Ultracentrifuge Study of the Self-Association of Fraction II of Procarboxypeptidase A*

David C. Teller

ABSTRACT: Sedimentation equilibrium experiments on fraction II of procarboxypeptidase A reveal that this protein self-associates in a buffer system consisting of 0.038 M KH_2PO_4–0.2 M KCl (pH 7.5) at 5°. Because the molecular weight averages do not correspond exactly to indefinite self-association, equations have been derived to predict the behavior of molecular weight averages when chain termination occurs in a self-associating system with identical equilibrium constants.

It has been found that chain termination and thermodynamic nonideality in indefinite self-associating systems have quite similar effects on molecular weight averages for such systems. It has been demonstrated that such a chain termination model does not fit the fraction II data. In order to investigate whether the observed deviations could be explained by the existence of a heterogeneous population of monomers which self-associate indefinitely, the equations for the ultracentrifuge behavior of the molecular weight averages have been derived for this model. Since this model predicts trends in the data which are not observed, it can also be excluded. Consequently, it is concluded that fraction II is an indefinitely self-associating system with a molecular weight of the monomeric species of about 22.9×10^3 g/mole, but the preparation is contaminated with a small amount of material which does not participate in the chemical equilibria.

Many proteins and enzymes undergo indefinite self-association. Included in this category are tobacco mosaic virus protein (Smith and Lauffer, 1967), catalase (Kiselev *et al.*, 1967), phosphorylase (Chignell *et al.*, 1968), and glutamine synthetase (Valentine *et al.*, 1968). The reaction mechanism appears to be dependent upon either of two subclasses of association: symmetrical monomers[1] with a single type of binding site as in the case of glutamine synthetase (Valentine *et al.*, 1968), or asymmetric monomers with a donor and acceptor site. In general the interaction of monomeric units is weak; proton fluctuation (Timasheff, 1966) among identical, symmetrical monomers may be sufficient to give rise to this type of self-association. The assembly of asymmetric structures into ordered arrays has been discussed by Caspar, Klug, and coworkers (Caspar and Klug, 1962; Caspar, 1966; Klug *et al.*, 1966; Kiselev *et al.*, 1968). Although two binding sites are sufficient for well-ordered structures, three or more sites would confer stronger binding in the formation of cylinders and icosahedra (Caspar and Klug, 1962). If the acceptor and donor sites of the two-site model occur on opposite sides of the molecule then a linear string is expected for the polymer; however, if the acceptor and donor sites have an angle deviating from 180° then a helix or ring will be formed (Khalil and Lauffer, 1967; Kiselev *et al.*, 1967). Formation of a helix does not differ chemically from a linear string unless subsidiary sites exist on the monomers; however, formation of a ring structure results in chain termination.

In this report, concerning the indefinite self-association of fraction II of procarboxypeptidase A, the equations for molecular weight behavior due to chain termination at the *n*-mer stage are presented as well as the computational method we have used to exclude this model for the self-association of fraction II. It has been found that it is very difficult to distinguish between chain termination and thermodynamic nonideality of the self-associating units. A second possible hypothesis for the deviations from ideal behavior is heterogeneity of the monomeric units participating in the chemical equilibria. These equations have been derived for the ultracentrifuge, and it can be shown that heterogeneity would result in a characteristic behavior which is readily distinguishable from chain termination and nonideality. Consequently it is concluded that fraction II is an indefinitely self-associating system with a molecular weight of the monomeric species of about 22,900 g/mole, but contaminated with a small amount of material not participating in the chemical equilibria.

Theory

As discussed by other authors (Van Holde and Rosetti, 1967; Elias and Bareiss, 1967; Adams and Lewis, 1968) the model for self-association of an indefinite type with equal free energy of association of a monomer unit to the noncovalent polymer can be represented chemically by the scheme

$$2F_1 \rightleftharpoons F_2 \qquad k = (F_2)/(F_1)^2$$

$$F_1 + F_2 \rightleftharpoons F_3 \qquad k = (F_3)/(F_2)(F_1)$$

$$F_1 + F_3 \rightleftharpoons F_4 \qquad k = (F_4)/(F_3)(F_1)$$

$$\cdots \qquad \cdots$$

$$F_{j-1} + F_1 \rightleftharpoons F_j \qquad k = (F_j)/(F_{j-1})(F_1)$$

* From the Department of Biochemistry, University of Washington, Seattle, Washington 98105. *Received July 28, 1969.* This research was supported by a U. S. Public Health Service Research grant from the National Institute of General Medical Sciences (GM 13401).

[1] The term monomer is used in this report to indicate the smallest unit participating in the chemical equilibria.

where (F_i) represent the activities of the various species. Provided that the activity coefficients γ_i, are given by $\gamma_i = \exp(iBM_1C)$ where B is the second virial coefficient, M_1 the monomer molecular weight, and C the total concentration in grams per liter (Van Holde and Rosetti, 1967), then the concentrations of the individual species are given by

$$C_2 = 2\left(\frac{k}{M_1}\right)C_1^2, \qquad C_3 = 3\left(\frac{k}{M_1}\right)^2 C_1^3,$$
$$\ldots, C_j = j\left(\frac{k}{M_1}\right)^{j-1} C_1^j, \ldots$$

This is the mechanism of reaction which will be assumed throughout most of this report. The further definition of the equilibrium constant on a grams-per-liter scale simplifies the presentation, $K = k/M_1$.

Ideal Self-Association of a Homogeneous Monomer. The total concentration of a system undergoing indefinite self-association with monomer molecular weight, M_1, is given by

$$C = C_1(1 + 2KC_1 + 3(KC_1)^2 + \ldots + n(KC_1)^{n-1} + \ldots) = C_1/(1 - KC_1)^2 \quad (1)$$

This result has been shown by many authors including Elias and Lys (1966), Van Holde and Rosetti (1967), and Adams and Lewis (1968). Several authors have commented on the constraint of KC_1 less than unity. This condition is a requirement of the second equality of eq 1. When KC_1 is greater than 1, very large macromolecular species will be formed until the solubility product of the aggregate is exceeded. The precipitate need not be amorphous, but may consist of tubules or long strings. An example of such a specific aggregation process is the crystallization of glutamine synthetase from Mn^{2+} at low ionic strength (Valentine *et al.*, 1968). Caspar (1966) has also discussed this point.

The number-average molecular weight at each concentration point is given by

$$M_n = M_1/(1 - KC_1) \quad (2)^2$$

as shown by Flory (1953), Elias and Lys (1966), and Van Holde and Rosetti (1967). If eq 2 is squared and M_1 times eq 2 subtracted from the result, then KM_1^2 times the second equality of eq 1 may be recognized. Hence, we may write

$$M_n = M_1 + KM_1^2C/M_n \quad (3)$$

This formula is exact for an ideal indefinite association. It is not necessary to make the approximations used by Elias and Lys (1966) as later pointed out by Elias and Bariess (1967). A plot of M_n *vs.* C/M_n is linear for such a reaction mechanism.

The weight-average molecular weight of such an indefinitely associating system is given by

$$M_w = M_1(1 + KC_1)/(1 - KC_1) \quad (4)$$

from which it has been shown (Smith and Lauffer, 1967; Van Holde and Rosetti, 1967; Elias and Bareiss, 1967),

$$M_w^2 = M_1^2 + 4KM_1^2C \quad (5)$$

This result shows that M_w^2 is linear function of concentration for such a "stacking" reaction.

Multiplication of eq 2 by two and subtraction of eq 4 produces the result

$$2M_n - M_w = M_1 \quad (6)$$

(Flory, 1953). This relation is significant for three reasons. First, it is an invariant function of concentration; second, it is sensitive to chain termination, nonideality, and heterogeneous monomers in predictable ways; and third, our computer programs from high-speed sedimentation equilibrium experiments plot $2M_n - M_w$ *vs.* C so that the constancy of this function is easy to recognize.

Combination of eq 6 and 3 gives two linear relations between the weight- and number-average molecular weights

$$M_w = M_1 + 2KM_1^2C/M_n \quad (7)$$

and the rearranged form

$$M_wM_n/C = M_1M_n/C + 2KM_1^2 \quad (8)$$

For a perfectly behaved, ideal system eq 7 and 8 would be equivalent. But, thermodynamic nonideality, random error, and other effects are incorporated in the data. Equation 7 distributes the points at large concentration widely while eq 8 does the reverse, emphasizing the points at low concentrations.

The Z-average molecular weight will be given by,

$$M_Z = M_1\frac{1 + 4KC_1 + (KC_1)^2}{(1 - KC_1)(1 + KC_1)} \quad (9)$$

(Adams and Lewis, 1968). This equation may be combined with the expressions for M_w and C to generate three relations which are linear. The first of these is

$$M_ZM_w/C = M_w^2/C + 2KM_1^2 \quad (10)$$

This equation predicts a unit slope and nonzero intercept for an indefinite self-association. In the analytical ultracentrifuge,

$$M_ZM_w - M_w^2 = \frac{2RT}{(1 - \bar{v}\rho)\omega^2}\frac{dM_w}{d(r^2)}$$

where R is the gas constant, T is the absolute temperature, $\bar{v}$ is the partial specific volume, ρ is the solution density, ω is the angular velocity, and r is the distance from the axis of rotation (Lansing and Kraemer, 1935; Goldberg, 1953). Thus, $2KM_1^2C$ characterizes the rate of change of the weight-average molecular weight with the square of the radial distance and is the variance of the weight distribution. Equation 10, like eq 9, has the property of representing low concentrations at high abcissa and ordinate values.

[2] Molecular weight averages used in this report are those values which are determined at each concentration. They are not mass averages (Adams, 1964).

By combining eq 5 and 10 we obtain,

$$M_Z M_w = M_1^2 + 6KM_1^2C \quad (11)$$

Multiplication of this equation by two and eq 5 by three followed by subtraction of the two expressions produces

$$3M_w - 2M_Z = M_1^2/M_w \quad (12)$$

This relation should be useful for the determination of the monomer molecular weight from low-speed sedimentation equilibrium experiments. In such experiments M_Z is readily calculated from the schlieren data while M_w at each point can be calculated by standard methods (Richards and Schachman, 1959; Richards *et al.*, 1968).

Chain Termination. Depending on the angle between aggregation sites, strings, helices, or rings may be formed. Since ring formation would constitute a chain termination at the n-mer stage with an additional binding constant, it is of interest to calculate the behavior of the molecular weight moments under chain termination conditions. The concentration will be given by

$$C = C_1[1 + 2(KC_1) + 3(KC_1)^2 + \ldots + n(KC_1)^{n-1}]$$

which is a closed series and not restricted to KC_1 less than unity. If we treat this series as an infinite series we obtain

$$C = \frac{C_1}{(1 - KC_1)^2} - C_1n(KC_1)^n[1 + KC_1 + (KC_1)^2 + \ldots] - C_1(KC_1)^n[1 + 2KC_1 + 3(KC_1)^2 + \ldots]$$

or, condensing the two negative infinite series yields

$$C = \frac{C_1}{(1 - KC_1)^2}\{1 - (KC_1)^n[1 + n(1 - KC_1)]\} \quad (13)$$

This equation has a leading term which is the same as previously observed for the indefinite association but also a negative term expressing chain termination.

The weight-average molecular weight of a ideal self-associating system is given by $M_w = M_1 \mathrm{d} \ln C/\mathrm{d} \ln C_1$ (Steiner, 1952). Thus the weight-average molecular weight may be derived from eq 13. The result is

$$\frac{M_w}{M_1} = \frac{1 + KC_1}{1 - KC_1} - \frac{n(n + 1) - (1 + KC_1)(KC_1)^n}{1 - (KC_1)^n[1 + n(1 - KC_1)]} \quad (14)$$

Once again, the first term describes the indefinite association while the second gives the chain termination expression.

The number-average molecular weight of a self-associating system is given by $M_n = M_1C/\int C \mathrm{d} \ln C_1$ (Steiner, 1954). For this thermodynamic mechanism we obtain

$$\frac{M_n}{M_1} = \frac{1}{1 - KC_1} - \frac{n(KC_1)^n}{(1 - (KC_1)^n]} \quad (15)$$

While these relations are rather complicated, an interesting set of equations exists at $KC_1 = 1$. At this point, the concentration is, $C = n(n + 1)/2K$ and the molecular weight moments are given by the equations,

$$M_n = M_1\frac{n + 1}{2}, \qquad M_w = \frac{M_1}{3}(2n + 1)$$

and

$$M_Z = \frac{3M_1}{2}\frac{n(n + 1)}{2n + 1}$$

These formulas arise from the well-known relations for sums of integers, sums of integers squared, etc. The equations for C, M_n, and M_w can also be obtained from eq 13 to 15 by application of l'Hospital's rule, which serves as a check of the correctness of eq 13–15.

In principle it is possible to treat KC_1 and $(KC_1)^n$ as unknowns and eliminate them by using eq 13, 14, and 15; however, the resulting relations are probably rather complex. It would appear that it is simpler to use an iteration scheme.

In order to calculate the best K and M_1 in the work reported below, we have used the iteration method outlined in Table I. In this flow chart the arrows indicate repeated operations. The principle of the iteration is variation of K at each M_1

TABLE I: Iteration of Best Monomer Molecular Weight and Equilibrium Constant for Various Termination Models.

Step No.	Operation
1	Assume polymer size, n, and assume $M_1 = 19 \times 10^3$ g/mole.
2	Define $K_0 = -0.1$ l./g and K increment, $\Delta K = 0.1$ l./g.
3	Define $K_\alpha = K_0 + \alpha \cdot \Delta K$, $\alpha = 1, 2, \ldots .20$.
4	From M_1 and K_α compute a table of $\tilde{M}_{x,i}$, $\tilde{C}_i$ where $x = n$, w, and z; $i = 1$–25 at increments in C_1 of 0.10 g/l. from $C_1 = 0$ to 2.5 g/l.
5	From observed concentration, C_j, find the closest value of $\tilde{C}_i$ in the table produced in step 4. Interpolate the predicted value of the molecular weight moment, $\tilde{M}_{x,j}$, by quadratic interpolation using C_j as the independent variable. Compute the average root-mean-square deviation of molecular weights, E_α $E_\alpha = \frac{1}{3}\sum_x \left\{\frac{1}{N}\sum_{j=1}^{N}(M_{x,j} - \tilde{M}_{x,j})^2\right\}^{1/2}$ where N is the total number of points.
7	Find the minimum of the 20 values of E_α. If this value is denoted by E_α (min), then replace K_0 by $K_{\alpha+1}$ for further iteration.
8	Replace ΔK by $\Delta K/10$ for next iteration.
9	Write the best K, $\pm M_n$, $\pm M_w$, $\pm M_Z$, E_α(min), where $\pm M_x$ are root-mean-square deviations.
10	Increment M_1 by 1×10^3 g/mole.
11	From E_α(min) *vs.* M_1 interpolate the best M_1, K, $\pm M_n$, $\pm M_w$, $\pm M_Z$.

until the best fit to the observed molecular weight moments is obtained. For calculations on fraction II of procarboxypeptidase A, M_1 was varied from 19×10^3 to 25×10^3 g per mole and the value of K was computed to three decimal digits. The values of $\tilde{C}_i$, $\tilde{M}_{n,i}$ and $\tilde{M}_{w,i}$ of Table I were calculated using eq 13–15. In order to obtain $\tilde{M}_{Z,i}$ the equation of Wales (1948) was employed, $\tilde{M}_{Z,i} = \tilde{M}_{w,i} + \tilde{C}_i(d\tilde{M}_w/d\tilde{C})_i$. The values of $(d\tilde{M}_w/d\tilde{C})_i$ were calculated from M_w by a first derivative, quadratic interpolation formula. The computation was time consuming since a very large number of possibilities were tested.

Nonideal Equations. While the equations for ideal systems are particularly simple, those which include the effect of activity coefficients on observed molecular weights are very complex (Van Holde and Rosetti, 1967; Adams and Lewis, 1968). Consistent with the previous assumption that $BM_i = iBM_1$ we may write

$$M_{n,a} = \frac{M_n}{1 + {}^1/_2BM_nC}$$

$$M_{w,a} = \frac{M_w}{1 + BM_wC}$$

and, since $M_Z = M_w(1 + d \ln M_w/d \ln C)$ (Wales, 1948), it may be shown that

$$M_{Z,a} = \frac{M_Z}{(1 + BM_wC)^2}$$

That is, all of the linear relations derived for ideal systems will show curvature. For example, eq 12 becomes

$$3M_{w,a} - 2M_{Z,a} = M_1^2\left(\frac{1}{M_{w,a}}\right) - (3M_{w,a}^2 + 2M_{Z,a}M_w + M_1^2)BC - (3M_{w,a}^3 + M_{Z,a}M_w^2)B^2C^2 - 3M_{w,a}^4B^3C^3 - \ldots . \quad (16)$$

This equation describes a zero intercept and initial slope of M_1^2 but will rapidly curve downward for a nonideal system, necessitating experiments at low concentrations in order to determine monomer molecular weights according to the method described for eq 12.

The best relation indicative of indefinite association with identical equilibrium constants is $2M_{n,a} - M_{w,a}$. The concentration dependence of this function is

$$2M_{n,a} - M_{w,a} = M_1 + {}^1/_2BM_1^2(4K - BM_1)C^2 + BM_1^2[K^2 - 7/2BM_1K + 241/256(BM_1)^2]C^3 + \ldots \quad (17)$$

As long as BM_1 is small $2M_{n,a} - M_{w,a}$ will be approximately constant; but for large virial terms it will be difficult to recognize a "stacking" mechanism. Nevertheless, the relation is important because it predicts the $2M_{n,a}$ *vs.* C graph should be initially flat and show upward parabolic behavior at intermediate concentrations ($4K > BM_1$ is assumed). Calculations made in this laboratory on the data published by Van Holde and Rosetti (1967) show that this upward quadratic relation is followed rather well. However, chain termination will result in quite similar behavior.

Monomeric Heterogeneity. Elias and Bareiss (1967) have shown that the weight concentration of material in an indefinite association among heterogeneous monomers is given by the equation

$$C = C_I/(1 - kC_I/M_{I,n})^2 \quad (18)$$

where C_I is the concentration of monomers and $M_{I,n}$ is the number-average molecular weight of the monomers. In the ultracentrifuge at sedimentation equilibrium $M_{I,n}$ and C_I will be functions of radial distance

$$\frac{dC_I}{d(r^2)} = C_IAM_{I,w,r} \quad (19A)$$

$$\frac{C_I}{M_{I,n,r}} - \frac{C_{I,m}}{M_{I,n,m}} = A\int_{r_m^2}^{r^2} C_I d(r^2) \quad (19B)$$

where the r subscript denotes that the variable is a function of distance from the center of rotation and the m subscript denotes that the variable is evaluated at the centripetal meniscus. The quantity A is $(1 - \bar{v}\rho)\omega^2/2RT$, where $\bar{v}$, the partial specific volume, is assumed identical for all species present in the solution $M_{I,w,r}$ denotes the weight-average molecular weight of the monomers at position r.

At any position the number-average molecular weight may be obtained from the sum of the molar concentrations (Elias and Bareiss, 1967)

$$M_{n,r} = M_{I,n,r}/(1 - kC_I/M_{I,n,r}) \quad (20)$$

From eq 18 and 20 it may easily be shown that

$$M_{n,r} = M_{I,n,r} + kM_{I,n,r}C/M_{n,r} \quad (21)$$

Upward curvature of a graph of $M_{n,r}$ *vs.* $C/M_{n,r}$ would be expected since both the slope and intercept will increase with increasing distance for heterogeneous monomer.

In order to obtain the weight-average molecular weight at each position we differentiate eq 18 with respect to r^2 and divide by eq 18

$$\frac{1}{C}\frac{dC}{d(r^2)} = AM_{w,r} = \frac{1}{C_I}\frac{dC_I}{d(r^2)} + \frac{2k}{1 - \frac{kC_I}{M_{I,n,r}}}\left.d\left(\frac{C_I}{M_{I,n,r}}\right)\right/d(r^2)$$

Equations 19A and 19B may now be employed to show that

$$M_{w,r} = M_{I,w,r} + \frac{2kC_I}{1 - \frac{kC_I}{M_{I,n,r}}}$$

Substitution of eq 18 and 20 into the second term of this expression yields

$$M_{w,r} = M_{I,w,r} + 2kM_{I,n,r}C/M_{n,r} \quad (22)$$

or, by rearrangement

$$M_{n,r}M_{w,r}/C = M_{I,w,r}M_{n,r}/C + 2kM_{I,n,r} \quad (23)$$

These two equations have the complimentary property of emphasizing different regions of the concentration distribution in the centrifuge cell as discussed for eq 7 and 8.

Multiplication of eq 21 by two and subtraction of eq 22 gives the result

$$2M_{n,r} - M_{w,r} = 2M_{In,r} - M_{I,w,r} \quad (24)$$

Since $M_{I,w,r} > M_{I,n,r}$ and $M_{I,w,r} < M_{I,w,r+\Delta r}$, we expect that a graph of $2M_{n,r} - M_{w,r}$ *vs.* r or C will show downward curvature if the monomer is heterogeneous. This is important since chain termination and nonideality produce upward curvature of $2M_{n,r} - M_{w,r}$ when plotted against concentration.

For a system with homogeneous monomers, $M_{w,r}^2$ is a linear function of concentration. This is also true for a heterogeneous monomer when measured by light scattering (Elias and Bareiss, 1967). However, in the ultracentrifuge this is no longer true. This may be shown by squaring eq 22 and combining the result with eq 21 to eliminate one of the $kC/M_{I,n,r}$ terms. This yields

$$M_w^2 = M_{I,w,r} + 4kCM_{I,n,r}\left[\frac{M_{I,w,r} - M_{I,n,r}}{M_{n,r}} + 1\right] \quad (25)$$

From this equation it is clear that as long as the monomers are homogeneous, the first term in brackets is zero which reduces to eq 5. However, for a heterogeneous monomer $M_{w,r}^2$ is not a linear function of concentration.

Other molecular weight relations may be derived for this model but increasing complexity is encountered for each average.

It should be noted that the equations of this section have been derived for the ultracentrifuge molecular weight moments determined at various values of r. They differ from the results of Elias and Bareiss (1967) due to the redistribution of monomers in the centrifugal field.

Materials and Methods

Fraction II of procarboxypeptidase A was purified by Dr. W. D. Behnke (Behnke and Neurath, 1970) who also carried out the experiment reported here. The high-speed sedimentation equilibrium experiments were performed in a buffer system consisting of 0.038 M KH_2PO_4–0.2 M KCl (pH 7.5) at 5°.

The Model E ultracentrifuge was focused at the two-thirds plane in the cell (Svensson, 1954, 1956) to ±0.01 in. and all lenses and components were centered about the optic axis (Richards *et al.*, unpublished data). Photographic plates (Kodak II-G) were read on a modified Nikon microcomparator (Teller, 1967). These data were processed by FORTRAN computer programs written in this laboratory for the IBM 7040–7094 IBSYS system according to the methods described by Teller *et al.* (1970). Most secondary analyses of the point-by-point molecular weight data were performed on an Olivetti Programma 101 desk top computer.

Results and Discussion

The most striking characteristic of an indefinite association with equal constants for association is the constancy of $2M_n - M_w$ as a function of concentration (eq 6). This is illustrated in Figure 1 for fraction II. In this figure it may be seen that M_w and M_n are nearly linear functions of C while $2M_n - M_w$ remains approximately constant.

The evidence that this is a chemical equilibrium interaction has been given by Behnke *et al.* (1970, Figures 3–5). Fraction II exhibits a single band on disc gel electrophoresis and has a sedimentation constant which increases with increasing concentration. In sedimentation velocity studies (Behnke *et al.*, 1970) the material displays a single, almost symmetrical boundary. This behavior, together with the sedimentation equilibrium results indicate an indefinite self-association (Gilbert, 1959). In the absence of this other evidence, such data as shown in Figure 1, only imply the random synthesis or random degradation of an unbonded polymer (Tanford, 1961).

Figure 2 illustrates the "molecular space" of the fraction II system. In this figure the M_Z *vs.* $1/M_w$ values exceed the space which would be occupied by a monomer–dimer–trimer–tetramer equilibrium (Yphantis and Roark, 1969; Teller *et al.*, 1970). The values of $2M_w - M_Z$ *vs.* $2/M_n - 1/M_w$ cluster about the monomer molecular weight in this graph and, since these numbers are essentially independent of virial terms, this behavior would also be a characteristic of a nonideal indefinite self-association. The lines drawn through the data points of Figures 1 and 2 are those predicted from M_1 and K calculated by the iterative procedure for an indefinite self-association as described in the theoretical section of this paper.[3]

Graphs of eq 3, 5, 7, 8, 10, and 11 are shown in Figure 3. Figure 3A demonstrates that M_n is a reasonably linear function of C/M_n but that the data from the two channels are different. Figure 3B presents M_w^2 as a function of concentration. This graph shows the same characteristics as Figure 3A as well as demonstrating slight curvature in the region of 1.4–2.8 fringe, a characteristic which is repeated in Figure 3C. Comparison of Figure 3C,D demonstrates the complimentary nature of eq 7 and 8. Abcissa values between 0.79 and 5.50 are the data obtained below 2.8 fringe. These same data occur below 0.09 fringe mole/g in Figure 3C. This figure (3C, eq 7) emphasizes the high concentration data while Figure 3D (eq 8) emphasizes the low concentration data. Figure 3E shows rather good superposition of the data from the two channels, but the slope of the line drawn through the points is 1.09 rather than the theoretical value of 1.0. Finally, Figure 3F (eq 11) shows quite a bit of deviation from linearity. Some of this deviation is almost certainly due to the greater uncertainty in values of M_Z relative to M_n and M_w, however.

The molecular parameters determined from the manipulations of data of Figure 3 are presented in Table II. As was

[3] The analysis of Steiner (1954) on the M_w and M_n data using a monomer molecular weight of 21×10^3 g/mole gave $k_2 = 17.72 \times 10^3$ l./mole, $k_3 = 8.64 \times 10^3$ l./mole, and $k_4 = 19.80 \times 10^3$ l./mole. The root-mean-square data fit was ±974 for M_n data and ±1318 for M_w data. These equilibrium constants are statistically identical so models with unique sets of constants have not been considered further for the fraction II system.

TABLE II: Determination of M_1 and K from the Data Presented in Figure 3.[a]

Function	Eq No.	Fig. No.	Both Cells		Cell 1		Cell 2	
			M_1	K	M_1	K	M_1	K
M_n *vs.* C/M_n	3	3A	22,150	0.437	23,320	0.397	21,020	0.486
M_w^2 *vs.* C	5	3B	23,420	0.349	24,320	0.328	22,604	0.363
M_w *vs.* C/M_n	7	3C	23,090	0.364	24,110	0.338	22,020	0.395
M_nM_w/C *vs.* M_n/C	8	3D	22,480	0.415	23,670	0.370	20,920	0.502
M_ZM_w *vs.* C	11	3F	23,780	0.341	24,890	0.310	22,660	0.376
Av			22,980	0.381	24,044	0.349	21,845	0.424
Std dev			668	0.042	632	0.035	838	0.065

[a] The units of K are liters per gram.

observed in Figure 3 the two channels differ in reaction characteristics and further, the five treatments of the data produce somewhat different results.

This variability could have several causes: a departure from the model due to chain termination, heterogeneity of monomeric species, or the presence of molecules not participating in the chemical reactions. The remainder of the discussion will be devoted to an attempt to resolve these possible causes of deviations of the data.

If the angle between donor and acceptor sites of the associating molecules are located 180° apart then a linear string will occur. Any angle deviating from 180° will form a helix or a ring depending on the coordinates of the donor and acceptor sites relative to the center of mass of the monomer. Helix formation with no chemical affinity between successive turns would be indistinguishable from a linear stack in its ultracentrifuge behavior. However, ring formation constitutes a chain termination mechanism, and from eq 13 to 15, would change the observed molecular weight distributions.

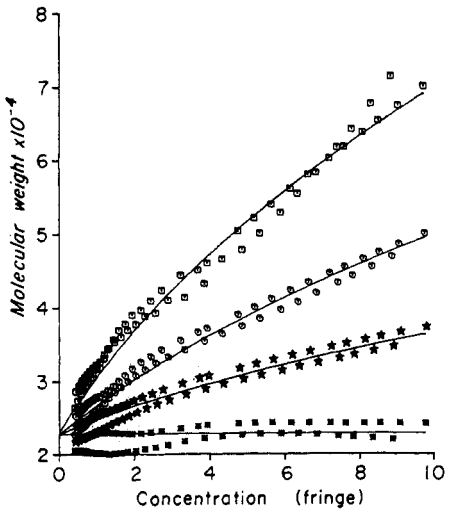

FIGURE 1: Point-by-point molecular weight distributions for fraction II. The solid lines are those predicted for an indefinite self-association with K = 0.376 l./g and M_1 = 22.87 × 10³ g/mole. Squares are $M_{Z,r}$ data, circles are $M_{w,r}$ data, stars are $M_{n,r}$ data, and the crosses are $2M_{n,r} - M_{w,r}$ data.

Equations 13–15 are only chain termination equations; ring formation would invoke approximately k^2 for the addition of the last monomer (neglecting entropy considerations) (Caspar and Klug, 1962). However, if the shape of the molecular weight distributions is not approximated by simple

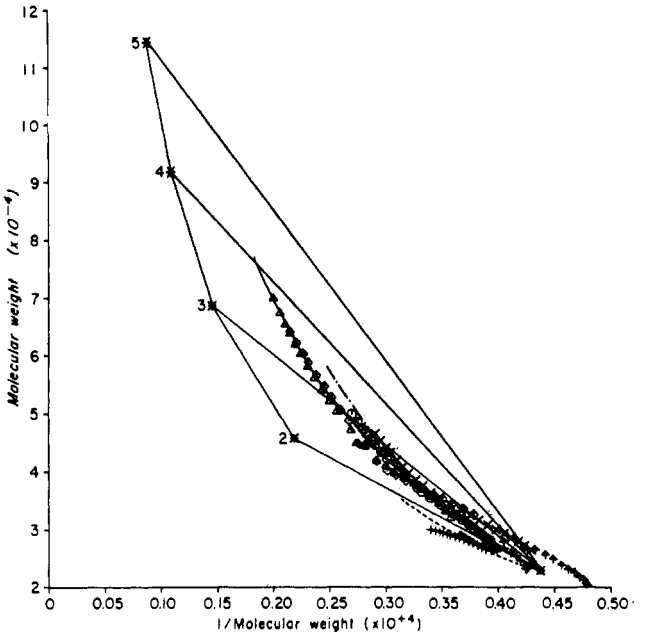

FIGURE 2: Molecular space for the fraction II system. The points labeled by numbers represent dimer, trimer, etc., of 22.87 × 10³ g/mole. The solid curved line represents predicted values of $M_{Z,r}$ *vs.* $1/M_{w,r}$. The short dashed line is the predicted curve of $2M_{w,r} - M_{Z,r}$ *vs.* $2/M_{n,r} - 2/M_{n,r} - 1/M_{w,r}$. Thd The values predicted for $M_{w,r}$ *vs.* $1/M_{n,r}$ are given by the broken line (-.-.-.-.). The equilibrium constant used for calculation of these curves was 0.376 l./g. Only data above one fringe displacement are plotted in the figure.

Table of Symbols

Channel No.	$M_{Z,r}$ *vs.* $1/M_{w,r}$	$M_{w,r}$ *vs.* $1/M_{n,r}$	$2M_{w,r} - M_{Z,r}$ *vs.* $2/M_{n,r} - 1/M_{w,r}$
1	△	○	+
2	◇	×	✦

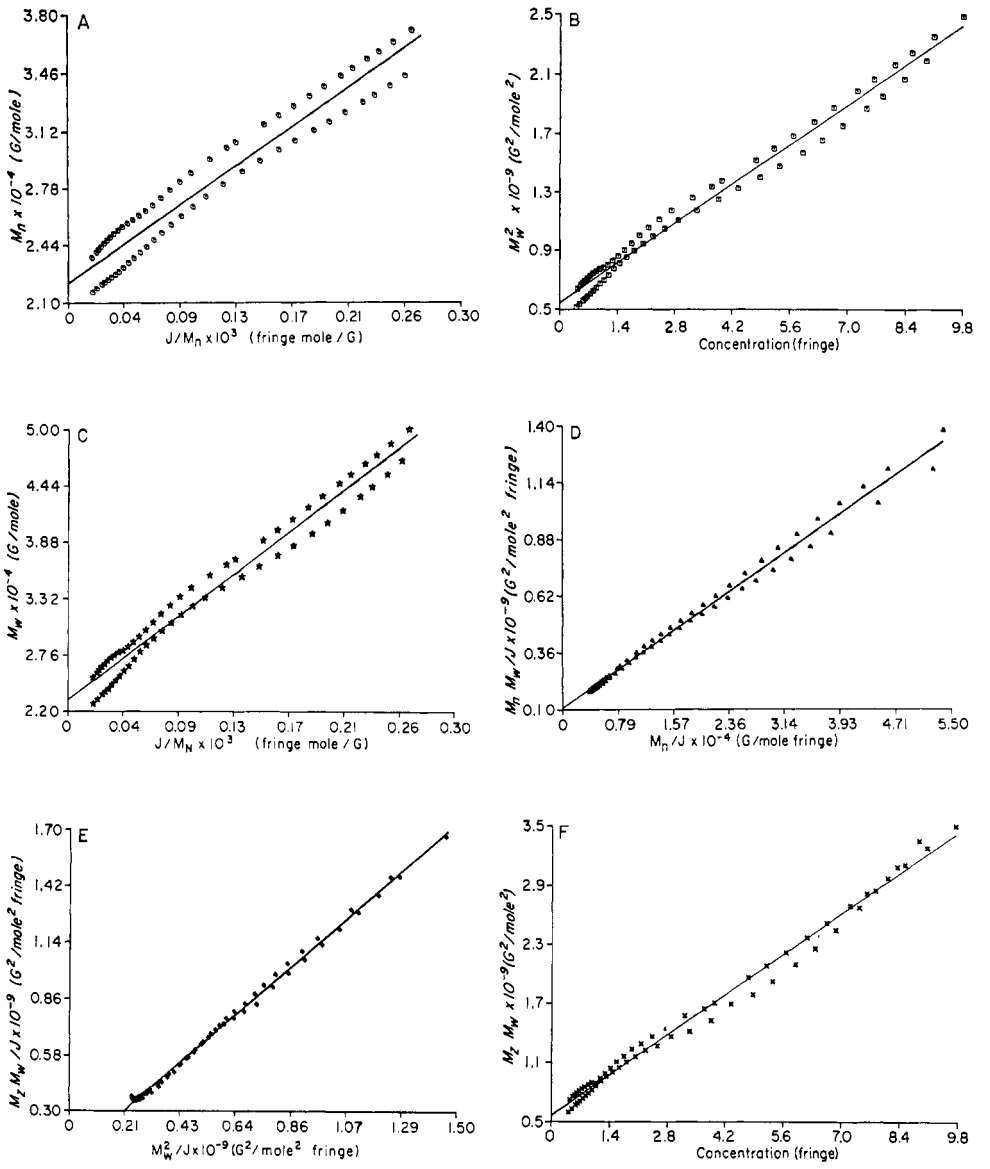

FIGURE 3: Graphs of the functions which are linear for indefinite self-association. (A) M_n *vs.* C/M_n, eq 3; (B) M_w^2 *vs.* C, C, eq 5; (C) M_w *vs.* C/M_n, eq 7; (D) M_nM_w/C *vs.* M_n/C, eq 8; (E) M_ZM_w/C *vs.* M_w^2/C, eq 10; and (F) M_ZM_w *vs.* C, eq 11. The molecular parameters computed from the slopes and intercepts are presented in Table II.

chain termination, then it is not necessary to consider ring formation rigorously. We shall attempt to determine to what extent chain termination can be excluded as a possible model for such self-associating systems. A calculation of the Gilbert (1959) schlieren pattern for simple chain termination indicates that only a single boundary would be observed for the sedimentation velocity studies (Behnke *et al.*, 1970).

From the computational method outlined in Table I of the Theory section the numbers presented in Table III were obtained. In this table the values of M_1 increase with increasing polymer size while the best constant decreases. The root-mean-square deviation of the M_n data is least for tetramers while the fit to the M_w is best for hexamers. Because the M_Z deviations always decrease with increasing degrees of polymerization, the average of the three molecular weight deviations follows the same trend. If the minimum average root-mean-square deviation is used as the criterion for selection of the model, then polymers greater than octamers cannot be excluded. However, the computation produces a dilemma because the three individual molecular weight moments fit three different models.

The difficulty is partially resolved by inspection of a graph of the predicted curves for polymers of 4 and 6. Figure 4 demonstrates that the M_w and M_Z curves predicted from the parameters for tetramers deviate substantially from the data at high concentration, while the M_n data are well described. On the other hand, the fit to the M_n data by the hexamer model (dashed lines) is not as good as the tetramer model because of the higher value of M_1 found in the computation. Presumably the higher M_1 arises from the influence of the

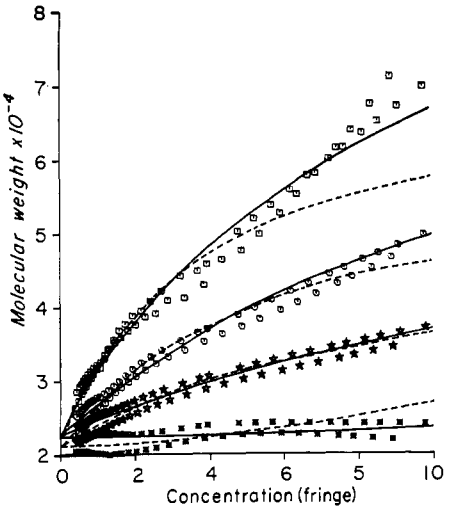

FIGURE 4: The data in this figure are the same as those of Figure 1. The lines are those predicted for chain termination at the tetrameric (dashed lines) and hexameric (solid lines) stages. The lines were computed using the molecular parameters given in Table III for these models. Squares represent M_Z data, circles represent M_w data, stars represent M_n data, and crosses represent $2M_n - M_w$ data.

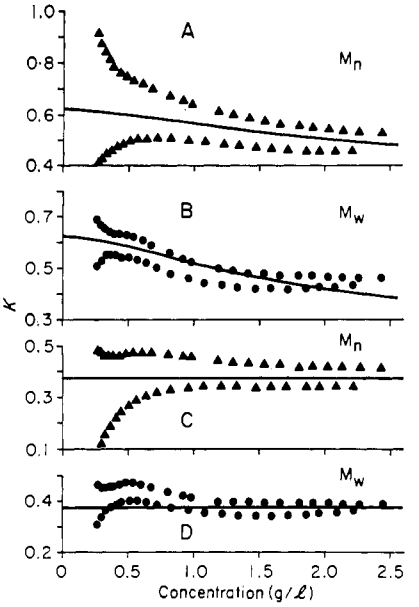

FIGURE 5: Calculated values of point-by-point equilibrium constants. The units of K are in liters per gram. (A) M_n data were used with $M_1 = 21.21 \times 10^3$ to calculate K. To produce the solid line in this figure, M_n values were predicted as a function of concentration for a tetrameric model with $K = 0.617$ l./g; these data were then treated as if the system were indefinitely associating. (B) M_w data were used with $M_1 = 21.21 \times 10^3$ g/mole to calculate K at each point. The solid line was calculated as in part A. (C) M_n data were used with $M_1 = 22.87 \times 10^3$ g/mole to calculate K at each point. The solid line is that for an indefinite association with $K = 0.376$ l./g. (D) M_w data were used with $M_1 = 22.87 \times 10^3$ g/mole to calculate K at each point. The solid line is that for $K = 0.376$ l./g.

M_w and M_Z data on the average root-mean-square error which was used as the criterion of the best K and M_1.

The predicted curves of M_w^2 *vs.* J, M_n *vs.* J/M_n, and M_nM_w/J *vs.* M_n/J for tetramers and hexamers are less sensitive than Figure 4 for deviations from these models. The amount of curvature in the predicted values is surprisingly small.

Figure 5 shows the calculation of point-by-point equilibrium constants from the relations: $K = (R_n^2 - R_n)/C$, where $R_n = M_n/M_1$, and $K = (R_w^2 - 1)/4C$, where $R_w = M_w/M_1$ (Van Holde and Rosetti, 1967). Figure 5A,B depicts the results calculated for $M_1 = 21.21 \times 10^3$ g/mole. To produce the solid lines shown in these two figures, M_n/M_1 and M_w/M_1 were calculated as a function of concentration for tetramers with $K = 0.617$ l./g. The data were then used to compute apparent equilibrium constants from the above R_n and R_w equations. Figure 5C,D shows the observed values of K for $M_1 = 22.87 \times 10^3$ g/mole. The flat lines correspond to a true indefinite association with $K = 0.376$ l./g. Apparently, the best model for a reaction is very dependent upon the assumed value of the monomer molecular weight. A salient feature of Figure 5A,B is that K determined from M_n data at each point is larger than K determined from M_w at the same concentration. This behavior mimics a virial coefficient. One may assume that the decrease of K with increasing concentration in these figures is due to nonideality and compute a value of BM_1 which makes K constant (Van Holde and Rosetti, 1967). Under these conditions a reasonably straight line is observed for these data, demonstrating further the similarity between chain termination and virial terms.

TABLE III: Iterative Computation of M_1 and K for Fraction II.

Polymer Size	M_1	Root-Mean-Square Deviations				
		K (l./g)	$\pm M_n$	$\pm M_w$	$\pm M_Z$	$\pm$ Av
4	21,210	0.617	955	1215	4286	2152
6	22,500	0.426	1036	1013	1960	1337
8	22,890	0.382	1094	1035	1769	1299
10	22,880	0.376	1084	1028	1751	1289
12	22,870	0.376	1083	1029	1733	1282
∞	22,870	0.376	1082	1029	1728	1280

The deviations of the data from the indefinite association model do not appear to be the result of chain termination. This conclusion is made because no single n-mer model provide a better data fit than the indefinite association. The iterative computation of K and M_1 for the indefinite association (Table III) produces the same result as the average of all the linear treatments of the data (Table II, row 6) and is not significantly better than polymers of 8 or above. The region of significant deviation from the theoretical behavior is at a concentration of 1–3 fringe. It was thought that heterogeneity of the monomeric species might account for these deviations and for this reason eq 21–25 were derived.

It was noted in the discussion of eq 21 that upward curvature of $M_{n,r}$ *vs.* $C/M_{n,r}$ would be observed for a heterogeneous monomer. From Figure 3A, it is apparent that exactly the opposite trend of data occurs. Further, Figure 3C,D should show upward curvature (eq 22 and 23), while $2M_{n,r} - M_{w,r}$ (eq 24) should curve downward. In all cases the reverse happens. For this reason, the curvature of data observed

in Figures 1 and 3 cannot be due to a heterogeneous population of monomers all of which participate in the chemical equilibria.

The third potential cause of curvature in Figure 3 is the existence of molecules which do not participate in the chemical equilibria. This is often observed in protein samples and particularly in proteolytic enzyme preparations (T. A. Horbett and D. C. Teller, unpublished data). This possibility is testable by variation of the initial concentrations but, unfortunately, this experiment was not performed. The presence of approximately 5% of small material with molecular weight in the range of 15,000–20,000 g/mole would give rise to the observed deviations from the model. At low concentration a large proportion of the molecules would be those of the contaminating protein while at high concentration most molecules would be of the self-associating type.

On the other hand, in view of the extreme purity of fraction II in disc gel electrophoresis (Behnke *et al.*, 1970), the deviations may be due to experimental and computational errors, but the appearance of the deviations at the same concentrations for both experiments would make this seem improbable.

The best present model for the self-association of fraction II is that of indefinite self-association with slight contamination by molecules not participating in the chemical reactions. The forces involved in the equilibria are relatively weak. Behnke and Neurath (1970) and Behnke *et al.* (1970) have shown that succinylated fraction II neither self-associates nor participates in complex formation with carboxypeptidase A. Since the formation of a complex between fraction II and carboxypeptidase A competes with the self-association of fraction II (Behnke *et al.*, 1970), we may conclude that the sites for self-association and complex formation are overlapping areas of the fraction II molecule, perhaps involving some of the same amino acid residues.

Acknowledgments

The author thanks Dr. W. D. Behnke for performing the experiment reported here and to thank Mr. R. D. Wade for assistance in these analyses. Discussions with Mr. C. E. Harris led directly to the idea for the chain termination equations.

References

Adams, E. T., Jr. (1964), *Proc. Nat. Acad. Sci. U. S. 51*, 509.

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

Adams, E. T., Jr., and Lewis, M. S. (1968), *Biochemistry 7*, 1044.

Behnke, W. D., Wade, R. D., and Neurath, H. (1970), *Biochemistry 9*, 4179.

Behnke, W. D., Teller, D. C., Wade, R. W., and Neurath, H. (1970), *Biochemistry 9*, 4189.

Caspar, D. L. D. (1966), *in* Principles of Biomolecular Organization, Wolstenholme, G. E. W., and O'Connor, M., Ed., Boston, Mass., Little, Brown, p 7.

Caspar, D. L. D., and Klug, A. (1962), *Cold Spring Harbor Symp. Quant. Biol. 27*, 1.

Chignell, D. A., Gratzer, W. B., and Valentine, R. C. (1968), *Biochemistry 9*, 1082.

Elias, H.-G., and Bareiss, R. (1967), *Chimia (Zurich) 21*, 53.

Elias, H.-G., and Lys, H. (1966), *Makromol. Chem. 96*, 64.

Flory, P. J. (1953), Principles of Polymer Chemistry, Ithaca, N. Y., Cornell Univ.

Gilbert, G. A. (1959), *Proc. Roy. Soc., Ser. A 250*, 377.

Goldberg, R. J. (1953), *J. Phys. Chem. 57*, 194.

Khalil, M. T. M., and Lauffer, M. A. (1967), *Biochemistry 6*, 2474.

Kiselev, N. A., De Rosier, D. J., and Klug, A. (1968), *J. Mol. Biol. 35*, 561.

Kiselev, N. A., Shpitzberg, C. L., and Vainshtein, B. K. (1967), *J. Mol. Biol. 25*, 433.

Klug, A., Finch, J. T., Leberman, R., and Longeley, W. (1966), *in* Principles of Biomolecular Organization, Wolstenholme, G. E. W., and O'Connor, M., Ed., Boston, Mass., Little, Brown, p 158.

Lansing, W. D., and Kraemer, E. O. (1935), *J. Amer. Chem. Soc. 57*, 1369.

Richards, E. G., and Schachman, H. K. (1959), *J. Phys. Chem. 63*, 1578.

Richards, E. G., Teller, D. C., and Schachman, H. K. (1968), *Biochemistry 7*, 1054.

Smith, C. E., and Lauffer, M. A. (1967), *Biochemistry 6*, 2457.

Steiner, R. F. (1952), *Arch. Biochem. Biophys. 39*, 333.

Steiner, R. F. (1954), *Arch. Biochem. Biophys. 49*, 400.

Svensson, H. (1954), *Opt. Acta 1*, 25.

Svensson, H. (1956), *Opt. Acta 3*, 164.

Tanford, C. (1961), Physical Chemistry of Macromolecules, New York, N. Y., Wiley.

Teller, D. C. (1967), *Anal. Biochem. 19*, 256.

Teller, D. C., Horbett, T. A., Richards, E. G., and Schachman, H. K. (1970), *Ann. N. Y. Acad. Sci. 164*, 66.

Timasheff, S. N. (1966), *Biopolymers 4*, 107.

Valentine, R. C., Shapiro, B. M., and Stadtman, E. R. (1968), *Biochemistry 7*, 2143.

Van Holde, K. E., and Rosetti, G. P. (1967), *Biochemistry 6*, 2189.

Wales, M. (1948), *J. Phys. Chem. 52*, 235.

Yphantis, D. A., and Roark, D. E. (1969), *Ann. N. Y. Acad. Sci.* (in press).