic g factor and the nuclear moment. The total shift is given by [13]

$$\frac{|\Delta\nu|}{\nu} = \frac{1}{3} pq \frac{S(S+1)}{kT} \left[\frac{\beta}{\gamma_H} Ag + \frac{1}{9}\beta^2 (g_{\parallel} - g_{\perp})(g_{\parallel} + 2g_{\perp}) \left\langle \frac{3\cos^2\theta - 1}{r^3} \right\rangle_q \right]. \qquad (3)$$

where p is the molality of the solution, S the electron spin quantum number, A the isotropic hyperfine coupling constant between proton and electron, γ_H the proton magnetogyric ratio, $g_{\parallel}$ and $g_{\perp}$ the components parallel and perpendicular to the symmetry axis of the molecule of the electronic Landé tensor of trace $3g$, β the Bohr magneton, θ and r the parameters defining the position of a water proton in the coordinates of the paramagnetic ion principal axis, the average being taken over all water molecules of the first hydration shell.

A constant proportionality upon binding supposes that both (i) A remains constant when part of the water of the hydration shell is replaced by another ligand, (ii) the pseudo-contact term remains negligibly small compared to the contact term as it is, due to the symmetry, for a complete hydration shell. The validity of this second assumption is questionable in the case of partial replacement of water. Indeed, calculations using the results of Luz and Shulman [11] for A and of La Mar [21] for the g-tensor anisotropy, show that the negligible influence of the pseudo-contact term on the shift of the coordinated water molecules is due to the averaging of the factor $(3\cos^2\theta - 1)$. Partial replacement by polyion ligands can, however, increase this contribution. Evidence for the pseudo-contact term is given by the shift (positive or negative) that we have observed for the TMA proton signals upon addition of Co^{2+} ions to TMA–polysalt solutions [10, 22]. Assuming a cylindrical polyelectrolyte model this effect is

understood in term of Manning's condensation of part of the TMA^+ ions in a cylindrical shell: the motion of TMA^+ ions in this limited volume does not average the geometric factor to zero. It should be theoretically possible, adopting a radial distribution of concentrations in agreement with the solution of the Poisson–Boltzmann equation to calculate the average factor $\langle(3\cos^2\theta - 1)/r^3\rangle$ for TMA ions. The TMA experimental shift could then be used to determine the pseudo-contact shift at any position, especially at the location of the coordinated water. Rough calculation, starting from the maximum shift observed in the case of TMA–PP (fig. 3) shows that in nonfavorable cases pseudo-contact interaction could contribute to half the total shift observed in Co^{2+} solution with polyelectrolyte. But the result is so sensitive to the details of the model (radius of the polyion, position of the paramagnetic center and orientation of its principal axis) that a general conclusion cannot be reached in this way concerning the role of the pseudo-contact term. However, the consistency of the density data with numbers of released water molecules obtained below in the case of PP and MA-MVE, neglecting the possible change of proportionality constant in eq. (3) upon binding, indicates, in this case, a negligible contribution of the pseudo-contact shift.

3. Results and discussion

As an illustration, figs. 2 and 3 give the variation of

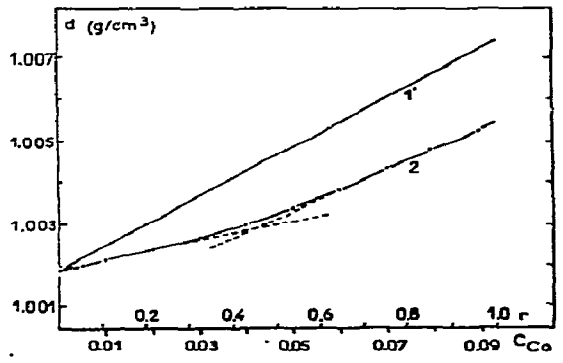


Fig. 2. Variation of the density of a TMA–PP solution upon addition of $CoCl_2$. Curve 1' represents the density of a pure $CoCl_2$ solution.

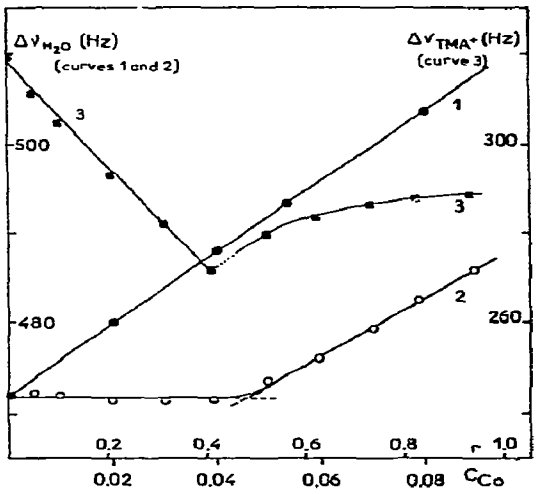


Fig. 3. Variation of the chemical shift (i) of water protons in solutions of $CoCl_2$ in absence (curve 1) and presence (curve 2) of TMA–PP, (ii) of TMA^+ protons (curve 3).

density d and chemical shift $\Delta\nu$ of a TMA–PP solution upon addition of $CoCl_2$, together with the reference curves as explained in section 2 (see fig. 1). The upper part of fig. 3 shows the variation of the shift of TMA protons. For all of the polyelectrolytes studied in this work the data have been plotted as function of r according to eq. (1) for ΔV (fig. 4) and eq. (2) for $6(X-x)/c$ (fig. 5). The results relative to PP and MA-MVE will be discussed first, as their interpretation is easier than for CMC's and PAA.

3.1. PP and MA-MVE

For polyphosphate (fig. 3) $\Delta\nu$ is independent of Co^{2+} concentration for $0 < r < 0.4$. For the region $0.6 < r < 1$, $\Delta\nu$ varies linearly with C_{Co} with a slope almost equal to that of the reference line. The two linear segments intersect at $r \approx 0.5$. A break in the variation of density at $r \approx 0.5$ can also be distinguished in fig. 2. In this simple case, these breaks can be interpreted as due to a nearly stoichiometric site binding of 0.5 equivalent Co^{2+} with complete loss of all rapidly exchangeable water molecules in the counterion hydration shell. The limiting value of r is smaller than the fraction 0.86 of condensed counterion calculated according to Manning's theory [1] for PP and a divalent counterion.

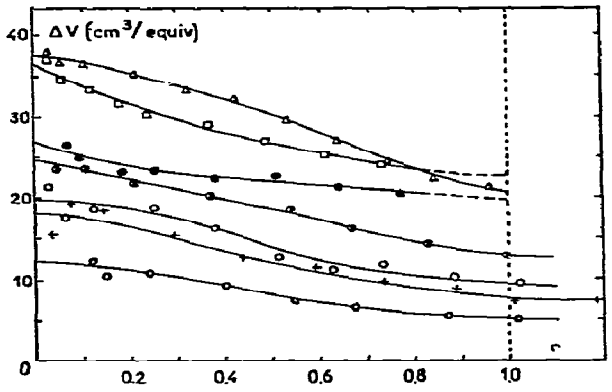


Fig. 4. Volume change per equivalent of added Co^{2+} to TMA–polysalt solutions as a function of r: (△) PP, (□) MA-MVE, (●) PAA, (◑) CMC DS 2.65, (○) CMC DS 2.1, (+) CMC DS 1.3, (◐) CMC DS 0.98.

The alternate copolymer MA-MVE displays a behaviour quite similar to PP but with a weaker site binding constant at higher r as can be seen in fig. 5. The curve relative to MA-MVE has an initial slope equal to 6 as for PP but departs more rapidly from its initial behaviour and precipitation occurs at large r before the horizontal asymptot is reached. It is therefore impossible to determine precisely the stoichiometry of site binding, the limiting value being in any case larger than $r = 0.5$.

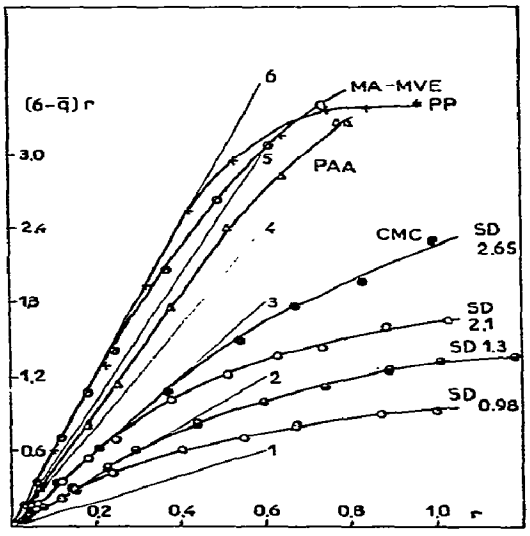


Fig. 5. Variation of $6-\bar{q}$ upon additions of $CoCl_2$ to solutions of TMA–polysalts.

Table 1

Polyelectrolyte	ΔV° (cm³/equiv.)	n	ΔV_p° (cm³/equiv.)	$-\phi_{E,p}$ (ref. [5])
PP	37.7	6.0	19.6	–
MA-MVE	36.5	6.0	18.4	19.9

Assuming a complete loss of hydration water upon binding, the total volume change ΔV^0 obtained from the extrapolation of ΔV to $r = 0$ (fig. 4, table 1) can be split into the individual contributions $\Delta V^0_{Co^{2+}}$ and ΔV_p^0 of the counterions and of the polyion charged sites respectively. Indeed, $\Delta V^0_{Co^{2+}}$ can be assumed to be equal to $-\frac{1}{2}\phi_E^0$, where $-\phi_E^0$ is the contribution of electrostriction to the apparent molal volume $\phi^0_{Co^{2+}}$ of Co^{2+} at infinite dilution. ϕ_E^0 can be written as the difference between $\phi^0_{Co^{2+}}$ and a geometric term $\phi_G^0 = 4.7_5\, r^3_{Co^{2+}}$ [23], where $r_{Co^{2+}}$ is the radius of the ion. Therefore

$$\Delta V^0_{Co^{2+}} = -\tfrac{1}{2}\phi_E^0 = -\tfrac{1}{2}(\phi^0_{Co^{2+}} - 4.7_5 r^3_{Co^{2+}}). \qquad (4)$$

With $\phi^0_{Co^{2+}} = -34$ cm³/mole [18, 23] and $r_{Co^{2+}} = 0.78$ Å one obtains $\Delta V^0_{Co^{2+}} = 18.1$ cm³/equiv. Subtracting this quantity from ΔV^0 yields the values of ΔV_p^0 listed in table 1 for PP and MA-MVE.

ΔV_p^0 can be identified with $-\phi_{E,p}$, contribution of electrostriction to the apparent molal volume of the polyion. For MA-MVE, $\phi_{E,p}$ has been obtained in a previous study [5] and compares extremely well with ΔV_p^0 as can be seen in table 1. The consistency of these two results gives a firm basis to the assumptions on which we have based our evaluation of the number of released water molecules from chemical shift measurements.

The comparison of NMR and density data at large r brings further interesting information. The average value of ΔV_p can be calculated from ΔV^1 and $\bar{q}^1$ at $r = 1$ according to

$$\overline{\Delta V}_p = \Delta V_p^1 + \Delta V^0_{Co^{2+}}(6 - \bar{q}^1)/6. \qquad (5)$$

For polyphosphate, fig. 4 shows that $\Delta V^1 = 20.5$ cm^3/mole* and application of eq. (5) yields $\overline{\Delta V}_p = 17.1$ cm^3/equiv. compared to $\Delta V_p^0 = 19.6$ cm^3/equiv. at $r = 0$. A similar calculation for MA-MVE yields 16.6 cm^3/equiv. for $\overline{\Delta V}_p$ compared to 18.4 cm^3/equiv. for ΔV_p^0. In both cases the precision is not sufficient for determining the decrease of electrostriction with the decrease in linear charge due to site binding. It must be emphasized that fig. 4 shows a decrease of ΔV_p for PP and MA-MVE even in the region of stoichiometric binding, likely because the electrostriction per site decreases upon binding [2].

3.2. Other polyelectrolytes

The results of density and chemical shift measurements have been analysed by means of eqs. (1) and (2). The results of these calculations are given in fig. 5 and table 2.

Although the average charge spacing is about the same (~ 2.5 Å) for PP, MA-MVE and PAA, the volume change at low r is smaller for PAA. This is in line with the NMR results (see fig. 5) which indicate only a partial dehydration of Co^{2+} upon binding on PAA. In fact, the number of released water molecules is a fractional number (4.5). This cannot be due to an equilibrium between different species since the variation of $\Delta\nu$ with r is linear up to $r \approx 0.5$. More likely the assumptions about the proportionality between $\Delta\nu$ and hydration number are not, in this case, strictly fulfilled. The value of n for PAA is probably overestimated since ΔV_p^0 is smaller than $-\phi_{E,p}$ (see table 2).

For the four CMC samples an additional difficulty arises from a possible modification of the binding site, as the DS is increased. This is indicated by the fact that the initial slopes in fig. 5 give $n = 2$ for DS 0.98 and 1.3 and $n = 3$ for DS 2.1 and 2.65. The values of ΔV_p^0 have been calculated assuming that each water molecule released by a cobalt ion contributes $18.1/6 \approx 3$ cm^3/mole to the total volume change. The results given in table 2 are in relatively good agreement with the values of $-\phi_{E,p}$ of ref. [5] for CMC's DS 0.98 and 2.65. The agreement is less satisfactory for the samples with DS 1.3 and 2.1.

Table 2

Polyelectrolyte	ΔV^0 (cm^3/equiv.)	n	ΔV_p^0 (cm^3/equiv.)	$-\phi_{E,p}$ (ref. [5])
PAA	27.2	4.5	13.6	17.6
CMC DS 2.65	24.7	3.0	15.7	17.1
CMC DS 2.1	19.8	3.0	10.8	8.2
CMC DS 1.3	18.0	2.0	12.0	8.1
CMC DS 0.98	12.0	2.0	6.0	5.1

The results listed in table 1 show that for PP and MA-MVE, ΔV_p^0 is close to $\frac{1}{2}\Delta V^0$. The same remark holds for the results of table 2 and also for those at $r = 1$ if we split ΔV^1 into ΔV_p^1 and $\Delta V_{Co^{2+}}^1$ by assuming an average volume change of 3 cm^3/mole per released water molecule. For instance for PP and CMC DS 0.98 $\Delta V^1 = 20.8$ and 5.5 cm^3/mole respectively and $\Delta V_p^1 = 10.3$ and 2.7 cm^3/mole. The agreement between $\frac{1}{2}\Delta V$ and ΔV_p is too general to be simply fortuitous. It may well indicate that counterion and polyion charged site contribute almost equally to the total volume change.

4. Conclusions

NMR shifts measurement appears to show that at low values of r, the binding of Co^{2+} by PP and MA-MVE is accompanied by the release of all water molecules from the inner hydration shell of Co^{2+}. This can be used to evaluate the individual contributions of counterions and polyions charged sites to the total volume change upon binding obtained from density measurements. Shift measurements also give an estimation of the fraction of Co^{2+} ions which can be bound with loss of exchangeable water molecules.

* A separate determination of ΔV^1 can be deduced from the apparent molal volumes at infinite dilution of TMA–PP and Co–PP, according to [5]

$$\Delta V^1 = \tfrac{1}{2}(\phi^0_{Co-PP} - \phi^0_{Co^{2+}}) - (\phi^0_{TMA-PP} - \phi^0_{TMA^+}).$$

All values are known from previous studies [5, 18, 23], except ϕ^0_{Co-PP} which has been measured in this work and found equal to 38.6 cm^3/mole. The above equation then yields $\Delta V^1 = 20.8$ cm^3 in excellent agreement with the value 20.5 cm^3 calculated from fig. 4 at $r = 1$. This indicates (i) that the binding of Co^{2+} is not affected by the presence of a small excess of TMA^+, and (ii) that the binding of TMA^+ is accompanied by a negligible volume change.

References

[1] G. Manning, J. Chem. Phys. 51 (1969) 924, 933 and 3249; Biopolymers 9 (1970) 1543.
[2] A. Ikegami, J. Polymer Sci. A2 (1964) 907; Biopolymers 6 (1968) 431.
[3] U.P. Strauss and Y. Po Leung, J. Am. Chem. Soc. 87 (1965) 1476.
[4] A. Begala and U.P. Strauss, J. Phys. Chem. 76 (1972) 254.
[5] C. Tondre and R. Zana, J. Phys. Chem. 76 (1972) 3451.
[6] C. Tondre and R. Zana, IUPAC Symposium Abstracts, Leiden (1970) vol. I, p. 387.
[7] R. Zana, C. Tondre, M. Rinaudo and M. Milas, J. Chim. Phys. Physicochim. Biol. 68 (1971) 1258.
[8] C. Tondre and R. Zana, J. Phys. Chem. 75 (1971) 3367.
[9] C. Tondre and R. Zana, Proc. NATO Inst. Advan. Study: Charged and Reactive Polymers, Forges les Eaux (1972), in press.
[10] P. Spegt and G. Weill, Compt. Rend. Acad. Sci. (Paris) C274 (1972) 587.
[11] Z. Luz and R.G. Shulman, J. Chem. Phys. 43 (1965) 3750.
[12] J.F. Hinton and E.S. Amis, Chem. Rev. 67 (1967) 367.
[13] D.R. Eaton and W.D. Phillips, Advances in magnetic resonance, Vol. I (Academic Press, New York, 1965).
[14] R. Zana and C. Tondre, submitted for publication.
[15] M. Eigen, Pure Appl. Chem. 6 (1963) 97.
[16] J.E. Desnoyers, R.E. Verall and B.E. Conway, J. Chem. Phys. 43 (1965) 243; 75 (1971) 3031.
[17] H. Stabinger, H. Leopold and O. Kratky: Digital Densimeter DMA 02, Institute for Physical Chemistry, University of Graz, Austria.
[18] F. Millero, Chem. Rev. 71 (1971) 147.
[19] S. Katz and J.E. Miller, J. Phys. Chem. 75 (1972) 1120.
[20] S. Lincoln, Coord. Chem. Rev. 6 (1971) 309.
[21] G.N. La Mar, J. Chem. Phys. 41 (1964) 2992.
[22] G. Weill and P. Spegt, Proc. NATO Inst. Advan. Study: Charged and Reactive Polymers, Forges les Eaux (1972), in press.
[23] R. Zana and E. Yeager, J. Phys. Chem. 71 (1967) 521 and 4241.

BIOPHYSICAL CHEMISTRY 1 (1973) 62–72. © NORTH-HOLLAND PUBLISHING COMPANY

SOME HYDRODYNAMIC AND OPTICAL PROPERTIES OF POLYRIBONUCLEOTIDES*

G. CHI CHEN and JEN TSI YANG

Department of Biochemistry and Biophysics and Cardiovascular Research Institute, University of California, San Francisco 94143, USA

Received 2 July 1973

The size and shape of four polyribonucleotides were studied by sedimentation and viscosity measurements. The results were correlated with their conformations based on optical activity studies. Polyriboadenylic acid in acidic solution dimerizes into two double-stranded forms, one with half-protonated bases and the other fully protonated. The fully protonated form is somewhat less asymmetrical than the half-protonated form. Polyriboguanylic acid in alkaline solution (near the second pK_a of guanine) undergoes a time-dependent disaggregation and the final form shows little base stacking. In acidic solution, it demonstrates a reversible transition near the first pK_a of guanine but without evidence of disaggregation. Polyribocytidylic acid undergoes a transition upon half protonation of the bases, but its molecular weight remains unchanged with pH. The results suggest that this polymer assumes a hairpin structure in acidic solution. Polyribouridylic acid has some degree of base stacking at room temperature. A transition to a hairpin structure occurs at low temperature.

1. Introduction

Synthetic polyribonucleotides can serve as simplified models of the structure of ribonucleic acids. It is known that poly(A)‡, poly(G), poly(C), and poly(U) can adopt different conformations under various experimental conditions such as changes in pH, solvent, and temperature. Structural findings in X-ray diffraction studies of fibers drawn from concentrated solutions of polyribonucleotides are often applied to studies of these polymers in dilute solution. The findings, however, may not necessarily be applicable, as we will show in the case of poly(C) in acidic solution.

In the present studies we analyzed the hydrodynamic properties (viscosity and sedimentation) of the four polyribonucleotides in dilute solutions. Our interest was in the relation of conformational changes resulting from variations in pH or temperature to the size and shape of these polymers. Some ORD and CD results are also presented for the purpose of discussion, even though voluminous data on this subject have been reported in the literature.

* Taken in part from a Ph.D. dissertation by G.C.C., University of California, San Francisco (1972).

‡ Abbreviations used in this work: poly(A), polyriboadenylic acid; poly(G), polyriboguanylic acid; poly(C), polyribocytidylic acid; poly(U), polyribouridylic acid; ORD, optical rotatory dispersion; CD, circular dichroism.

2. Experimental

2.1. Materials

Poly(A), poly(G), poly(C), and poly(U) were purchased from Miles Laboratory, Inc. and Schwarz Bio-Research, Inc. All chemicals were of reagent grade. Double-distilled water was used.

All polyribonucleotides were deproteinized by phenol extraction [1], and stored at −20°C before use. Solutions were prepared in 0.1 M NaCl and dialyzed exhaustively against appropriate buffers at 4°C. The dialysis tubing had been successively boiled in water, 0.001 M EDTA, 0.001 M HCl, and water. Solvents were filtered through Millipore type PH filters (pore size 0.3μ) before use and polymer solutions through type HA (0.45μ), SM (5μ), or SC (8μ) filters. The pH of the solutions was measured with a Radiometer 25 pH meter equipped with an expanded scale, which had been calibrated against standard buffers.

The polyribonucleotide concentrations were determined spectrophotometrically with a Zeiss PMQII spectrophotometer. We used the following molar absorptivities, $\epsilon_{max} \times 10^{-3}$ in ℓ mole^{-1} cm^{-1}, per nucleotide residue on the basis of phosphorus analysis [2, 3]: poly(A), 10.1 in 0.08 M NaCl and 0.02 M sodium citrate plus citric acid (pH 7.0); poly(G), 10.3 in 0.01 M sodium citrate–HCl buffer (pH 7.0) or 9.8 in 0.1 M NaCl–0.05 M sodium cacodylate buffer (pH 7.0); poly(C), 6.2 in 0.08 M NaCl and 0.02 M sodium citrate plus citric acid (pH 6.2) or 7.1 (pH 4.8); poly(U), 8.7 in 1 M KF (pH 7.0).

2.2. Methods

2.2.1. Viscosity

Viscosity measurements were made in two suspension-type Ubbelohde two-bulb viscometers at 25 ± 0.05°C. The flow times for water were about 1000 and 270 seconds for the lower bulbs and 700 and 200 seconds for the upper bulbs. The intrinsic viscosity, $[\eta]$, was determined from the Huggins equation:

$$\eta_{sp}/c = [\eta] + k'[\eta]^2 c.$$

2.2.1. Sedimentation velocity and equilibrium

Sedimentation measurements were carried out in two Spinco model E ultracentrifuges, one equipped with a schlieren optics and the other with an absorption optics and a photoelectric scanner. The wavelength of light was set at 265 nm for the absorption optics. The AN-D and AN-F rotors were used for speeds higher than 10000 rpm and the heavy AN-J rotor for lower speeds.

2.2.3. Optical activity

ORD was measured on a Cary 60 spectropolarimeter and CD on a Durrum–Jasco J-10 spectropolarimeter, both under constant nitrogen flush. The data are expressed in terms of mean residue rotation, $[m]$, and mean residue ellipticity, $[\theta]$, with dimensions in deg cm^2 dmole^{-1}.

3. Results and discussion

3.1. Polyriboadenylic acid

3.1.1. Viscosity and sedimentation

Fig. 1 shows the change in intrinsic viscosity of poly(A) with pH in acidic solution. The intrinsic viscosity increased abruptly near the pK_a of the base in the polyribonucleotide (5.87 in 0.1 M NaCl [4]), reached a plateau between pH 5.6 and 5.2, and then dropped sharply to another plateau near pH 5.0–4.8. Holcomb and Tinoco [3] have studied the temperature effect on the specific viscosity of poly(A) in neutral and acidic pH's. Replotting their data at a concentration of 0.024% at four pH's also led to a bell-shaped curve; however, we could not explain their finding that the specific viscosity at pH 4.6 approaches zero, which they attributed to extensive aggregation. We studied two batches of poly(A), both of which demonstrated the same profile. Also plotted in fig. 1 are the $[m]_{290}$ and $[\theta]_{275}$ of the polymers against pH. Their profiles parallel the hydrodynamic measurements, suggesting that the poly(A) molecule undergoes two conformational transitions in acidic solution, one at pH 5.7–5.9 and the other at pH 5.0–5.2 at 0.1 M salt concentration.

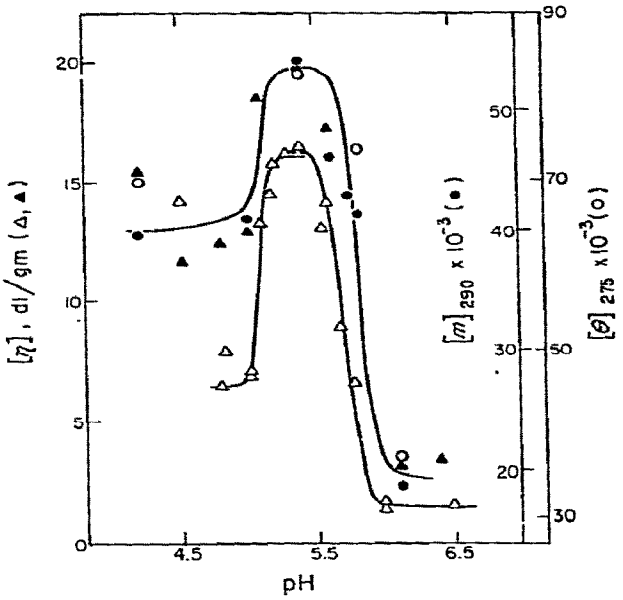


Fig. 1. Variations in the intrinsic viscosities and the CD and ORD extrema of poly(A) with pH in 0.08 M NaCl and 0.02 M sodium citrate plus citric acid at 25°C. All symbols, except the open triangles, represent measurements on a single lot.

X-ray diffraction studies of poly(A) fibers drawn from a concentrated acidic solution indicated that the

Table 1

Hydrodynamic properties and molecular weights of poly(A) at 25°C[a]

Solvent pH	$[\eta]$ (dl/g)	s^0 (S)	$M_{sd.eq.}$ [b] ($\times 10^{-5}$)	M_β [c] ($\times 10^{-5}$)	$M_\Phi^{1/3}\rho^{-1}$ [c] ($\times 10^{-5}$)	L_η(rod) [d] (Å)	L_s(rod) [e] (Å)
6.5	3.5	8.5	3.1	3.3	4.8	1800	1700
6.1	3.3	–	–	–	–	–	–
5.6	17.3	8.7	–	6.3	10.7	3800	3200
5.4	19.7	9.7	5.9	7.9	13.5	4000	2900
5.0	12.5	9.3	5.6	6.3	10.7	3500	2900
4.8	12.5	10.3	5.9	7.3	12.0	3500	2600

a) Schwarz lot No. w-2065; solvent, 0.08 M NaCl and 0.02 M sodium citrate plus citric acid.
b) $\bar{v}$ = 0.546 ml/g [8] was assumed unchanged with pH.
c) Estimated from the relation: $\beta(\text{or } \Phi^{1/3}\rho^{-1}) = N_0 s^0 [\eta]^{1/3}\eta_0/M^{2/3}(1-\bar{v}\rho)$.
d) $L(\text{in Å}) = 6.82\,([\eta]M)^{1/3}\,(\rho^2/\nu)^{1/3}$.
e) $L(\text{in Å}) = 1.76 \times 10^{-17}\,[M(1-\bar{v}\rho)/\eta_0 s^0][p^{2/3}/(f/f_0)]$.

structure is a double-stranded helix [5]. The same structure is believed to exist in acidic solution of poly(A) under proper experimental conditions. This helix is thought to be more extended than the single-stranded monomer, which could account for the sharp increase in the intrinsic viscosity near pH 5.9 (fig. 1). The sedimentation coefficients of poly(A) rose during the acid-induced transition (table 1); however, the changes were not as drastic as those of $[\eta]$. An increase in s^0 indicates an increase in the molecular weight of the polymer, or a decrease in its frictional coefficient, or both. In the present case the increase was largely due to the dimerization of poly(A) in acidic solution (see below). According to potentiometric titration [4], the two acid-induced transitions (fig. 1) probably resulted in a half-protonated and a fully protonated double-stranded helix. With full protonation of the bases poly(A) has a net charge of zero; therefore, it could be less asymmetrical than the half-protonated form, causing the intrinsic viscosity to drop below pH 5. The occurrence of aggregation at high concentration or low pH's made precise viscosity measurements difficult. The solutions were filtered through Millipore filters before use. The extent of aggregation was checked by the loss of absorbance of the filtered solutions. In spite of these difficulties, our conclusions drawn from the hydrodynamic measurements remain valid.

3.1.2. Optical activity

The ORD and CD of poly(A) have been extensively

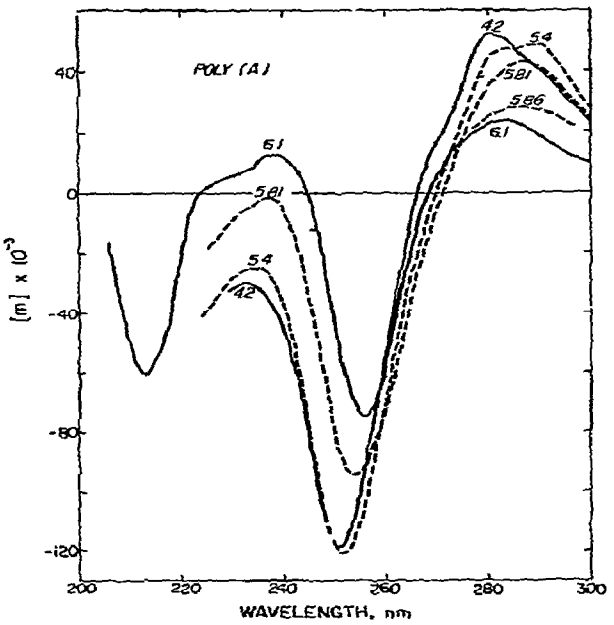


Fig. 2. The ORD of poly(A) at several pH's in 0.08 M NaCl and 0.02 M sodium citrate plus citric acid at 25°C.

described in the literature. The ORD peak near 285 nm splits into two peaks under proper experimental conditions [3, 4]; for example, the ORD of the poly(A) solution at pH 5.0 shows a double maximum, which was designated as a mixture of the two acidic forms A and B [4]. As can be seen in fig. 2, the ORD peak on the long wavelength side showed a red-shift with increasing magnitude and the trough a blue-shift with increasing magnitude when the pH was changed from 6.1 to 5.85 or 5.81. At pH 4.2 the peak reversed to a blue-shift, but with further increase in magnitude. We confirmed a double maximum at intermediate pH's, but found that its magnitude at, for example, pH 5.4, was always larger than that at pH 5.7–5.9 and smaller than that at pH 4.2, whereas Adler et al. [4] reported that the magnitude of the double peak at pH 5.0 was less than that at either pH 5.81 (B form) or 4.0 (A form). Our CD data also led to the same conclusions as the ORD. The enhancement of the multiple Cotton effects at acidic pH's indicates that the adenine bases in both acidic forms are highly stacked. This finding is in accord with the observation that both forms show strong hypochromicity (about 20%) compared with poly(A) at neutral pH.

Adler et al. [4] reported that the ratio of the twin peak, $[m]_{280}/[m]_{287}$, of poly(A) first dropped sharply at pH 5.81 and then increased continuously with decreasing pH in the range of 5.8 to 4.0. They proposed only acidic B form at pH 5.81 and a gradual conversion of the B form to the A form between pH 5.8 and 4.0. We found that the ratio $[m]_{280}/[m]_{287}$ first decreased with pH, then reached a plateau, and finally increased again at lower pH's. The changes occurred within the same range as those shown in fig. 1. We conclude that the B form is stable between pH 5.2 and 5.7 and that the A form exists below pH 5.2 under our experimental conditions.

The Cotton effects of poly(A) were reported to depend at times on the previous handling of the sample; for example, pre-heating for 5 to 10 minutes at 95–100°C could increase the magnitude of the extrema [4, 6]. We therefore studied samples obtained from two sources, Miles and Schwarz, both of which were unaffected by pre-heating at 95–100°C for 5 to 10 minutes (after correction of the increase in concentration by about 5% through evaporation). The ORD of our unheated poly(A) was comparable to the ORD of heated poly(A) as reported by Fasman and coworkers [4, 6].

3.1.3. Size and shape

We determined the molecular weight of poly(A) at several pH's either by the method of sedimentation equilibrium or by a combination of intrinsic viscosity and sedimentation coefficient which gave a reasonable estimate. The hydrodynamic properties also provide additional information about the shape of the molecules. The data in table 1 show that poly(A) dimerized when the pH of the solution was lowered from 6.5 to 5.4 or lower. The transition between the two acidic forms (fig. 1), however, was not accompanied by any change in molecular weight.

The poly(A) molecule in neutral solution is neither an exact rigid ellipsoid nor, because of base stacking, a flexible coil. Whether it can be represented by either of the two models is open to question. Combination of $[\eta]$ and s^0 leads to a β-function for a hydrodynamically equivalent ellipsoid of revolution [7]:

$$\beta = N_0 s^0 [\eta]^{1/3} \eta_0 / M^{2/3} (1 - \bar{\upsilon}\rho). \qquad (1)$$

which is a function of axial ratio, p, only, and for a prolate ellipsoid, β varies from 2.12×10^6 at $p = 1$ to 3.6×10^6 at $p = 300$. Here N_0 is the Avogadro number, η_0 the solvent viscosity, M the molecular weight of the polymer, $\bar{\upsilon}$ the partial specific volume of the polymer, and ρ the density of the solvent. We assumed $\bar{\upsilon}$ unchanged with pH and used a value of 0.546 ml/g for poly(A) [8]. The β-function is rather insensitive toward p, and this insensitivity to shape provides a simple means for estimating the molecular weight of the particles. The p, and therefore the β-value, can be estimated from $[\eta] = \nu\bar{\upsilon}/100$, where the viscosity increment, ν, defines p. If hydration is taken into consideration, according to Oncley's formula [9], the axial ratio so determined will be slightly smaller than the ratio determined without hydration. For highly asymmetrical particles neglect of hydration will amount to only a few percent difference in the estimated molecular weight. If the polymer molecules resemble flexible coils, we can use Flory's $\Phi^{1/3}p^{-1}$-function to replace β in eq. (1); experimentally $\Phi^{1/3}p^{-1}$ is usually about 2.5×10^6 [10].

As shown in table 1, the molecular weight of poly(A) at pH 6.5 estimated from the β-function agreed more closely to the molecular weight determined by the method of sedimentation equilibrium than to that based on the $\Phi^{1/3}p^{-1}$-function, suggesting that even for a single-stranded polyribonucleotide chain the ellipsoid

model is a better representation. In acidic solution, where poly(A) is believed to form a double-stranded helix, a flexible coil would be a poor model. The results in table 1 support this conclusion. The higher M_β than that of a dimer at pH 5.4 and 4.8 is probably due to the presence of some aggregation.

Eisenberg and Felsenfeld [8] determined the molecular weight of poly(A) by light scattering and by its intrinsic viscosity and sedimentation coefficient at various temperatures and concluded that the β-function changed with temperature. They questioned the common practice of using an assumed fixed β-value for the estimation of molecular weight. Indeed, there is no justification for assuming a fixed β-value, unless one adopts a flexible coil model and uses Flory's $\Phi^{1/3}p^{-1}$-function. Eisenberg and Felsenfeld [8] found that the intrinsic viscosity of poly(A) decreased with rising temperature. This observation suggests that the poly(A) molecule becomes more flexible at higher temperature and can be represented by an equivalent prolate ellipsoid of smaller axial ratio. Consequently, the β-value diminishes gradually with increasing temperature, as would be expected. In our studies we estimated the axial ratio from the intrinsic viscosity at each pH rather than from a constant β-value.

Next we estimated the length of the equivalent ellipsoid of poly(A) in acidic solution. From the definition of $[\eta]$ and s^0, the length, L, of a prolate ellipsoid (= the major axis) is given by [11]:

$$L \text{ (in Å)} = 6.82([\eta]M)^{1/3}(p^2/\nu)^{1/3} \qquad (2)$$

and

$$L \text{ (in Å)} = 1.76 \times 10^{-17} \frac{M(1-\bar{\nu}\rho)}{\eta_0 s^0} \frac{p^{2/3}}{f/f_0}, \qquad (3)$$

(f/f_0) being the frictional ratio, a function of p only. For very asymmetrical particles, the two factors involving p in eqs. (2) and (3) are rather insensitive to the axial ratio assumed. Thus, the length of a hydrodynamically equivalent prolate ellipsoid can be estimated with reasonable confidence, even though p is not known accurately. According to X-ray diffraction studies of poly(A) fibers drawn from an acidic solution [5], the helical molecule consists of two parallel intertwined chains, and successive residues are related by a translation of 3.8 Å per base pair. If the same conformation exists in acidic solution, the double-stranded helix of our poly(A) sample, which had a degree of polymerization of about 900, would have a contour length of about 3500 Å. The length of the equivalent ellipsoid of our sample as calculated from eqs. (2) and (3) (table 1) was close to this value. Such close agreement could be fortuitous; however, qualitatively, we could conclude that the two acidic forms were very extended.

Two points are worth mentioning. First, the estimated length based on viscosity measurements was greater than that based on s^0 values. The disadvantage of viscometry is that it requires a moderately high polymer concentration, whereas with absorption optics it is possible to measure the s^0 value of poly(A) in extremely dilute solution. If there was aggregation of poly(A) in the viscosity measurements, both the length and the molecular weight based on $[\eta]$ would be greater than the values without aggregation (table 1). According to eq. (2), L varies with $([\eta]M)^{1/3}$. Thus, we can still obtain a reasonable estimate of L, even though $[\eta]$ is somewhat inaccurate and M larger than the true value. For the same reason, the difference between the calculated lengths of the two acidic forms is small, despite the sharp drop in $[\eta]$ during the second transition near pH 5 (fig. 1). Second, the double helix of poly(A) would be expected to behave like a stiff chain and the length of its equivalent ellipsoid should be much less than the contour length of the helix. Since the molecular weight of our poly(A) sample was fairly low, the molecule must be stiffer than the molecule having a molecular weight of several millions. Our estimated values, however, are unexpectedly close to the contour length of the double helix. Although the effect of polydispersity on the calculated length in eqs. (2) and (3) is not known, we believe that it cannot account for the results in table 1. One plausible explanation is that the double strands of poly(A) are not exactly matched, but are overlapped, with portions of single-stranded chains protruding at both ends of the helix. Such a staggered array could conceivably increase the equivalent length of the particles.

3.2. Polyriboguanylic acid

3.2.1. Optical studies

Fig. 3 shows the CD of poly(G) in acidic solution. We used 0.01 M sodium citrate–HCl buffer because the acidic pK_a of guanine is higher at lower ionic strength, thus enabling us to follow conformational changes in a convenient pH range. Significant changes in the CD spectrum were observed below pH 3.5; the positive

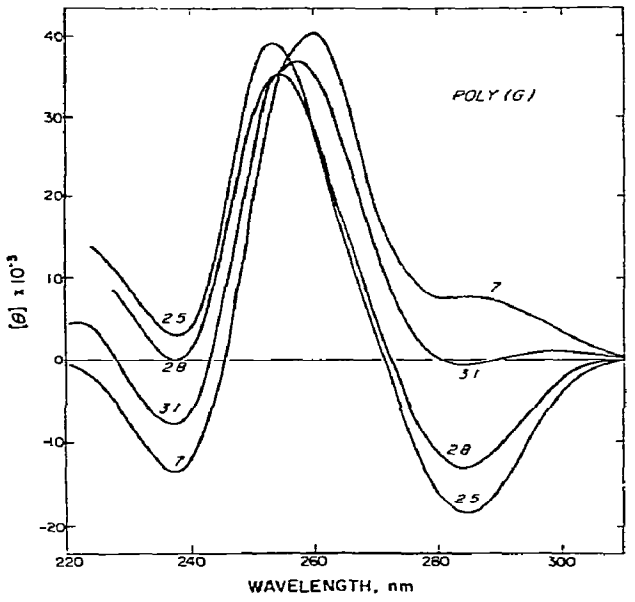


Fig. 3. The CD of poly(G) at several pH's in 0.01 M sodium citrate-HCl buffer at 25°C.

shoulder near 280–285 nm turned into a negative minimum at pH less than 3 (see also refs. [12, 13]), the negative minimum near 240 nm became positive below pH 2.8, and the maximum near 260 nm showed a blue-shift. In addition, a large negative minimum appeared near 195 nm. The corresponding ORD revealed similar changes. However, Ulbricht et al. [14] reported that in 0.15 M salt solution the rotations of poly(G) changed little as the pH was reduced to 1 and the positions of the extrema showed a red-shift.

Plotting the CD extrema of poly(G) against pH revealed a sharp transition between pH 2.5 and 3.5 with a midpoint at pH 3 under our experimental conditions (fig. 4). Since the pK_a of poly(G) in 0.01 M NaCl was 2.75 [13], this conformational change was probably caused by the protonation on N(7) of the guanine base. This acid-induced change in optical properties was reversible when the acidic poly(G) solution was dialyzed back to neutral pH.

At pH 2 in 0.01 M sodium citrate–HCl buffer, the ORD and CD of poly(G) showed no marked changes between 25 and 70°C. The magnitude of the ORD extrema decreased by about 15–20% at 90°C, and the polymer did not precipitate even at temperatures higher than 90°C. In contrast, the poly(G) molecules at pH 3.3 without added salt precipitated upon heating to 45°C. Thus, the observed changes in the CD spectra (fig. 3)

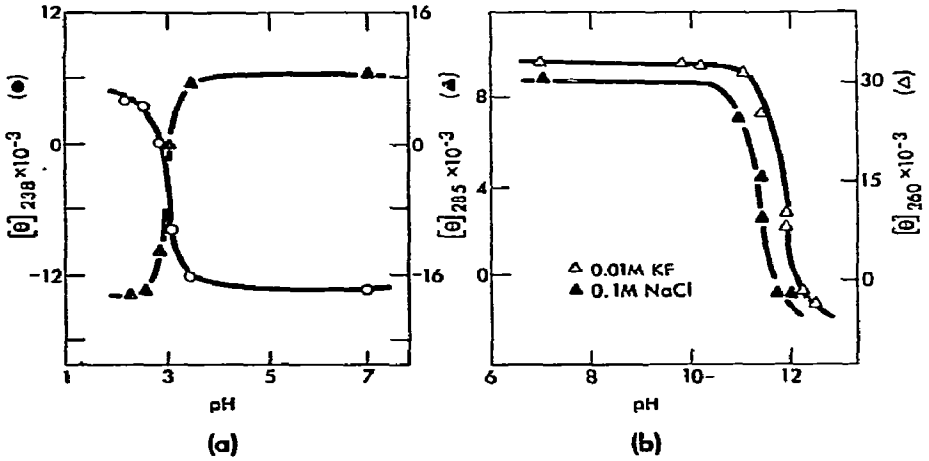


Fig. 4. Changes in the CD extrema of poly(G) with pH at 25°C. (a) Left: in 0.01 M sodium citrate-HCl buffer; (b) right: in 0.01 M KF and 0.1 M NaCl.

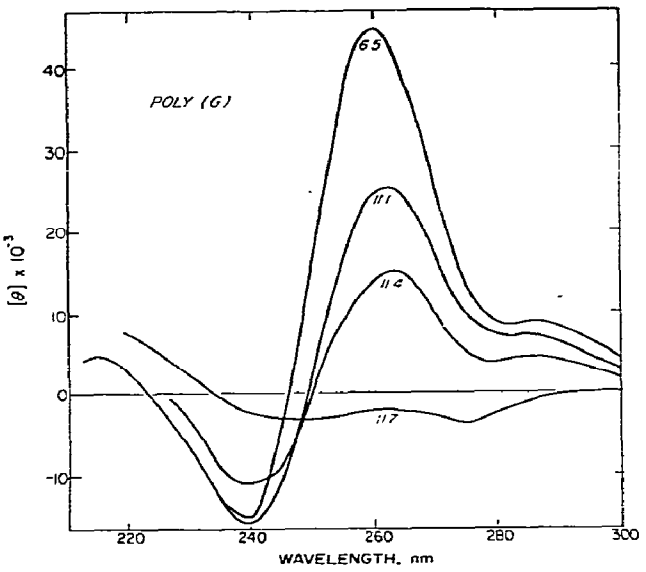


Fig. 5. The CD of poly(G) at several pH's in 0.1 M NaCl at 25°C.

could not be attributed to possible disaggregation of poly(G) at low pH's, as had been proposed by Wolfe et al. [12].

Fig. 5 shows the CD of poly(G) in alkaline solution. The general features of the spectrum at neutral pH agreed qualitatively with those reported previously [12, 15]. There was no detectable change in the CD of poly(G) up to pH 10. At higher pH's the magnitude of the Cotton effects began to decrease and at pH 11.7 the CD actually became slightly negative over the range of 230–300 nm studied (above pH 11.7 the ORD showed only small negative Cotton effects, in agreement with the finding of Ulbricht et al. [14]. Of the polyribonucleotides studied, only poly(G) in alkaline solution at pH's higher than the second pK_a of guanine had little, if any, base stacking, as revealed by loss of multiple Cotton effects.

The variation in the magnitude of the CD extrema with alkaline pH indicated a marked change in the conformation of poly(G) between pH 11 and 12 (fig. 4). The change was probably due to the deprotonation of the guanine bases and each nucleotide residue would carry two negative charges. As a result, the electrostatic repulsion would be so strong that the bases would be forced to unstack.

The conformational changes shown in fig. 5 were time dependent. The multiple Cotton effects at pH 7 could be recovered by back titration from pH 12, provided that the solution was kept at pH 12 for less than 30 minutes at room temperature or less than 3 hours at 4°C. Gel filtration of poly(G) in both neutral and alkaline (pH 12.4) solutions indicated that the polyribonucleotide aggregates gradually disintegrated after exposure to alkali and that the extent of disaggregation depended on the duration of exposure at alkaline pH.

3.2.2. Hydrodynamic studies

Both the intrinsic viscosity and sedimentation coefficient of poly(G) dropped significantly when the pH of the solution was raised from neutral to above 10 (table 2), suggesting a reduction in the size of the polymer. Like the optical properties, the s^0 values of poly(G) in alkaline solution (pH > 11) were both time and temperature dependent. The molecular weight of poly(G) was determined by a combination of sedimentation and diffusion coefficients or by use of the β- and $\Phi^{1/3}p^{-1}$-functions. The limited results (table 2) make it difficult to choose between a prolate ellipsoid of revolution and a flexible coil model. However, for our sample, the β-value happened to be close to 2.5×10^6; thus, the estimated molecular weight was little affected by the chosen model. As shown in table 2, a marked decrease in molecular weight of poly(G) did occur when the pH of the solution was raised from 7.5 to 11.3 and the decrease was even more pronounced at higher pH's, a fact indicative of disaggregation of poly(G). Fresco and Massoulie' [16] first reported that poly(G) at neutral pH tended to aggregate and form a multiple-stranded helix with extremely high thermal stability, a conclusion questioned by Pochon and Michelson [17] because of lack of evidence for more than two strands for the poly(G) complex. Michelson et al. [18] found that a 7 mg/ml solution of oligoguanylic acid was extremely viscous and suspected that the preparation of Fresco and Massoulie' [16] might be of relatively short chain length and behave differently from poly(G). However, our poly(G) sample with a relatively high molecular weight showed only moderate intrinsic viscosity at neutral pH. Our findings strongly support the possibility that poly(G) at neutral pH is in a highly aggregated form.

In acidic solution in contrast to alkaline solution, the viscosity of poly(G) was only slightly reduced and

Table 2

Hydrodynamic properties of poly(G) at 25°C

Solvent pH	$[\eta]$ (dl/g)	s^0 (S)	$D^0 \times 10^7$, (cm^2/sec)	M_{s-D} [d] ($\times 10^{-5}$)	M_β [d] ($\times 10^{-5}$)	$M_\Phi^{1/3} p^{-1}$ [d] ($\times 10^{-5}$)
Lot: M-11-08-314						
7.5 [a]	0.33	9.5	2.6	2.0	1.6	1.8
11.3 [b]	0.11	4.5			0.3	0.3
11.6 [b]	–	1.3				
Lot: 11-06-314 [c]						
7.0	0.29	8.1				
3.5	–	8.2				
2.8	–	9.0				

a) In 0.1 M NaCl plus 0.05 M cacodylate buffer.
b) In 0.15 M Na_2HPO_4 after dialysis for about 12 hours.
c) In 0.01 M sodium citrate – HCl buffer.
d) $\bar{v}$ = 0.55 ml/g [19] was assumed unchanged with pH.

the sedimentation coefficient was modestly increased. For instance, the specific viscosity of a 0.04% solution dropped from 0.34 dl/g at pH 7 to 0.30 dl/g at pH 3.1 at 25°C. Because not enough of the sample remained, we were unable to determine its intrinsic viscosity, but it was probably somewhat lower than at neutral pH. The change in s^0 value was small between pH 7 and 3.5, a region with no significant changes in optical properties. The s^0 value did increase at pH's below 3. These hydrodynamic measurements ruled out any marked change in the size and shape of poly(G) in the acid-induced transition region. As the pH was further lowered, the s^0 value continued to increase, and the solution showed significant aggregation below pH 2 as judged by the decrease in its absorbance upon filtration. Since the structure of the aggregates in both neutral and acidic solutions is not known, it is difficult to speculate on the kind of conformational changes that took place as a result of acidification. Nevertheless, our results demonstrate marked differences in the conformation and molecular shape of poly(G) in acidic and alkaline solutions.

3.3. Polyribocytidylic acid

3.3.1. Viscosity and sedimentation

Table 3 lists the intrinsic viscosities and sedimentation coefficients of poly(C) at several pH's. The $[\eta]$, $[\theta]$, and $[m]$ of poly(C) in acidic solution (fig. 6) showed a sharp transition, suggesting a marked change in conformation as well as in shape. The mid-point of the transition occurred at pH 5.7–5.9 at 0.1 M salt concentration, which is identical with the pK_a of 5.7 for

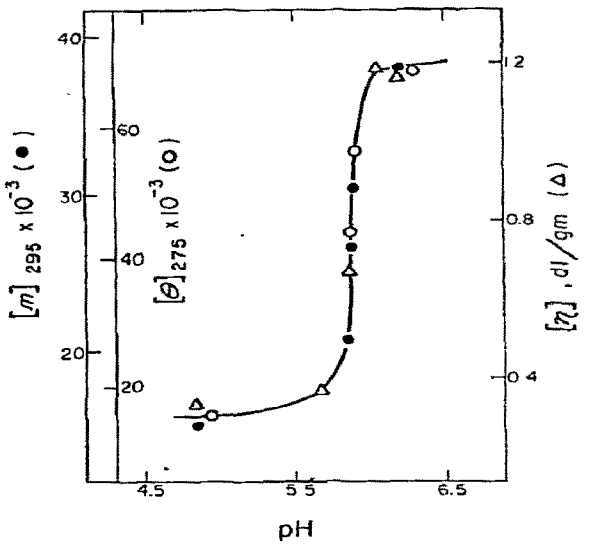


Fig. 6. Variations in the intrinsic viscosities and CD and ORD extrema of poly(C) with pH in 0.08 M NaCl and 0.02 M sodium citrate plus citric acid at 25°C.

Table 3
Hydrodynamic properties of poly(C) at 25°C [a]

Solvent [a] pH	$[\eta]$ (dl/g)	s^0 (S)	M_β [b] ($\times 10^{-5}$)	L_η(rod) [b] (Å)	L_s(rod) [b] (Å)
6.2	1.17	5.46	1.1	840	810
6.05	1.18	–			
5.87	0.66	–			
5.65	0.37	–			
4.8	0.32	7.52	1.1	500	500
3.6	–	7.47			

[a] Schwarz lot No. S-6701; solvent. 0.08 M NaCl and 0.02 M sodium citrate plus citric acid.
[b] $\bar{v}$ = 0.505 ml/g [25] was assumed unchanged with pH; for equations for β-function and for rod length, see footnote c, d and e of table 1.

the cytosine base in the polymer [20]. Both the decrease in $[\eta]$ and increase in s^0 of poly(C) in acidic solution (see also refs. [20, 21]) suggest that the molecule is less asymmetrical at acidic pH than in neutral solution.

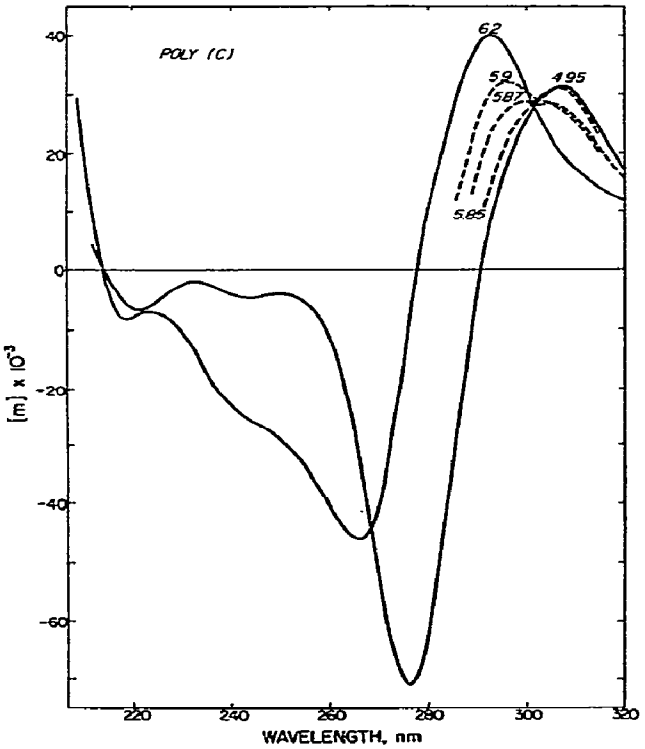


Fig. 7. The ORD of poly(C) at several pH's in 0.08 M NaCl and 0.02 M sodium citrate plus citric acid at 25°C.

3.3.2. Optical activity

Fig. 7 shows the ORD of poly(C) at various pH's. Lowering the pH from 6.2 to 4.95 produced a red-shift in the spectrum and also decreased the magnitude of the Cotton effects. Two peaks were observed in the intermediate pH's and there was an isorotational point near 300 nm, a fact indicative of the presence of only two forms, one at neutral pH and the other in acidic solution. The general features of the ORD spectrum (and CD spectrum) agreed well with those described in the literature (see, for example, ref. [22]).

Unlike poly(A), poly(C) cannot have two acidic forms, one with half-protonated bases and the other fully protonated. According to the results of X-ray diffraction studies of poly(C) fibers [23], the double-stranded helix has a positive charge on N(3) upon half protonation. In contrast, full protonation of the cytosine bases would put two positive charges on the two N(3) positions adjacent to each other and the resultant electrostatic repulsion could disrupt the double-stranded helix.

The pH-induced transition was in such a narrow range (fig. 6) that the structure in the transition region was very unstable; for example, the ORD spectrum at pH 5.85 could be shifted to that at pH 5.9 merely by shaking the poly(C) solution or allowing it to stand. However, we have no evidence that this instability near the pK_a value causes the structure of poly(C) to collapse, as reported by the disappearance of the rotations for solutions at pH slightly above the pK_a value [24].

3.3.3. Size and shape

As in the case of poly(A), we can adopt an equivalent prolate ellipsoid model for poly(C) and estimate its molecular weight at pH 6.2 and 4.8 from the β-function. With $\bar{v} = 0.505$ ml/g [25], the molecular weight of our sample was 1.1×10^5 daltons at both pH's (table 3). Unlike poly(A), the acid-induced conformational transition of poly(C) does not involve a change in the size of the polymer. With the poly(C) sample used, we found that the β-values were not too different from Flory's $\Phi^{1/3}p^{-1}$ of 2.5×10^6. Therefore, even if we had chosen a flexible coil model, the estimated molecular weight would not have differed much from that listed in table 3. Our results clearly rule out the possibility of a dimerization of poly(C) in acidic solution. A possible explanation is that the poly(C) molecule adopts an antiparallel hairpin-lime structure, although the results of X-ray diffraction studies suggest a parallel double-stranded dimer [23]. Conceivably, fibers drawn from a concentrated solution could differ in conformation from polymers in dilute solution. Furthermore, during the drawing two hairpin-like molecules could be stretched into a parallel double-stranded complex.

In neutral solution poly(C) has one negative charge per phosphate group and the molecule is thought to be rather extended, although flexible. Upon half protonation the molecule would tend to contract if it adopted a hairpin structure, which would account for the marked drop in intrinsic viscosity, as contrasted with poly(A), and the gradual increase in the sedimentation coefficient. As in the case of poly(A) we can estimate the length of an equivalent ellipsoid for poly(C) according to eqs. (2) and (3). Our sample had a degree of polymerization of about 360, so that the hairpin helix would have a maximum length of about 560 Å, assuming a translation of 3.1 Å per base pair [23]. The equivalent length of about 500 Å at pH 4.8 (table 3) supports the hairpin structure, although the helix must be bent or distorted to account for the shorter equivalent length. Even the single-stranded poly(C) at pH 6.2 appears to be very extended, since its equivalent length was about two thirds the length of a stretched polynucleotide chain.

3.4. Polyribouridylic acid

3.4.1. Optical and hydrodynamic studies

Poly(U) has little secondary structure at room temperature on the basis of hypochromicity studies [26]. However, its ORD and CD do display multiple Cotton effects, although much smaller in magnitude than those of other polyribonucleotides, suggesting the existence of some base stacking. Lowering the temperature of the solution toward 0°C increased the magnitude of the Cotton effects, in association with a significant blue-shift (5–10 nm) of the positive CD band near 265–270 nm, a fact indicative of the formation of base pairing [27]. In our studies the variations in the CD and ORD extrema of poly(U) with temperature gave a melting temperature of 8°C in 1 M KF, but in 0.1 M KF the transition was not complete even at 0°C (cf. ref. [28]).

The intrinsic viscosity of poly(U) decreased with the lowering of temperature (table 4) and the Huggins' constants were less than one, suggesting the absence of aggregation. The sedimentation coefficient, s_f^0, also decreased with decrease in temperature, but when converted to standard conditions, $s^0_{20,w}$, actually increased when the temperature was lowered. This fact, together with the decrease in $[\eta]$, suggests a smaller hydrodynamic volume or a less asymmetrical structure for poly(U) at lower temperatures.

The molecular weight estimated from the β- or $\Phi^{1/3}p^{-1}$-function was essentially unaffected by temperature, but

Table 4

Hydrodynamic properties of Poly(U) [a]

Properties	Temperature (°C)	
	25	3
$[\eta]$ (dl/g)	0.70	0.50
s_f^0 (S)	4.76	3.01
$s^0_{20,w}$ (S)	4.27	4.85
$M_{sd.eq.} \times 10^{-4}$ (at 20°C) [b]	8.0	
$M_\beta \times 10^{-4}$ b)	7.8	8.3
$M_{\Phi^{1/3}p^{-1}} \times 10^{-4}$ b)	9.5	9.3

a) Solvent: 0.08 M KF and 0.02 M sodium citrate plus citric acid at pH 6.2.

b) $\bar{v} = 0.564$ ml/g [31] was assumed unchanged with temperature.

the results based on β-function showed better agreement with the value determined directly by sedimentation equilibrium (table 4).

The temperature-induced transition of poly(U) is dependent on the nature of the salts used, as well as on their concentration. For instance, we found that 0.1 M KF was more effective than 0.1 M NaCl for inducing the thermal transition. We studied the hydrodynamic properties of poly(U) in 0.08 M KF and 0.02 M sodium citrate plus citric acid because some evidence indicates that higher salt concentrations cause the molecules to aggregate [29, 30]. The disadvantage of using low salt concentrations is that the transition is not complete even at 0°C. Nevertheless, our data are sufficient to detect any change in the molecular size of poly(U) in the transition region. The fact that poly(U) had a constant molecular weight over the temperature range studied (0–25°C) supports the current concept of poly(U), that is, a hairpin-like structure at low temperature with the polyribonucleotide chain folded back upon itself, but a single-stranded chain with little secondary structure at room temperature.

In summary, measurements of the hydrodynamic properties of the four polyribonucleotides can provide information about the size and shape of these polymers under various experimental conditions. In association with measurements of optical activities, they are useful for the characterization of the conformations and conformational changes of these polyribonucleotides as a result of changes in pH and temperature.

Acknowledgement

This work was aided by grants from the U.S. Public Health Service (GM-10880 and HL-06285).

References

[1] H. Saito and K. Muira, Biochim. Biophys. Acta 72 (1963) 619.
[2] P.S. Chen, T.Y. Toribara and H. Warner, Anal. Biochem. 28 (1956) 1756.
[3] D.N. Holcomb and I. Tinoco Jr., Biopolymers 3 (1965) 121.
[4] A.J. Adler, L. Grossman and G.D. Fasman, Biochemistry 8 (1969) 3846.
[5] A. Rich, D.R. Davies, F.H.C. Crick and J.D. Watson, J. Mol. Biol. 8 (1961) 71.
[6] B. Davidson and G.D. Fasman, Arch. Biochem. Biophys. 144 (1971) 650.
[7] H.A. Scheraga and L. Mandelkern, J. Am. Chem. Soc. 75 (1953) 179.
[8] H. Eisenberg and G. Felsenfeld, J. Mol. Biol. 30 (1967) 17.
[9] J.L. Oncley, Ann. N.Y. Acad. Sci. 41 (1941) 121.
[10] P.J. Flory, Principles of polymer chemistry (Cornell Univ. Press, Ithaca, 1953) ch. 14.
[11] J.T. Yang, Advan. Protein Chem. 16 (1961) 323.
[12] F.H. Wolfe, K. Oikawa and C.M. Kay, Can. J. Biochem. 47 (1969) 637.
[13] D. Thiele and W. Guschlbauer, Biopolymers 10 (1971) 143.
[14] T.L.V. Ulbricht, R.J. Swan and A.M. Michelson, Chem. Commun. (1966) 63.
[15] S.K. Arya, Ph.D. Dissertation, Univ. of Calif., San Francisco (1968).
[16] J.R. Fresco and J. Massoulie', J. Am. Chem. Soc. 85 (1963) 1352.
[17] F. Pochon and A.M. Michelson, Proc. Natl. Acad. Sci. U.S. 53 (1965) 1425.
[18] A.M. Michelson, J. Massoulie' and W. Guschlbauer, Progr. Nucl. Acid Res. Mol. Biol. 6 (1967) 84.
[19] H.G. Tennent and C.F. Vilbrandt, J. Am. Chem. Soc. 65 (1943) 424.
[20] K.A. Hartman Jr. and A. Rich, J. Am. Chem. Soc. 87 (1965) 2033.
[21] C. Akinrimisi, C. Sander and P.O.P. Ts'o, Biochemistry 2 (1963) 340.
[22] G.D. Fasman, C. Lindblow and L. Grossman, Biochemistry 3 (1964) 1015.
[23] R. Langridge and A. Rich, Nature 198 (1963) 725.
[24] W. Guschlbauer, Proc. Natl. Acad. Sci. U.S. 57 (1967) 1441.
[25] A. Gulik, H. Inoue and V. Luzzati, J. Mol. Biol. 53 (1970) 221.
[26] R.C. Warner, J. Biol. Chem. 229 (1957) 711.
[27] J.T. Yang and T. Samejima, Progr. Nucl. Acid Res. Mol. Biol. 9 (1969) 223.
[28] J.C. Thrierr, M. Dourlent and M. Leng, J. Mol. Biol. 58 (1971) 815.
[29] R.A. Brown, Arch. Biochem. Biophys. 115 (1966) 102.
[30] D.B. Millar and M. Mackenzie, Biochim. Biophys. Acta 204 (1970) 82.
[31] L.D. Inners and G. Felsenfeld, J. Mol. Biol. 50 (1970) 373.

BIOPHYSICAL CHEMISTRY 1 (1973) 73–86. © NORTH-HOLLAND PUBLISHING COMPANY

HUMAN PROCOLLAGEN I. AN ANIONIC TROPOCOLLAGEN PRECURSOR FROM SKIN FIBROBLASTS IN CULTURE

Edward BAŃKOWSKI * and William M. MITCHELL

Departments of Microbiology and Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA

Received 25 June 1973

Tropocollagen is derived from an extracellular precursor, procollagen. Conversion to tropocollagen is accomplished by one or more tissue proteases dependent *in vitro* on the presence of serum in the culture medium. Twenty-four hour cultures in which serum has been excluded yield an apparently undegraded precursor, procollagen I. The latter is approximately twice the size of tropocollagen, possesses an acidic pI in contrast to the alkaline pI of tropocollagen, and shares secondary structural characteristics common to tropocollagen. Procollagen I exhibits a sharp thermal transition point at 39° with a ΔT of 2° indicating that the collagenous portion of the molecule is in the triple helical configuration prior to proteolytic excision from the parent molecule. The amino acid composition is remarkable when compared to tropocollagen in the large quantity of acidic residues, decreased glycine and imino acids, and the presence of cysteine. Three models of procollagen I structure are presented and discussed relative to the available experimental evidence.

1. Introduction

Based on the insolubility of collagen under physiological conditions, Schmitt [1] originally postulated the existence of a soluble, precursor form of tropocollagen. The rapidity of triple helix formation *in vivo* as contrasted to that encountered *in vitro* prompted Speakman [2] to propose recently the existence of registration peptides on a precursor form of tropocollagen. Although no data have been reported to date which specifically validate the Speakman hypothesis, laboratories have recently reported the demonstration of a tropocollagen precursor [3–10], procollagen, and its conversion to tropocollagen by tissue protease [11, 12]. The variety of *in vitro* culture conditions employed by various investigators has resulted in confusion regarding the molecular properties of procollagen and has prompted Church et al. [5] to suggest a nomenclature in which procollagen I represents a nondegraded state while further numerical designations represent sequential cleavage products. This report describes the effect of serum on the procollagen product from human skin fibroblasts in culture. Conditions of culture are described which yield apparent procollagen I in quantities amenable to detailed physical and chemical analysis.

* Present address, Department of Biochemistry, Academy of Medicine, Bialystok, Poland.

2. Experimental procedure

2.1. Materials

Eagle's minimal essential medium and fetal calf serum were the products of Grand Island Biological. Tricine, Tris base, trypsin, pepsin, and pronase were obtained from Sigma Chemical, tosyllysine chloromethyl ketone from Calbiochem, β-aminoproprionitrile fumarate from Aldrich Chemical, ascorbic acid from Fisher Scientific, hydrogen peroxide from Merck, [5-^{3}H] proline (25 C/mmole) from Schwarz–Mann, Spectrafluor PPO-POPOP from Amersham/Searl and Triton X-100 from Research Products International. Bacterial collagenase was purified from Worthington Biochemicals' CLS crude collagenase by a chromatographic method [13] with inhibition of residual clostripain activity by hydrogen peroxide [14] or the active site inhibitor, tosyllysine chloromethyl ketone

[15]. Carboxymethyl cellulose and agarose (Biogel A 0.5 M and electrophoretic agarose) were the products of BioRad. Lathyritic rat skin collagen was the kind gift of Dr. Hector Aguilar.

2.2. Methods

2.2.1. Culture conditions

Adult human fibroblasts were obtained by skin biopsy and maintained as a continuous culture at 37° in Eagle's minimal essential medium (buffered with 20 mM Tricine) at pH 7.36 and supplemented with 15% heat inactivated fetal calf serum. Procollagen products were harvested during a 24-hour period every 4 days from contact inhibited cells at confluence in 1000 cm^2 roller flasks containing 60 ml of culture fluid rotating at 1 rev per 5½ minutes. For the harvest cycle, fresh medium either with or without 15% fetal calf serum was added containing β-aminoproprionitrile fumarate (50 μg/ml) and ascorbic acid (50 μg/ml) plus 25 μC/ml [5-^{3}H] proline. The harvest medium was extensively dialyzed against 0.02 M Tris–0.5 M NaCl at pH 7.4 to remove non-incorporated radioactivity. This procedure results in no visible precipitate or of precipitated counts within the dialysis bag. Concentration was achieved by vacuum dialysis at 4° against 1 M NaCl–0.05 M Tris, pH 7.4 or by precipitation of procollagen I (see text) which results on acidification of the medium by dialysis against 0.15 M Na acetate, pH 4.

2.2.2. Chemical methods

Protein concentration – All protein concentrations were determined by a microbiuret method [16] using bovine plasma albumin solution, prepared by nitrogen content, as a standard.

Hydrolysis by proteases – Various proteases of widely variant specificities were used as structural molecular probes. Limit digests employed 20:1 or greater mass excesses of protease to [*Pro*-5-^{3}H] procollagen to ensure maximum hydrolysis. Trypsin and pronase hydrolysis was accomplished in 50 mM Tris – 0.9% NaCl at pH 8.0 and 7.2, respectively. Collagenase digestion utilized the same buffer plus 10 mM CaAc at pH 7.2. Acetic acid (0.5 M) at pH 2 was used for pepsin digestion.

Amino acid analysis – Samples were hydrolyzed under vacuum in 6 N HCl at 110° for 20 hours. Analysis was achieved using a Technicon single column system with norleucine employed as an internal standard. De novo conversion of [^{3}H] proline to hydroxyproline was analyzed on the Technicon analyzer by use of a stream splitter and unlabelled proline and hydroxyproline as internal standards.

2.2.3. Physical methods

Chromatography – Molecular sieve columns at 8° of Sephadex G-200 employed a variety of solvents. Agarose (38 cm x 2.5 cm glass columns) utilized an eluting buffer of 0.5 M NaCl – 0.02 M Tris, pH 7.4 and a flow rate of 15 ml per hour. Carboxymethyl cellulose ion exchange chromatography at 40° was utilized for the identification of collagen subunits using acid soluble lathyritic rat skin collagen as a standard according to the method of Piez (17).

Gel electrophoresis – Discontinuous gel electrophoresis was performed on 1% agarose gels in 10 cm x 5 mm cylinders. The agarose was melted in a buffer of 0.12 M Tris at pH 8.1 and layered into a Bio-Rad precision gel forming apparatus (Model 200) and allowed to polymerize by cooling. Samples for electrophoresis were dialyzed against the above buffer, mixed with sucrose (6% final concentration) and bromphenol blue, and layered on the agarose column under an upper electrode buffer of 0.043 M Tris, 0.046 M glycine at pH 8.9. The lower chamber buffer was 0.12 M Tris, pH 8.1. Electrophoresis utilized a constant current of 3mA per tube. Gels were removed, sliced into 1.2 mm sections, dissolved in 0.5 ml hot water, mixed with scintillation cocktail, and counted for radioactivity.

Isoelectric focusing – pH gradients of 3 to 10 were formed in a LKB 110 ml electrofocusing column and run at 500 volts for 2 days at 8° following the general method of Vesterberg and Svensson [18] using 3% Ampholine (v/v) in sucrose. Fractions of 2.2 ml were collected and assayed for pH at 20° using a Radiometer TTT 1 standardized against pH 7 and pH 4 buffers. 0.2 ml volumes were diluted to 1ml with water prior to counting. To eliminate non-specific fluorescence at alkaline pHs all samples above pH 8 was neutralized with 0.5 M acetic acid prior to counting.

Scintillation counting – Assessment of radioactivity of various aqueous samples was made in a Beckman LS spectrometer at ≈ 50% efficiency. The scintillation cocktail consisted of 29.5% Triton X-100, 2.4% Spectraflour, and 68.1% toluene (v/v); a volume of 1 ml sample to 10 ml cocktail was used.

Sedimentation equilibrium – Meniscus depletion sedimentation equilibrium [19] was accomplished with a Spinco model E analytical ultracentrifuge equipped with a Raleigh interference optical system. 3 mm columns of procollagen I in 1 M NaCl – 0.02 M Tris, pH 7.4, were layered over an inert base of perfluorotributylamine. During the run, the temperature was maintained at 5° and measurements were made at 11000 and 14000 rpm using double sector cells equipped with quartz windows. Fringe positions were recorded on Spectroscopic II G photographic plates and analyzed with a Nikon model 6 microcomparator at 50 x magnification. Runs were made for 5 days, equilibrium being assured after the third day by a lack of change in the fringe concentration gradient on sequential photographs.

The apparent weight average molecular weight ($M_{w,app}$) over the liquid column was evaluated by the relationship:

$$M_{w,app} = \frac{2RT}{(1-\bar{v}\delta)\omega^2} \frac{d \ln c}{d x^2}, \quad (1)$$

in which R is the gas constant, T the absolute temperature, $\bar{v}$ the partial specific volume of the protein, δ the solution density, ω the angular velocity in radians per second, x the radial distance in cm, and c the concentration in fringes. The value for the solvent density was obtained with a 50 ml Leach pycnometer. The partial specific volume of procollagen I (0.712 g/ml) was calculated from amino acid composition by the use of the known specific volume of the component amino acids [20].

Circular dichroism – Measurements of the circular dichroism from 250 nm to 195 nm were performed on a Cary model 60 spectropolarimeter equipped with a model 6001 CD accessory. Experiments were carried out using a 1 mm path length cell with procollagen I dissolved in 1 M NaCl – 0.05 M Tris, pH 7.4, at a concentration of 60 μg per ml. The circular dichroism data were recorded by the instrument directly in terms of degrees ellipticity, θ_λ; the mean specific ellipticity, $[\theta']_\lambda$, having units of degree cm^2 per decimole at wavelength λ, was calculated according to the relationship:

$$[\theta']_\lambda = \theta_\lambda (MRW)/10lc, \quad (2)$$

where MRW is the mean residue weight of the sample, c the concentration in grams per cm^3, and l the path length of the sample solution in cm. A mean residue weight of 106 was calculated from the amino acid composition.

3. Results

3.1. Collagenous nature of the [^{3}H]-proline labeled product from serum free culture medium

In a 24-hour pulse with [^{3}H] proline in which no serum is present in the medium, 1–2% of the counts are incorporated into non-dialyzable products with a proline:hydroxyproline ratio of approximately 3:1 of which approximately one half is rendered dialyzable by limit digests with bacterial collagenase at 23°. On standard carboxymethyl cellulose ion exchange columns used for collagen chain analysis, the product is retained on the column and requires base for elution (fig. 1A). Digestion with pepsin at 15° for 6 hours results in approximately one half of the counts eluting with the salt gradient. Using a column calibrated with lathyritic rat skin collagen, a small amount of radio-

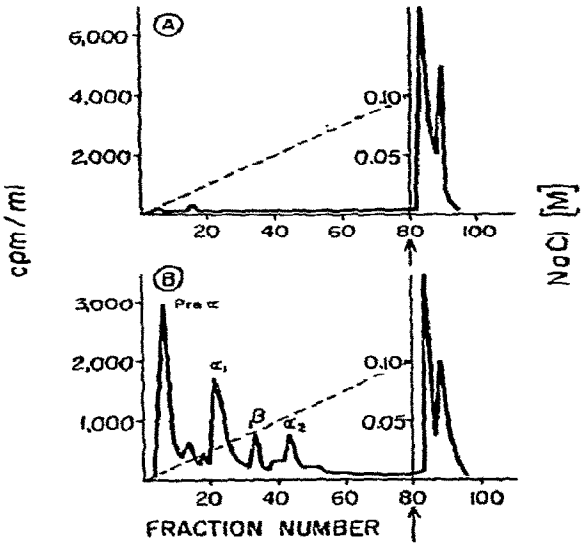


Fig. 1. Carboxymethyl cellulose chromatography of serum free medium at 40°. NaCl gradient is indicated by the dashed line and right ordinate. Arrow on abcissa indicates step salt change to 0.4 NaOH + M NaCl. A, medium without pepsin treatment; B, following pepsin digestion at 15°.

activity elutes in the β chain region despite the presence of β-aminoproprionitrile in the serum free pulse medium (fig. 1B). Apparent α 1 chains are in a roughly 2:1 excess over α 2 as expected. However, almost one half of the counts liberated by pepsin digestion elute prior to the standard α chain elution volumes (labeled pre α) suggesting a relative acidic composition.

3.2. The influence of serum in the culture medium on the procollagen product

Quantitatively, the presence of serum in the culture medium generally results in approximately a twofold increase in the level of [^{3}H]-proline incorporation. However, as demonstrated below, the presence of serum results in significant qualitative changes in the nature of the products obtained.

3.2.1. Molecular sieve chromatography

Preliminary experiments on Sephadex G-200 under a variety of eluting conditions (0.06 M acetate pH 4.8, or 0.05 M Tris pH 7.4, or 6 M urea) resulted in virtually total loss of counts on the column. However, by using agarose columns in 0.5 M NaCl – 0.02 M Tris pH 7.4, 100% recovery is routinely obtainable. The majority of the non-dialyzable [^{3}H] proline from serum free labeling experiments elutes on agarose (exclusion limit of approximately 400000 daltons for globular proteins) as one major peak in the void volume with only minor smaller components is routinely observed (fig. 2A). However, cultures pulsed with radioactive proline in the presence of serum yield not only a major peak in the void volume of the column but significant quantities of apparently smaller components (fig. 2B).

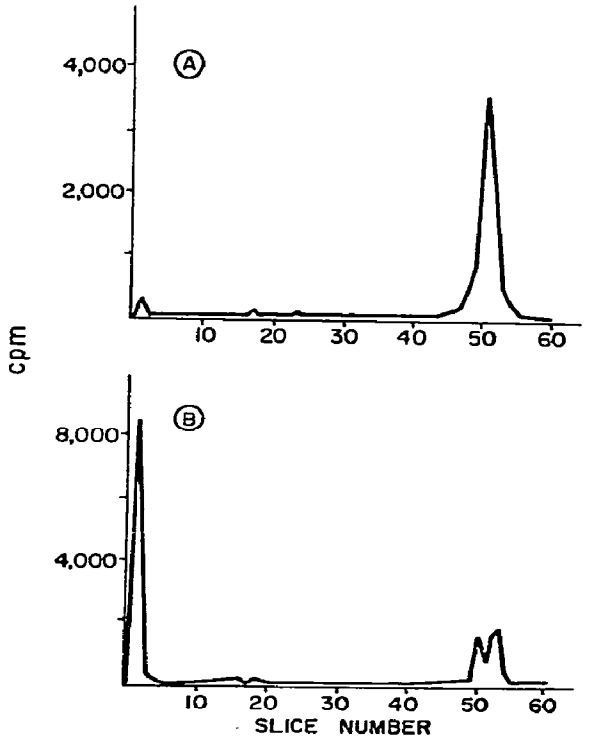


Fig. 2. Molecular sieve chromatography on agarose of human fibroblast culture fluid following extensive dialysis. A, serum free culture medium; B, serum containing culture medium.

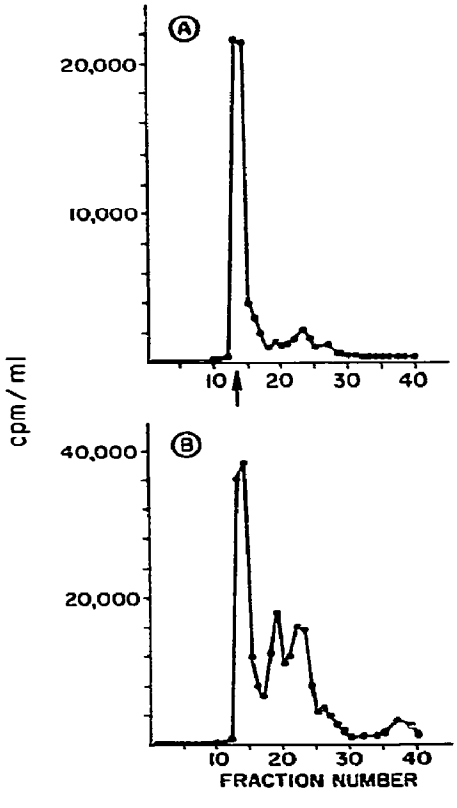


Fig. 3. Electrophoresis of human fibroblast culture fluid following extensive dialysis. A, serum free culture medium; B, serum containing culture medium.

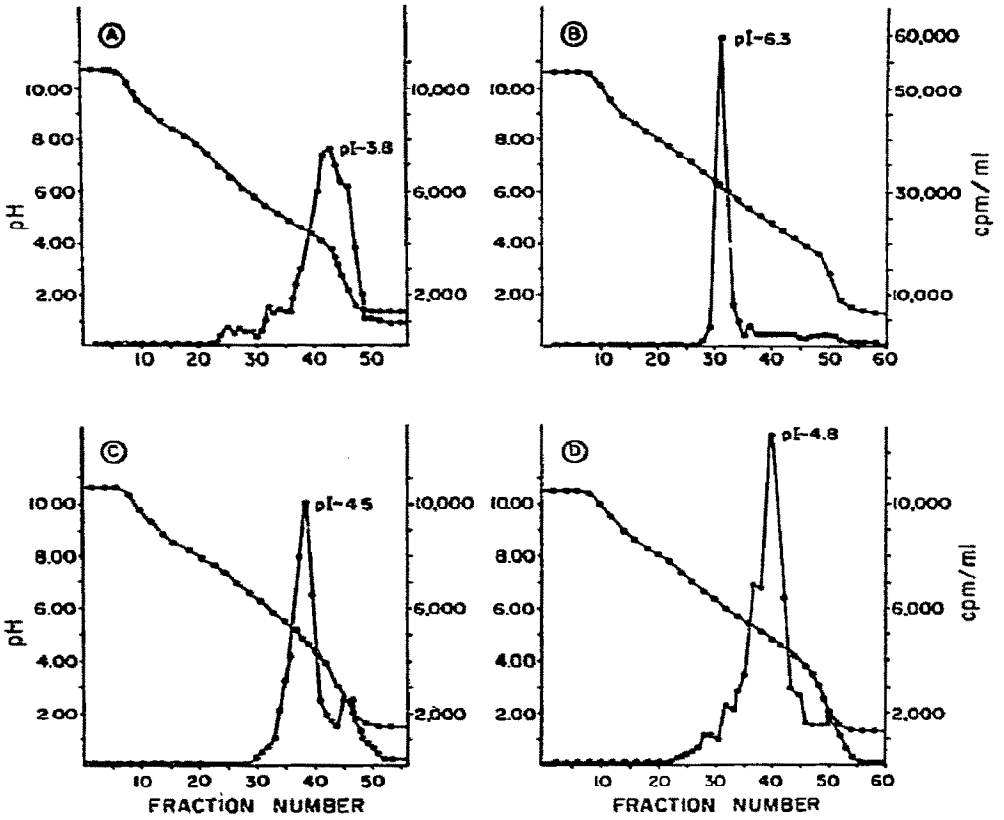


Fig. 4. Isoelectric focusing of the major components of human fibroblast culture medium obtained by gel filtration. Left ordinate and solid circles indicates pH; right ordinate and open circles indicates observed cpm at the indicated fraction. A, serum free culture medium (agarose peak I); B, serum containing culture medium (agarose peak I); C, serum containing culture medium (agarose peak II); D, serum containing culture medium (agarose peak III).

3.2.2. Disc electrophoresis

In standard 7.5% polyacrylamide systems (both pH 9 and 4) [21], procollagen fails to penetrate the gel. For this reason agarose was employed as a more porous support material. Fig. 3A demonstrates that the de novo synthesized, secreted high molecular weight product from serum free cultures migrates as a single band. However, in serum containing cultures, the rapidly moving anionic peak is lost and is replaced by a small double peak migrating at the same rate and a major peak which possesses a much slower rate of migration (fig. 3B). In all cases at least 80% of the applied radioactivity was recovered.

3.2.3. Isoelectric focusing

Fig. 4 illustrates results obtained from the isoelectric focusing of major peaks from agarose molecular sieve chromatography. The major peak of radioactivity from serum free cultures (fig. 4A) has an isoelectric point of pH 3.9 ± 0.1 in striking contrast to the cationic properties of tropocollagen [22]. A single symmetrical peak of radioactivity is observed at pH 4.0 when the gradient is acid at the bottom of the column. When the gradient is reversed (as illustrated), precipitation occurs, yielding asymmetry on elution from the column with an apparent isoelectric point of pH 3.8. In serum containing cultures (fig. 4B), a major shift in isoelectric point to pH 6.3 occurs for the material eluting from

agarose columns in the void volume. The peak is sharp and symmetrical with no tendency to precipitate at the isoelectric point. The smaller molecular weight components in serum containing cultures possess anionic isoelectric points, do not precipitate, and are heterogenous (figs. 4C and 4D).

3.2.4. Conversion of procollagen to collagen

As previously illustrated (Fig. 1B), pepsin digestion of procollagen from serum free culture media liberates not only α chains but an apparently acidic species which elutes on carboxymethyl cellulose prior to the earliest α chain elution. Chromatography of serum containing culture medium on CM cellulose at 10° yields a pre α peak eluting at an identical position plus a major residual activity requiring base for elution (fig. 5A). Following incubation at 45° for 30 min. and chromatography at 40°, counts elute in the α_2 position as well as in pre α (fig. 5B). In order to correlate the origin of pre α and α_2 eluting counts to the components separated by molecular sieve chromatography on agarose (fig. 2B), each major agarose fraction was subjected to chromatography on carboxymethyl cellulose, hydrolysis by bacterial collagenase, and proline hydroxylation analysis. Agarose peak I is more extensively hydroxylated and digestible by collagenase than either peaks II or III (table 1). The lability of counts in the peak II control following a 37° incubation period suggests the presence of a protease in this fraction. On carboxymethyl cellulose peak I behaves similarly to that obtained from serum free cultures in that base is required for elution (fig. 6A). Fraction II elutes in the pre-α position with no residual counts remaining on the column (fig. 6B) while fraction III precipitates under the conditions used (0.06 M acetate, pH 4.8, and 40°). Recombination of the three fractions in solvents used for gel filtration, concentration, and dialysis against the standard carboxymethyl cellulose solvents, and incubation at 45° followed by elution at 40°, results in the appearance of α_2 and pre-α peaks (fig. 6C). Thus, presumptive evidence for a protease in serum containing cultures which elutes in the agarose peak II region and which utilizes procollagen as a substrate is apparent.

In an effort to ascertain conditions in which the proteolytic effect occurs, culture medium from serum containing cultures were incubated at three different

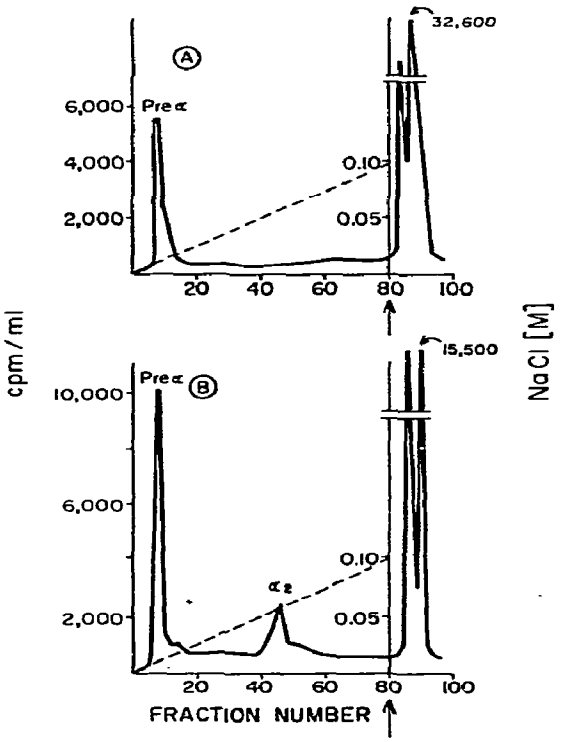


Fig. 5. Carboxymethyl cellulose chromatography of serum containing culture medium. Ordinates as in fig. 1. A, chromatography at 10°; B, chromatography at 40° following 30 min incubation at 45°.

Table 1

Proline hydroxylation and susceptibility to collagenase of procollagen products from serum containing cultures separated on agarose. All collagenase hydrolyses at 37°. See fig. 2B for details of isolation

Agarose peak	$\left(\frac{[^3H]\text{-HO-proline}}{[^3H]\text{-proline}}\right)$	% dialyzable[a] [^{3}H]-imino residues	
		control	collagenase
I	0.38	3.7	64.5
II	0.16	25.6	40.4
III	0.11	10.9	20.5

[a] Dialysis at 5°.

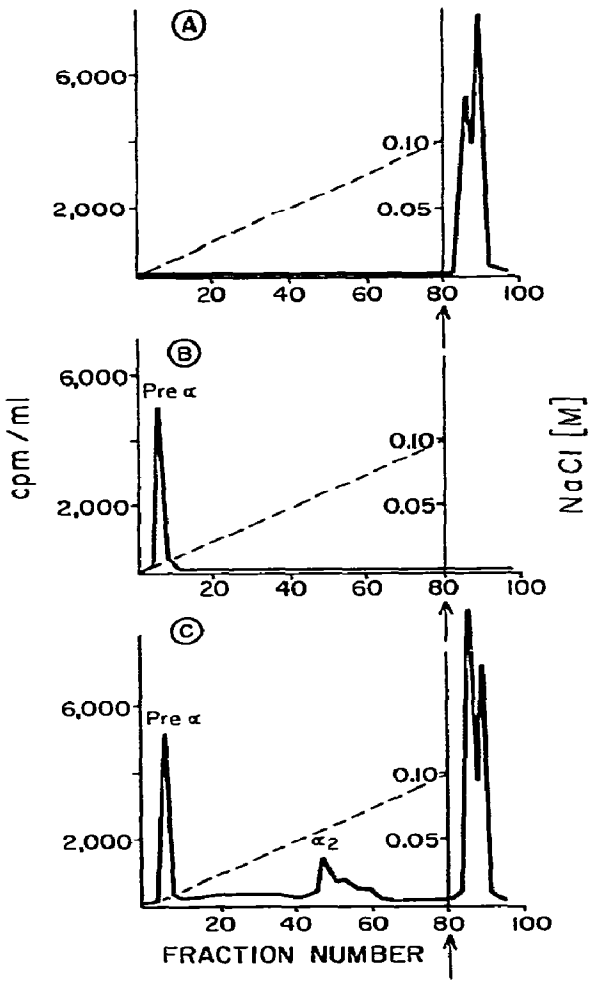


Fig. 6. Carboxymethyl cellulose chromatography at 40° of the major components of human fibroblast culture medium obtained by gel filtration as illustrated in fig. 2B. Ordinates as in fig. 1. A, fraction I; B, fraction II; C, recombination of fractions I, II, and III and incubation at 45° for 30 min prior to chromatography.

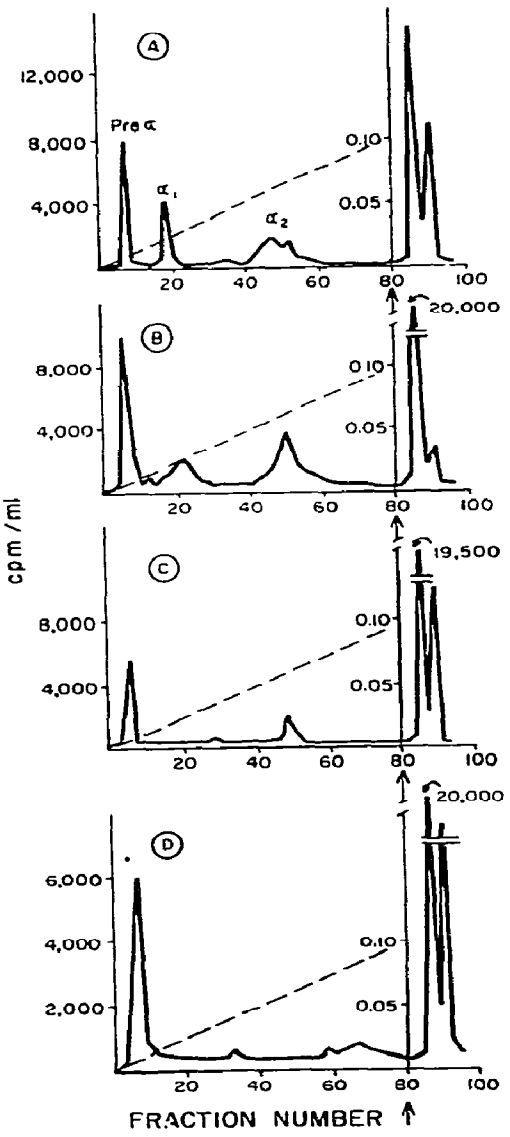


Fig. 7. Carboxymethyl cellulose chromatography at 40° of serum containing culture medium incubated at 37° for 24 hours at various hydrogen ion concentrations. A, pH 4.8 (0.06 M Na acetate); B, pH 7.36 (culture media); C, pH 8.5 (0.05 M Tris – 0.9% NaCl); D, pH 7.36 (culture media) containing 1 mM EDTA-Na_2.

pHs for 24 hours at 37° and subjected to carboxymethyl cellulose chromatography. In acetate buffer (0.06 M, pH 4.8) both α_1 and α_2 peaks appear in approximately 2:1 ratio as well as pre α (fig. 7A). At pH 7.4 (culture medium only), the previously noted α_2 and pre α can be seen, but there is a new peak of radioactivity eluting

in the α_1 position (fig. 7B). Nevertheless, α_2 is greater in quantity than α_1. At pH 8.5 (0.05 M Tris – 0.9% NaCl), there is a reduction in α_2 and pre-α content with no α_1 peak (fig. 7C). In a manner analogous to the non-incubated situation (fig. 5A), inclusion of 1 mM EDTA-Na_2 in a 24-hour incubation at pH 7.4 results in pre-α material only (fig. 7D). Incubation of procollagen isolated from serum free cultures with fresh, heat inactivated fetal calf serum fails to yield α_1, α_2, or pre-α components. Thus the observed EDTA inhibited proteolytic activity is not indigenous to serum but is an apparent product of fibroblasts in culture, induced by one or more factors in serum.

3.3. Partial characterization of putative procollagen I

Evidence has been presented which indicates that an anionic precursor of tropocollagen is secreted by human skin fibroblasts in culture. The data suggest that one or more factors in serum either activate or induce synthesis and/or secretion of protease which convert procollagen to tropocollagen yielding a variety of products in serum containing cultures. In serum free cultures the absence of the acidic pre α peak on carboxymethyl cellulose chromatography argues against significant degradation and provisionally qualifies this material for the designation of procollagen I. The physical properties of high salt solubility, high molecular weight, and precipitation at the isoelectric point (fig. 8) were used to isolate the putative procollagen I used in the studies described below.

3.3.1. Adsorptive properties of procollagen I in low salt

Procollagen I in a variety of solvents (0.06 M acetate,

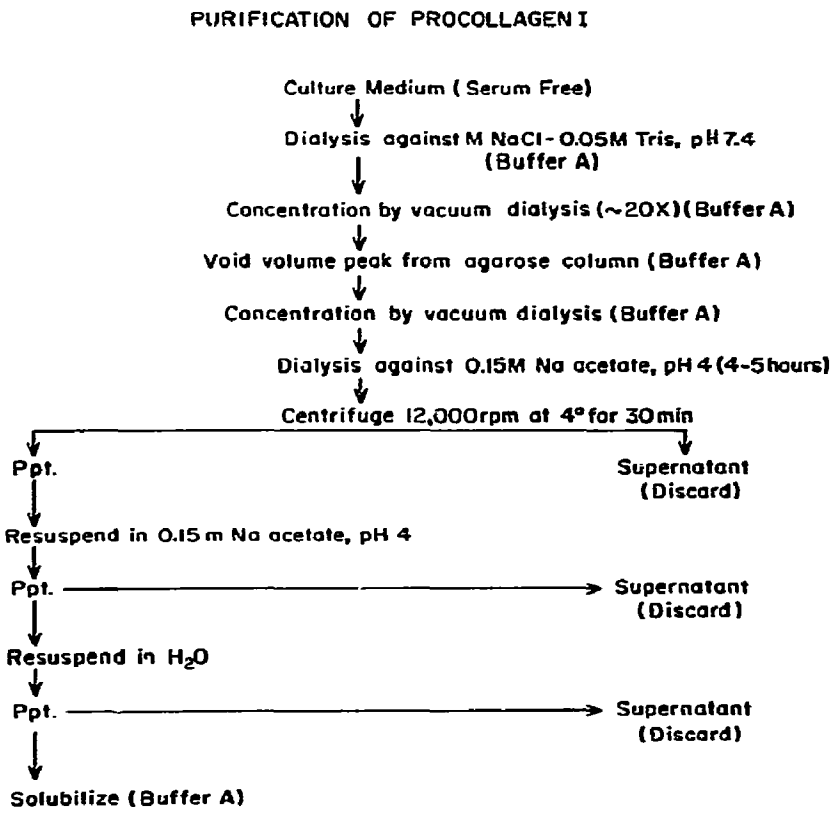


Fig. 8. Isolation procedure for procollagen I.

Table 2

Comparison of the amino acid contents (residues/1000 residues, uncorrected for hydrolytic losses) of human procollagen I, human skin collagen, human basement membrane collagen, and human complement Clq

Amino acid	Human skin procollagen I	Human skin collagen[a)]	Human basement membrane collagen[b)]	Clq of human complement[c)]
aspartic acid	93.0	47.2	67.7	85.0
threonine	49.5	18.3	36.7	52.7
serine	65.7	36.9	48.9	55.2
glutamic acid	110.0	77.7	98.3	91.9
proline	114.2	125.1	57.9	60.3
glycine	154.6	324.4	221.0	173.2
alanine	70.2	114.5	59.8	44.2
half-cystine	17.3	tr.	22.8	24.6
valine	51.5	24.5	35.6	55.2
methionine	13.5	7.0	13.2	25.4
isoleucine	27.6	10.4	31.3	38.1
leucine	55.1	24.8	65.3	60.0
tyrosine	20.0	3.5	15.3	30.2
phenylalanine	24.0	12.6	24.9	42.4
lysine	38.7	26.6	19.5	36.8
histidine	14.4	5.4	14.2	14.0
arginine	42.4	49.0	38.0	42.9
hydroxyproline	31.8	90.9	103.5	38.9
hydroxylysine	5.9	5.9	26.1	18.8
tryptophan	ND	ND	ND	9.1
total acidic amino acids	203.0	124.3	166.0	176.9
total basic amino acids	87.0	81.5	83.6	98.5
total imino acids	146.0	216.0	161.4	99.2

a) As reported by Fleischmajer and Fishman [25].
b) As reported by Westberg and Michael [26].
c) As reported by Yonemasu et al. [27].

pH 4.8; 0.05 M Tris HCl, pH 7.4 or pH 8.5; < 0.9% NaCl) exhibits a profound ability to adsorb to glass and plastic surfaces. This property can be largely inhibited with NaCl at a molar concentration of 0.5 or greater.

3.3.2. Amino acid composition

The anionic behavior of procollagen I as compared to that of tropocollagen can be readily explained on the acidic amino acid composition (table 2). Although no data is available on the degree of amidation, human skin procollagen I has approximately 1.6 times the content of aspartic and glutamic residues than that of human skin collagen, despite a relatively equivalent content of basic amino acids. As in tropocollagen, the predominant amino acid is glycine. Roughly 1/6 of the total sites are occupied by glycine in procollagen I while every third residue in tropocollagen is glycine. Thus the minimum weight of native procollagen I is approximately twice that of tropocollagen assuming few glycines in the "non-collagenous" portion of the precursor. The total imino acid content of procollagen I is also diminished but to a lesser extent than that of glycine. Of consider-

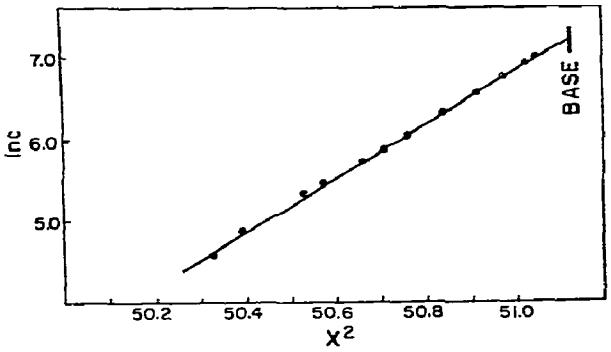


Fig. 9. Typical ln c versus X^2 plots of procollagen I. Speed at 11000 rpm after 72 hours.

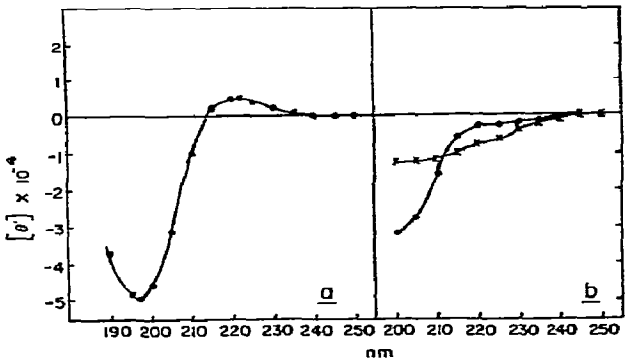


Fig. 11. Circular dichroic spectra. A, rat skin collagen in acetic acid (1%) at 5°; B, human skin procollagen in 1 M NaCl – 0.05 M Tris, pH 7.4. •——• at 5°, x——x at 50°.

able interest is the identification of significant quantities of cysteine in procollagen I. Since this analysis was made on material in which the cysteine residue was not derivatized, the reported value probably represents a low value, due to the relative instability of cysteine under standard acid hydrolysis conditions.

3.3.3. Sedimentation equilibrium

Sedimentation equilibrium of procollagen I yielded linear ln c versus X^2 plots at each of the four protein concentrations analyzed with no apparent evidence of size heterogeneity or aggregation (fig. 9). Nevertheless, a marked concentration dependence was observed between cells with large loading concentration differences (fig. 10). Extrapolation of this apparent linear regression slope to zero concentration yielded a molecular weight of 539000 ± 40000 in agreement with the conclusions drawn from amino acid analysis. Due to low absorptivity at 280 nm, analysis at lower protein concentrations using the UV scanner of the Model E were impractical.

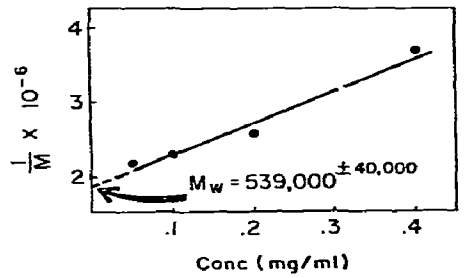


Fig. 10. Apparent weight average molecular weight as a function of initial cell loading concentration with extrapolation to infinite dilution.

3.3.4. Comparison of the circular dichroic spectra of procollagen I with skin collagen

Due to the insolubility of procollagen I in acid buffer the usual measurement of optical activity of collagen derived material in acetate is impossible for procollagen I. Fig. 11A illustrates the typical spectrum of native rat skin collagen in dilute acetic acid. The characteristic negative absorption peak at 197 nm and positive absorption at 222 nm is present. In contrast procollagen I in 1 M NaCl – 0.05 M Tris, pH 7.4 fails to exhibit the 222 peak (fig. 11B). Circular dichroism scans below 200 nm were impossible due to adsorption by the buffer, but the general curve configuration suggests a minimum approximately 60% that of rat skin lathyritic tropocollagen in acetate.

3.3.5. Stability of the secondary structure

Although the circular dichroism spectra suggest features analogous to tropocollagen, procollagen I is sensitive to proteolytic digestion below the denaturation temperature by non-collagenolytic enzymes which ordinarily cannot attack the polyproline II configuration (table 3). Procollagen at 23° is readily hydrolyzed to a majority of dialyzable peptides by pronase and to

Table 3

Procollagen I susceptibility to proteolytic hydrolysis

	% dialyzable[a] [^{3}H]-imino residues	
	23°	37°
control	0.5	3.1
collagenase	48.4	56.1
pronase	77.7	89.9
pepsin	24.1	52.2
trypsin	22.4	51.6

[a] Dialysis at 5°.

a more limited extent by trypsin and pepsin. At 37°, more of the procollagen structure is in a form susceptible to proteolytic attack by pepsin and trypsin. Approximately 1/2 of the [^{3}H]-imino acids are rendered dialyzable by bacterial collagenase in support of the collagenous nature of the material and is consistent with amino acid analysis and ultracentrifuge data on the minimum molecular weight of procollagen I.

The thermal stability of procollagen I was analyzed by monitoring the circular dichroic negative extremum as a function of temperature (fig. 12). A sharp transition occurs at 39°, closely paralleling native tropocollagen denaturation behavior.

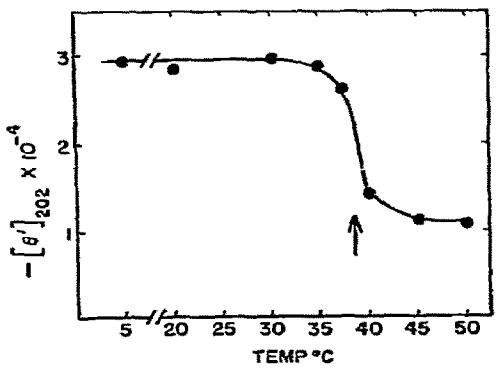


Fig. 12. $[\theta']_{202}$ as a function of temperature.

4. Discussion

The existence of an extracellular precursor of collagen has been well substantiated by reports from a number of independent laboratories [3–10]. There is common agreement that the precursor is larger than tropocollagen and is converted to the latter by the proteolytic activity of an enzyme(s) commonly referred to as procollagen peptidase [11, 12]. It is the existence of such proteolytic activities which probably account for variances in the reported molecular properties of procollagen. In the studies reported here, the omission of serum from roller flasks of human skin fibroblasts in culture results in the production of a tropocollagen precursor of approximately 540000 daltons. The glycine content of 15.5% and imino acid content of 14.6% supports the sedimentation equilibrium data and compares reasonably closely with the respective values of 20.8% and 15.5% for the intracellular collagen precursor recently reported by Tsai and Green [23]. In contrast, Bornstein et al. [24] have reported a value of 115000 for the pro-α-1 chain from chick calvaria with a 29% and 16.7% content of glycine and imino acids, respectively. Thus the difference of 60000–70000 daltons per chain between the data reported here and that reported by several laboratories may be due to partial proteolysis with the main precursor moiety demonstrating a comparatively low glycine content. The addition of serum to our system results in a heterogeneity of [^{3}H]-proline labeled products, an alkaline shift of isoelectric point in the collagenous portion of procollagen, and the appearance of an acidic peak on carboxymethyl cellulose which we refer to as pre α. This pre-α component is a common feature of published chromatograms in which the molecular weight of procollagen is substantially less than that reported here. The proteolytic activity is not present in fresh serum but must be exposed to fibroblasts, an observation in agreement with Tsai and Green [23]. The ability to liberate α-1 chains preferentially at low pH as compared to α-2 at neutral pH suggests a multiplicity of proteases.

Comparison of the amino acid composition of procollagen I to that reported for human skin collagen [25], human basement membrane collagen [26], and human complement C1q [27] suggests that procollagen I may function as the precursor of the latter three (table 2). The greatest deviation exists between procollagen I and skin collagen. As previously noted there is 1/2 the glyc-

ine content in procollagen I, but the total acidic amino acids are increased by 60% resulting in the low isoelectric point observed for procollagen I. Both basement membrane collagen and Clq component of complement have an acidic residue and glycine content intermediate between procollagen I and collagen. Thus partial proteolysis of the precursor moiety could result in a molecule with these properties. The specificity of end product production would then depend on protease specificity and/or degree of lysine hydroxylation and carbohydrate content. If this hypothesis proves to be valid, then the low imino acid content of Clq of complement would suggest a proline clustering in the more distal regions of procollagen I, while the common cysteine content of procollagen I, basement membrane collagen, and Clq of complement suggest a more proximal location for the residue in the precursor moiety.

The secondary structure of procollagen I as judged by circular dichroism in the far ultraviolet exhibits features in common with acid soluble collagen. Although penetration below 200 nm in NaCl was not possible, the configuration of the procollagen curve suggests a minimum at the same wavelength as collagen (197 nm) at 60% of the observed ellipticity. The spectral data, as well as the susceptibility of procollagen I to proteolytic attack by a variety of proteases, clearly indicates that the precursor mouity is not completely in the polyproline II configaration [28]. Nevertheless, procollagen I exhibits a sharp thermal transition point at 39° with a ΔT of 2° which is closely comparable to that of native tropocollagen [29]. Since the thermal stability of native tropocollagen is directly related to the total imino acid content [30], it is of interest to note that the lowered total imino acid content of procollagen I did not lower the melting temperature (fig. 13), which suggests that the observed change in optical activity represents thermal denaturation of a collagenous unit component. Moreover, the small ΔT observed is strong evidence that the collagenous portion of the molecule is in the triple helical configuration prior to proteolytic excision from the parent molecule.

Electron micrographs of procollagen from chick embryo cell culture [7] as well as kinetic analysis of procollagen biosynthesis [10] indicates that the precursor moiety is located at the amino terminus. The recent demonstration of size distribution of procollagen dependent on disulfides [9, 31], suggests that procollagen is synthesized as separate chains and aligned intracellularly by disulfide pairing. Such a mechanism should result in rapid chain registration and formation of the native tropocollagen structure as originally suggested by Speakman [2].

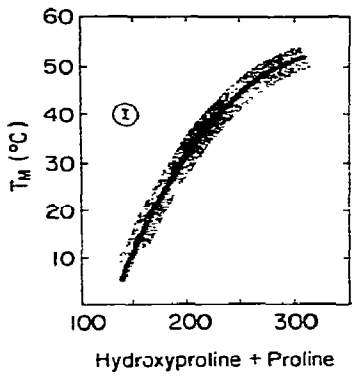


Fig. 13. Thermal denaturation temperature versus imino acid content for a variety of collagens and procollagen I. The shaded area represents the observed distribution around the mean (solid line) [30] while (I) indicates the relationship of procollagen I to the melting profile. All data corrected to theoretical values expected in water [29].

The studies reported here strongly suggest that procollagen I is considerably larger than generally held. Fig. 14 illustrates three possible molecular configurations. In each, the collagenous regions are depicted as helical and are associated in a cooperative structure as indicated by the small ΔT observed in this study. Church et al. [5] originally postulated a single chain containing both α-1 and α-2 elements. Fig. 14A illustrates this model. Since tropocollagen chains are aligned NH_2 to COOH terminal in parallel, connecting links which would allow the triple helical configuration for the collagenous portion, would necessarily need to be of approximate equal size. Evidence against this model is formidable. Kinetic labelling of tropocollagen by Vuust and Piez [10] indicate separate initiation points for α-1 and α-2 chains while the polysomes involved in procollagen synthesis [32] would appear to be too small to support an mRNA yielding a molecule of more than 500000 daltons. The identification of cysteine in procollagen I and the demonstration by Burgeson et al. [9] of altered chromatographic properties of a collagen precursor following reduction by β-mercaptoethanol or $NaBH_4$ leads credence to two additional models. In each (figs. 14B and C) di-

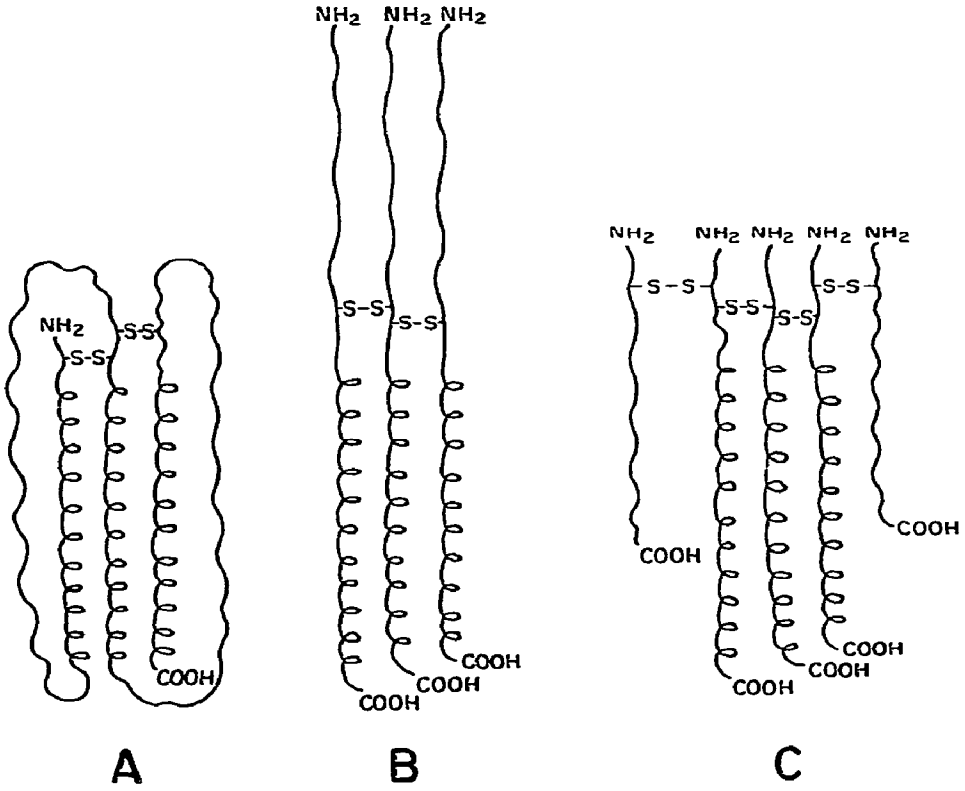


Fig. 14. Structural models for procollagen I. A, single chain; B, three chain; C, multiple chain.

sulfides hold the three chains together by covalent bonding as well as that to be expected through non-covalent forces of the triple helix. In fig. 14B each chain is approximately twice the length of α chains. In fig. 14C the major procollagen extension is represented as approximately 10–15% of the main chain while additional peptides conceivably of use in transport are covalently bonded to the main chain by disulfides, or alternatively, as tightly bound non-covalent acidic polypeptides which are not dissociated under the conditions employed in this study. The occurrence of a small amount of apparent β chain in our material as well as the small quantities available, has precluded us from arriving at an unambiguous answer. We currently are culturing human fibroblasts in which β-aminopropionitrile is present continuously rather than only in the serum free 24-hour collection period. The availability of mg quantities of procollagen I free of contaminating aldehyde cross links should allow the proper selection of the correct model.

Acknowledgement

This study was supported by United States Public Health Service Grant AM10833 and Career Development Award Grant 1 K4 AM34798.

References

[1] F.O. Schmitt, Bull. N.Y. Acad. Med. 36 (1960) 725.
[2] P.T. Speakman, Nature 229 (1971) 241.
[3] D.L. Layman, E.B. McGoodwin and G.R. Martin, Proc. Natl. Acad. Sci. US 68 (1971) 454.
[4] G. Bellamy and P. Bornstein, Proc. Natl. Acad. Sci. US 68 (1971) 1138.

[5] R.L. Church, S.E. Pfeiffer and M.L. Tanzer, Proc. Natl. Acad. Sci. US 68 (1971) 2638, 3241.
[6] P.B. Ramaley and J. Rosenbloom, FEBS Letters 15 (1971) 59.
[7] P. Dehm, S.A. Jiminez, B.R. Olsen and D.J. Prockop, Proc. Natl. Acad. Sci. US 69 (1972) 60.
[8] E. Bankowski and W.M. Mitchell, Federation Proc. 31 (1972) 3513.
[9] R.E. Burgeson, A.W. Wyke and J.H. Fessler, Biochem. Biophys. Res. Commun. 48 (1972) 892.
[10] J. Vuust and K.A. Piez, J. Biol. Chem. 247 (1972) 856.
[11] C.M. Lapiere, A. Lenaers and L.D. Kohn, Proc. Natl. Acad. Sci. US 68 (1971) 3054.
[12] P. Bornstein, H.P. Ehrlich and A.W. Wyke, Science 175 (1972) 544.
[13] W.M. Mitchell and W.F. Harrington, J. Biol. Chem. 243 (1968) 4683.
[14] W.M. Mitchell, Johns Hopkins Med. J. 127 (1970) 192.
[15] W.H. Porter, L.W. Cunningham and W.M. Mitchell, J. Biol. Chem. 246 (1971) 7615.
[16] S. Zamenhof, in: Methods in enzymology, Vol. 3, eds. S.P. Colowick and N.O. Kaplan (Academic Press, New York, 1957) p. 702.
[17] K.A. Piez, E.A. Eigner and M.S. Lewis, Biochemistry 2 (1963) 58.
[18] O. Vesterberg and H. Svensson, Acta Chem. Scand. 20 (1966) 820.
[19] D.A. Yphantis, Biochemistry 3 (1964) 297.
[20] E.J. Cohn and J.T. Edsall, Amino acids and proteins (Reinhold, New York, 1943) p. 370.
[21] W.M. Mitchell, Biochem. Biophys. Res. Commun. 147 (1967) 171.
[22] J.E. Eastoe, in: Treatise on collagen, Vol. 1, ed. G.N. Ramachandran (Academic Press, New York, 1967) p. 30.
[23] R.L. Tsai and H. Green, Nature New Biol. 237 (1972) 171.
[24] P. Bornstein, K. von der Mark, A.W. Wyke, H.P. Ehrlich and J.M. Monson, J. Biol. Chem. 247 (1972) 2808.
[25] R. Fleischmajer and L. Fishman, Nature 205 (1965) 264.
[26] N.G. Westberg and A.F. Michael, Biochemistry 9 (1970) 3837.
[27] K. Yonemasu, R.M. Stroud, W. Niedermeier and W.T. Butler, Biochem. Biophys. Res. Commun. 43 (1971) 1388.
[28] W.F. Harrington and P.H. von Hippel, Advan. Prot. Chem. 16 (1961) 1.
[29] P.H. von Hippel and K. Wong, Biochemistry 2 (1963) 1399.
[30] J. Josse and W.F. Harrington, J. Mol. Biol. 9 (1964) 269.
[31] B.D. Smith, P.H. Byers and G.R. Martin, Proc. Natl. Acad. Sci. US 69 (1972) 3260.
[32] E. Lazarides and L.N. Lukens, Nature New Biol. 232 (1971) 37.

BIOPHYSICAL CHEMISTRY 1 (1973) 87–96. © NORTH-HOLLAND PUBLISHING COMPANY

A CHEMICAL INSTABILITY MECHANISM FOR ASYMMETRIC CELL DIFFERENTIATION†

Peter ORTOLEVA and John ROSS
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Received 20 July 1973

We propose a mechanism of asymmetric mitosis applicable for cells even without inherent architectural asymmetry and in the presence of symmetric conditions such as a homogeneous environment. The theory is based on the instability of the symmetric development in time of the prospective daughters of a cell in mitosis. The macroscopic equations of chemical reaction, diffusion, and permeation of the various chemical species in the cell are given formal expression, and are then linearized about the symmetric development in order to test the stability to asymmetric perturbations. Instability to such perturbations indicates the deterministic onset of asymmetric division (differentiation). Only small external gradients of concentration, temperature, light, etc. are necessary to polarize the asymmetry for the purpose of a particular morphology. The theory is compared qualitatively with experiments on melanocytogenesis, is used to suggest new experiments, and is proposed as a possible alternative to other mechanisms. Possible application of the theory to the experiments on the first division in the egg of Fucus is considered. Finally, a simple model of a product-enhanced reaction mechanism is developed in detail which shows that the history of the initial preparation of the cell prior to mitosis may play a role in determining the possibility of asymmetric division.

1. Introduction

There are many examples of cell division which yield progeny having qualitatively different characteristics from those of the parent. A number of different events may occur. In undifferentiated division two cells of type A are formed from a cell of the same type. In symmetric cell differentiation a cell of type A produces two daughters of type B. If a cell of type A produces two cells, one of type B and one of type C (where B may be identical to A), then this process will be called asymmetric cell differentiation. The process may occur by means of different mechanisms. First, the origin of the asymmetry of the differentiation may be an inherent architectural asymmetry of the parent cell, as occurs for instance in the case of the grasshopper neuroblasts. Here the division results in a ganglion and a neuroblast as daughter cells [1]. Second, the process may require the presence of an external gradient (of temperature, chemical composition, pressure, light intensity, electric field, etc.) sufficiently strong to be significant over cellular dimensions as may be of importance in regeneration in association with mitosis, i.e., epimorphosis [2]. Third, the asymmetric differentiation may be initiated on one of a random or ordered distribution of sites, either within the cell or in the membrane, due to (irreversible) activation by a hormone or light for instance [3]. This mechanism requires that the time scale between successive initiations of the asymmetric differentiation on two sites, or localized set of sites, be longer than the time necessary for signals to be sent to prevent similar activation at the remaining sites. Fourth, the differentiation may occur by an inexact replication process which yields two daughters of different genotypes, or by chromatin diminution as in the bifurcation of the germ cell from the somatic cell lines [4].

In this article we discuss another possible process of asymmetric cell differentiations, one applicable even to the most stringent condition of a parent cell without

† This work was supported in part by the National Science Foundation and the Camille and Henry Dreyfus Foundation. A brief account has been submitted as a Note to the journal Developmental Biology.

architectural asymmetry immersed in a homogeneous environment. In an abstraction, parent cells without essential asymmetry are termed symmetric. The mechanism we discuss could be operative in asymmetric parent cells but is applicable to repetitious symmetric cells, such as in the first division of the egg of Fucus [5].

In a pioneering work Turing [6] showed how the homogeneous evolution of certain chemical reaction mechanisms with feedback and maintained far from chemical equilibrium may become unstable to the onset of spatially inhomogeneous perturbations due to the coupling of diffusion and reaction. He proposed this as a possible chemical basis for biomorphogenesis, giving an explanation of how a homogeneous cell field may break up into differentiated regions. The onset of symmetry-breaking instability has been shown theoretically [7] and experimentally [8] † to lead to stable dissipative spatial structures. The role of heterogeneous catalysis in the formation of stable, inhomogeneous structures is discussed in ref. [9]. The importance of chemical instabilities to the formation and maintenance of structure in biological systems has been discussed by a number of authors [6, 7, 9–12].

The analogy between symmetry breaking in chemical and biological systems suggests a mechanism for asymmetric cellular differentiation. The theory is based on the equations for chemical reaction rates and transport of the chemical species in the differentiating system, i.e., those which participate in the development of the asymmetry. The possibility for the growth of the asymmetry is posed as a stability question and the theory predicts that symmetric differentiation may occur. If the parent cell at some point becomes unstable to an asymmetric concentration fluctuation, then asymmetric differentiation is initiated. The course of this asymmetric development is determined by the non-linearity of the equations and the asymmetry may become stable due to this non-linearity. Thus the system becomes committed to a well-defined course of events leading to asymmetric mitosis. The equations of reaction, diffusion, and permeation yield unique solutions for given initial and boundary conditions and hence the course of differentiation is deterministic. For spherically symmetric parent cells in a homogeneous environment only the orientation of the differentiation is random. In order that chemical symmetry-breaking mechanism may provide a theory for cell differentiation, the time scales of the reactions and transport (diffusion and permeation) must be comparable.

The theory of asymmetric cell differentiation presented here may be extended to the analysis of the creation of additional localization and structure in initially asymmetric cells.

A distinction must be made between asymmetric cell differentiation which occurs (1) in a homogeneous medium and (2) in a medium which has gradients. The second case may be illustrated in blastoma (epimorphic development) where there are nutritional, hormonal, and temperature gradients. In the eggs of Fucus [5], the polarity of the differentiation of the first division may be fixed by external gradients of salinity and voltage or direction of illumination. In the absence of these gradients, however, polarity still develops but is randomly oriented. Our analysis provides a mechanism for the development of asymmetry in the absence of an external gradient. Hence in this theory the possible presence of vanishingly small gradients serves only to orient the asymmetry along the direction of the gradients. If, contrary to our mechanism, an external gradient proves to be essential for the development of asymmetry, as may well be the case for certain systems, then the magnitude of the gradient must be sufficiently large to constitute an appreciable concentration difference over a distance of cellular dimensions.

Reversal of differentiation (B → A) is known; see, for example, the case of reversion of non-proliferating, multi-nucleated fibers into proliferating mono-nucleated cells during wound healing in muscle [13]. Clearly, in such cases, irreversible genetic changes do not occur.

In order to stress further the importance of cell components other than the nucleus in asymmetric differentiation, we refer to the case of mitosis in certain neuroblasts [3]. The architecturally asymmetric neuroblast divides into a ganglion and daughter neuroblast. If the mitotic spindle is rotated 180° so that the nucleus normally targeted for the pre-ganglion pole will enter the preneuroblastic region, then an apparently normal neuroblast and ganglion are nevertheless formed from the respective precursor region. Thus no genetic asymmetry is required to induce the asymmetric differentiation.

† See also ref. [7], p. 261.

2. Theory

2.1. Statement of model

We begin consideration of the onset of asymmetry in mitosis where the development is still symmetric with respect to the prospective daughters (halves). For instance, in the eggs of Fucus [5] this initial state is spherically symmetric and asymmetry arises prior to nuclear division. The asymmetry may, however, arise after nuclear division from a state symmetric with respect to the plane of division. In such cases the asymmetry may be initiated either before or after the formation of the phragmoplast in plant cells or invagination of the parent cell plasma membrane in animal cells. At such stages the two halves of the cell, prospective daughters, are in a state of strong coupling. Thus, as a simplifying convenience, in discussing the mass transfer we substitute a model of two compartments for the actual geometry. After formation of a separating membrane, but prior to isolation of one daughter from the other, the daughter cells are in a state of weak coupling. In this state, transport through tight or gap junctions, particularly in the embryo, may play an important role in maintaining the differentiation. During the strong coupling phase each prospective daughter is described by its average concentrations of the various substituents (ions, hormones, enzymes, etc.) and the state of activation of the genome. We assume that the concentrations in the parent are uniform, and hence so are the initial concentrations in the complex. (The additional but not crucial problem of the distribution of material in each prospective daughter is not considered here.) Concentration differences between the prospective daughters occur by fluctuations. If the fluctuations decay, then symmetric development continues. However, it is possible that the system may become unstable to these asymmetric perturbations, in which case concentration differences will grow. Such instabilities provide a mechanism of the initiation of asymmetric differentiation.

Transport between prospective daughters is due to permeation through incompletely formed dividing membranes (in plant cells), diffusion through the region of common cytoplasm and diffusion within each daughter. (Streaming may also play a role in distributing material.) We assume that the transport from one prospective daughter to the other may be approximated by an effective permeation flux which vanishes with the difference of average concentrations. Thus if $c_\alpha^{(i)}$ is the average concentration of species α in each prospective daughter ($i = 1, 2$) the rate of increase J_α of species α in prospective daughter 1 due to transport from daughter 2 is given by

$$J_\alpha = h_\alpha [c_\alpha^{(2)}(t) - c_\alpha^{(1)}(t)] . \tag{2.1}$$

The effective permeability h_α decreases in time as the cell halves pass from strong to weaker coupling as a result of decreased contact area or completion of the dividing membrane. (Note that more complex processes such as active transport may be included by taking more general force-flux relationships.)

2.2. Analysis

It is convenient to introduce a vector notation. Thus, for example, $\boldsymbol{c}^{(i)}$ represents the column vector of average concentrations of all species in cell half i. Intercellular transport is therefore given by $\boldsymbol{J} = \mathbf{H}(\boldsymbol{c}^{(1)} - \boldsymbol{c}^{(2)})$, where $\mathbf{H}$ is a matrix of effective permeabilities. With this notation we also may now simply include cross diffusional or permeative effects (i.e., the diffusion or permeation of species α driven by a gradient in concentration of species β). Direct transport indicated in (2.1) is described by a diagonal matrix $\mathbf{H}$.

Exogenous agents may be introduced into the dividing cell from the environment due to the permeability of the plasma membrane; loss to the environment may also play a role. Thus we introduce an external permeability (matrix) $\mathbf{G}$. In an external gradient the prospective daughters may in general be subjected to different external concentrations $\boldsymbol{c}^{(i)0}(t)$, and hence the increase in concentration of agents in the half i due to interaction with the environment is given by

$$\boldsymbol{J}^{(i)0} = \mathbf{G}[\boldsymbol{c}^{(i)0}(t) - \boldsymbol{c}^{(i)}(t)] . \tag{2.2}$$

In general the permeability of the plasma membrane, $\mathbf{G}$, may depend on the composition, both inside and outside the cell, but will take the same functional form for both halves, at least in the early stages of differentiation. That is, we assume the prospective daughters to have the same plasma membrane initially; permeation may, however, depend on conditions inside and outside each half of the complex.

The overall rate of change in composition is due both

to transport and reaction. The reaction rates in the two halves are taken to be a function of the *average* concentrations. By our assumption of initial symmetry the functional form $F[c]$ of the chemical rate equations is the same for both halves. Thus, the dynamics of the system is described according to

$$dc^{(1)}/dt = G[c^{(1)0} - c^{(1)}] + H[c^{(2)} - c^{(1)}] + F[c^{(1)}],$$
$$dc^{(2)}/dt = G[c^{(2)0} - c^{(2)}] + H[c^{(1)} - c^{(2)}] + F[c^{(2)}]. \quad (2.3)$$

Undifferentiated (symmetric) division and symmetric differentiation in a homogeneous environment corresponds to the case $c^{(i)}(t) = c(t)$ $(i = 1, 2)$. The remainder of the analysis will be devoted to investigating the stability of this symmetric development to asymmetric perturbations in a homogeneous environment $c^{(i)0} = c^0$. This constitutes the most stringent conditions for the onset of asymmetry. Let $\delta c^{(i)}(t)$ represent the deviation in one half of the complex from symmetric development:

$$c^{(i)}(t) = c(t) + \delta c^{(i)}(t). \quad (2.4)$$

If perturbations satisfying $\delta c^{(1)} - \delta c^{(2)} \neq 0$ grow, then asymmetric differentiation is initiated; otherwise, a small asymmetric fluctuation decays and such differentiation does not occur.

The stability analysis proceeds by linearizing eqs. (2.3) in a small perturbation $\delta c^{(i)}$ [see (2.4)], with the result

$$d\delta c^{(1)}/dt = \mathbf{E}^* \delta c^{(1)} + H[\delta c^{(2)} - \delta c^{(1)}] + \mathbf{R}^* \delta c^{(1)}, \quad (2.5)$$

where the superscript * implies that the quantity is calculated with values of the concentrations in the symmetric development $[c(t)]$, and the elements of $\mathbf{E}^*$ and $\mathbf{R}^*$ are

$$E_{\alpha\beta} = -G_{\alpha\beta} + \sum_{\gamma} (\partial G_{\alpha\gamma}/\partial c_\beta)(c^0_\gamma - c_\gamma), \quad (2.6)$$

$$R_{\alpha\beta} = \partial F_\alpha/\partial c_\beta. \quad (2.7)$$

An equation similar to (2.5) holds for the other half labelled 2.

Let us introduce symmetric and asymmetric perturbations

$$s \equiv \tfrac{1}{2}(\delta c^{(1)} + \delta c^{(2)}), \qquad a \equiv \tfrac{1}{2}(\delta c^{(1)} - \delta c^{(2)}). \quad (2.8)$$

If we define a matrix Ω according to

$$\Omega(t) = \mathbf{E}^* + \mathbf{R}^*, \quad (2.9)$$

we obtain

$$ds/dt = \Omega(t)s, \qquad da/dt = [\Omega(t) - 2H(t)]a. \quad (2.10)$$

We see that the symmetry breaking (a) and symmetry preserving (s) perturbations are not coupled in a homogeneous medium. The possibility for growth of asymmetric perturbations is determined by the solutions of (2.10). If the coefficients of that equation are slowly varying in time then the stability to asymmetric perturbations is given qualitatively by the eigenvalues $z(t)$ of the matrix $[\Omega - 2H]$. If the real parts of all z are negative, asymmetric perturbations will decay; if $\mathrm{Re}(z)$ for at least one eigenvalue is positive for a sufficiently long period of time the asymmetric perturbations will grow.

2.3. Discussion

(1) Instability to asymmetric perturbations may occur due to many causes: (a) reaction mechanisms in the cytoplasm which involve auto- or cross catalysis, or product inhibition, or other homogeneous chemical feedback [14]; (b) periodic precipitation phenomena [15]; (c) coupling of stable and unstable reactions to concentration dependent plasma membrane permeabilities [16] and active transport; (d) heterogeneous feedback mechanisms involving genome activation and repression, membrane and particle (i.e., ribosomes) bound reactions and reactions in well-defined local regions (mitochondria) [17].

(2) In the earliest stages of the mitotic complex, diffusion is likely to be much more important than permeation, simply because the dividing membrane between the two halves of the complex may be just forming. In that case the length scale associated with unstable, inhomogeneous perturbations is of the order of $(D/k)^{1/2}$, where k is the rate coefficient in the rate-determining step of the mechanism responsible for the development of the asymmetry (in the differentiative

system). This length scale must be of cellular dimensions as a necessary condition for the onset of asymmetric differentiation in the early stages dominated by diffusion. A calculation for one part of a glycolytic reaction mechanism involving product enhanced enzymatic activity shows the length scale of unstable perturbations to be of the order of 10^{-4} to 10^{-2} cm [14], consistent with cellular dimensions.

(3) The effective permeability **H** must be limited so that a concentration difference between the two halves of the complex can develop and persist. If the effective permeabilities **H** are very large then the eigenvalues of $[\Omega - 2\mathbf{H}]$, (2.10), are approximately those of $-2\mathbf{H}$ (plus a correction term involving the chemical rates and the exogenous permeability **G**). Therefore the asymmetric perturbations decay quickly as $\exp[-2\mathbf{H}t]$. On the other hand, if the rates of the differentiative system (when stable) and the exogenous permeation are very large, then the decay of the asymmetric perturbations is again fast, approximately $\exp[\Omega t]$. Symmetry breaking requires a delicate balance between **H** and the symmetric properties of Ω.

(4) For differentiation to occur, the time to set up an appreciable asymmetric distribution must be shorter than that during which the prospective daughters are decoupled, either by cleavage or other cytoplasmic decoupling mechanism, or by separation of equivalent local sites of reaction such as the daughter nuclei themselves.

(5) The model provides criteria for the *initiation* of asymmetric perturbations and these are obtained from the *linearized equations* (2.10). The states to which the system evolves after that initiation are given by the nonlinear equations (2.3). For instance, if the differentiative reaction system is unstable to symmetric and asymmetric perturbations, competitive events are possible: symmetric differentiation ($A \rightarrow 2B$); asymmetric differentiation ($A \rightarrow B + C$); and reproduction without differentiation ($A \rightarrow 2A$). The outcome among such competitive routes is determined by non-linear effects in the permeation and reaction mechanisms and is beyond the present analysis.

(6) The developed asymmetry in the two halves of the complex may be sustained in a number of different ways. (a) If the daughter cells remain sufficiently strongly coupled after division, then the same mechanism leading to the growth of asymmetric perturbations may maintain the two daughters in a dissipative structure, a stable time independent asymmetric state. (b) If the final coupling is weak the onset of asymmetry may drive the daughters to different stable states of the isolated cell system. Indeed the same types of mechanisms leading to symmetry-breaking instability may lead to multiple stable steady states (see appendix). (c) If the onset of weak coupling leaves one of the daughters with an excess of a chemical species (enzyme or RNA) which persists for a long time, then the reaction mechanisms in one daughter may be altered, compared to that of the parent, during that time. In particular, some reactions may be essentially irreversible. (d) Chemical composition changes due to the growth of asymmetry may bring about conformational transitions, such as in cell membranes which result in changes in transport and catalytic properties. (e) Asymmetric concentration differences may lead to irreversible genetic changes with attending changes in reaction mechanisms.

(7) The inclusion of external gradients in the analysis can be made without difficulty. Unlike random fluctuations, such gradients provide a well-defined axis and polarity for the development. If the development of an asymmetric mitosis is due to a symmetry breaking instability, then only vanishingly small gradients are required for including polarity.

(8) In the appendix we illustrate the theory with a simple, product-enhanced, reaction mechanism and find conditions for the existence of three symmetric states. One of the symmetric states is stable over its entire range of existence. Another, however, while stable to symmetric perturbations, may become unstable to asymmetric perturbations as the interdaughter permeative coupling is decreased. Hence the possibility for the initiation of asymmetric differentiation depends on the initial symmetric state in which the complex has been prepared.

3. Evidence for proposed theory

3.1. Melanocytogenesis

The extensive studies of hormone-induced differentiation of melanoblasts into melanocytes [18] (in explants from xanthic goldfish tailfin) provides in our view a set of data for which the proposed theory serves in qualitative, but broad terms. A brief summary of the

findings on melanocytogenesis (the transformation of a melanoblast into a melanocyte) in ref. [18] are:

(i) Melanoblasts are stem cells which may undergo asymmetric mitosis upon stimulation by several different hormones; a melanocyte and a melanoblast are the result of asymmetric mitosis. (ii) Melanocytes do not appear in pairs. (iii) Mitosis is essential for hormone-induced melanocytogenesis. (iv) Melanoblasts cannot synthesize melanin, but melanocytes can. (v) The synthesis of melanin requires tyrosine, oxygen, and tyrosinase; protein synthesis after mitosis is not necessary for the synthesis of melanin. (vi) "Since the phenotypic appearance of active tyrosinase does not occur in melanoblasts which have been treated with hormone but have not undergone mitosis, one can postulate further that hormone ... causes a segregation of tyrosinase and its inhibitor during mitosis. The end result of such action is that one of the daughter cells now possesses active tyrosinase ..." "... the primary effect of the hormone is not at the genetic, but at the cytoplasmic level, and leads to the release or segregation of tyrosinase from its inhibitor during mitosis". There is evidence that the functioning of the enzyme tyrosinase may not be due to the removal of an inhibitor but due to the activation of a proenzyme [19]. (vii) As the hormone concentration is increased the probability of occurrence of melanocytes increases sharply.

What is the concordance of these experimental results and the theory? The result of the asymmetric mitosis is two cells, one containing tyrosinase but no inhibitor (or proenzyme plus activator), the other containing inhibitor (or proenzyme and no activator) (vi). Thus a tyrosinase inhibitor (activator) concentration difference is established during mitosis. Since hormone stimulation induces asymmetric mitosis (i), and the absence of hormone leads to symmetric mitosis, then there exists no asymmetry in the architecture of the cell prior to hormone stimulation. The hormone induces or completes the differentiative chemical reaction system. Upon mitosis, this system, coupled with diffusion and permeation, becomes unstable to asymmetric perturbations which grow and bring about asymmetric differentiation (i, ii). If mitosis is inhibited melanocytogenesis is absent (iii). During the asymmetric differentiation active tyrosinase appears in one daughter only (vi), a case of an asymmetric concentration growth from an initial symmetric distribution. The daughter with a low concentration of tyrosinase inhibitor (high concentration of protyrosinase activator) is the melanocyte (iv, v).

To what extent do the other theories of asymmetric differentiation (cited in the Introduction) fit these experimental findings? The fact that symmetric mitosis occurs in the absence of hormone indicates the lack of inherent architectural asymmetry. Second, the experiments were carried out in the absence of external gradients. The third mechanism is possible but not likely because of the sharp increase of melanocytogenesis at hormone concentrations (10^{-3} mg/ml) which exceed by many orders of magnitude the concentrations of stem cells (vii). Finally, a hormone-induced genetic irreversible mechanism may initiate the asymmetric differentiation in which case one of the daughter nuclei activates the tyrosinase system. This usually requires protein synthesis, however, which is contrary to the available evidence (v).

The theory based on chemical symmetry breaking fits the available experimental findings but more thorough substantiation is required. We may use the theory to make some predictions and pose some questions which lead to suggestions for new experiments. First, if a chemical symmetry-breaking mechanism is the basis of certain types of asymmetric differentiation, then a gradient of a chemical activator or inhibitor in one direction within the dividing cell is accompanied by a gradient of another substance in the opposite direction. Second, the differentiative system, which becomes unstable during mitosis, can be stabilized by chemical changes which affect the differentiative system without necessarily interfering with mitosis itself. Thus we suggest that melanocytogenesis may be arrestable by chemical means other than mitotic inhibitors. Such experiments may give some information on the chemical species and reaction mechanisms of the differentiative system. Third, the possibility of the occurrence of a symmetry breaking instability in a system with reaction and diffusion suggests the search for such instabilities in cell extracts upon hormone stimulation [6–8]. Fourth, a crucial test of the theory is an experiment which questions the possibility of asymmetric cell division in a homogeneous environment (such as in a suspension of cells at low density cultured in a gradient-free medium). Fifth, a vanishingly small external gradient should polarize the asymmetry.

3.2. Eggs of Fucus

Cell differentiation in the eggs of Fucus provides another example for which the theory may possibly be useful. The initial state of a Fucus zygote is nearly spherically symmetric. The first division is asymmetric and yields an elongated rhizoid and a more or less spherical thallus. Some important observations [5] made on this system are: (1) In the absence of imposed gradients the orientation of the polarity of the differentiation is random. (2) The asymmetry of the differentiation is *polarizable* by the imposition of gradients of chemical composition and voltage, or light. (3) More than one rhizoid may be formed; this was accomplished by illuminating zygotes with two antiparallel sources. (4) As the zygote matures a potential develops across cell membrane (inside negative) in association with the development of large concentration differences across the cell membrane. After the membrane potential reaches a critical value of around 80 mV there develops a current loop going through the cell from rhizoidal pole to thallus, and is then completed by an extracellular pathway. The onset of current is accompanied by a 5–10 mV drop in trans-membrane potential.

It is difficult to reconcile these experimental results with any of the four mechanisms not based on chemical symmetry breaking cited in the Introduction. A detailed analysis of this system within the framework of the theory proposed here seems warranted.

Appendix: A model system

We consider a simple model system having multiple steady states, one of which may become unstable to asymmetric perturbations. The reaction mechanism is a modification of one used [7] to demonstrate chemical oscillations and symmetry breaking instability. Our system consists of species S and P participating within the cell in the reaction mechanism:

$$A \rightarrow S, \tag{A.1}$$

$$2P + S \rightarrow 3P, \tag{A.2}$$

$$S \rightarrow D, \tag{A.3}$$

$$P \rightarrow E. \tag{A.4}$$

The trimolecular reaction is an analogue for a product-enhanced enzymatic process[‡]. Concentrations of species A, D and E are kept constant within the dividing cell and the reverse reactions are neglected. The permeability of S and P through the plasma membrane is taken to be zero. A and P are supplied from an internal pool or biosynthetic pathway. The rate equations describing the prospective daughter 1 may be written

$$dS^{(1)}/dt = -kS^{(1)}[P^{(1)}]^2 + q_1 A - q_2 S + h_s(S^{(2)} - S^{(1)}), \tag{A.5}$$

$$dP^{(1)}/dt = kS^{(1)}[P^{(1)}]^2 - q_3 P + h_p(P^{(2)} - P^{(1)}), \tag{A.6}$$

where k and q_i ($i = 1, 2, 3$) are rate coefficients, and h_s, h_p are permeability coefficients for exchange between prospective daughters. A similar equation holds for prospective daughter 2. An equivalent model system is one in which reaction (A.2) takes place in a membrane bound volume immersed in a medium maintained at fixed S^0 and P^0. We have

$$dS^{(1)}/dt = -kS^{(1)}[P^{(1)}]^2 + g_s(S^0 - S^{(1)}) + h_s(S^{(2)} - S^{(1)}), \tag{A.7}$$

$$dP^{(1)}/dt = kS^{(1)}[P^{(1)}]^2 + g_p(P^0 - P^{(1)}) + h_p(P^{(2)} - P^{(1)}), \tag{A.8}$$

where g_s, g_p are plasma membrane permeability coefficients. It is seen that (A.7) is equivalent to (A.5). In the limit $P^0 \rightarrow 0$, (A.8) is equivalent to (A.6). We now consider all rate coefficients and permeabilities to be constant and work with the notation in (A.7) and (A.8) choosing units such that $k \equiv 1$. We limit ourselves to the case $P^0 = 0$. We shall find the symmetric and asymmetric steady states of (A.7) and (A.8) and shall analyze their stability as a function of the source concentration S^0 and the interdaughter coupling h_s, h_p.

The symmetric steady states may be found from (A.7) and (A.8) by equating the concentrations in the prospective daughters $S^{(i)} = S$, $P^{(i)} = P$. Eliminating S we obtain

$$P[P^2 - \delta P + g_s] = 0. \tag{A.9}$$

‡ Consider an enzyme E to occur in three states, E, EP and EP_2. If EP_2 is the active form and if the equilibrium between the three forms of enzymes are fast then the effective rate law for the process $S \xrightarrow{[EP_2]} P$ takes the trimolecular form in the low P limit.

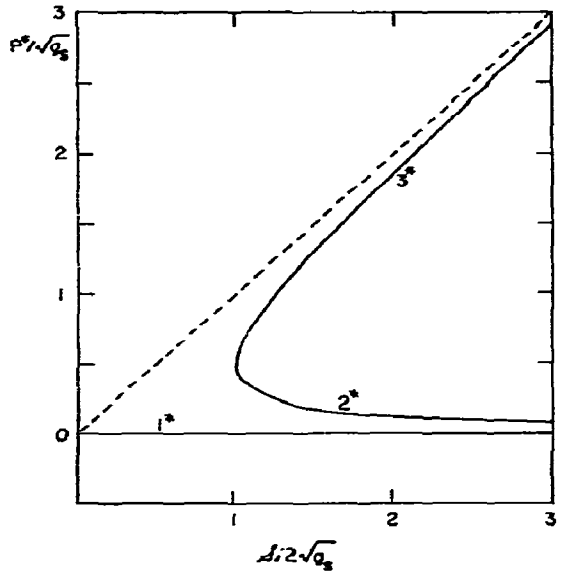


Fig. 1. Branches of multiple stationary states calculated for the system (A.7), (A.8). The steady-state concentration of P, denoted by P^*, is given as a function of the external concentration σ, see (A.10).

where σ is a normalized external concentration,

$$\sigma = g_s S^0/g_p. \tag{A.10}$$

We find three symmetric steady states which are given by

$$1^*;\quad P^* = 0,\quad S^* = S^0, \tag{A.11}$$

$$2^*;\quad P^* = \tfrac{1}{2}\sigma - \tfrac{1}{2}(\sigma^2 - 4g_s)^{1/2},\quad S^* = g_p/P^*, \tag{A.12}$$

$$3^*;\quad P^* = \tfrac{1}{2}\sigma + \tfrac{1}{2}(\sigma^2 - 4g_s)^{1/2},\quad S^* = g_p/P^*, \tag{A.13}$$

and shown in fig. 1. From (A.12) and (A.13) it may be seen that unless

$$\sigma > \sigma_c \equiv 2(g_s)^{1/2}, \tag{A.14}$$

there is only one steady state solution. In fig. 1, the three steady states are shown as functions of σ. If P^0 is increased from its assumed value of zero, then the steady-state values of P for the 1^* branch are positive and increase slowly with σ; the 1^* branch therefore intersects the 2^* branch at non-zero σ when $P^0 > 0$.

The stability of the steady states may be tested with respect to asymmetric as well as symmetric perturbations. The analysis is essentially the same as that outlined in section 2. Here we have the simplification that the matrices Ω and H in (2.10) are constant.

State 1^* is always stable. State 2^* is always unstable over the entire range of its existence $\sigma_c < \sigma < \infty$. An upper critical value of σ, σ_u, defined by

$$\sigma_u(H) = \frac{2g_s(g_p + 2h_p) - \Delta_0}{\{(g_p + 2h_p)[g_s(g_p + 2h_p) - \Delta_0]\}^{1/2}}, \tag{A.15}$$

where

$$\Delta_0(H) = 2[2h_p(h_s + g_s) - g_p h_s], \tag{A.16}$$

may be shown to exist such that for $\sigma_c < \sigma < \sigma_u$, 3^* is unstable to asymmetric perturbations. If now we introduce

$$\alpha \equiv h_p/h_s;\qquad h_s = h \tag{A.17}$$

we find that at a value of h, denoted by h_c, such that $\sigma_u = \sigma_c$, the 3^* branch is always stable for $h > h_c$. Noting that $\sigma_u = \sigma_c$ when $\Delta_0 = 0$, we find that h_c is given by

$$h_c = (g_p/2\alpha) - g_s. \tag{A.18}$$

Since the symmetric steady state 3^* is unstable to asymmetric perturbations for $\sigma_c < \sigma < \sigma_u$ we might expect to find asymmetric steady-state structures. The condition for $\sigma_u > \sigma_c$ is equivalent to $\Delta_0(H) < 0$ or

$$0 < h < h_c. \tag{A.19}$$

With this we proceed to calculate the asymmetric steady states. For arbitrary h_s, h_p the calculation is straightforward but tedious. Therefore we consider the simple case $h_p = 0$ for which $h_c = \infty$, (A.19); this retains some interesting features.

Setting the r.h.s. of (A.7) and (A.8) equal to zero we may eliminate $S^{(i)}$ and obtain

$$P^{(2)*} = hP^{(1)*}[(P^{(1)*})^2 - \sigma P^{(1)*} + g_s + h]^{-1}, \tag{A.20}$$

and a similar equation with the superscript 1 and 2 interchanged. Combining these two equations we obtain a quintic equation in $P^{(1)}$ which factors into a term iden-

tical to the one on the left-hand side of (A.9), corresponding to the three possible symmetric steady states and a quadratic factor corresponding to two asymmetric steady states. The asymmetric steady-state concentration of $P^{(1)*}$ is given by

$$P_{\pm}^{(1)*} = \tfrac{1}{2}\mathcal{J}' \pm \tfrac{1}{2}[\mathcal{J}^2 - 4(g_s + 2h)]^{1/2}, \qquad \text{(A.21)}$$

where

$$\mathcal{J}' = [(g_s + 2h)/(g_s + h)]\,\mathcal{J} > \mathcal{J} \qquad \text{(A.22)}$$

and the corresponding $P_{\pm}^{(2)*}$ is then obtained by symmetry or from (A.20). It is seen that real solutions can be found only for the condition

$$\mathcal{J} > \mathcal{J}_L = 2(g_s + h)/(g_s + 2h)^{1/2}. \qquad \text{(A.23)}$$

For the present case $h_p = 0$ we note that $\mathcal{J}_u = 2\mathcal{J}_L$.

The stability analysis of the symmetric states is summarized in fig. 2 with a phase diagram in the normalized external concentration $\mathcal{J}$ (A.10) and the interdaughter coupling strength $h(\equiv h_s)$ at fixed $\alpha(\equiv h_p/h_s)$. We consider only the case $P^0 = 0$. State 1^* is stable in all domains, I–III. State 2^* is unstable in all domains where it exists, II, III. State 3^* is stable in domain II.

In the initial state of mitosis, the prospective daughters are strongly coupled, that is h is large, and the cell complex is taken to be in a symmetric stable state, corresponding to conditions such as point A in fig. 2. As mitosis progresses, the permeability h decreases. If the system at point A is in stable state 3^*, then on decreasing h at constant $\mathcal{J}$ the system may pass into domain III where it is unstable to asymmetric fluctuations and therefore asymmetric differentiation may begin. If, however, the system at point A had been in the symmetric state 1^* then the mitosis would have developed symmetrically, since the state 1^* is stable with respect to asymmetric as well as symmetric perturbations. The course of the mitosis is thus seen to depend on the preparation of the initial state of the cell.

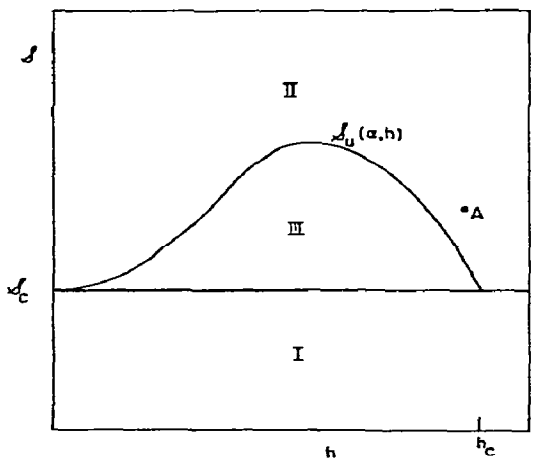


Fig. 2. Schematic phase diagram in the external concentration, $\mathcal{J}$, (A.10) and interdaughter coupling, h, (A.17). The domains labelled by Roman numerals represent regions of stability of the symmetric states as described in the text near (A.23).

Acknowledgement

We are most grateful to Professor Eugene Bell, Department of Biology, Massachusetts Institute of Technology, for a number of essential discussions and a critical reading of the manuscript.

References

[1] J.G. Carlson, Chromosoma 5 (1952) 199.
[2] L. Wolpert, Current Topics Devel. Biol. 6 (1971).
[3] C.H. Waddington, Principles of embryology (George Allen and Unwin, London, 1956) p. 348 ff.
[4] E.B. Wilson, The cell in development and heredity, 3rd revised Ed. (MacMillan, New York, 1928) p. 323 ff.
[5] L.F. Jaffe, Develop. Biol. Suppl. 3 (1969) 83.
[6] A.M. Turing, Phil. Trans. Roy. Soc. (London) 237B (1952) 37.
[7] P. Glansdorff and I. Prigogine, Thermodynamic theory of structure, stability and fluctuations (Wiley–Interscience, New York, 1971).
[8] M. Herschkowitz, Compt. Rend. Acad. Sci. (Paris) 270C (1970) 1049.
[9] P. Ortoleva and J. Ross, J. Chem. Phys. 56 (1972) 4397.
[10] G. Nicolis, Advan. Chem. Phys. 19 (1971) 209.
[11] B.C. Goodwin and M.H. Cohen, J. Theoret. Biol. 25 (1969) 49.
[12] P. Ortoleva and J. Ross, J. Chem. Phys. 56 (1972) 287.
[13] D. Yaffe, in: Cell differentiation, eds. R. Harris, P. Allin and D. Visa (Munksgaard, Copenhagen, 1972) p. 263.
[14] I. Prigogine, R. Lefever, A. Goldbeter and M. Herschkowitz, Nature 223 (1969) 913.
[15] M. Flicker and J. Ross, A Model of Chemical Instability for Liesegang Rings, unpublished.
[16] H. Hahn, P. Ortoleva and J. Ross, J. Theoret. Biol. 41 (1973) 503.

[17] E.K. Bimpong-Bota, P. Ortoleva and J. Ross, Localized Chemical Instabilities, unpublished.
[18] T.T. Tchen, R.N. Ammeraal, K.-H. Kim, C.M. Wilson, W. Chavin and F. Hu, U.S. National Cancer Institute Monograph No. 13, ed. W.J. Rutter (1963) p. 67.
[19] J.S. McGuire, Biochem. Biophys. Commun. 40 (1970) 1085.

BIOPHYSICAL CHEMISTRY 1 (1973) 97–103. © NORTH-HOLLAND PUBLISHING COMPANY

THE REVERSIBLE TITRATION OF TYROSYL RESIDUES IN HUMAN DEOXYHEMOGLOBIN *

Enrico BUCCI

Department of Biological Chemistry, University of Maryland School of Medicine, Baltimore, Maryland 21201, USA

Received 14 June 1973

The reaction of hemoglobin with N-acetyl imidazole at neutral pH indicated that in carboxyhemoglobin 1.80 residues per heme were acetylated while in deoxyhemoglobin only 1.15 residues were available to the reagent. The reversible titration of these residues in alkali was followed by difference spectrophotometry at 245 nm. Hill plots of the titration data, assuming 2 residues titrable per heme and $\Delta\epsilon = 10500$ per tyrosyl residue upon ionization, showed a slope of 1.5 and a pH 1/2 near 11. The average pK of these groups in carboxyhemoglobin was previously found to be near 10.5. Also, by difference spectrophotometry it was shown that exposure of deoxyhemoglobin to alkaline pH was accompanied by a modification of the Soret region of the absorption spectrum, which might indicate the appearance of liganded conformation in the deoxyhemoglobin system. The sedimentation velocity of deoxyhemoglobin demonstrated that at alkaline pH dissociation into dimers occurred at pH's lower than 10, where no ionization of tyrosines was detectable. The titration of tyrosines was independent from protein concentration.

The low availability of tyrosyl residues to acetylation in deoxyhemoglobin, the cooperativity of proton binding of these residues and the change in conformation of hemoglobin concomitant with their titration are all consistent with results of Simon et al., Moffat, and Moffat et al., and with the model proposed by Perutz for explaining the heme–heme interaction. The free energy of the pK shift of the tyrosyl residues in carboxy and deoxyhemoglobin can be included in the free energy of the heme–heme interaction.

1. Introduction

Recent papers of Perutz [1], Moffat et al. [2], Simon et al. [3] and Moffat [4], indicate the relevance of the penultimate tyrosyl residue of the hemoglobin subunits to the mechanism of the heme–heme interaction. It appears that deoxygenation of hemoglobin is accompanied by the hiding of that residue inside a hydrophobic pocket where it is stabilized by a hydrogen bond with the peptide backbone of the chains [1]. This change in conformation should be accompanied by a decrease in availability of tyrosyl residues to acetylation, and by an increase of their pK value. Also it is expected that the negative charges acquired upon deprotonation would expel those residues from the hydrophobic pocket described by Perutz [1]. This might produce a conformational change which should mimic the restoration of some liganded structure in the deoxygenated system.

The results here presented are consistent with this picture. The free energy of the pK shift of the tyrosyl residues in liganded and unliganded hemoglobin should probably be included in the free energy of the heme–heme interaction.

2. Experimental procedure

Human hemoglobin was prepared from toluene hemolyzates [5] and purified from organic and inorganic ions by recycling for one hour in cold through a mixed bed resin column.

Heme concentration was measured spectrophotometrically on the basis of the following extinction coefficients: $14000\,M^{-1}\,cm^{-1}$ for the CO derivative at 540 nm, $190000\,M^{-1}\,cm^{-1}$ for the CO derivative at 420 nm and $13000\,M^{-1}\,cm^{-1}$ for the deoxy derivative at 555 nm.

* This project was supported in part by PHS grants HL13164 and HL13178. Computer time was supported through the facilities of the Health Science Computer Center of the University of Maryland School of Medicine in Baltimore, Maryland, and of the Computer Science Center of the University of Maryland in College Park, Maryland.

Removal of the heme was obtained by titrating the deionized protein solution to pH 3 with 1N HCl and pouring it into 10 to 30 volumes of acetone at room temperature. The precipitate was washed 3 times with acetone, dissolved in 0.004 M HCl and dialyzed for 24 hours or more against at least two changes of a large volume of 0.004 M HCl in order to completely remove the acetone. The final solutions were clarified by filtration through Millipore membranes and used for amino acid analysis and spectrophotometric measurements. On the basis of the yield of arginines and lysines the following extinction coefficients per $\alpha\beta$ dimer were calculated for apohemoglobin in 0.004 M HCl: 28500 $M^{-1} cm^{-1}$ at 278 nm and 16800 $M^{-1} cm^{-1}$ at 290 nm.

Amino acid analyses were performed with a Beckman 120 c autoanalyzer using the standard 4 hour run. Hydrolyzates were prepared heating the protein in 6 N HCl at 110°C for 24 hours in sealed evacuated tubes. The ratio between the yields of lysines and arginines always corresponded to that expected from the amino acid composition of human apohemoglobin. The yield of histidines was 5–10% lower than expected. Different hydrolysis times, 18 or 30 hours, did not modify appreciably the yields of lysines and arginines.

Reaction of CO or deoxyhemoglobin with acetylimidazole was performed by adding 75 mg of the reagent to 40 mg of protein in a final volume of 10 ml, in 0.04 M veronal buffer at pH 7.5 [6]. After one hour at room temperature the solution was filtered through Sephadex G25 equilibrated with water. When deoxyhemoglobin was acetylated, carbon monoxide was rapidly bubbled through the solutions immediately before the gel filtration, which required approximately 5 min. Parallel acetylation of CO and deoxyhemoglobin was obtained by transferring part of the deoxygenated mixture of hemoglobin and acetylimidazole into another container which was flushed with carbon monoxide. To ensure a complete anaerobiosis and to avoid formation of ferric hemoglobin, dithionite was added in amounts varying from 0.02 to 1 mg/ml. This did not seem to affect the number of tyrosyl residues acetylated in either CO or deoxyhemoglobin.

Acetylimidazole was purchased from K & K and recrystallized. N-acetyltyrosine and N,O-acetyltyrosine were purchased from Cyclo Chemical and used as such. Dissolved in 0.004 M HCl, they showed a negligible absorption at 290 nm and an extinction coefficient at 273 nm of 1270 $M^{-1} cm^{-1}$ and 210 $M^{-1} cm^{-1}$, respectively. Following the procedure of Riordan et al. [6] the difference between the two coefficients, $\Delta\epsilon = 1060\ M^{-1} cm^{-1}$, was referred to the O-acetylation and used to calculate the number of tyrosyl residues acetylated in apohemoglobin.

Deoxygenation was obtained by flushing the protein solutions with nitrogen or bubbling nitrogen through the various solvents and/or reagents. Prepurified nitrogen was used, passed through copper at high temperature and humidified with water. The gas was sent into the various containers through polyethylene capillary tubings inserted in the vessels through soft rubber caps. Anaerobic transfer of the deoxygenated solutions from one container to another was accomplished forcing the solutions through the tubings with nitrogen pressure. To avoid turbidity hemoglobin solutions were deoxygenated at a minimum concentration of 1 mg/ml, dilution with the appropriate solvent was done after deoxygenation. To test the degree of deoxygenation the ratio between the absorptions at 555 and 540 was measured. This ratio is 1.24 in deoxyhemoglobin [7]. During the course of acid base titrations ratios of 1.23 to 1.25 were considered acceptable. Ratios lower than 1.23 indicated incomplete deoxygenation, while ratios higher than 1.25 indicated the presence of ferrous hemochromogen, which shows an absorption near 558 nm approximately 2 times higher, and an absorption at 540 nm slightly lower, than the correspondent absorptions of deoxyhemoglobin.

Deoxygenated solutions of hemoglobin were stable for at least one day when stored inside spectrophotometric cuvettes. However, when stored inside ultracentrifuge cuvettes spectrophotometric measurements, conducted directly on the cells, demonstrated that a partial reoxygenation was already present after 30 min, and within 2 hours it was practically completed. Adding dithionite to these solutions retarded the reoxygenation and 0.012 mg/ml of dithionite appeared to ensure a complete deoxygenation for several hours. The reported amounts of dithionite refer to weighted quantities of the reagent.

Difference spectrophotometry of deoxyhemoglobin solutions at different pH's was performed preparing large amounts, 100 to 400 ml, of a stock solution at the desired reference pH and, after filling the reference cuvette, readjusting the pH of the solution to the value desired for the sample cuvette, by addition of 0.1 to 0.5 N NaOH or HCl. In this way continuous forward and backward titrations could be followed by difference spectrophotometry. The total dilution of the initial

solution produced by the reagents was kept below 1% and disregarded. To adjust the pH, deoxygenated solutions of NaOH and HCl were used with gas-tight, gas-flushed syringes, whose needles were inserted through the soft rubber caps into the titration vessel. All of the experiments were performed in 0.25 M NaCl and 0.004 M glycine at 23°C. Addition of ethylenediaminetetraacetic acid up to a concentration of 10^{-3} M was without effect on the titrations. The titration vessel was flushed with nitrogen for the period of the titration.

A Cary 14 instrument was used for spectrophotometric measurements.

Measurements of pH were performed with a Radiometer M26 instrument equipped with a GK2302B combination electrode. This was inserted through the rubber stopper of the titration vessel and the KCl solution inside the electrode was bubbled with nitrogen and kept under a stream of nitrogen during the anaerobic titrations.

Sedimentation velocity was measured with a Beckman Model E ultracentrifuge. The absorption optics was used selecting the mercury band near 405 nm with a Farrand interference filter. The films were scanned with a Joyce and Loebl microdensitometer, and the tracings digitized with a Thompson PF/10 digitizer. The boundary position was calculated from the second moment of the boundary distribution

$$\bar{r}^2 = r_p^2 - (2/C_p)\int_{r_m}^{r_p} r c dr,$$

where C_p is the protein concentration in the plateau region, r_m and r_p are the radial position of the meniscus, and of an arbitrary point in the plateau region, respectively. These measurements agreed very well with the positions determined by the interception of the boundary with the horizontal line drawn at half the optical density difference between the meniscus and the plateau region, indicating the symmetry of the boundaries. This was confirmed by experiments of superimposition at the illuminated table. After assembly, the ultracentrifuge cells were immersed in water with dithionite and all the operations of gas flushing and filling with protein solutions were carried out under water. Deoxygenation of the samples was checked spectrophotometrically at the beginning and end of each run. All runs were performed in 0.25 M NaCl, 0.01 M glycine and 10^{-5} to 10^{-3} M ethylenediaminetetraacetic acid at temperatures comprised between 20 and 23°C.

Table 1

Number of tyrosines per heme acetylated by acetyl imidazole in CO and deoxyhemoglobin

Preparation number	Derivative of hemoglobin	acetyl tyrosines per heme
2	deoxy	1.18
5	deoxy	1.10
2	CO	1.82
5	CO	1.82

3. Results

3.1. Acetylation of tyrosyl residues in presence and absence of ligand

Acetylation was performed in parallel on CO and deoxyhemoglobin following a procedure very similar to that described by Riordan et al. [6]. After removal of the heme, the apohemoglobins were analyzed spectrophotometrically in 0.004 M HCl. Quantitative amino acid analyses indicated that the acetylation did not modify the extinction coefficient of apohemoglobin at 290 nm. Therefore, the ratios between the absorptions at 278 and 290 nm, were used for calculating the number of tyrosyl residues acetylated.

Table 1 reports the results obtained in two different acetylations. It is possible that the values slightly greater than unity, obtained for the tyrosyl residues acetylated in deoxyhemoglobin, were due to the addition of carbon monoxide just before the filtration through Sephadex. The addition was indispensable for avoiding extensive denaturation of the protein during the process. In CO hemoglobin approximately one more residue per heme was available to acetylation than in the absence of ligand. The values reported here for CO hemoglobin correspond very well with those reported by Riordan and Vallee for bovine CO hemoglobin [6].

3.2. Titration of tyrosyl residues

The titration was followed by difference spectrophotometry. The reference solution was kept between pH 7 and 9, and the difference spectrum scanned from 340 to 240 nm. Only when the sample solution was above pH 10 meaningful difference spectra were re-

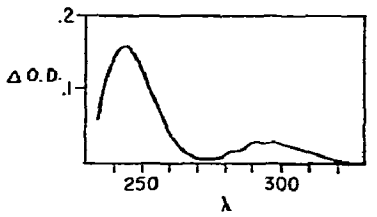


Fig. 1. Difference spectrum of deoxyhemoglobin at pH 9.4 (reference compartment) and at pH 10.61 (sample compartment). Heme concentration 19.2 × 10^{-6} M. 2 cm path cell.

corded. A typical spectrum is shown in fig. 1. It is very similar to that obtained when alkaline and neutral solutions of free tyrosines are compared [8]. The only difference was in the region near 290 nm where some fine structure was present. This might suggest a contribution of tryptophyl residues to the difference spectrum [9, 10]. The OD at 245 nm was plotted against pH. In fig. 2, the data of several experiments are collected and normalized to 20.5 x 10^{-6} M heme for a 2 cm path cuvette. The experiments were performed in three ranges of concentration, near 20, 40 and 80×10^{-6} M heme. The titrations appeared independent from the concentration of the protein. All of the titrations not found to be reversible within the scattering shown in fig. 2, were discarded. For this reason no data are reported above pH 10.7.

The reaction with acetyl imidazole at neutral pH indicated that nearly 2 residues per heme were exposed to the solvent in CO hemoglobin and nearly 1 per heme in deoxyhemoglobin. This was because, as seen by Perutz [1], Moffat et al. [2], Simon et al. [3], and Moffat [4], one of them becomes inaccessible inside a hydrophobic pocket upon deoxygenation. However, this residue should be expelled from that pocket upon deprotonation because of the acquired negative charge. Therefore it was assumed that in deoxyhemoglobin 2 residues per heme were titrated in alkali. From data obtained in a previous investigation [11] $\Delta\epsilon$ = 10500 M^{-1} cm^{-1} per tyrosyl residue at 245 nm was assumed.

Fig. 3 shows the usual plot of $\log[\alpha/(1-\alpha)]$ against pH for calculating the apparent pK of ionizable groups. This plot is a "Hill plot" as defined by Wyman [12], and a slope higher than 1.0 would indicate cooperativity of binding. The line traced in fig. 3 was obtained with the least squares procedure and shows a slope very near 1.5. Assuming only 1 residue titratable per heme produced a non-linear plot with a slope increasing to 3 with the ionization. Assuming 3 titratable groups per heme slightly lowered the slope of the plot to near 1.4. The value of 1.5 found for the slope of the plot was a minimum because: (a) it was calculated in the initial region of the titration of the groups, where the slopes of "Hill plots" tend to slow down to a value of 1 [12];

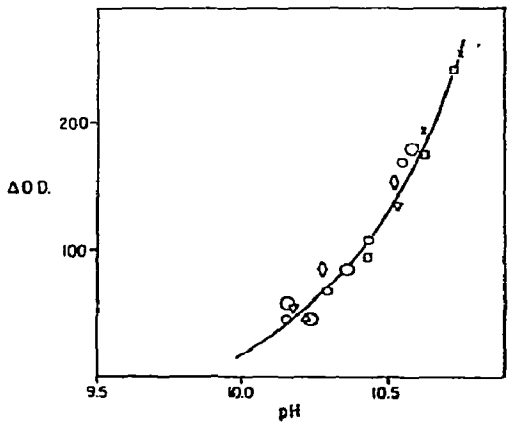


Fig. 2. Spectrophotometric titration of deoxyhemoglobin as described in the text. Different symbols refer to different experiments. All data have been normalized to 20.5 × 10^{-6} M heme. 2 cm path cell.

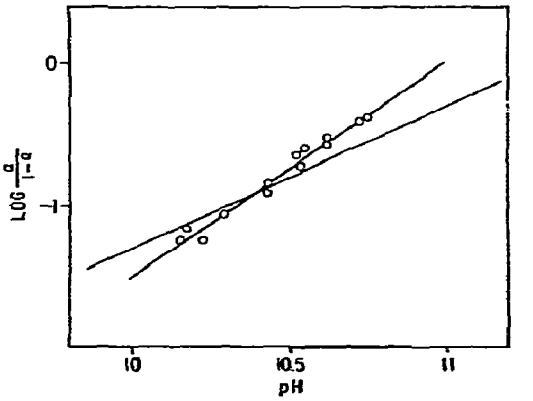


Fig. 3. Hill plot of the data shown in fig. 2. It was assumed that 2 tyrosyl residues were titrated. A line with a slope = 1 has been traced for comparison. The interpolating line was obtained with the least squares procedure.

Table 2

Sedimentation velocity of deoxyhemoglobin at alkaline pH in presence of about 0.012 mg/ml of dithionite. Runs were performed at the indicated pH in 0.25 M NaCl, 0.004 M glycine and 10^{-3} M ethylenediaminetetraacetic acid. The weight fraction of the dimeric species was calculated by the method of Kirschner and Tanford [13], using values of $S_{20,w}$ 4.55 and 2.85 for the tetramer and the dimer, respectively [14]

Run	pH	Heme conc. $\times 10^6$	$S_{20,w}$	Weight fraction of dimers
98	6.7	4.9	4.56	–
117	9.62	20.7	4.12	0.25
113	9.80	22.4	3.94	0.36
126	9.82	17.9	3.89	0.38
116	10.11	20.03	4.08	0.27
123	10.25	18.84	4.02	0.31
112	10.45	21.80	3.80	0.44

(b) the pK of the two groups involved were probably not identical and not all of them participated to the positive interaction; (c) no correction for the negative electrostatic interaction predicted by the Lindstrom–Lang model was applied to the titration data. Their slopes seem to indicate a cooperative protonation of tyrosines in deoxyhemoglobin.

3.3. Sedimentation of hemoglobin at alkaline pH

Table 2 shows the results obtained. It can be assumed that the low protein concentration and the presence of 0.25 M NaCl made the virial coefficients too small to affect the sedimentation data. Also, previous investigations demonstrated that at 20°C in 0.25 M NaCl, the hemoglobin molecule expands above pH 10.0, the expansion being reversible at least up to pH 10.4 [15]. On this basis it appears that, under conditions similar to those used for the spectrophotometric titrations, a substantial dissociation of deoxyhemoglobin was present at pH's where no ionization of tyrosyl residues was detectable.

3.4. Changes in conformation associated with the titration

Fig. 4 shows the difference spectrum in the Soret region obtained comparing solutions of deoxyhemoglobin at pH 9.1 and 10.6. A similar result is reported by Andersen et al. [16]. The spectrum very much resembles the one reported by Brunori et al. [17] as the difference spectrum between the isolated chains and the product of their recombination in the absence of ligands. It indicates a change in the conformation of the heme pocket and, if we accept the idea that the isolated chains are still in their liganded form when deoxygenated [18], this might indicate that the exposure of hemoglobin to alkaline pH produced the appearance of a detectable amount of liganded conformation in the system.

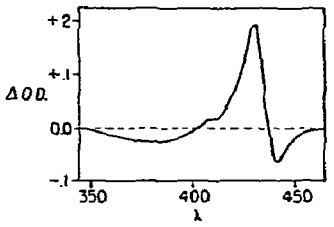


Fig. 4. Difference spectrum in the Soret region of hemoglobin at pH 9.1 (sample compartment) and at pH 10.6 (reference compartment). Heme concentration 20.2×10^{-6} M. 1 cm path cell.

Because of the very high sensitivity of this region of the absorption spectrum to formation of ferric forms and to reoxygenation of hemoglobin, these experiments had to be performed in the presence of dithionite (1 mg/ml). When the difference spectrum was scanned between 480 and 240 nm it was apparent that the modifications of the absorption in the Soret region were strictly concomitant with the ionization of tyrosyl residues. Attempts to measure titration curves in the Soret region were frustrated by poor reversibility.

4. Discussion

Previous work indicated that under conditions very similar to those used here the average pK of tyrosyl residues in liganded hemoglobin is near 10.5 [11]. This pK was found to be considerably greater in deoxyhemoglobin, consistent with findings of Nagel et al. [19]. However, they do not report high values for the slopes of their titrations. The higher ionic strength used here

might have produced a different overlapping of the titration of the groups involved, resulting in a more clear interaction. Also, they did not test the reversibility of their titrations. Andersen et al. [16] claim that the titration of tyrosines in deoxyhemoglobin is linked to the dissociation of the protein into dimers and this would explain the steepness of the titrations. It was not possible to confirm their conclusions. According to the results shown in table 2, dissociation of deoxyhemoglobin in alkali preceded the ionization of tyrosines. The sedimentation values obtained here are lower than those reported by Andersen et al. [16]. The discrepancy might be due to the higher ionic strength used here. In fact, Briehl [20] has shown that a detectable dissociation of deoxyhemoglobin is produced by 1 M NaCl already at pH's near 9. Also, the reversible titration of those residues was independent from protein concentration. On this basis it is believed that the dissociation of deoxyhemoglobin in alkali and the ionization of tyrosyl residues are two independent phenomena.

According to the model proposed by Perutz [1], the transition from oxy to deoxyhemoglobin is accompanied by the hiding of the penultimate tyrosine of the subunits inside a pocket where it is stabilized by a hydrogen bond with the peptide backbone of the chains. Nevertheless, it should be expelled from that hydrophobic pocket, upon deprotonation, because of the acquired negative charge. A mutual interdependence, between the conformation of the protein and the position of those tyrosines, would require that some oxy conformation be restored in the deoxygenated system upon their titration. This would explain the modifications of the Soret region of the absorption spectrum of deoxyhemoglobin, concomitant with the titration of tyrosines. Indeed those modifications suggested the appearance of a detectable amount of oxy conformation in the deoxygenated system. The different conformations of oxy and deoxyhemoglobin produce a fine structured difference spectrum near 290 nm [9, 10]. As shown in fig. 1, this fine structure was present in the difference spectra used to follow the ionization of the tyrosines. Qualitatively it was proportional to their ionization and was strictly concomitant with it. This also supports the hypothesis that some oxy conformation appeared in the deoxygenated system upon titration of the tyrosines. The modification of the kinetic of the reaction of hemoglobin with carbon monoxide noted by Andersen et al. [16] above pH 10 is also consistent with a change in the conformation of deoxyhemoglobin at that pH. Finally, a conformational change produced by the ionization of tyrosines is consistent with their cooperative protonation in deoxyhemoglobin.

The question is raised how much the conformational change detectable with the spectrophotometer is due to the ionization of the tyrosines and how much was produced by the titration, of the salt bridges between the $141\alpha_1$ arginine and the $126\alpha_2$ aspartate, and between the $146\beta_1$ histidine and the $40\alpha_2$ lysine. The other salt bridges described by Perutz [1] are not considered because they are completely titrated near pH 9. According to Perutz and TenEyck [21] the removal of the 141α arginines and 146β histidines does not modify the positions of the respective penultimate tyrosyl residues in deoxyhemoglobin. This appears consistent with the finding that the dissociation of hemoglobin in dimers, presumably of the $\alpha_1\beta_1$ type [22], did not effect the titration of tyrosines. On this basis and on that of the very strict concomitance of the spectrophotometric modifications in the Soret and UV regions described, it is proposed that the observed conformational change was dependent on the ionization of tyrosyl residues in deoxyhemoglobin.

As a consequence, the free energy of the pK shift of the tyrosyl residues in CO and deoxyhemoglobin should be included in the free energy of the heme–heme interaction. In the model of Perutz [1], the salt bridges, formed by the COOH terminals in deoxyhemoglobin, stabilize the hydrophobic and electrostatic interactions, established by the adjacent tyrosyl residues with the F and H helixes of the subunits, and vice versa. In his calculations, Perutz [1] considered only the contributions to the heme–heme interaction of the salt bridges. Bunn and Guidotti [23], raise the point that, in this case, the strong pH dependence of many of those bonds, in the neutral or slightly alkaline region of pH, would produce a pH dependence of the heme–heme interaction much higher than what can be measured. Including the contribution of the tyrosyl residues will decrease the relative weight of the contribution of the salt bridges to the heme–heme interaction. The extrapolated value of pH 1/2 in the Hill plot shown in fig. 3, would indicate an average $pK \geq 11.0$ for tyrosines in deoxyhemoglobin. This corresponds to a pK shift of about 0.5 from that in CO hemoglobin and to a free energy difference of approximately 0.8 kcalories/heme.

This represents a very crude value, nevertheless it indicates that tyrosines might contribute a non-negligible fraction of the overall free energy of the heme–heme interaction, which was calculated to be 1.5–3.0 kcalories/heme [12].

References

[1] M.F. Perutz, Nature 228 (1970) 726.
[2] J.K. Moffat, S.R. Simon and W.H. Konigsberg, J. Mol. Biol. 58 (1971) 89.
[3] S.R. Simon, D.J. Arnot and W.H. Konigsberg, J. Mol. Biol. 58 (1971) 69.
[4] J.K. Moffat, J. Mol. Biol. 58 (1971) 79.
[5] D.L. Drabkin, J. Biol. Chem. 164 (1946) 703.
[6] J.F. Riordan, W.E. Wacker, B.L. Vallee, Biochemistry 9 (1965) 1755.
[7] R. Benesch, G. MacDuff and R.E. Benesch, Anal. Biochem. 11 (1965) 81.
[8] D.B. Wetlaufer, Advan. Protein Chem. 17 (1962) 303.
[9] R.W. Briehl and J.F. Hobbs, J. Biol. Chem. 245 (1970) 544.
[10] K. Imai, Biochemistry 12 (1973) 128.
[11] Y.K. Yip and E. Bucci, J. Biol. Chem. 243 (1968) 5948.
[12] J. Wyman, Advan. Protein Chem. 19 (1964) 223.
[13] A.G. Kirschner and C. Tanford, Biochemistry 3 (1964) 291.
[14] S.J. Edelstein, M.J. Rehmar, J.S. Olson and Q.H. Gibson, J. Biol. Chem. 245 (1970) 4372.
[15] E. Bucci, C. Fronticelli and B. Ragatz, J. Biol. Chem. 243 (1968) 241.
[16] M.F. Andersen, J.K. Moffat and Q.H. Gibson, J. Biol. Chem. 245 (1971) 2796.
[17] M. Brunori, E. Antonini, J. Wyman and S. Anderson, J. Mol. Biol. 34 (1968) 357.
[18] E. Antonini, E. Bucci, C. Fronticelli, J. Wyman and A. Rossi-Fanelli, J. Mol. Biol. 12 (1965) 375.
[19] R.L. Nagel, H.M. Ranney and L. Kucinskis, Biochemistry 5 (1966) 1934.
[20] R.W. Briehl, J. Biol. Chem. 245 (1970) 538.
[21] N.F. Perutz and L.F. TenEyck, Cold Spring Harbor Symp. Quant. Biol. 36 (1971) 295.
[22] C. Fronticelli and E. Bucci, in preparation.
[23] H.F. Bunn and G. Guidotti, J. Biol. Chem. 247 (1972) 2345.

BIOPHYSICAL CHEMISTRY 1 (1973) 104–111. © NORTH-HOLLAND PUBLISHING COMPANY

CATION BINDING TO β-CASEIN. A COMPARISON OF ELECTROSTATIC MODELS

Charles W. SLATTERY
Department of Biochemistry, Loma Linda University, Loma Linda, California 92354, USA

and

David F. WAUGH
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Received 9 May 1973
Revised manuscript received 14 August 1973

The binding of cations of β-casein at pH 6.6 was considered previously. Available for three sodium concentrations, $I = 0.04$, 0.08, or 0.16 M are: (1) proton releases between I and (2) for each I, as calcium activity is increased, correlated sequences of monomer net charge, proton release, site bound calcium and protein solvation. Models for ion binding are examined. Critical considerations are the intrinsic binding constants between hydrogen(H), calcium(Ca) and sodium(Na) ions and phosphate(P) and carboxylate(C) sites, and the effects of electrostatic interaction between sites as influenced by spatial fixed charge distribution, ionic strength and dielectric constant (D). Anticipated intrinsic binding constants are $k^0_{H,P} = 3 \times 10^6$, $k^0_{Ca,P} = 120$, $k^0_{Na,P} = 1$, $k^0_{H,C} = 7 \times 10^4$, and $k^0_{Ca,C} = 5.6$.

Distributed charge models, either surface or volume, are inadequate since any reasonable monomer size yields fixed charge densities requiring $k^0_{H,P}$ and $k^0_{Ca,P}$ which are too low when the maximum in D is 75. Also, with increasing calcium binding, calculated proton release is only 0.4 to 0.5 of that observed.

Discrete charge models accept anticipated k^0 and yield calculated sequences of calcium binding and proton release which are in good agreement with those observed provided that: (1) using the known amino acid sequence of the phosphate-containing acidic peptide portion of the molecule, peptide fixed charge is distributed at the lowest I so as to minimize electrostatic free energy; (2) in the region of fixed charge, D is approximately 5; (3) the distances between peptide fixed charges decrease with increasing ionic strength or calcium binding and (4) while protein is in solution, the acidic peptide and the remainder of the molecule are essentially electrostatically independent.

1. Introduction

An analysis of the binding of ions to charged macromolecules is complicated by the fact that more than one ion is often competing for a particular binding site, and neighboring charges, through electrostatic interaction, affect the degree of binding at this site. Critical considerations are the intrinsic binding constants, and the effects of electrostatic interaction as influenced by spatial fixed charge distribution, ionic strength and dielectric constant. The complete expression yielding the extent of binding of a particular ion to a macromolecule is

$$\bar{\nu}_i = \sum_{j=1}^{n} \frac{k^0_{\bar{q}} a_i \exp[-\phi_{\bar{q}}(\bar{Z})]}{1 + \sum_{q=1}^{m} k^0_{qj} a_q \exp[-\phi_{\bar{q}}(\bar{Z})]}. \quad (1)$$

In eq. (1), n is the total number of binding sites, m is the number of different ion species competing for these sites. a_q is the activity of the qth ion species and $\bar{\nu}_i$ is the binding of the ith ion species, in moles ion per mole protein. The k^0_{qj} are the intrinsic binding constants of the qth ion species to the jth site. Each term of the sum over j represents the average binding of the ith ion species to the jth site. In eq (1), $\phi_{\bar{q}}(\bar{Z})$ is equal to $(\delta W_{el}/\delta Z_j)(NZ_i/RT)$ (Tanford [1]), where $\delta W_{el}/\delta Z_j$

is the change in the molecular electrostatic free energy as binding of the ith ion of charge Z_i alters the charge on the jth site. In these, N is Avogadro's number, R is the gas constant and T is the absolute temperature. The general procedure for calculating binding by means of eq. (1) is to use a particular model to calculate W_{el} and $\delta W_{el}/\delta Z_j$ and then obtain $\phi_{\bar{j}}(\bar{Z})$ from the latter.

Where the spatial distribution of fixed charge is known, the most complete accounting of ion binding can be made by applying electrostatic theory, as is done in the approach of Harris and Rice [2, 3] and Rice and Harris [4]. In most studies (see Tanford [1]), fixed charge distribution is not known. However, molecular positive and negative charge is known, along with information concerning molecular weight, molecular shape or radius of gyration and the nature of the binding sites. In such cases, estimates of the numbers of binding sites having particular intrinsic binding constants are obtained through models for which it is assumed that the net charge is distributed uniformly: either over the surface of an impenetrable sphere (Scatchard et al. [5]) or cylinder (Hill [6]), or throughout a selected volume as in the models of Hermans and Overbeek [7] and Tanford [8].

The binding of cations to β-casein has been examined (Waugh et al. [9]). Proton release is given as the monovalent ionic strength alone is increased from 0.04 to 0.08 and 0.16. For each monovalent ionic strength, correlated sequences are determined as calcium concentration is increased: site bound calcium ($\bar{\nu}_{Ca,S}$), the total ionic strength (I_t), the ionic strength in the region of the fixed charges (I_d), observed proton release (δH^+_{obs}), monomer net charge (Z), and solvation (G). These values are essentially independent of specific models. In addition, as will be indicated below, information is also available concerning the structure of the β-casein monomer.

The information available is sufficiently extensive to provide an opportunity to compare models designed to account for ion binding, to examine the requirements for intrinsic binding constants and to reveal certain system characteristics for which any specific model must account.

Intrinsic binding constants: In the following calculations, hydrogen, calcium, and sodium ions are considered to bind to carboxylate and phosphate groups present on the macromolecule. A set of anticipated intrinsic binding constants was established by the following considerations.

For the second binding of H^+ to orthophosphate, $k^0_{H,P}$ is about 7×10^6. However, for alkyl phosphates, Kumler and Eiler [10] found values as low as 3×10^6, and Alberty et al. [11] report values of an apparent association constant, $k_{H,P}$, near 1×10^6 for AMP-5, α-AMP-3 and β-AMP-3. The ionic strengths used by these authors and the data of Kielland [12] lead to activity coefficients which increase this to a value of $k^0_{H,P} = 2 \times 10^6$. It is anticipated that the serine or threonine phosphate monoesters of the caseins bind near $k^0_{H,P} = 3 \times 10^6$.

Davies and Hoyle [13] obtain apparent binding constants for Ca^{2+} to HPO_4^{2-} (orthophosphate) ranging from 400 to 600, while the results of Smith and Alberty [14] give an apparent constant of 50 ± 2 at 25° and $I = 0.2$. Calculation of activity coefficients for the latter gives a value of $k^0_{Ca,P} = 560 \pm 22$. The value of $k^0_{Ca,P}$ is not known for organic phosphates. To proceed, it was assumed that the ratio $k^0_{H,P}/k^0_{Ca,P}$ for orthophosphate (1.25×10^4) would be multiplied by two for organic phosphates in order to account for the statistical difference in the number of ionizable hydrogen atoms compared to the number of oxygen atoms to which they can return [10]. The other factors (inductive, steric and solvation) should maintain the same ratio. This gives anticipated $k^0_{Ca,P} = 120$.

For sodium binding, measurements of pH were used by Smith and Alberty [15] to determine an apparent constant, $k_{Na,P} = 4$. Strauss and Ross [16], using conductance and electrophoresis methods which are considered to be more accurate when the binding is low, obtain $k^0_{Na,P} = 1$, and this value was used for all calculations.

For the binding of protons to carboxylate groups, the anticipated constant is that for organic acids: $k^0_{H,C} = 7 \times 10^4$. Values of $k^0_{Ca,C}$ and $k^0_{Na,C}$ have not been measured. The latter is certainly negligible. The assumption that the orthophosphate ratio of 1.25×10^4 holds also for $k^0_{H,C}/k_{Ca,C}$ yields anticipated $k^0_{Ca,C} = 5.6$. The anticipated set of constants is then $k^0_{H,P} = 3 \times 10^6$, $k^0_{Ca,P} = 120$, $k^0_{Na,P} = 1$, $k^0_{H,C} = 7 \times 10^4$ and $k^0_{Ca,C} = 5.6$.

2. Distributed charge models and binding

The simplest model for a polyion which permits calculation of the effect of net charge on binding is that

used by Scatchard et al. [5] and Scatchard and Yap [17]. In this model, the protein molecule is viewed as a sphere, impenetrable to mobile ions, with the net charge uniformly distributed over its surface. In this case

$$W_{el} = \frac{\bar{Z}^2\epsilon^2}{2DR_p}\left(1 - \frac{\kappa R_p}{1+\kappa A}\right) \quad (2)$$

and

$$\phi_i(\bar{Z}) = \frac{N Z_i \bar{Z}\epsilon^2}{RTDR_p}\cdot\left(1 - \frac{\kappa R_p}{1+\kappa A}\right). \quad (3)$$

Here, ϵ is the protonic charge, D is the dielectric constant, κ is the Debye parameter using the solution ionic strength (I_t), R_p is the effective radius of the macroion and A is the sum of R_p and the radius of the mobile ion being bound.

Ion binding at equilibrium was obtained by determinating a set of $\bar{\nu}_i$, using eq. (1), which were compatible with $\phi_i(\bar{Z})$ as given by eq. (3). It is noted firstly, that eq. (1) simplifies considerably for distributed charge models since only the net charge is recognized, and calculations thus involve a single $\phi_i(\bar{Z})$, and secondly, that $\Sigma\bar{\nu}_i$ determines $\bar{Z}$.

An examination was made of the related values of $\bar{\nu}_{Ca,S}$, $\bar{Z}$ and G as given by Waugh et al. [9]. These are recorded in the first three columns of table 1. Using the surface distributed charge model, the large electrostatic effect which results from a high surface charge density emerged as a problem. For example, if $\bar{Z}$ is distributed over a sphere whose radius yields the volume of the anhydrous molecule, D is taken at its maximum of 75 and no binding is attributed to the carboxylate groups, the value of $k^0_{Ca,P}$ required to account for initial calcium binding is extraordinarily small ($k^0_{Ca,P} = 2$ and $k^0_{H,P} = 5 \times 10^4$). On the other hand, if it is assumed that $\bar{Z}$ is distributed over the outer surface of a sphere equivalent to a solvated macroion, D is taken as 75 and anticipated k^0 are used, the value of G required to account for initial binding is about 15. This would require more than one-fifth of the solution volume to be water of solvation. For the data of table 1, it is assumed that $\bar{Z}$ are distributed over the outer surfaces of solvated spheres determined by anhydrous molecular volume and observed G values. The maximum D was taken as 75 and allowed to vary (column 4 of table 1) to account for binding. This required the low values of $k^0_{H,P} = 5 \times 10^5$ and $k^0_{Ca,P} = 20$. As indicated, the series of D is non-monotonic. Calculated values of $\bar{\nu}_H$ (bound protons) are given in column 5. This column predicts that 0.90 protons will be released between $\bar{\nu}_{Ca,S} = 0.31$ and 6.82. The observed release is 2.42 protons.

Distribution of charge in a volume has been examined by using the model of Tanford [8]. In this model, part

Table 1

A comparison at $I = 0.04$ of variables for reproducing calcium binding using distributed surface charge and distributed volume charge models. For surface charge $k^0_{H,P} = 5 \times 10^5$ and $k^0_{Ca,P} = 20$. For volume charge $k^0_{H,P} = 1 \times 10^6$ and $k^0_{Ca,P} = 40$

			Surface		Volume	
$\bar{\nu}_{Ca,S}$	G	$\bar{Z}$	D	$\bar{\nu}_{H,calc}$	D	$\bar{\nu}_{H,calc}$
0.31	7.29	−13.8	61	1.93	26	2.38
0.93	6.75	−12.8	37	2.42	1	2.82
1.29	6.39	−12.2	41	2.19	4	2.60
1.74	5.97	−11.5	38	2.18	3	2.57
2.09	5.63	−11.0	38	2.12	3	2.52
3.09	4.60	− 9.5	49	1.66	24	1.96
4.03	3.45	− 8.1	75	1.28	75	1.51
5.39 [a]	2.80	− 8.3	56	1.15	46	1.33
5.80	2.30	− 5.9	56	1.09	52	1.26
6.82	1.98	− 5.2	31	1.03	7	1.17

[a] First precipitate.

of the protein molecule is present as an impenetrable sphere of radius R_0 and the remainder, which carries distributed $\bar{Z}$, occurs in a concentric shell penetrable by mobile ions. In present calculations, it is assumed that all of the volume is penetrable except that directly occupied by the anhydrous protein and its tightly bound solvent (0.5 g water/g protein; see Waugh et al. [9]). The relation for W_{el} is given in Tanford [1, p. 472, eqs. (26-42)]. This was differentiated with respect to $\bar{Z}$ and multiplied by NZ_i/RT to give $\phi_i(\bar{Z})$.

A number of examinations of this model have been made. For example, employing observed G values and maximum $D = 75$ requires the low values of $k^0_{H,P} = 1 \times 10^6$ and $k^0_{Ca,P} = 40$. At these values, non-monotonic variations in D are required as indicated in column 6 of table 1. Proton binding, shown in the last column, predicts a release of 1.21 protons between $\bar{\nu}_{Ca,S} = 0.31$ and 6.82, which is to be compared with 2.42 observed.

Both distributed charge models (and the model of Hermans and Overbeek [7]) require $k^0_{H,P}$ and $k^0_{Ca,P}$ which are well below anticipated values. Further, D first decreases and then must increase to account for experimental data. This behavior of D is contrary to expectation. In addition, these models predict an unacceptably small proton release. An examination of all calculations suggests that, so long as distributed charge is assumed, there is no acceptable way by which electrostatic interaction can be reduced to permit choice of reasonable k^0 values, to yield adequate proton release and to yield an acceptable sequence in D. With respect to the latter, it is noted that distributed charge models differ from discrete charge models in that they predict a monotonic increase in W_{el} as D is decreased. Clearly, distributed charge models require all sites having the same k^0 to be equivalent and do not recognize effects requiring local electrostatic interactions for their production.

3. Discrete charge model

To the extent that eq. (1) is applicable, local electrostatic interactions are taken into account. The accompanying expressions for W_{el} and $\phi(\bar{Z})$ are:

$$W_{el} = \frac{\epsilon^2}{D} \sum_{j=1}^{n} Z_j \sum_{k>j}^{n} Z_k \frac{\exp(-\kappa r_{jk})}{r_{jk}}, \tag{4}$$

$$\phi_{\bar{q}}(\bar{Z}) = \frac{\epsilon^2 N Z_i}{DRT} \sum_{\substack{k=1 \\ k \neq j}}^{n} Z_k \frac{\exp(-\kappa r_{jk})}{r_{jk}}. \tag{5}$$

These equations require a model specifying the spatial position of each fixed charge from which the distance, r_{jk}, between a charge Z_j at site j and another charge Z_k at site k can be calculated.

β-Casein contains an acidic peptide comprising the first 25 amino acid residues at the N-terminal end. The amino acid sequence for this acidic peptide portion of the molecule has been given by Manson [18] in an oral communication and by Ribadeau-Dumas et al. [19]. The latter sequence has been used here. It places four phosphates in positions 15, 17, 18, and 19, seven carboxylates in positions 2, 4, 5, 11, 14, 20 and 21 and two arginines in positions 1 and 25. The rest of the molecule contains one additional phosphate and 17 additional carboxylate groups, which are assumed to be able to bind calcium ions, and 18 groups capable of carrying positive charge.

Examinations by Waugh et al. [9, 20], suggest that the acidic peptide is solvent accessible and that it carries essentially all of the monomer net charge. The remainder of the molecule, which has a compact anisometric form (the body), has a net charge near zero. It is evident that major electrostatic effects are due to the acidic peptide. Except where noted, in calculations it is assumed that ions bind to the body phosphates and carboxylates at their k^0 values.

In an extended peptide chain, the distance between amino acid residues is 3.62 Å [21]. The acidic peptide would thus form a rodlet approximately 90 Å long with side chains carrying the charged groups. The average side chain length is 7 Å.

With respect to acidic peptide geometry, preliminary calculations revealed two important model independent requirements. Firstly, irrespective of body groups, if acceptable k^0 values are to be used to account for initial calcium binding the peptide at the lowest monovalent ionic strength ($I = 0.04$) must be expanded in such a way as to minimize electrostatic interactions. Secondly, for any selected k^0 values, conformation cannot remain constant: the peptide must progressively collapse to account for calcium and proton binding sequences.

Model for β-casein: The following model was used to

reveal certain system characteristics for which any specific model must account. All charges were placed on rings having radii, σ, at the average side chain length of 7 Å. From bond angles and lengths the unstrained distance between rings, Δ, is 3.62 Å. The distance, r_{jk}, between charges placed at j and k is given by

$$r_{jk} = [(P_j - P_k)^2\Delta^2 + 4\sigma^2 \sin^2 \tfrac{1}{2}(\Psi_j - \Psi_k)]^{1/2}, \quad (6)$$

where P_j is the position of the amino acid carrying the jth group in the acidic peptide and Ψ_j is the angle between a radius to the charge and an arbitrary plane which includes the rod axis. Examination of the effects on binding produced by alterations in the density of fixed charge was made by varying the distance, Δ, between rings.

4. Discrete charge binding

To obtain a charge distribution for initial minimization of peptide electrostatic free energy at $I = 0.04$, the unstrained maximum of $\Delta = 3.62$ Å was chosen, and charged groups were placed on their corresponding rings in sequence by a process which successively minimizes rodlet electrostatic free energy. The first charge (Arg) was placed on its ring at $\Psi_1 = 0°$. The second (Glu) was then placed on its ring and trial values of r_{12} and W_{el} calculated at each one degree increment around the ring. The position, Ψ_2, giving the minimum in W_{el} was selected. Successive charges were introduced in sequence and minimization carried out using all preceding charges.

Given this acidic peptide structure when fully charged, eq. (5) readily yields the $\phi_{\bar{y}}(\bar{Z})$ for each site. However, binding alters the average charge at each binding site, and this will affect electrostatic interaction. To obtain the bound ion distribution at equilibrium maximum charge reduction at each peptide binding site was first obtained for selected k^0 values by using $\exp[-\phi_{\bar{y}}(\bar{Z})] = 1$ in eq. (1). Using these charge reductions, eq. (5) was used to calculate minimum $\phi_{\bar{y}}(\bar{Z})$. These values and eq. (1) were used to obtain minimum binding for reduction of site charge. Eq. (5) was again used to obtain new $\phi_{\bar{y}}(\bar{Z})$. Convergence was facilitated by using $\phi_{\bar{y}}(\bar{Z})$ midway between the two preceding values. Iteration was continued until $\phi_{\bar{y}}(\bar{Z})$ and $\bar{\nu}_i$ were constant.

The following method was used to correlate experimental data. For each $\bar{\nu}_{Ca,S}$ (column 1, table 2) the Debye parameter, κ, was calculated using the total solution ionic strength given in the second column. Body groups were assumed to bind at their k^0 values. Calcium ion activity in the acidic peptide region was calculated using the data of Kielland [12] and the ionic strength in this region, I_d (column 3). A trial set of k^0 was selected. For each $\bar{\nu}_{Ca,S}$, a series of Δ was chosen, and for each Δ, eqs. (5) and (1) were solved for $\bar{\nu}_i$ at equilibrium. The Δ yielding $\bar{\nu}_{Ca,S}$ was selected, and this was compared with the unstrained maximum. In addition, proton releases, obtained from calculated binding were compared with observed values.

The choice of dielectric constants is important in determining the $\phi_{\bar{y}}(\bar{Z})$. A value of $D = 75$ was first examined and, as with distributed charge models, a major problem was to reduce electrostatic interaction sufficiently so as to account for initial calcium binding at $I = 0.04$. For example, using anticipated k^0, the required value of Δ was far too large to be accomodated by straining of the acidic peptide. For another set of calculations, Δ was taken as a reasonable maximum of 3.85 Å (a difference between unstrained and strained conformations of $\Delta W_{el} = 0.25$ kcal per mole, the strain to be produced by electrostatic repulsion). Keeping other k^0 at anticipated values, initial calcium binding required the low values of $k^0_{Ca,P} \approx 40$ (and $k^0_{H,P} \approx 1 \times 10^6$). In addition, permitting all k^0 to vary, and including the possibility that ratios of binding constants to phosphate may vary, does not allow selection of significantly higher k^0. Generally, at even slightly higher k^0, the initial Δ becomes far too large and calculated proton release is unacceptably small.

The discrete charge model constructed here exhibits a dependency on dielectric constant which is absent in models assuming distributed charge. As D is decreased, electrostatic free energy reaches a maximum at $D = 15$ and then decreases. This results from the influence of the two arginine positive charges, which eventually extends to negate the effects of neighboring negative charges. As D is reduced below 15, progressively higher values of $k^0_{Ca,P}$ and $k^0_{H,P}$ can be used, and the differences between calculated and observed proton release progressively decrease. Accepting anticipated k^0 and a maximum value of $\Delta = 3.85$ Å requires a value of $D = 5$.

These values have been used in calculations, the results of which are recorded in columns 4 to 7 of table 2. The specification that D be as low as 5 is in accord

Table 2
Calcium binding, proton binding and proton release for β-casein

$\bar{\nu}_{Ca,S}$	I_t	I_d	Δ	$\bar{\nu}_H$	δH^+ calc.	δH^+ obs.
($I = 0.04$, $\bar{\nu}_{Na,P} = 0.15$)						
0.00				3.68		
0.56	0.0403	0.0581	3.85	3.38	–	–
0.95	0.0405	0.0590	3.61	3.17	0.21	0.40
1.32	0.0408	0.0607	3.61	2.90	0.48	0.56
1.73	0.0411	0.0622	3.30	2.71	0.67	0.75
2.02	0.0414	0.0640	3.14	2.55	0.83	0.91
3.06	0.0432	0.0616	2.99	1.98	1.40	1.38
4.07	0.0479	0.0948	2.99	1.58	1.80	1.89
4.50 [a]				1.43	1.95	2.00
5.41	0.0579	0.1081	1.73	1.12	2.26	2.27
5.82	0.0635	0.1267	1.49	1.03	2.35	2.41
6.78	0.0933	0.1473	0.86	0.91	2.47	2.58
($I = 0.08$, $\bar{\nu}_{Na,P} = 0.15$)						
0.00				3.10		
0.11	0.0802	0.0995	2.99	3.09	–	–
0.46	0.0806	0.1009	2.20	3.08	0.01	0.17
0.89	0.0809	0.1019	1.96	3.03	0.06	0.30
0.98	0.0816	0.1039	2.28	2.80	0.29	0.34
1.33	0.0819	0.1051	2.04	2.69	0.40	0.45
1.86	0.0841	0.1116	2.20	2.33	0.76	0.68
2.47	0.0850	0.1162	1.73	2.08	1.01	0.86
3.74	0.0940	0.1402	1.73	1.57	1.52	1.37
4.15	0.0988	0.1542	1.73	1.45	1.64	1.50
4.50 [a]				1.35	1.75	1.63
4.83	0.1050	0.1502	1.10	1.24	1.85	1.73
($I = 0.16$, $\bar{\nu}_{Na,P} = 0.49$)						
0.00				2.66		
0.03	0.1602	0.1809	2.04	2.66	–	–
0.22	0.1610	0.1825	1.79	2.65	0.01	0.04
0.30	0.1611	0.1836	1.73	2.65	0.01	0.15
0.75	0.1616	0.1843	1.37	2.65	0.01	0.21
0.77	0.1622	0.1856	1.41	2.65	0.01	0.23
1.45	0.1645	0.1907	1.34	2.25	0.41	0.44
1.61	0.1696	0.2018	1.73	2.14	0.52	0.53
2.66	0.1751	0.2134	1.26	1.78	0.88	0.82
3.20	0.1812	0.2259	1.18	1.57	1.09	1.00
4.50 [a]				1.26	1.38	1.31
4.52	0.1858	0.2347	0.71	1.27	1.39	1.35

[a] First precipitate.

with Harris and Rice [2]. They select a value of 5.5 as an upper limit on D in a situation where charges are close together (polymethacrylic acid) with the lower limit at about 2 for charges interacting through organic matter.

Table 2 records results of calculations up to and past the precipitation point, but not including situations involving apparent charge reversal. Specified in table 2 are $\bar{\nu}_H$ at $\bar{\nu}_{Ca,S} = 0$ (obtained by extrapolation), the maximum calculated sodium binding for each I, and the interpolated values of $\bar{\nu}_H$ at $\bar{\nu}_{Ca,S} = 4.5$. Maximum sodium binding at each I is small. In calculations, each $\bar{\nu}_{Ca,S}$ is matched by a selection of Δ (column 4), and this selection fixes a particular calculated proton binding (column 5). Sequential binding values give calculated proton release (column 6), and these are to be compared with the observed proton release values of column 7. The two sets are in good agreement. On the average, observed release is 0.04 proton greater than that calculated and the standard deviation of the single difference is ± 0.11 proton. The specific discrete charge model presented here appears to be a reasonable approximation.

The values of Δ given in column 4 of table 2 reveal acceptable trends. While protein is in solution, Δ is nearly constant or decreases slightly with increasing $\bar{\nu}_{Ca,S}$, and at a particular $\bar{\nu}_{Ca,S}$, Δ decreases as I is increased. When precipitate forms, Δ decreases sharply. Since Δ is the single parameter used to account for all changes in site density, it must include effects due to local conformational changes and molecular association which can affect the proximity of sites to each other. A decrease in Δ correlates with a decrease in G, with $(G-U)$ always sufficiently large to accomodate the acidic peptide structure as required by Δ.

The assumption has been made that with respect to calcium binding the body and peptide are essentially electrostatically independent. That this is the case is indicated by the following. The $\bar{\nu}_{Ca,S}$ are arbitrarily matched by selecting site density on the acidic peptide while maintaining that on the body constant. If there were body–peptide electrostatic interactions which affected calcium binding, the choice of Δ would compensate for these in the calcium binding sequence. However, the electrostatic effect, $\exp[-\phi(\bar{Z})]$, for proton binding at a site is the square root of that for calcium binding. In the complex summation of eq. (1), any significant compensation to match calcium binding would produce a progressive divergence between calculated and observed proton release sequences.

Calculations suggest that there may be body groups which bind protons, but not calcium, and are sensitive to electrostatic interaction with other fixed charges. This is indicated by the following. Starting at $I = 0.04$, an increase to $I = 0.08$ produces an observed release of 0.98 proton and an increase to $I = 0.16$, a release of 1.49 protons. From table 2, corresponding calculated releases are, respectively, 0.58 and 1.02 proton. The differences between calculated and observed values are larger than experimental error. This monovalent ionic strength effect is not observed up to $\bar{\nu}_{Ca,S} = 4.5$, where the increase in total ionic strength due to calcium chloride addition is still small (I_t of table 2). However, at $\bar{\nu}_{Ca,S} \approx 11$ (0.08 M $CaCl_2$) release is independent of I and is 4.93 protons [9]. It is noted that the release accompanying an increase in I alone occurs when protein is in solution and in a state of minimum association. If, as indicated above, calculations account adequately for the peptide, but not for the body, then the monovalent ionic strength effects are due to body group interactions. One attractive possibility implicates the five histidines of β-casein [19]. These bind protons but not calcium and have pK^0 near the ambient pH of 6.6. If, on the average, they are closer to fixed negative than positive charge, increasing I will produce proton release. What is required is an average release of about 0.10 proton per histidine.

Acknowledgement

This work was partially supported by Grants GM 0.5410 and GM 00778 of the U.S. Public Health Service, and by a grant from the Carnation Company Research Laboratories. Computational assistance was received from the Loma Linda University Computation Facility supported in part by NIH Grant RR-276-08.

References

[1] C. Tanford, Physical chemistry of macromolecules (Wiley, New York, 1961).
[2] F.E. Harris and S.A. Rice, J. Phys. Chem. 58 (1954) 725.
[3] F.E. Harris and S.A. Rice, J. Chem. Phys. 25 (1956) 955.
[4] S.A. Rice and F.E. Harris, J. Phys. Chem. 58 (1954) 733.
[5] G. Scatchard, J.C. Coleman and A.L. Shen, J. Am. Chem. Soc. 79 (1957) 12.
[6] T.L. Hill, Arch. Biochem. Biophys. 57 (1955) 229.

[7] J.J. Hermans and J.Th.G. Overbeek, Rec. Trav. Chim. 67 (1948) 761.
[8] C. Tanford, J. Phys. Chem. 59 (1955) 788.
[9] D.F. Waugh, C.W. Slattery and L.K. Creamer, Biochemistry 10 (1971) 817.
[10] W.D. Kumler and J.M. Eiler, J. Am. Chem. Soc. 65 (1943) 2355.
[11] R.A. Alberty, R.M. Smith and R.M. Bock, J. Biol. Chem. 193 (1951) 425.
[12] J. Kielland, J. Am. Chem. Soc. 39 (1937) 1675.
[13] C.W. Davies and B.E. Hoyle, J. Chem. Soc. (1953) 4134; (1955) 1038.
[14] R.M. Smith and R.A. Alberty, J. Am. Chem. Soc. 78 (1956) 2376.
[15] R.M. Smith and R.A. Alberty, J. Phys. Chem. 60 (1956) 180.
[16] U.P. Strauss and P.D. Ross, J. Am. Chem. Soc. 81 (1959) 5295.
[17] G. Scatchard and W.T. Yap, J. Am. Chem. Soc. 86 (1964) 3434.
[18] W. Manson, Report presented at the 18th International Dairy Congress, October, Sydney, Australia (1970).
[19] B. Ribadeau-Duamas, G. Brignon, F. Grosclaude and J. Mercier, European J. Biochem. 25 (1972) 505.
[20] D.F. Waugh, L.K. Creamer, C.W. Slattery and G.W. Dresdner, Biochemistry 9 (1970) 786.
[21] R.B. Corey and L. Pauling, Proc. Roy. Soc. B141 (1953) 10.

BIOPHYSICAL CHEMISTRY 1 (1973) 112–116. © NORTH-HOLLAND PUBLISHING COMPANY

A PROTON MAGNETIC RELAXATION STUDY OF FERRIMYOGLOBIN IN AQUEOUS IONIC SOLUTIONS

Greta PIFAT
Institute "Rudjer Bošković", Zagreb, Yugoslavia

and

S. MARIČIĆ
Institute of Biology, University of Zagreb, Zagreb, Yugoslavia

Received 8 June 1973
Revised manuscript received 9 July 1973

Proton magnetic longitudinal (T_1) relaxation times have been measured for acid (horse) ferrimyoglobin solutions (0.1 M NaCl and KH_2PO_4, 2 M NaCl and 1 M $MgCl_2$) from 5°C to 35°C in dependence on myoglobin concentration up to 6 mM. The enhancement of the relaxation rate due to the paramagnetic haem iron, which is observed in this temperature range is compared with analogous data for the ferrihaemoglobin solution. The conclusion is that the protons exchanging from the haem pocket with bulk solvent are not those from the water molecule at the sixth ligand site of haem iron. The exchanging protons are more than 4 Å away from the haem iron being closer to it in ferrimyoglobin than in ferrihaemoglobin. This distance becomes larger in solutions with higher salt concentration, the largest difference between 0.1 M NaCl and 1 M $MgCl_2$ being over one Angstrom unit. This indicates a conformational change of the haem pocket, possibly its tightening.

1. Introduction

Myoglobin differs from haemoglobin in that it is a monomeric protein in solutions up to concentrations of 2 mM [1]. This simplifies the study of the conformational state(s) of the ferric haem pocket by proton magnetic relaxation (PMR) both with regard to the self-consistency of the parameters derived from the existing theory and the influence of varying solvent conditions such as ionic strength. Information on the latter is needed for elucidating the effect of salts on the association of haemoglobin subunits which is also reflected in the PMR measurements.

Mildvan et al. [2] discussed similar measurements assuming the mechanism of a thermally activated exchange of the water protons from the sixth coordination site at the haem iron and the bulk solvent. On the other hand, Derzhanski et al. [3] interpreted their own measurements by invoking a thermally activated change of the protein conformation with no exchange of water molecules from the haem pocket. Experimental evidence has recently been presented [4] supporting the former mechanism and excluding the latter one. It is the purpose of this paper to corroborate the hydration model of the haem pocket advanced earlier for ferrihaemoglobin solutions [5], by comparing the two sets of data† on the basis of the structural similarity of the ferric haem pockets in the two proteins [6, 7].

2. Experimental

2.1. Materials

The horse myoglobin samples were lyophilized, commercial products of "Serva" and "Koch & Light", and were used as obtained, i.e., in the ferric form and with-

† Part of this discussion was presented by S.M. at the Royal Society Meeting on Haemoglobin, in London, on 23 February, 1973.

out any further purification except for centrifuging off the solutions before the measurements. All the solutions were adjusted close to pH 6 and the concentrations were determined by spectrophotometry using the extinction coefficients quoted in the literature [1].

2.2. Methods

The measurements of the longitudinal proton magnetic relaxation times, T_1, were done at 29 MHz with subsequent evaluation of the results as described earlier [5].

3. Results

The concentration dependence of $\Delta(1/T_1)$ as shown in fig. 1 for the different salt solutions at 10 and 34°C was obtained from the usual Arrhenius plots of the relaxation rates, $\Delta(1/T_1)$, induced by the presence of myoglobin in the solutions. The slopes of the least-squares straight lines in fig. 1, i.e., the relaxation rates at 10 and 34°C normalized per unit Mb-concentration are given in table 1.

Table 1
The proton magnetic relaxation rates per unit (horse) ferrimyoglobin concentration, $\Delta(1/T_1)$, sec^{-1} $M^{-1}\ell$, in different salt solutions at pH 6

	$\Delta(1/T_1)$		$\Delta(1/T_1)_{34-10°C}$
	34°C	10°C	
0.1 M KH_2PO_4	1337 ± 25	518 ± 13	819 ± 25
0.1 M NaCl	1203 ± 27	463 ± 6	740 ± 27
2 M NaCl	643 ± 16	248 ± 4	395 ± 16
1 M $MgCl_2$	814 ± 25	609 ± 15	205 ± 25

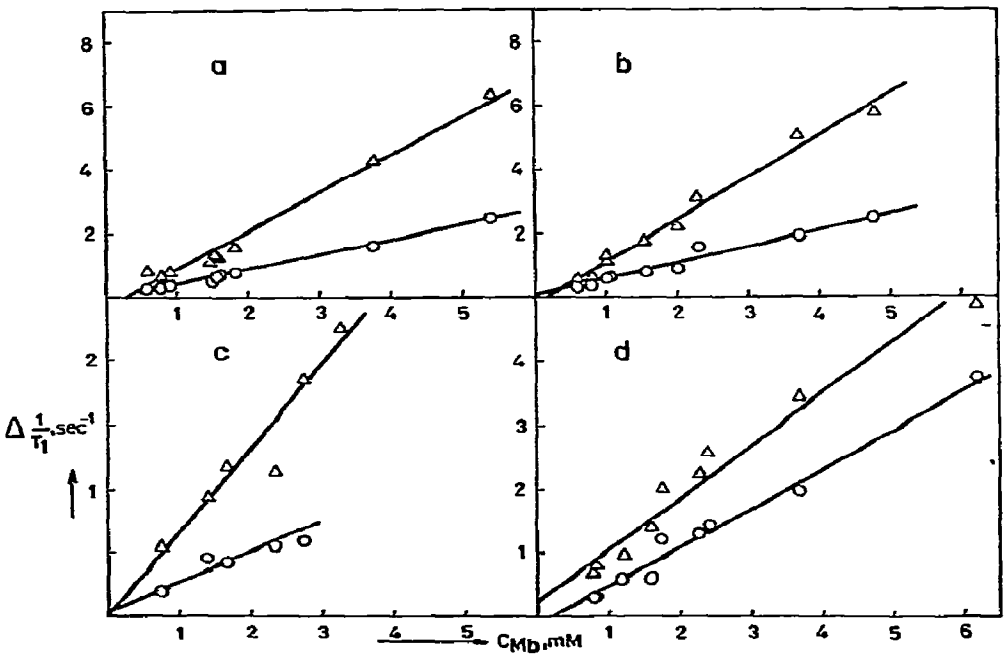


Fig. 1. The concentration (mM Mb) dependence of the proton magnetic relaxation rates induced by the presence of (horse) ferrimyoglobin in different salt solutions at pH 6. Circles, data at 10°C; triangles, data at 34°C. The solutions were: (a) 0.1 M NaCl; (b) 0.1 M KH_2PO_4; (c) 2 M NaCl, and (d) 1 M $MgCl_2$. The slopes of the least-squares best straight lines are given in table 1.

4. Discussion

The following model of the proton magnetic relaxation mechanism in haemoprotein solutions was put forward in ref. [5]: (a) fast exchange of (water) protons between the haem pocket and bulk solvent; this was recently verified [4]; (b) the sixth ligand water molecule is not participating in this exchange mechanism, but there is another (next neighbour) water molecule in the haem pocket whose exchange with bulk solvent results in the observed temperature enhancement of the relaxation rates. The present data for Mb solutions are also consistent with this second aspect of the proposed model.

The increase in the myoglobin induced relaxation rates between 10 and 34°C is related to the distance, r, between the electron spins of the ferric haem iron interacting by dipole–dipole coupling with the proton spins of the nearest water molecule(s) which exchange with bulk solvent:

$$\{\Delta(1/T_1)\}_{34-10^\circ C} \propto r^{-6}. \qquad (1)$$

(For details of the pertinent equations see ref. [5]). The differential relaxation rates, $\{\Delta(1/T_1)\}_{34-10^\circ C}$, in table 1, change considerably with various salts and ionic strength of the solutions. This implies a change in r and it is difficult to rationalize how a stereochemically precisely defined water molecule at the sixth ligand site of the haem iron could change its position without an alteration of the spectra in the visible range. From a quantitative point of view the calculation of the corresponding r's reveals that these distances are much larger than required by the position of the sixth ligand.

In order to obtain r one must know the proportionality constant in eq. (1), which includes the correlation time, τ_C, during which the spins interact unperturbed. The perturbation of this interaction depends on the time characteristic of the myoglobin rotation, τ_R, the residence time of the exchanging proton(s), τ_B, and the longitudinal electronic relaxation time, τ_S:

$$\frac{1}{\tau_C} = \frac{1}{\tau_R} + \frac{1}{\tau_B} + \frac{1}{\tau_S}. \qquad (2)$$

$1/\tau_C$ will be determined by the fastest of the three rates, i.e., by the shortest time constant. The contribution of $1/\tau_B$ is negligible [5]. The relative importance of the remaining two rates may now be tackled by comparing the curves in fig. 2, which show the dependence of r for Mb and Hb on the correlation time, τ_C, for the measuring Larmor frequency of 29 MHz, calculated as before [5]. The Hb-curve is taken from the latter reference, while the Mb-curve was calculated using the value $740 \pm 27\ \text{sec}^{-1}\,\text{M}^{-1}\text{l}$ from table 1 (0.1 M NaCl solution) with the theoretical magnetic moment of the haem iron, 5.92 BM. The shaded uncertainty limits (larger than the experimental error) are given by assuming the exchange of one water molecule (i.e., two protons), or only one proton.

The dielectric correlation times, τ_D, for both haemoglobin and myoglobin have been measured by the dielectric frequency dispersion technique [8]. If in eq. (2) $1/\tau_C$ were equal to $1/\tau_R$, then the theory [9] requires that $\tau_R = \tau_D/3$. Hence, using the $\tau_{D,Hb}$ and $\tau_{D,Mb}$ values [8], $\tau_{C,Hb} = (1.3 \pm 0.1) \times 10^{-7}\text{sec}/3 \approx 4 \times 10^{-8}$sec, and $\tau_{C,Mb} \approx 1 \times 10^{-8}$sec. With these two correlation times one obtains from fig. 2 the iron-to-proton distance close to 6 Å for both ferriMb and ferriHb, which would be expected because of the similarity of the haem pockets in these two haemoproteins. However, it follows from eq. (2) with the neglect of $1/\tau_B$, that for $\tau_C \approx 10^{-8}$sec the relaxation time of the unpaired heam-iron electrons, τ_S, should be longer than 10^{-8}sec. This is not acceptable on the basis of the comparison of hyperfine shifted lines in high resolution proton magnetic resonance spectra of low and high-spin (acid) ferric haemoproteins [10]. The value for τ_S has been accordingly set to be 10^{-10}sec. PMR dispersion data of Koenig et al. are quoted in the latter reference as being in accord with $\tau_C = \tau_S = 10^{-10}$sec. Our unpublished [11] dispersion measurements in the range 1.5 to 60 MHz for Hb and Mb (ferri- and CO-form) at 35 and 5°C do not show a proper dispersion step, but they do indicate that it may be expected between 200 and 300 MHz Larmor frequency corresponding to τ_C of about 6×10^{-10}sec. With this value for τ_C one obtains from fig. 2 the following interspin distances, r (Å):

	number of exchanging protons	
	2	1
acid ferriMb in 0.1 M NaCl	4.8	4.3
acid ferriHb in 0.1 M NaCl	5.9	5.3

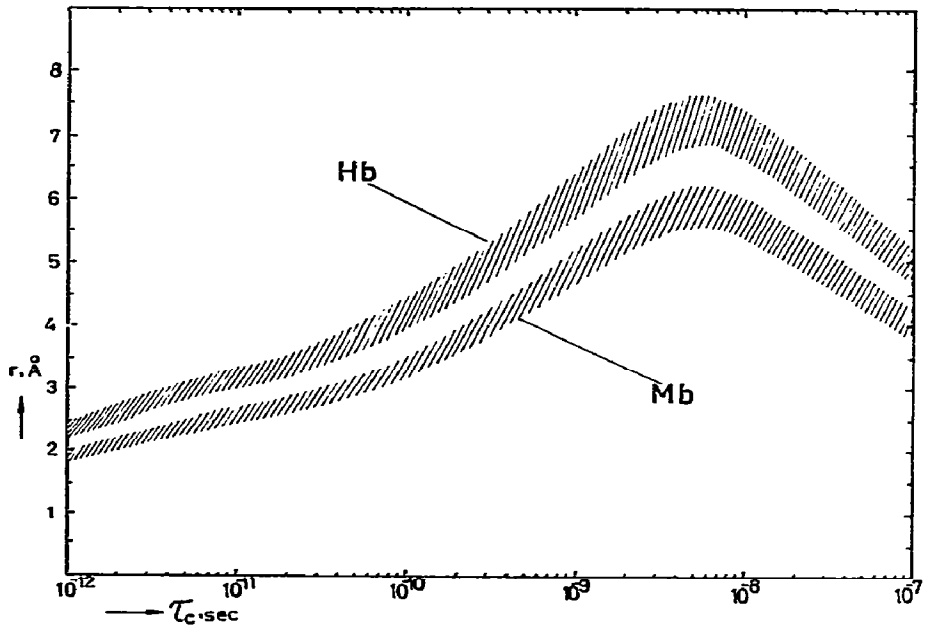


Fig. 2. The distance, r (Å), between the ferri-haem iron in Mb and Hb and the closest exchangeable proton(s). in dependence on the correlation time, τ_C(sec), for this dipole–dipole spin interaction. The upper edges of the shaded curves are calculated by assuming two protons and the lower edges by assuming only one proton exchanging.

With all the similarity of the haem pocket in Mb and Hb [6, 7] expressed as well in their electronic absorption spectra, magnetic properties, and the NMR resonances related to the haem iron, there is no reason for expecting different τ_S's for the two haemoproteins, which is necessary if their r's ought to be equal. (If $\tau_{C,Hb} \approx 6 \times 10^{-10}$ then, for an equal r, $\tau_{C,Mb} \approx 2 \times 10^{-9}$, which again is too long [10].) Hence, the difference in the interspin distances between Mb and Hb seems to be real. On the other hand, in order to satisfy the X-ray value ($r = 2.8$ Å [6, 7]) for the distance between the heme iron and the protons of the water in the 6th coordination site of the heme, τ_C for Mb should be 1.2×10^{-11} sec or 4×10^{-11} sec, which is too low [10], while for Hb (with $r = 2.8$ Å) τ_C should even be an order of magnitude shorter (10^{-12} sec), which is quite improbable.

A rather firm conclusion therefore emerges: there is a water molecule (or, more exactly, there are exchangeable protons) in the haem pocket of acid ferrihaemoproteins next to that water molecule which is liganded directly to the haem iron at its sixth coordination site: these exchangeable protons are closer to the haem iron in myoglobin than in haemoglobin. Another water molecule, hydrogen-bonded by the available hydrogen from the sixth site water molecule will have its protons about 5 Å from the haem iron, but there are, of course, other possibilities for the location of the exchanging protons.

The exchangeable (water) protons may be regarded as a probe for the overall conformational state(s) of the haem pocket. (Were the sixth site water molecule involved only in the exchange mechanism no such information could be expected because of the strictly defined stereochemistry of this water molecule.) This is well borne out by the data in table 1.

The relaxation rates induced by the paramagnetic haem-iron, $\{\Delta(1/T_1)\}_{34-10°C}$, differ in the 0.1 M KH_2PO_4 and in 0.1 M NaCl solutions, although the resulting difference in r's (4.7_6 and 4.8_4 Å) is negligible. These changes are, however, quite large for solutions of high ionic strength, yielding $r_{(2\,M\,NaCl)} = 5.4$ and

$r_{(1\,M\,MgCl_2)}$ = 6.0 Å, with all these values obtained assuming 2 exchangeable protons. Judging by the elongation of the r's, there is a tightening of the haem pocket in more concentrated ionic solutions, the largest change in the location of the exchanging protons with respect to the haem-iron being more than one Ångstrom unit.

It has been shown by nuclear quadrupole relaxation measurements [12] that chloride ions bind to haemoglobin, while sodium ions apparently do not. It is not known whether magnesium ions interact with haemoproteins directly. Additional measurements concerning the hydration shell around myoglobin in different salt solutions would be necessary in order to estimate the relative importance of specific-binding versus solvent-structure mechanism [13] in bringing about the conformational changes around the haem pocket as evidenced by the results presented in this paper. One observation, however, seems worth a conclusion: the differential relaxation rate for 1 M $MgCl_2$ solution is only one half that for the 2 M NaCl solution, although the overall chloride ion concentration is practically the same in both instances. Besides, the corresponding relaxation rates at 10 and 34°C differ substantially between the two solutions. It therefore follows that the magnesium ions themselves bring about considerable conformational changes around the haem pocket, though the mechanism by which it takes place and the very nature of those changes remain unknown.

The method of proton magnetic relaxation senses the subtle changes in the environment of the paramagnetic haem iron in acid ferric haemoproteins. In order to use this method to its full advantage, i.e., by deriving the structural and kinetic parameters with an even greater confidence level, more elaborate measurements are required, as discussed elsewhere [14]. This becomes of primary importance when one wishes to compare small changes in these parameters on an absolute scale, especially for haemoproteins of different origin.

Acknowledgement

This work was done under the contracts 3112/I, 4032/I, 07-2134/I, and 233/S from the Yugoslav Research Funds, and in part with the financial help from the PL-480 contract 02-004/1 (N.I.H., U.S.A.). We thank Miss Lj. Kolesar for performing some of the measurements.

References

[1] E. Antonini and M. Brunori, Hemoglobin and myoglobin in their reactions with ligands (North-Holland, Amsterdam, 1971).
[2] A.S. Mildvan, N.M. Rumen and B. Chance, in: Probes of structure and function of macromolecules and membranes, Vol. 2 (Academic Press, New York, 1971) p. 205.
[3] A. Derzhanski, A. Georgieva, K. Kotev and B.P. Atanasov, Biochim. Biophys. Acta 214 (1970) 83.
[4] S. Vuk-Pavlović, S. Maričić and B. Benko, submitted for publication.
[5] G. Pifat, S. Maričić and Š. Grandja, Biopolymers 12 (1973) 905.
[6] H.C. Watson, in: Progress in stereochemistry, Vol. 4 (Butterworth, London, 1968).
[7] M.F. Perutz, H. Muirhead, J.M. Cox and L.G. Goaman, Nature 219 (1968) 131.
[8] P. Schlecht, A. Mayer, G. Hettner and H. Vogel, Biopolymers 7 (1969) 963.
[9] A. Abragam, The principles of nuclear magnetism (Clarendon Press, Oxford, 1961). p. 300.
[10] K. Wüthrich, Structure Bonding 8 (1970) 53.
[11] G. Lahajnar, I. Zupančič, R. Blinc, G. Pifat and S. Maričić, to be published.
[12] T.E. Bull, J. Andrasko, E. Chiancone and S. Forsén, J. Mol. Biol. 73 (1973) 251.
[13] P.H. von Hippel and T. Schleich, in: Structure and stability of biological macromolecules, Vol. 2, eds. S.N. Timasheff and G.D. Fasman (Dekker, New York, 1969) p. 417.
[14] J. Brnjas-Kraljević, S. Maričić and S. Vuk-Pavlović, Croat. Chem. Acta 45 (1973) in press.

BIOPHYSICAL CHEMISTRY 1 (1973) 117–124. © NORTH-HOLLAND PUBLISHING COMPANY

THE CHARACTERISTICS OF THE INTERMEDIATES IN COLLAGEN-FOLD FORMATION ‡

Leon YUAN and Arthur VEIS

Department of Biochemistry, Northwestern University Medical School, Chicago, Illinois 60611, USA

Received 9 July 1973

The kinetics of collagen-fold formation has been analyzed in terms of a series of consecutive reactions, rather than the usual set of parallel reactions. Data on the regain of optical rotation and viscosity for denatured ichthyocol, formaldehyde cross-linked ichthyocol, isolated α1-chains, α2-chains and the β_{12}-component of rat skin collagen have been utilized. The Guggenheim point averaging procedure was used for data analysis for those reaction steps where final equilibrium measurements cannot be made. By considering the data over appropriate time ranges, optical rotation constants characteristic of each sequential intermediate have been determined. Comparisons of the constants for the first nucleated intermediate in the cases of varying degrees of cross-linking and differing pyrrolidine ring contents, show that both hydrogen-bonding interactions and pyrrolidine ring content contribute to the stabilization of the multi-stranded nucleated fold intermediate. The presence of single chain nucleated intermediate seems highly unlikely.

1. Introduction

When a gelatin solution is cooled below 40°C, most of its physical properties change in ways characteristic of a partial return of the disordered gelatin molecules to ordered collagen-like structures [1, 2]. This partial reordering, variously called collagen-fold formation or renaturation, is very strongly dependent on the reaction conditions. The significant variables are temperature, ionic strength, pH, gelatin concentration and the degree of cross-linking between gelatin chains. Several schemes have been proposed to describe the mechanism of collagen-fold formation, all of them based on studies of the kinetics of the fold-formation [3–12], but no definitive understanding of the renaturation mechanism has yet been reached.

All mechanisms proposed presume a multi-step process involving nucleation of the fold, growth of fold-units over extensive lengths of the system, and in the case of severe undercooling, an annealing or rearrangement process leading to perfection of the folded structure. The development of the initial three-stranded collagen helix is itself considered to be a multi-step process. Von Hippel [1] has summarized the earlier proposals of von Hippel and Harrington [4, 5] and Flory and Weaver [6] in terms of the scheme:

$$\mathrm{RC} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{NC} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \mathrm{I} \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} \tfrac{1}{3}\mathrm{H}. \qquad (1)$$

Von Hippel's [1] molecular interpretation of this scheme was that RC denotes chain regions in random conformations; NC, nucleated chain elements, presumably pyrrolidine-rich regions, which have assumed local conformations typical of the collagen-fold; I, single chain segments in which the fold-conformations extend over several residues; and H, a hydrogen-bonded, three-stranded unit of the collagen helix. The rate limiting step, in both the von Hippel–Harrington and the Flory–Weaver models, is the formation of I, leading to an apparent first-order dependence of the rate on gelatin concentration. This is also the step which can supposedly be followed by optical rotation measurements. Most recently, Harrington and Karr [12] treated the overall kinetic data as a sum of parallel first-order reactions. Their analysis required three concurrent reactions, proceeding at markedly different rates. The regain of optical rotatory power on fold-formation was followed

‡ Supported by a grant from the National Institutes of Health. AM-13921.

and the experimental data were analyzed as:

$$[\alpha]_t - [\alpha]_0 = \sum_{n=1}^{i} ([\alpha]_{\infty,i} - [\alpha]_0)\exp(-k_i t), \qquad (2)$$

where $[\alpha]_t$ was the observed specific rotation at time t; $[\alpha]_0$, the specific rotation at $t = 0$ and $[\alpha]_{\infty,i}$, the specific rotation of the completed reaction of type i, with rate constant k_i. The inherent disadvantages of this approach are in the assumption that the refolding process consists of concurrent rather than consecutive reactions and in the proper identification of a value for $[\alpha]_{\infty,i}$.

In the present work we have attempted to analyze the refolding in terms of a set of consecutive first-order reactions, as implicit in scheme (1), and to determine the appropriate constants for each stage. Our use of a sequence of first-order reactions, however, does not mean that we apply these specifically to the molecular interpretation of the steps of scheme (1).

2. Data treatment and results

Kinetic studies of collagen-fold formation are available from previous work. Drake and Veis [9] examined the refolding process at 15°C for denatured ichthyocol (DI) and for formaldehyde cross-linked ichthyocol (synthetic γ-gelatin) (CLI) all in 0.1 M HAc, 0.1 M KCl at pH 2.8. Piez and Carrillo [10] carried out similar studies on the isolated α1, α2 and β_{12} chains from rat skin collagen. Optical rotation and viscosity measurements were made in both studies. The refolding kinetics are dependent upon the total pyrrolidine ring content [13] as well as on the degree of cross-linking. Since ichthyocol and the α1 and α2 chains of rat skin collagen have quite different pyrrolidine ring contents, the two sets of data should be viewed separately before correlations are made.

A typical experimental result for CLI is shown in fig. 1a, where the specific rotation, $-[\alpha]^{15}_{365}$, is plotted as a function of time. The absolute value of $[\alpha]^{15}_{365}$ increases sharply during the first half hour after quenching the solution from 40 to 15°C. After the first rapid change, the optical rotation change slows and finally reaches an apparent plateau value after standing overnight. This value keeps changing slightly, however, over a period of many days. As a result, it is difficult to measure a final equilibrium value. As seen in eq. (2), the data es-

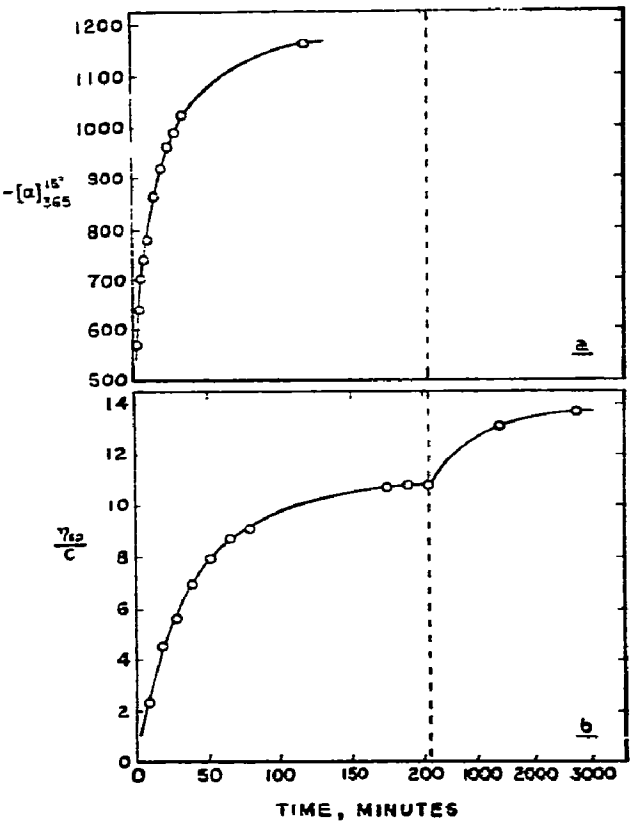


Fig. 1. Recovery of specific rotation and reduced viscosity as a function of time for cross-linked ichthyocol γ-gelatin (CLI). All measurements in pH 2.8, 0.1 M KCl and 0.1 M acetic acid at 15°C.

sentially are to be analyzed as a set of first-order equations. The rate constants of the first order steps are not sufficiently separated so that a plot of log $[\alpha]$ versus t does not provide individual straight line segments. The Guggenheim procedure [14] has been used to rectify these data. In this procedure, one plots the logarithm of $([\alpha]_t - [\alpha]_{t+\Delta})$ versus time, where Δ is a constant small time increment. Such a plot relating to the data from fig. 1a is shown in fig. 2. After about 15 minutes of renaturation, the data plot yields a straight line over the entire period of the terminal reaction. For the case in question, this straight line gives a first-order rate constant of 7.35×10^{-4} sec^{-1}. The values extrapolated from this line to earlier time periods are then subtracted from the corresponding observed values and the differences plotted in the same semi-log fashion. A second straight

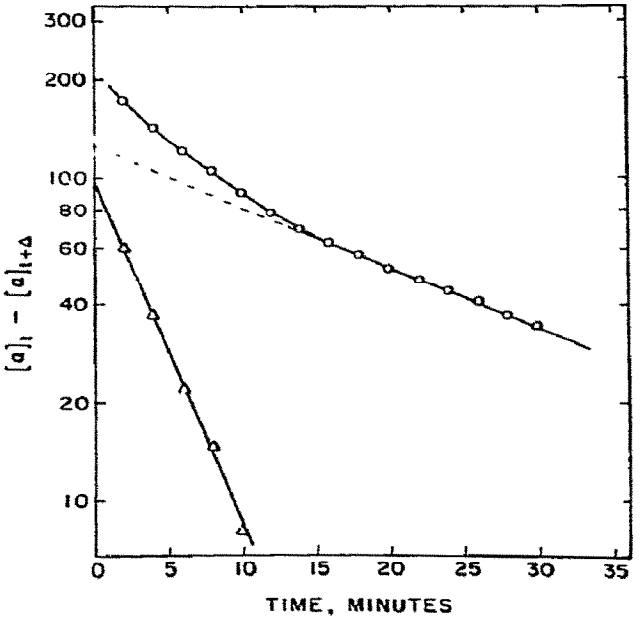


Fig. 2. The Guggenheim plot for the data of fig. 1. The constant time increment is six minutes.

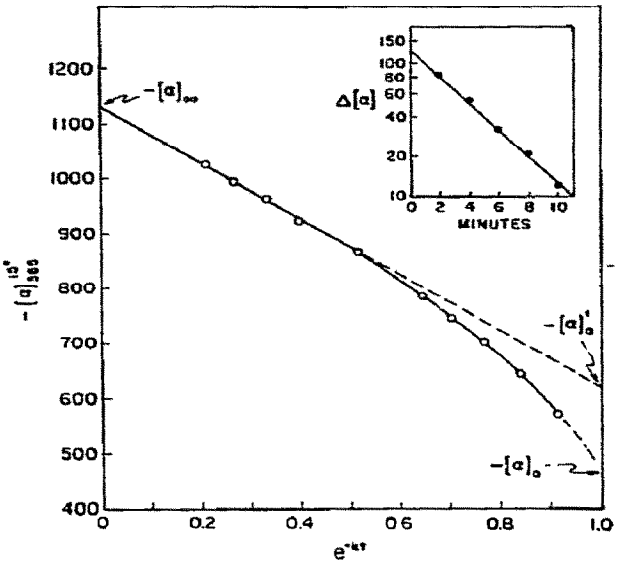


Fig. 3. Plot of specific rotation versus $\exp(-k_s t)$ for determining the zero and infinite time rotation values. Inset, plot of the deviation from linearity, $\Delta[\alpha]_t$, against time.

line is obtained for this difference plot, with a first-order rate constant value of 42×10^{-4} sec^{-1}. Over the time range studied, the optical rotation regain data for CLI appear to be well represented by considering only two first order processes. The choice of Δ is crucial in plotting the data; for best results, different values of the constant time increment, Δ, should be tried.

The zero-time optical rotation value at the refolding temperature T, $[\alpha]_0^T$, is very close to the value for the completely denatured CLI at $T > T_M$. However, a more precise value, at T, can be obtained by plotting $[\alpha]_t$ versus $\exp(-k_s t)$, k_s being the rate constant of the slowest process (fig. 3). The zero-time value can be extrapolated if enough data points are available near zero time, or a still better value can be determined by plotting the deviation from linearity against time in a semi-log plot (see inset in fig. 3). The slope of this line also gives the rate constant for the fast reaction as well as the intercept, $[\alpha]_0^T$. Extrapolation of the $[\alpha]_t$ versus $\exp(-k_s t)$ plot to infinite time yields a good value of $[\alpha]_\infty^T$. This value of $[\alpha]_\infty^T$ may not correspond to a true value of $[\alpha]^T$ at infinite time because of other complicating processes, nor does it correspond to the value for completely native collagen. In the present case, $[\alpha]^{15}_{\text{native},365} = -1350°$, while as seen in fig. 3, $[\alpha]^{15}_{\infty,365} = -1140°$. The use of two methods of plotting the data to obtain the rate constants of the two reaction steps allows one to refine the data by a reiterative process. The first use of the k_s as obtained from the raw data Guggenheim plot in the $[\alpha]_t$ versus $\exp(-k_s t)$ plot shows up departures from linearity at long reaction times. One can adjust k_s, then recalculate k_f, and so on until a best fit is obtained in both plots.

With values of $[\alpha]_0^T$ and $[\alpha]_\infty^T$ determined, one can normalize the experimental data by computing $([\alpha]_t - [\alpha]_\infty)/([\alpha]_0 - [\alpha]_\infty)$, which provides the per cent of the polypeptide in the random-coil form. When the normalized data are plotted as $\ln\{([\alpha]_t - [\alpha]_\infty)/([\alpha]_0 - [\alpha]_\infty)\}$ versus t, fig. 4, only two straight line regions are obtained, another indication that two first-order rate constants are sufficient to characterize the regain of optical rotation upon the refolding of CLI. The observed data fit the equation

$$\frac{[\alpha]_t - [\alpha]_\infty}{[\alpha]_0 - [\alpha]_\infty} = (1 - Y_\alpha)\exp(-k_1 t) + Y_\alpha \exp(-k_2 t), \quad (3)$$

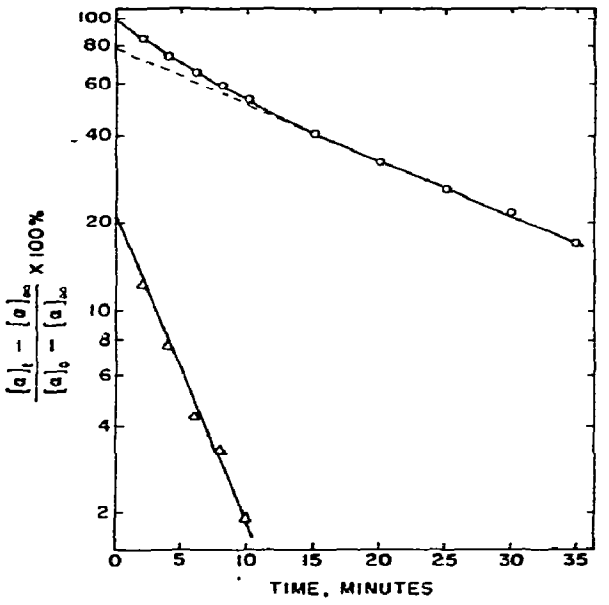


Fig. 4. Fraction of the polypeptide in the denaturated form as a function of time for cross-linked γ-gelatin.

where k_1 and k_2 are the rate constants for the fast (k_1) and slow (k_2) reaction processes. The parameter Y_α represents the fraction of the total rotation change related to the slower process, described by k_2.

The viscosity increases during collagen-fold formation in much the same way as the optical rotation, but as indicated in fig. 1b, the time scale for the viscosity change extends over a much longer period. It is evident that the fast process seen in the first step of the optical rotation regain does not contribute appreciably to the solution viscosity. In spite of this time scale difference, the viscosity recovery data can be treated formally by the same type of normalized equation, viz:

$$\frac{(\eta_R)_t - (\eta_R)_\infty}{(\eta_R)_0 - (\eta_R)_\infty} = (1 - Y_\eta)\exp(-k_I t) + Y_\eta \exp(-k_{II} t). \quad (4)$$

The (η_R) terms represent the reduced viscosity at times t, zero and infinity with the zero and infinity time values being obtained by extrapolation as with the optical rotation. Y_η, k_I and k_{II} have the same meaning as in eq. (3) but the subscript differences imply that these constants might refer to different processes than those seen by the optical rotation.

Table 1

Kinetic constants for the regain of optical rotation during collagen-fold formation of cross-linked and non-cross-linked ichthyocol. Conditions: $T = 15°C$, pH 2.8, 0.1 M HAc, 0.1 M KCl. Specific rotation measured at 365 nm. $[\alpha]^{15}_{365,\text{native}} = -1350°$

Sample	Rate constants $\times 10^4$ sec^{-1}				
	k_1	k_2	Y_α	$-[\alpha]^{15}_0$	$-[\alpha]^{15}_\infty$
CLI	41.9	7.4	0.79	465	1138
DI	2.57	0.40	0.75	465	815

The kinetic parameters for rotation recovery, analyzed according to the above procedure and eq. (3), are shown in table 1. The zero time values, $-[\alpha]^{15}_0$, agree well with the values measured directly on the denatured solutions at 40°C. The infinite time values, $-[\alpha]^{15}_\infty$, while consistent for the two samples, are much lower than that of native, undenatured cross-linked ichthyocol (−1350°). This discrepancy suggests that either some non-cross-linked chains are present, not refolding over the time span of the rotation measurements, or that some intermolecular interactions providing out-of-register chain alignments occur preventing complete renaturation of each individual γ-gelatin molecule. This may be the case because a very slow but constant change in optical rotation is observed as the solution is tempered at 15°C for long periods.

The viscosity parameters analyzed according to eq. (4) are shown in table 2. The sample chosen for analysis

Table 2

Kinetic constants for the regain of viscosity during collagen-fold formation of cross-linked and non-cross-linked ichthyocol. Conditions: $T = 15°C$, pH 2.8, 0.1 M HAc, 0.1 M KCl. Concentration: 0.074 g/dl (CLI); 0.170 g/dl (DI)

Sample	Rate constants $\times 10^4$ sec^{-1}				
	k_I	k_{II}	Y_η	$[\eta]^{15}_0$	$[\eta]^{15}_\infty$
CLI	4.75	0.20	0.30	0.72	13.85
DI	0.42	–	–	0.60	3.58

was the same as one used in the rotation recovery experiment but at a lower concentration to minimize the intermolecular component of the viscosity change. Again, two rate constants are sufficient to express the data. The zero time value obtained by extrapolation, $(\eta_R)_0^{15}$, was 0.72 dl/g, whereas the 40°C value $(\eta_R)^{40}_{\text{denatured CLI}}$, was 0.82 dl/g. The reduced viscosity of the non-denatured ichthyocol was 29 dl/g at c = 0.074 g/dl, while the intrinsic viscosity, $[\eta]_{\text{native}}$, was 15 dl/g. This value agrees well with the $(\eta_R)_\infty$ determined by extrapolation. The formal similarity between viscosity and optical rotation data ends when one compares the values of the rate constants and Y for the two properties. The rate constant k_I has a value less than one-fifth that of k_1 and, in fact, is on the order of k_2. The slower viscosity change rate constant, k_{II}, has no counter-part in the rotation change and is less than one-twentieth of the value of k_I. The fast reaction in viscosity accounts for about 70% of the total (Y_η), whereas the fast reaction in optical rotation accounts for only 20% of the total reaction (Y_α).

During the optical rotation studies, the fast reaction, monitored continuously, takes place within the first ten minutes and quickly decays to zero, whereas the slow process continues for a few hours. On the other hand, one cannot obtain viscosity data in a reliable fashion much before 8 to 10 minutes after initiating fold-formation and it is difficult to obtain enough accurate data points for any time within the first ten minutes of refolding. The fast reaction measured by viscosity covered the first 300 to 400 minutes, the same range as that of the slower optical rotation stage. From the overlapping time ranges for the reactions characterized by k_2 and k_I, the similarity of values of these rate constants, and the similarity of the fraction of the total change ascribable to these reactions, it seems reasonable to conclude that the slow process measured by optical rotation and the fast process measured by viscosity represent the same event. Thus, the experimental data for the cross-linked ichthyocol, as viewed by two very different kinds of measurements, indicate three reaction steps. As depicted by the equation:

Optical rotation change

$$A \xrightarrow{k_1} B \underset{k_I}{\xrightarrow{k_2}} C \xrightarrow[k_{II}]{} D, \qquad (5)$$

viscosity change

the time scales and rate constants (tables 1 and 2) are such that only the first two steps are involved in the optical rotation change, while only steps two and three are within the range of the viscosity measurements‡.

Up to this point, the determination of the rate constants has been independent of whether the reactions considered were parallel or consecutive. If we now assume that the set of events in eq. (5) are consecutive, and further accept that for all practical purposes the final step C → D does not contribute significantly to the optical rotation change at the early stage of the renaturation, then the concentrations of components A, B and C are related to the optical rotation constants by

$$[A] = [A]_0 \exp(-k_1 t), \qquad (6)$$

$$[B] = [A]_0 [k_1/(k_1 - k_2)] [\exp(-k_2 t) - \exp(-k_1 t)], \qquad (7)$$

$$[C] = [A]_0 - [A] - [B], \qquad (8)$$

where $[A]_0$ is the initial concentration of component A. The total optical rotation is the sum of the optical rotation values from all three components, each calculated as the product of concentration and a characteristic specific rotation constant, α_i. Thus,

$$[\alpha]_{\text{total}} = [\alpha]_A + [\alpha]_B + [\alpha]_C$$

$$= \alpha_A [A] + \alpha_B [B] + \alpha_C [C]. \qquad (9)$$

Substituting eqs. (6), (7) and (8) into eq. (9), one obtains the following relationship:

$$\frac{[\alpha]_{\text{total}} - \alpha_C}{\alpha_A - \alpha_C} = (1 - Y_\alpha)\exp(-k_1 t) + Y_\alpha \exp(-k_2 t), \qquad (10)$$

where

$$Y_\alpha = \frac{\alpha_B - \alpha_C}{\alpha_A - \alpha_C} \frac{k_1}{k_1 - k_2}. \qquad (11)$$

Eq. (11) is an important formulation because it permits the evaluation of α_B, the rotation constant for the first intermediate. The terms Y_α, k_1 and k_2 come from direct analysis of the kinetics (table 1), while α_A is equivalent to $[\alpha]_0$ and α_C to $[\alpha]_\infty$.

‡ This situation is not true in the case of the rat skin data to be discussed later. There the rotation data cover a more extended time range and the situation is more complex.

Table 3

Kinetic constants for the refolding of isolated α1, α2 and β_{12} components from rat skin collagen

Rate constants $\times 10^4$ sec^{-1}	Optical rotation		
	α1	α2	β_{12}
k_1	7.46	1.16	12.4
k_2	0.986	0.065	1.48
k_3	0.15	–	0.10
Y_α(A→C)	0.31	0.87	0.40

Table 4

Specific rotation constants for the three stage consecutive reaction mechanism for collagen-fold formation

	α1-(Rat)[a]	α2-(Rat)	β_{12}-(Rat)	CLI	DI
α_A	497	509	497	465	465
α_B	880	584	776	700	593
α_C	1021	931	928	1138	815
α_D	1131	–	1078	–	–
$\sum$imino acids, res/1000	226	201	213	197	197

[a] Optical rotations were measured at 313 nm and converted to 365 nm readings for a direct comparison with ichthyocol collagen.

The kinetic parameters based on the refolding data for (DI), the non-cross-linked denatured ichthyocol, are also shown in tables 1 and 2. Aside from the difference in magnitude of the optical rotation rate constants, the main difference is that in the time period covered by the available data, only a single rate constant is obtained. The data of Piez and Carrillo [10] on the refolding kinetics of rat skin α1, α2 and β_{12} molecules are all summarized in table 3. These data differ from the ichthyocol data in that the rotation was measured over a longer period of time and a more detailed analysis of the late stages of refolding was possible. Thus, the complete transition, A → D [eq. (5)], is viewed by the specific rotation changes. The viscosity data of Piez and Carrillo [10] still cover too short a time period to provide all of the rate constants.

The calculated rotation constants are presented in table 4 for all of the rotation data. The differences between constants can be considered in two ways, as a function of intrinsic chain composition, and as a function of chain registration arising from cross-linking.

Piez and Gross [13], in a study typical of many, have shown that the denaturation temperature of soluble collagens from a variety of sources is a linear function of the total pyrrolidine ring content. Comparing the rotation constants, α_B, of α1, α2 and DI in table 4 and the pyrrolidine ring compositions, it appears that α_B is directly dependent on the content of imino acids. Rat skin α2 has very nearly the same imino acid composition as the denatured ichthyocol, DI, and the rate constants and rotation parameters are quite similar. The rat skin α1 chain, on the other hand, has a much higher imino acid content and shows a markedly increased rate constant for process A → B and the characteristic rotation of intermediate B. The characteristic rotation of B can be interpreted as indicating a marked enhancement of structural order in B, and hence, in the stability of B. If, as Flory and Weaver [6] have suggested, the stability of the nucleated intermediate determines the minimum number of residues, n^*, required to participate cooperatively in a conformational fluctuation leading to further propagation of the fold-unit, and if the rate of the first step is inversely proportional to n^*, then this heightened value of α_B for the α1-component as compared to α2 is responsible for the sevenfold difference in the rate constants k_1 for the two cases.

From the argument above, it would appear that α_B for DI is somewhat high. However, Veis and Drake [15] showed DI to contain about 20% β- and only 80% α-component. Based on the imino acid content of rat skin β_{12}, we would expect an α_B value of 730. The value of 776 shown in table 4, thus indicates that cross-linking also provides chain regions with enhanced stability in which fold-formation may be nucleated. This is even more strikingly demonstrated in comparing α_B values for DI and CLI where the only composition difference is in the presence of cross-linkages in CLI. The fact that k_1 and α_B for DI are higher than in rat skin α2 can be attributed to the presence of the β-component.

Veis and Schnell [16], Veis et al. [17] and Bensusan and Nielsen [18] have all shown by hydrogen–deuterium exchange experiments that the first stage of fold-formation does involve the formation of inter-chain hydrogen bonds. The very high values for k_1 for the cross-

linked collagens, compared to the lesser effects of change in α_B (e.g., compare α_B and k_1 for rat skin $\alpha 1$ and β_{12}) indicate that chain registration to provide inter-chain hydrogen-bonding can also supply nucleation sites for fold-formation, or add markedly to the stability of the nucleated regions, again reducing the size of n^*. Veis and Legowik [19] used equilibrium data close to the melting temperature to estimate n^* for bovine α and γ-gelatins and found n^* to be less for γ than α.

3. Conclusions

The data analysis discussed above presents a rather more detailed insight into the mechanism of collagen-fold formation than heretofore possible. By considering the refolding processes as consecutive rather than parallel reactions, one may determine constants characteristic of the state of order in the hypothetical nucleated intermediate. Analysis of these constants, along with the rate constants for the first rapid reaction step, show that two factors can lead to the formation of stable nuclei. First of course, is the content of imino acids which determines the structural stability of a substantial part of each gelatin chain. Equally important, however, as shown by the data on cross-linked gelatins, is the fact that hydrogen-bonding interchain interactions also are involved in the stabilization of the initial nuclei, reducing the number of peptide units which must interact cooperatively before chain growth can occur. Chain registration, by cross-linking in the native state, leads to the most favorable hydrogen-bonding alignment.

This interpretation of the consecutive reaction sequence suggests that the molecular interpretation of eq. (1) presented earlier is inappropriate. One cannot distinguish between a single chain nucleated form, NC of eq. (1), and a hydrogen-bonded multi-stranded nuclei, B of eq. (5). In fact, the hydrogen–deuterium exchange studies during fold-formation [17, 18] imply that nucleation and hydrogen-bonding are concurrent processes. Thus, it seems likely that, as depicted in the scheme represented in eq. (12)

$$\text{Single chain random coils} \xrightarrow[\text{nucleation step}]{k_1} \text{helix nuclei, } n^* \text{ residues, three chains (may start in } \textit{either} \text{ cross-link regions or pyrrolidine-ring rich regions)}$$

$$\xrightarrow[\text{helix propagation}]{k_2} \text{lengthen helical segments}$$

$$\xrightarrow[\text{annealing reordering}]{k_3} \text{perfected triple helical regions} \qquad (12)$$

the triple helix nuclei indistinguishably originate in either pyrrolidine residue rich sequences or cross-link containing sequences, but that in either case the ordered crystallite helix nucleus is a compound helix, hydrogen-bonded structure.

The second reaction stage, contributing to the major change in both optical rotation and viscosity is the propagation of helix from the nuclei. The final step is that of annealing or tempering, perfecting the triple helix by securing the most stable interchain chain alignments. In the highly cross-linked γ-gelatins, this most stable alignment is facilitated and one achieves nearly complete renaturation with small contribution from step 3.

References

[1] P.H. von Hippel, in: Treatise on collagen, Vol. 1, ed. G.N. Ramachandran (Academic Press, New York, 1967) p. 253.
[2] A. Veis, The macromolecular chemistry of gelatin (Academic Press, New York, 1964).
[3] H. Boedtker and P. Doty, J. Phys. Chem. 58 (1954) 968.
[4] P.H. von Hippel and W.F. Harrington, Biochim. Biophys. Acta 36 (1959) 427.
[5] W.F. Harrington and P.H. von Hippel, Arch. Biochem. Biophys. 92 (1961) 100.
[6] P.J. Flory and E.S. Weaver, J. Am. Chem. Soc. 82 (1960) 4518.
[7] K. Altgelt, A.J. Hodge and F.O. Schmitt, Proc. Natl. Acad. Sci. US 47 (1961) 1914.
[8] J. Engel, Arch. Biochem. Biophys. 97 (1962) 150.
[9] M.P. Drake and A. Veis, Biochemistry 3 (1964) 135.
[10] K.A. Piez and A.L. Carrillo, Biochemistry 3 (1964) 908.
[11] G. Beier and J. Engel, Biochemistry 5 (1966) 2744.
[12] W.F. Harrington and G.M. Karr, Biochemistry 9 (1970) 3725.
[13] K.A. Piez and J. Gross, J. Biol. Chem. 235 (1960) 995.
[14] E.A. Guggenheim, Phil. Mag. 2 (1926) 538.

[15] A. Veis and M.P. Drake, J. Biol. Chem. 238 (1963) 2005.
[16] A. Veis and J. Schnell, in: Symposium on fibrous proteins, ed. W.G. Crewther (Butterworths, London, 1967), p. 193.
[17] A. Veis, S. Ghosh and F. Jacobs, Conn. Tiss. Res. 1 (1972) 135.
[18] H.B. Bensusan and S.O. Nielsen, Biochemistry 3 (1964) 1367.
[19] A. Veis and J.T. Legowik, in: Structure and function of connective and skeletal tissue, eds. S. Fitton-Jackson et al. (Butterworths, London, 1965) p. 70.

BIOPHYSICAL CHEMISTRY 1 (1973) 125–129. © NORTH-HOLLAND PUBLISHING COMPANY

ENERGY TRANSFER BETWEEN CHLOROPHYLL C AND CHLOROPHYLL A

Danuta FRACKOWIAK and D. WRÓBEL

Institute of Physics, Poznań Technical University, 61-138 Poznań Piotrowo 5, Poland

Received 18 April 1973
Revised manuscript received 10 July 1973

Spectral properties of solutions containing mixtures of chlorophyll a and chlorophyll c are investigated. The yield of excitation energy migration from chlorophyll c to chlorophyll a is obtained ranging from 23 to 48% dependent on the used dye concentrations. The back transfer from chlorophyll a to chlorophyll c is negligible. The shape of the polarization excitation spectrum of chlorophyll c in the Soret band region is less composed than that of chlorophyll a. Depolarization of chlorophyll a fluorescence by chlorophyll c is in agreement with the conclusion drawn from fluorescence quenching that excitation energy migrates from chlorophyll c to chlorophyll a.

1. Introduction

The mechanism responsible for energy transfer between accessory pigments and chlorophyll a (chl a) *in vivo* has not as yet been definitely established. Many authors presume that this transfer follows the inductive resonance mechanism, whereas there is also some evidence [1, 2] that interaction between pigments could be much stronger than that following the "very weak" Förster [3, p. 67] mechanism.

In the present paper, the solution of a mixture of chlorophyll a and chlorophyll c (chl c) is investigated.

In the inductive resonance mechanism, the efficiency of energy transfer is proportional to the overlapping of the donor emission band and acceptor absorption spectrum [3, p. 69]. Therefore, in the case of a mixture of chl a and chl c, only inefficient transfer of energy from chl c and chl a can be predicted from this mechanism.

The Soret bands of chl a and chl c overlap to a high degree [4, p. 154, p. 160], suggesting that "prerelaxational" [5] or "fast" [2] transfer of excitation can be, at least partially, involved in the efficient energy migration between those chlorophylls observed *in vivo*. Such "fast" energy transfer (intervening before Boltzmann distribution of vibrational energy has been reached) was proposed by Bauer et al. [2] to explain the energy migration between chlorophyll b (chl b) and chl a.

The energy transfer from chl c to chl a plays a role in the photosynthesis of organisms dispersed in water. Certain types of organisms exhibit an equivalent photosynthesis efficiency in yellow (500 – 600 nm) and red (600 – 690 nm) light. From the absorption spectra of chl a and chl c it follows that in this case chl c plays the role of accessory pigment responsible for efficient photosynthesis [6; 4, p. 513]. Inoué et al. [7] measured the yield of energy migration between chl c and chl a *in vivo* and obtained a value of about 55%.

The purpose of the present paper is to obtain information as to what type of mechanism is responsible for this high yield of energy transfer.

2. Material and methods

Pigments were extracted from *Fucus* using a mixture of methanol, acetone and ethyl ether. Chl c was separated from chl a and carotenoids on a column filled with cellulose. The column was washed with petroleum ether with 0.5% of propanol and again with ethyl ether containing 5% of methanol. The chl c zone was located on the top of the column. It was cut from the column and eluated with ethyl ether. The procedure of purification of chl c was similar to that used by Jeffrey [8]. Separation of chl a from carotenoids was performed in a column filled with sugar.

Dye concentrations were obtained from absorption

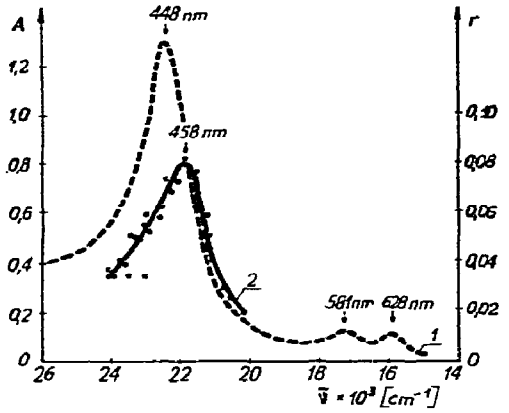


Fig. 1. (1), Absorption spectrum of chl c solution in ethyl ether; (2), emission anisotropy versus wavenumber.

spectra using absorption coefficients of chl c and chl a given in the literature [8, 9].

Details concerning the measurements of absorption, fluorescence and excitation spectra are given in an earlier paper [10].

The polarization of emission was measured with the device used in fluorescence spectra measurements but equipped with polaroid sheets.

Spectral bandwidths were, in the case of excitation spectra: in the exciting beam 6 nm, in the fluorescence beam 20 nm; and for polarization measurements: 20 nm and 20 nm, respectively.

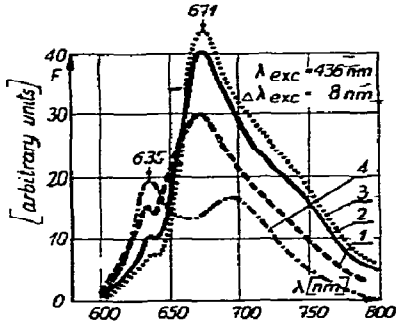


Fig. 2. Fluorescence spectra of chl c and chl a mixture excited at 436 nm. (1), $C_{chl\ c}$ = 23.5 μM, $C_{chl\ a}$ = 8 μM; (2), $C_{chl\ c}$ = 23.5 μM, $C_{chl\ a}$ = 26 μM; (3), $C_{chl\ c}$ = 23.5 μM, $C_{chl\ a}$ = 53 μM; (4), $C_{chl\ c}$ = 12.3 μM, $C_{chl\ a}$ = 0 μM.

3. Results

Fig. 1, curve 1 shows the absorption spectrum of chl c solution in ethyl ether.

Fig. 2 shows the fluorescence spectrum of chl a and chl c mixtures for various concentrations of the dyes. The fluorescence is excited with 436 nm, a wavelength belonging to the Soret band of both pigments. The molar absorption coefficients of both pigments in this region are comparable, but as follows from fig. 2, the contribution from chl a fluorescence predominates in the emission spectrum (the maximum at 671 nm belongs to chl a emission, that at 635 nm to chl c).

Fig. 3 shows the fluorescence spectra of the same set of samples but excited with a 450 nm beam, absorbed predominantly by chl c. Though in this case the contribution to emission from chl c is higher, some of the energy is nevertheless transferred to chl a.

The yields of energy migration between chl c and chl a listed in table 1 are calculated from fluorescence and absorption spectra using the formula given in ref. [2]. As can be seen from table 1 the yield of energy migration *in vitro* is slightly lower than that obtained *in vivo* [7], but *in vivo* the average distances are possibly shorter than in solution and in organisms the dye molecules are mutually oriented in lamellar system. The migration yield from chl a to chl c is, in the limit of accuracy of our experiments, negligible. If the "pre-relaxational" mechanism of transfer was involved, one would expect a higher efficiency of "back transfer" from chl a to chl c than that obtained. The differences between the yields of energy transfer from chl c to chl a under various excitations are within limits of accuracy of the method (table 1), whereas on the ground of "pre-relaxational" mechanism much bigger differences can be expected [11].

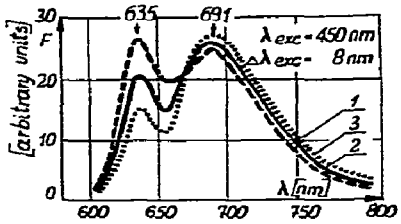


Fig. 3. Fluorescence spectra of chl c and chl a mixture excited at 450 nm. $C_{chl\ c}$: 23.5 μM; $C_{chl\ a}$: (1) 8μM, (2) 26μM, (3) 53μM.

Table 1
Yield of energy migration between chl c and chl a

Wavelength of excitation (nm)	Concentration (μM)		Yield of energy migration from chl c to chl a	Yield of energy migration from chl a to chl c
	chl a	chl c		
450	26	23.5	0.23 ± 0.12	0.04 ± 0.12
436			0.30 ± 0.10	0.08 ± 0.04
450	53	23.5	0.37 ± 0.11	0.05 ± 0.15
436			0.49 ± 0.09	0.09 ± 0.08

From the plot of $\log[f/(1-f)]$ (f – efficiency of energy migration) versus $\log R$ (R – mean distance between donor and acceptor), the exponent x of the donor–acceptor distance R in the function [2] describing the yield of energy migration is obtained. This x-value is about 6 what suggests "very weak" Forster [3, p. 67] interaction between chlorophylls, at least at the donor and acceptor concentrations used by us.

The same method of measurement and calculation for the case of biliproteins and chlorophyllin mixtures reveals values between 3 and 5 for x [12]. It is possible that the presence of proteins has influences on the mechanism of interaction between pigment groups.

In fig. 4, the excitation spectrum of chl c emission measured for a chl a + chl c mixture is shown. The maximum of excitation spectrum is located at 453 nm, i.e., near the chl c absorption maximum.

The excitation spectrum of chl a emission of the same sample (fig. 5) presents a much higher peak at the absorption maximum of chl a (at 431 nm) than at the maximum of chl c fluorescence (at 453 nm), because the yield of chl c fluorescence is about 5 times lower than that of chl a. Therefore, the excitation spectrum results are in agreement with those presented in figs. 2 and 3 as well as in table 1.

Fig. 1, curve 2 shows the emission anisotropy, $r = (J_{\parallel} - J_{\perp})/(J_{\parallel} + 2J_{\perp})$, versus the wavenumber of exciting light for the solution of chl c with chl a added (concentration ratio 8 : 1). Fluorescence was observed at the maximum of chl c emission (at 635 nm); therefore, this graph presents practically the shape of the chl c excita-

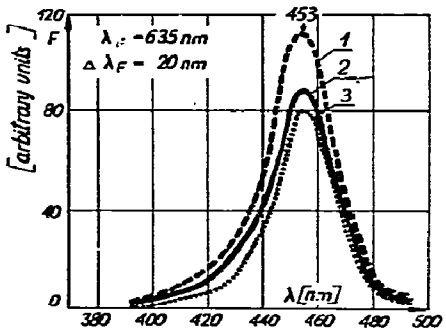


Fig. 4. Excitation spectra of chl c and chl a mixture observed at 635 nm. $C_{\text{chl c}}$: 23.5 μM; $C_{\text{chl a}}$: (1) 8μM, (2) 26μM, (3) 53μM.

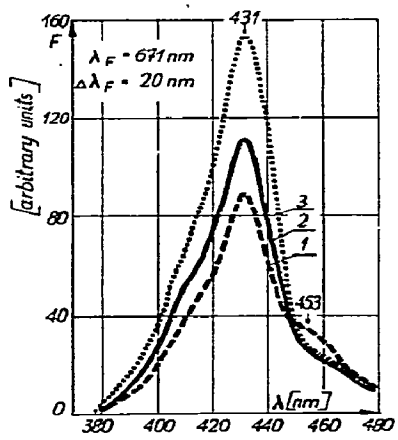


Fig. 5. Excitation spectra of chl c and chl a mixture observed at 671 nm. $C_{\text{chl c}}$: 23.5 μM; $C_{\text{chl a}}$: (1) 8μM, (2) 26μM, (3) 53μM.

Table 2
Emission anisotropy of chl c with chl a mixtures

Wavelength of excitation (nm)	Wavelength of fluorescence (nm)	Concentration ratio $C_{chl\ c}/C_{chl\ a}$	Ratio of absorption coefficients divided by ratio of fluorescence yields $\frac{\epsilon_{chl\ c}/\epsilon_{chl\ a}}{\eta_{chl\ c}/\eta_{chl\ a}}$	Emission anisotropy r
432	635	8 : 1		0.050 ± 0.01
(max of chl a absorption)	(max of chl c fluorescence)	3 : 1	≈ 4	0.030 ± 0.01
		1 : 1		0.028 ± 0.01
452	635			
(max of chl c absorption)	(max of chl c fluorescence)	8 : 1	≈ 60	0.076 ± 0.01
452	671	3 : 1		0.007 ± 0.01
(max of chl c absorption)	(max of chl a fluorescence)		≈ 60	
		1 : 1		−0.015 ± 0.01

tion polarization spectrum. This spectrum of chl c in the Soret band region is much less complex than that of chl a [13; 4, p. 156; 14, p. 257]. The Soret band of chl a is composed of two transitions with different polarization and energy, whereas chl c seems to be more like porphyrine with symmetry higher than chl a and therefore with two completely degenerate transitions responsible for the Soret band.

The fluorescence polarization measurements of chl c are difficult because of the low fluorescence efficiency of this dye. Therefore, the spectral region used in the exciting beam had a width of 20 nm, and the spectral resolution of the excitation polarization spectrum was worse than that obtained previously in chl a measurements [13].

However, fig. 1 suggests that the transitions are "almost" degenerate, or at least that the degeneracy is removed to a lesser extent than in the case of chl a.

The addition of chl c to a chl a solution causes strong depolarization of chl a fluorescence (table 2), in agreement with the conclusion drawn from fluorescence quenching that energy migrates from chl c to chl a. The range of dye concentrations used excludes self-quenching phenomena. Thus the observed changes in fluorescence polarization are related with the following three effects:

(1) an increase in contribution from chl a to absorption and emission of the mixture with increasing chl a concentration;
(2) depolarization of fluorescence by energy migration; and
(3) a decrease in the amount of light absorbed by chl c due to absorption of the exciting beam by chl a.

Evaluation of these three effects shows that in the case of excitation at 452 nm (at the maximum of chl c absorption) and observation of emission at 671 nm (maximum of chl a emission), the most important is the second effect – depolarization by energy migration. An increase in chl a concentration influences also, though to a lesser extent, the anisotropy of chl c emission, but in this case the effects 1 and 3 play an important role because of the overlap in this excitation region of both absorption spectra.

From our results it follows that in model systems energy migrates from chl c to chl a with an efficiency similar to migration *in vivo*. The mechanism of this migration seems to follow the "slow" inductive resonance mechanism.

Acknowledgement

This study was carried out under Project No. 09.1.7, coordinated by the Institute of Ecology of the Polish Academy of Sciences.

References

[1] D. Frąckowiak and J. Grabowski, Photosynthetica 4 (1970) 236.
[2] R.K. Bauer, L. Szalay and E. Tombacz, Biophys. J. 12 (1972) 731.
[3] Th. Förster, in: "Comprehensive biochemistry, Vol. 22, eds. M. Florkin and E.H. Stotz (Elsevier, Amsterdam, 1967).
[4] J.C. Goedheer, in: The chlorophylls, eds. L.P. Vernon and G.R. Seely (Academic Press, New York, 1966).
[5] R.K. Bauer and H. Cherek, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 20 (1972) 961.
[6] R.C. Dougherty, H.H. Strain, W.A. Svec, R.A. Uphaus and J.J. Katz, J. Am. Chem. Soc. 92 (1970) 2826.
[7] H. Inoué, K. Sugahara and M. Nishimura, Proceedings of the Seventh International Seaweed Symposium, Sapporo, Japan (1971) p. 515.
[8] S.W. Jeffrey, Biochem. J. 86 (1963) 313.
[9] A.S. Holt and E.E. Jacobs, Am. J. Bot. 41 (1954) 710.
[10] D. Frąckowiak and S. Surma, Photosynthetica 2 (1968) 75.
[11] Z. Salamon, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys., to be published.
[12] D. Frąckowiak and J. Grabowski, Photosynthetica 7 (1973), to be published.
[13] D. Frąckowiak and Z. Trumpakaj, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 12 (1964) 183.
[14] G.P. Gurinovich, A.N. Sevchenko and K.N. Soloviev, in: Spektroskopiya khlorofilla i rodstvennykh soyedinieni (Nauka i Tekhnika, Minsk, 1968).

BIOPHYSICAL CHEMISTRY 1 (1973) 130–133. © NORTH-HOLLAND PUBLISHING COMPANY

SOME ASPECTS OF COOPERATIVITY IN HUMAN HEMOGLOBIN

Lambert H.M. JANSSEN and Simon H. DE BRUIN
Department of Biophysical Chemistry, University of Nijmegen, Nijmegen, The Netherlands

Received 14 May 1973
Revised manuscript received 30 August 1973

The cooperativity in hemoglobin can be described by the Hill parameter n, the free energy of interaction ΔF_I and the allosteric free energy ΔF_A. By this latter is meant here the free energy change associated with the transition from the deoxy to the oxy conformation in hemoglobin. In this paper some general relations between n, ΔF_I and ΔF_A are given. A method is presented by which ΔF_A can be calculated from oxygenation data.

1. Introduction

In the cooperative behaviour of the oxygen binding by hemoglobin an important part is played by the salt bridges which constrain the deoxy form and which break up upon going to the relaxed oxy conformation [1, 2]. The change in free energy associated with the transition from the deoxy to the oxy conformation will be defined here as the allosteric energy ΔF_A. This free energy change was called the inter-subunit bonding energy by Noble [3] who calculated a value of 8 kcal/tetramer for it from the difference in ligand affinity between deoxyhemoglobin and the α and β chains, whereas Perutz [1] estimated a value between 6 and 12 kcal/tetramer, based on the presence of six salt bridges. On the other hand, Wyman [4] has introduced the concept of free energy of interaction, ΔF_I, which is quite different from ΔF_A [5, 6]. Besides ΔF_A and ΔF_I, the Hill parameter n is frequently used to describe the cooperative effects in hemoglobin. The object of this paper is to show some relations between n, ΔF_A and ΔF_I and further to give an estimate of the magnitude of ΔF_A, based on oxygenation data of human hemoglobin.

2. The free energy of interaction ΔF_I

Our discussion will partially be based on experimental oxygenation curves and therefore we need parameters which satisfactorily describe such a curve. We chose the Adair scheme [7] for this purpose. In this model the fractional saturation Y is given by:

$$Y=\frac{k_1p+3k_1k_2p^2+3k_1k_2k_3p^3+k_1k_2k_3k_4p^4}{1+4k_1p+6k_1k_2p^2+4k_1k_2k_3p^3+k_1k_2k_3k_4p^4}, \quad (1)$$

where k_1, k_2, k_3 and k_4 are the intrinsic association constants for the reaction $Hb(O_2)_{i-1} + O_2 \rightleftharpoons Hb(O_2)_i$ ($i = 1$ to 4) and p the partial oxygen pressure. It should be realized that eq. (1) will be used in this paper merely as a mathematical description of an oxygen saturation curve. The Adair scheme has been chosen since the available experimental oxygenation curves are analyzed according to this scheme. The fact that equivalent binding sites are assumed in the derivation of eq. (1) provides no impediment to the use of it and we are not concerned with the physical meaning of k_1 to k_4. The only limits we impose on the Adair model in this section is that $k_4 \geq k_3 \geq k_2 \geq k_1$. The models mostly used for a description of the interactions in hemoglobin, viz., the Monod–Wyman–Changeux and the Koshland–Némethy–Filmer model fulfil this condition [6].

The Hill plot is defined as $\log[Y/(1-Y)]$ against $\log p$ and the slope n at $Y = 0.5$ at the half saturation pressure $p_{1/2}$ is called the Hill parameter.

We now will discuss the relation between n and ΔF_I. Wyman [4] has shown that when Y approaches 0 or 1, n should become 1 independently of any model. Furthermore, he defined ΔF_I as the free energy of interaction per heme which can be calculated from the distance

between the asymptotes of a Hill plot. It has been shown that in terms of the Adair model $\Delta F_I = RT\ln(k_4/k_1)$ [6, 8]. The cooperativity, however, as measured by n will certainly also depend on k_2 and k_3, so it will be evident that ΔF_I can only partially describe allosteric effects. To illustrate this point we investigated the dependence of n on ΔF_I.

Let $k_2 = ak_1$, $k_3 = bk_2$, $k_4 = c^{-1}k_1$ and $k_1 p_{1/2} = x$. Then it follows from the definition of n and from eq. (1):

$$n = (6x - 2a^2bx^3 + 4)/(3x + 3ax^2 + a^2bx^3 + 1), \quad (2)$$

where x follows from

$$a^2bc^{-1}x^4 + 2a^2bx^3 - 2x - 1 = 0. \quad (3)$$

Thus the Hill parameter is determined only by the relative magnitudes of the Adair constants. For each value of ΔF_I, n may take a range of values, depending on k_2 and k_3. However, n reaches a maximum when $k_1 = k_2$ and $k_3 = k_4$. In that case eq. (2) reduces to the simple equation:

$$n = 4(c^{1/2} + 1)/(4c^{1/2} + 3c + 1). \quad (4)$$

In fig. 1, curve A gives this relation between n and ΔF_I. On the other hand, n becomes minimal for a fixed value of ΔF_I when $k_1 = k_2 = k_3$ or $k_2 = k_3 = k_4$. In fig. 1, curve B represents this minimal n as a function of ΔF_I. Thus as long as $k_4 \geq k_3 \geq k_2 \geq k_1$ which seems to hold for human hemoglobin (at least in the absence of allosteric effectors as 2,3-diphosphoglyceric acid) [8,9], any observed combination of n and ΔF_I should fall in the region between curves A and B in fig. 1. For example n for human hemoglobin at neutral pH is found mostly near 2.7. From the curves presented in fig. 1 it follows that in that case ΔF_I should vary between 2.2 and 3.3 kcal per heme as indeed usually is found. The figure further shows that an increase in ΔF_I does not necessarily correspond with an increase in n and vice versa.

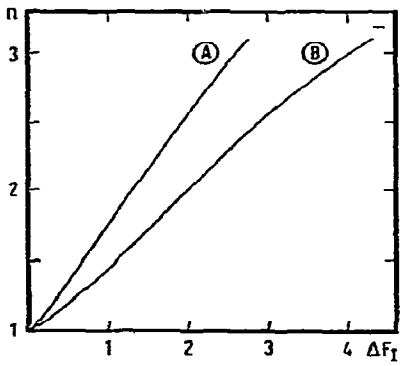


Fig. 1. Dependence of n on ΔF_I (kcal/heme). Curve A represents the maximum and curve B the minimum value of n for a given ΔF_I under the restrictions mentioned in the text. In the calculations a temperature of 25° was assumed.

3. The allosteric free energy ΔF_A

As stated in the introduction ΔF_A stands for the standard free energy change of the transition

$$[\alpha_2\beta_2]_T \rightarrow [\alpha_2\beta_2]_R, \quad (5)$$

where $\alpha_2\beta_2$ represents the non ligated hemoglobin tetramer and T and R the tensed deoxy and the relaxed oxy conformation, respectively. When ligand binding can be described by the two-state model [8], ΔF_A is equal to $RT\ln L$, L being the allosteric constant. Our aim is to obtain an expression for ΔF_A in terms of Adair constants by considering the reactions

$$[\alpha_2\beta_2]_T + 4O_2 \rightarrow [\alpha_2(O_2)_2\beta_2(O_2)_2]_R \quad (6)$$

and

$$[\alpha_2\beta_2]_R + 4O_2 \rightarrow [\alpha_2(O_2)_2\beta_2(O_2)_2]_R. \quad (7)$$

The free energy change associated with reaction (6) is the total free energy of oxygenation ΔF_O. The free energy change of reaction (7) will be represented by ΔF_R. It will be clear that ΔF_A is equal to $\Delta F_O - \Delta F_R$. ΔF_O is simply given by $4RT\ln p_m$, p_m being the median oxygen pressure [4] whereas p_m is related to the Adair constants by the relation $p_m^4 = (k_1k_2k_3k_4)^{-1}$ [6, 8]. For the determination of ΔF_R we will assume that before the last ligation step takes place, the molecule has switched over from the T to the R state. In the model proposed by Perutz [1] a conformational change has been suggested after the binding of the second ligand. Kinetic experiments [10, 11], non-linear relationships

between fractional saturation and structural changes of hemoglobin as observed by electron paramagnetic resonance [12], nuclear magnetic resonance [13], ultraviolet spectroscopy [14] and the release of 2,3-diphosphoglyceric acid [15, 16] seem to support indeed that at least the binding of the last ligand takes place when hemoglobin is in the relaxed conformation. Theoretical considerations support this view [6, 8, 17–19]. Indicating the affinity constants of the reactions

$$[\alpha_2(O_2)\beta_2(O_2)_2]_R + O_2 \rightarrow [\alpha_2(O_2)_2\beta_2(O_2)_2]_R, \quad (8)$$

$$[\alpha_2(O_2)_2\beta_2(O_2)]_R + O_2 \rightarrow [\alpha_2(O_2)_2\beta_2(O_2)_2]_R \quad (9)$$

by a_4 and b_4, respectively, ΔF_R becomes $-RT\ln a_4^2 b_4^2$ and so

$$\Delta F_A = RT\ln a_4^2 b_4^2 p_m^4. \quad (10)$$

Unfortunately, values of a_4 and b_4 are lacking for oxygen as a ligand. Only with *n*-butyl isocyanide as a ligand have these constants been determined [20]. Here it was found that a_4 is about two times b_4.

However, for k_4 we can write $k_4 = 2a_4b_4/(a_4 + b_4)$ and with $a_4 = fb_4$ eq. (10) becomes

$$\Delta F_A = RT\ln(k_4^4 p_m^4) + \Delta F_f, \quad (11)$$

where $\Delta F_f = 2RT\ln[(f+1)^2/4f]$. It should be noted that ΔF_f is always positive. Eq. (11) reduces to

$$\Delta F_A = RT\ln(k_4^4 p_m^4) = RT\ln(k_4^3/k_1k_2k_3) \quad (12)$$

when $f = 1$. There are several arguments suggesting that a_4 and b_4 do not differ more than a factor 2–3. First, the isolated α and β chains have equal or nearly equal oxygen affinity [21, 22], although an unequivocal interpretation is obscured by the different association behaviour of the isolated chains; moreover, the affinity of the isolated β chains has been reported to be equal to k_4 [22]. It is also pertinent that on modification of hemoglobins or on total dissociation of modified hemoglobins the cooperativity mostly reduces but *n* never becomes smaller than 1 [23–25], whereas a difference of a factor two in affinity would result in an *n* of about 0.9 in absence of any cooperativity.

Studies on artificial cyano- or aquomet intermediates of human hemoglobin indicate a small non-equivalence of the chains. From reported values of *n* and $p_{1/2}$ one can easily calculate the affinity constants a_2 and b_2 for the binding of the second molecule of oxygen of the intermediates $\alpha_2\beta_2^+$ and $\alpha_2^+\beta_2$. It was found that b_2 is about 1.5 to 2.5 times a_2 [21, 26]. On the other hand, the measurements of Maeda et al. [27] on cyanomet intermediates result in values for a_2 and b_2 of 3.9 and 3.4 $mmHg^{-1}$, respectively, whereas $k_4 = 4.0$ $mmHg^{-1}$ under identical experimental conditions [8]. In the case of non-equivalence ΔF_A calculated using eq. (12) will be too small, but ΔF_f amounts to only 0.14 kcal when $f = 2$ and to 0.34 kcal when $f = 3$. In view of the absolute magnitude of ΔF_A this is not a serious error so we will use eq. (12) as a very good approximation.

We now want to make some remarks about ΔF_A. By rewriting eq. (12) in the form

$$\Delta F_A = \Delta F_I + RT\ln(k_4^2/k_2k_3), \quad (13)$$

it follows that $\Delta F_A = \Delta F_I$ when $k_2 = k_3 = k_4$ and $\Delta F_A = 3\Delta F_I$ when $k_1 = k_2 = k_3$. So $\Delta F_I \leqslant \Delta F_A \leqslant 3\Delta F_I$. The limits $\Delta F_A = \Delta F_I$ and $\Delta F_A = 3\Delta F_I$ correspond with a minimal value of *n*, whereas the maximal value of *n* may be observed when $\Delta F_A = 2\Delta F_I$. Since ΔF_I ranges from 2.2 to 3.3 (see above), ΔF_A should vary between 2.2 and 9.9 kcal/tetramer. It should be noted here that ΔF_A has been defined per tetramer and ΔF_I per heme.

It should further be recognized that the occurrence of Adair parameters in the several equations does not mean that the validity of these equations depends on the validity of the Adair equation (1). If a different starting model was used other formulae would appear. However, applied to the same experimental data, the same value of ΔF_A should be found. In other words, the value of ΔF_A is of course independent of any model. The difference between the way of calculating ΔF_A as presented here and that of Noble [3] is evident. The approximation introduced by Noble is that in eq. (10) p_m is equated to $p_{1/2}$ and a_4 to b_4, taking for the last two the affinity of the isolated chains, thus assuming that this affinity does not change if an isolated chain is embodied in the tetrameric relaxed hemoglobin molecule.

It should be realized that ΔF_A changes when experimental conditions as pH, temperature, ionic strength, protein concentration and concentration of allosteric

effectors are varied. We briefly want to discuss the temperature and pH effect. The temperature effect is relatively small: in the region 20–30°, ΔF_A does not change significantly according to the results of Imai and Tyuma [28]. The pH dependence is rather large and results from the change in $p_{1/2}$ with pH. As long as k_4 remains constant, which seems to be the case in pH region 7.0 to 7.8 [22], this change in ΔF_A is given by $4RT(\Delta \ln p_m)$. This means that ΔF_A decreases continuously from pH 7 to higher pH. For example, from the data of Bunn and Guidotti [29] one finds a decrease in ΔF_A of about 1.7 kcal going from pH 7.0 to 7.8 (0.03 M bis–tris, 0.01 M Cl^-, equating p_m to $p_{1/2}$).

Applying eq. (12) to the Adair constants reported by Roughton and Lyster [9] for human hemoglobin (pH 7.0, 0.6 M phosphate buffer, 19°) and to the data of Imai [8] (pH 7.4, 0.05 M bis–tris buffer, 25°) one can easily calculate a value for ΔF_A of 5.7 kcal and 4.7 kcal, respectively. The difference of 1 kcal between the two values can be attributed mainly to the difference in pH of the two sets of data.

References

[1] M.F. Perutz, Nature 228 (1970) 726.
[2] M.F. Perutz, Nature 237 (1972) 495.
[3] R.W. Noble, J. Mol. Biol. 39 (1969) 479.
[4] J. Wyman, Advan. Protein Chem. 19 (1964) 223.
[5] H.A. Saroff and A.P. Minton, Science 175 (1972) 1253.
[6] L.H.M. Janssen and S.H. de Bruin, Intern. J. Peptide Protein Res. 5 (1973) 27.
[7] G.S. Adair, J. Biol. Chem. 63 (1925) 529.
[8] K. Imai, Biochemistry 12 (1973) 798.
[9] F.J.W. Roughton and J. Lyster, Hvalradets Skr. 48 (1965) 185.
[10] U. Schmelzer, R. Steiner, A. Mayer, T. Nedetzka and H. Fasold, European J. Biochem. 25 (1972) 491.
[11] J.M. Salhany, D.H. Mathers and R.S. Eliot, J. Biol. Chem. 247 (1972) 6985.
[12] T. Asakura, Biochem. Biophys. Res. Commun. 48 (1972) 517.
[13] W.H. Huestis and M.A. Raftery, Biochem. Biophys. Res. Commun. 49 (1972) 1358.
[14] K. Imai and T. Yonetani, Biochem. Biophys. Res. Commun. 50 (1973) 1055.
[15] R. MacQuarrie and Q.H. Gibson, J. Biol. Chem. 246 (1971) 5832.
[16] R. MacQuarrie and Q.H. Gibson, J. Biol. Chem. 247 (1972) 5686.
[17] J.J. Hopfield, R.G. Shulman and S. Ogawa, J. Mol. Biol. 61 (1971) 425.
[18] J. Herzfeld and H.E. Stanley, Biochem. Biophys. Res. Commun. 48 (1972) 307.
[19] R.T. Ogata and H.M. McConnell, Biochemistry 11 (1972) 4792.
[20] J.S. Olson and Q.H. Gibson, J. Biol. Chem. 246 (1971) 5241.
[21] M. Brunori, G. Amiconi, E. Antonini, J. Wyman and K.H. Winterhalter, J. Mol. Biol. 49 (1970) 461.
[22] I. Tyuma, K. Shimizu and K. Imai, Biochem. Biophys. Res. Commun. 43 (1971) 423.
[23] E. Antonini, J. Wyman, R. Zito, A. Rossi-Fanelli and A. Caputo, J. Biol. Chem. 236 (1961) PC 60.
[24] J.V. Kilmartin and J.A. Hewitt, Cold Spring Harbor Symp. Quant. Biol. 36 (1972) 311.
[25] J.A. Hewitt, J.V. Kilmartin, L.F. ten Eyck and M.F. Perutz, Proc. Natl. Acad. Sci. US 69 (1972) 203.
[26] R. Banerjee and R. Cassoly, J. Mol. Biol. 42 (1969) 351.
[27] T. Maeda, K. Imai and I. Tyuma, Biochemistry 11 (1972) 3685.
[28] K. Imai and I. Tyuma, Biochem. Biophys. Res. Commun. 51 (1973) 52.
[29] H.F. Bunn and G. Guidotti, J. Biol. Chem. 247 (1972) 2345.

BIOPHYSICAL CHEMISTRY 1 (1973) 134–139. © NORTH-HOLLAND PUBLISHING COMPANY

NUCLEOSIDE CONFORMATIONS. 12. AN INFRARED STUDY OF THE POLYMORPHISM OF GUANINE NUCLEOSIDES IN THE SOLID STATE†

Jean-Pierre LEICKNAM and Claude CHAUVELIER
Département de Recherche et Analyse, Centre d'Etudes Nucléaires de Saclay, F-91190-Gif-sur-Yvette, France

and

Jean François CHANTOT* and Wilhelm GUSCHLBAUER
Service de Biochimie, Département de Biologie, Centre d'Etudes Nucléaires de Saclay, F-91190-Gif-sur-Yvette, France

Received 14 June 1973

IR spectroscopy was used to define the previous observations of Shimanouchi et al. on the polymorphism of guanosine in the solid state. It is shown that form II is guanosine in its crystal form while form I is the helical tetrameric gel. This work was extended to 8-bromoguanosine which also shows polymorphism. It is shown that one of the forms (I) is also that of the helical gel.

1. Introduction

In the field of nucleoside chemistry guanylic compounds are probably unique [2]. The high degree of rotational freedom of the base around the glycosidic linkage makes the guanosine derivative conformation very sensitive to environmental changes and in particular to protonation [3, 4]. Also, only guanosine and its analogues are known to form gels in concentrated aqueous electrolytic solutions [5, 6] and there is now increasing evidence [7–12] which supports the tetrameric arrangement between guanosine residues proposed by Gellert et al. [6].

In solution, infrared spectroscopy would afford a major contribution to the structural investigation of these helical aggregates, if the technique were not strongly restricted by the presence of water in the gel formation. In this respect, Raman spectroscopy appeared to provide a more sensitive method [13] which could give much useful information on the base stacking and the backbone conformation of these pseudo-polymers [14].

Most of the experiments with monomeric guanosine have been carried out in deuterium oxide [15–18] or in Nujol [19, 20]. In the solid state [21], two forms were observed in Nujol for neutral guanosine with quite different spectra in the range 1500–1800 cm^{-1}. Infrared spectroscopy is used here in a first attempt to precise the observations made by Shimanouchi et al. [21] and to study the polymorphism of guanosine and 8-bromo-guanosine in the solid state when included in potassium chloride matrices and to compare its infrared spectra with those of the gels drawn from aqueous solutions.

2. Material and methods

Commercial guanosine (Merck, Darmstadt, Germany) was recrystallized from water to yield guanosine dihydrate. 8-bromo-guanosine dihydrate was prepared as

† Part 11 is ref. [1].
* Boursier de Thèse du C.E.A. (1968–1973). This work is part of the Thesis of J.F. Chantot, which has been presented to l'Université de Paris-Sud (Centre d'Orsay).

previously reported [8]. The gels were formed according to the usual procedure [9] and the fibers drawn as described in a preceding paper [7]. Infrared spectra were recorded with a 225 Perkin–Elmer double beam spectrometer (slit width: 1 cm^{-1}; accuracy in the measurement: 2 cm^{-1}). In the solid state, the samples were examinated either in Nujol mulls or by inclusion in KCl pellets which gave well resolved spectra. Infrared measurements on solid gels were made either from viscous gels which were allowed to dry on nickel grids or CaF_2 plates or from ground solid fibers [7] included in KCl matrices. In solution, spectra were recorded in ^2H_2O and dimethyl sulphoxide (DMSO).

3. Results

3.1. Guanosine

Early in this study, it was observed that depending on the conditions of preparation of the sample, different spectra were obtained for guanosine (Guo). A suspension of Guo dihydrate in Nujol gave the spectrum a of fig. 1 which was identical to those of the films of sample deposited on grids or CaF_2 plates. Also the spectra did not vary markedly when these mulls or films were subjected to large temperature changes from −150 to 25°C. In KCl pellets, the most interesting feature is that the spectral response was strongly dependent on the thermal treatment of the sample. In fig. 1 is shown the spectrum c at room temperature of Guo dihydrate in a KCl pellet without previous heating (form II). If the same pellet was first heated for a few minutes at 120–130°C, then spectrum d of fig. 1 was obtained (form I) with an enlarged absorption intensity around 1695 cm^{-1} and disappearance of the peaks at 1635 and 1735 cm^{-1}. On the other hand, if Guo dihydrate was heated to 120–130°C before making the KCl pellet, form II (spectrum c) was observed again, even if the disk was subjected to further heating. In return, a sample which had yielded form I could be readily reconverted to form II by dissolving in water, drying in air and then making a new pellet. This excluded any thermal degradation. The conditions are summarized in the legend of fig. 1.

If gels of Guo were allowed to dry slowly on grids at 5°C, no precipitation occurred. Samples prepared in this way also yielded spectra identical to form I (spectrum d).

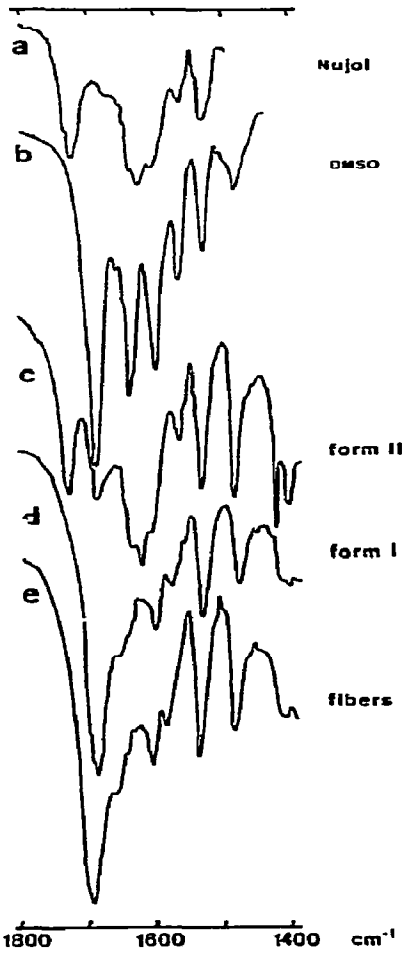


Fig. 1. Infrared spectra of Guo at room temperature. Conditions: (a), Nujol mull of Guo; (b), DMSO solution of Guo; (c), crystalline Guo in KCl disk; (d), Guo in KCl disk. The pellet was heated to 120–130°C for a few minutes before taking the spectrum; (e), fibers drawn from Guo gel solutions included in a KCl pellet.

In a further attempt to clarify this point, fibers of Guo were drawn from aqueous gels as for X-rays fiber diffraction [7]. The fibers were ground with KCl to make a pellet (spectrum e). Again, form I was obtained. On some occasions mixed forms I and II were observed. Since Guo gels sometimes give rise to local precipitations, this result is not surprising.

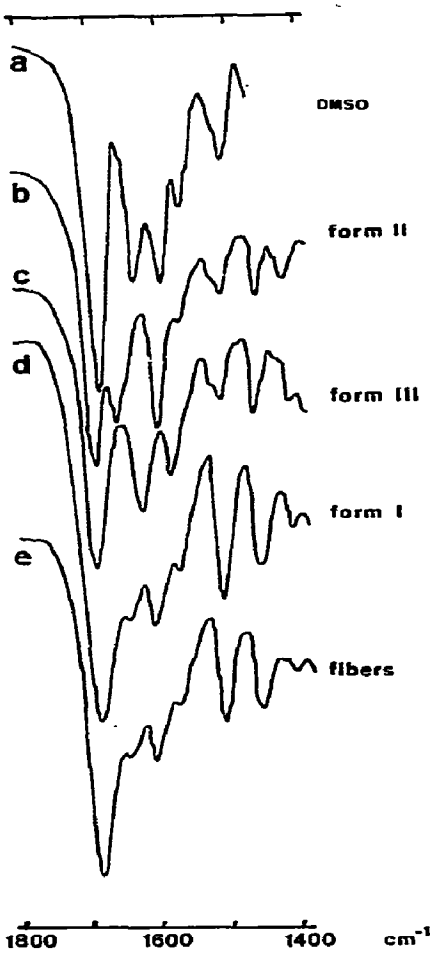


Fig. 2. Infrared spectra of 8BrGuo at room temperature. Conditions: (a), DMSO solution of 8BrGuo; (b), crystalline 8BrGuo in KCl disk; (c), 8BrGuo in KCl disk. The sample was heated to 120–130°C a few moments before making the pellet; (d), 8BrGuo in KCl disk. The pellet was first heated for a few minutes before taking the spectrum; (e), fibers drawn from 8BrGuo gel solutions included in a KCl pellet.

3.2. 8-bromo-guanosine

In fig. 2, the spectral results on 8-bromo-guanosine (8BrGuo) are presented. Three forms can be clearly distinguished. Form II (spectrum b) was obtained from KCl pellets of the crystalline product. If the pellets were heated to 120°C, then a transformation into form I (spectrum d) took place. But in contrast with Guo, a third form (form III, spectrum c) was observed when 8BrGuo hydrate was heated a few moments to 120°C before making the pellet. The disk, in turn, did not give rise to noticeable spectral changes on further heating. Conversely, pellets which gave form I or III could be redissolved in water and dried in air. New disks prepared from these powders gave spectra characteristic of form II. Visibly, no thermal degradation was involved in these transformations. As for Guo, gels of 8BrGuo deposited on grids or dried gel fibers pelleted with KCl yielded spectrum e, characteristic of form I. It can be noted that a mixture of KCl and guanine nucleosides in water, dried to 120°C and then pelleted also gave form I.

Gels formed in ^2H_2O in the presence of KCl (fig. 3) showed very similar shape to those of the dried fibers (fig. 2). Increasing the temperature, a gradual shift of the bands at 1690 cm^{-1}, ν(C=O), to 1670 cm^{-1} and of the ring vibration at 1592 cm^{-1} to 1580 cm^{-1} was observed in analogy to these observations on guanosine-3′ and 5′-phosphate [15].

4. Discussion

Shimanouchi et al. [21] had already mentioned the existence of two forms for neutral Guo in Nujol mulls which gave two very different infrared absorption spectra without studying this dimorphism in detail. Unfortunately, no details about the obtainment of the different forms were published which would allow the comparison with our results. The two sets of data may be correlated, however, by a close resemblance in the reported spectra.

If one compares, in the region 1500–1800 cm^{-1}, where the ν(C=O) and $\delta(NH_2)$ bands are located, the spectrum of 8BrGuo (Guo is very similar) in DMSO with that in ^2H_2O (figs. 2 and 3), the disappearance of the 1635 cm^{-1} band in DMSO by deuteration is particularly noteworthy. This band corresponds therefore to the angular deformation vibration of the NH_2 group.

The band at 1688 cm^{-1} in DMSO (1692 cm^{-1} for Guo) is displaced to 1680 cm^{-1} in KCl-containing ^2H_2O (gel). It is probably due to the valence vibration of the C=O group. The displacement results from the variation of the dielectric constant and from the formation of hydrogen bonds with the carbonyl group (tetramer formation) which do not exist in DMSO.

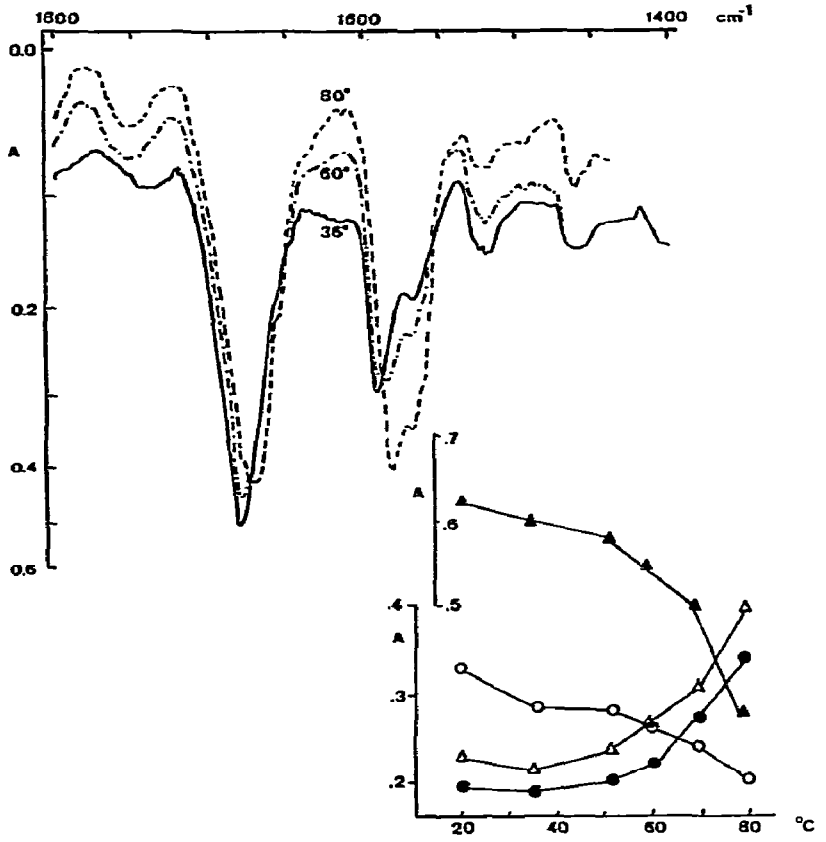


Fig. 3. Infrared spectra in ^2H_2O solution of 8BrGuo (0.1M; pD 7; 50μ path length; 0.1M KCl) as a function of the temperature. Insert: melting curves of 8BrGuo. ▲: 1680 cm^{-1}; △: 1580 cm^{-1}; ○: 1592 cm^{-1}; ●: 1518 cm^{-1}. The lack of a plateau at higher temperatures results from incomplete melting due to the high thermal stability of this gel [8].

The bands below 1600 cm^{-1} can be attributed to in-plane ring skeleton vibrations of the purine rings and show very little displacement if the solvent is changed.

In case of Guo or 8BrGuo gel formation (figs. 1 and 2, spectra d and e) the absorbance at 1690 cm^{-1} is greatly increased (compared with crystalline form or in DMSO). It appears reasonable to suppose that the vibration ν(C=O) and $\delta(NH_2)$ are superimposed. This implies that the wavenumber of the latter vibration increased, if hydrogen bonds are formed. This is represented in the heating curves (fig. 3) where the wavenumber of the band at 1680 cm^{-1} decreased by 10 cm^{-1} between 20 and 80°C. Similar observations were made by Miles and Frazier [15] for the melting of guanylic acid gels. The ring vibration at 1592 cm^{-1} at 20°C is displaced and split into two bands indicating the liberation of the purine rings from the tetramer stack at higher temperatures.

Comparison of the solid state spectra of Guo and 8BrGuo (figs. 1 and 2) shows very significant differences on the ν(C=O) level (band at 1735 cm^{-1} in Guo, absent in 8BrGuo). This indicates that less (or weaker) hydrogen bonds are formed in the case of Guo than in 8BrGuo crystals. Inspection of the published crystal structures of these two compounds [22, 23], which are considerably different in many respects, confirm this idea. While in 8BrGuo [22] the carbonyl group forms two hydrogen bonds (with NH of its neighbour

and with a H_2O molecule), mainly one hydrogen bond (with the amino group of its neighbour) seems to exist in Guo, while one of the hydrogen atoms of one water molecule is shared by two carbonyl oxygen atoms [23], i.e., only one carbonyl of every two has a water hydrogen bond.

In DMSO the two compounds show great similarity, indicating that the suppression of the hydrogen bonds in the crystal renders the C=O group quite similar in both cases. On the other hand, the 1530 cm^{-1} band in Guo (fig. 1) is displaced to 1512 cm^{-1} in 8BrGuo (fig. 2). This substitution effect makes it probable that we deal with a ring deformation vibration of the purines. A similar displacement is observed in the spectra of form I of Guo and 8BrGuo which is the main difference between the two compounds.

As stated above, the spectra observed in KCl pellets of Guo and 8BrGuo in form I and in dried gels or fibers are quite similar (see figs. 1 and 2, spectra d and e). The following additional points can be made:

– Gels of 8BrGuo prepared in ^2H_2O (0.1M KCl) at 20°C gave spectra characteristic of form I taking into account the isotopic effect of deuteration (fig. 3). Increase of the temperature from 20° to 80°C (at which monomers are mostly present) produced profound changes and the high temperature spectra showed essentially the characteristics of form II.

– In a previous study [9], it had been observed that certain Guo analogues only formed gels. While 2′, 3′ substituted Guo derivatives easily led to aggregation, this structuration did not take place in 5′-O-acetylated Guo. 5′-O-acetyl-Guo and 2′,3′,5′-tri-O-acetyl-Guo were therefore subjected to the same treatment as Guo. Despite repeated attempts, only one spectral form was observed, similar to form II (considering the effects of acetylation). Deoxyguanosine has been found to aggregate in concentrated aqueous solutions. This compound showed a clear shift towards the spectral form I for the helical structure. Heating of this nucleoside, however, produced irreversible changes, probably due to partial pyrolysis.

– Although circular dichroism is not well indicated for studies in the solid state, we have performed measurements on very carefully prepared pellets in order to reduce the heterogeneity of the system. Compared with form II, the signal for nucleoside pellets in the form I is notably increased, indicating ordered asymmetric structuration.

– Debye–Scherrer X-ray diagrams have been taken from various powders drawn from ground pellets corresponding to the different forms of the two Guo derivatives. In form I, unit repeat at 10.5 Å for Guo and at 7.2 Å for 8BrGuo was observed which cannot belong to the corresponding hydrates in form II. Comparison of this result with those found by X-ray fiber diffraction [7] shows that the above periodicities can be roughly related to the distances between the two nearest molecules of Guo or 8BrGuo in two adjacent tetramers.

It appears therefore reasonable to assign form I of Shimanouchi et al. [21] to the helical tetrameric form of Guo and 8BrGuo while form II appears its crystalline form.

After terminating this work, a paper of Howard et al. [18] came to our attention on the aggregation of 8-bromo-guanosine-5′-phosphate and its interaction with poly(C). While their conclusions on the band assignment resemble in many respects those observed here, these authors seem to be unaware of the four papers published on many Guo analogues in the last four years [8–11], among that the gel formation of 8BrGuo [8].

Acknowledgement

Thanks are due to Mme P. Charpin from the Laboratory of Crystallography who has performed Debye–Scherrer X-ray diagrams and to Mr. J.M. Delabar for helpful discussions.

References

[1] J.M. Delabar and W. Guschlbauer, J. Am. Chem. Soc. 95 (1973) 5729.
[2] W. Guschlbauer, Jerusalem Symposium 4 (1972) 297.
[3] W. Guschlbauer and Y. Courtois, FEBS Letters 1 (1968) 183.
[4] Tran-Dinh Son, W. Guschlbauer and M. Guéron, J. Am. Chem. Soc. 94 (1972) 7903.
[5] I. Bang, Biochem. Z. 26 (1910) 293.
[6] M. Gellert, M.N. Lipsett and D.R. Davies, Proc. Natl. Acad. Sci. US 48 (1962) 2013.
[7] P. Tougard, J.F. Chantot and W. Guschlbauer, Biochem. Biophys. Acta 308 (1973) 9.
[8] J.F. Chantot and W. Guschlbauer, FEBS Letters 4 (1969) 173.
[9] J.F. Chantot, M.T. Sarocchi and W. Guschlbauer, Biochimie 53 (1971) 347.
[10] W. Guschlbauer, J.F. Chantot, M.T. Sarocchi and D. Thiele, Studia Biophys. 24/25 (1970) 29.

[11] J.F. Chantot and W. Guschlbauer, Jerusalem Symposium 4 (1972) 205.
[12] J.F. Chantot, Arch. Biochem. Biophys. 153 (1972) 347.
[13] R.C. Lord and G.J. Thomas, Spectrochim. Acta 23A (1967) 2551.
[14] E. Small and W.L. Peticolas, Biopolymers 10 (1971) 1377.
[15] H.T. Miles and J. Frazier, Biochim. Biophys. Acta 79 (1964) 216.
[16] F.B. Howard and H.T. Miles, J. Biol. Chem. 240 (1965) 801.
[17] H.T. Miles and J. Frazier, Biochem. Biophys. Res. Commun. 49 (1972) 199.
[18] F.B. Howard, J. Frazier and H.T. Miles, J. Biol. Chem. 247 (1972) 6733.
[19] C.L. Angell, J. Chem. Soc. (1961) 504.
[20] A.T. Tu and C. Friederich, Biochemistry 7 (1968) 4367.
[21] T. Shimanouchi, M. Tsuboi and Y. Kyogoku, Advan. Chem. Phys. (1965) 435.
[22] S.S. Tavale and H.M. Sobell, J. Mol. Biol. 48 (1970) 109.
[23] U. Thewalt, C.E. Bugg and R.E. Marsh, Acta Cryst. B26 (1970) 1086.