An Equilibrium Ultracentrifuge Study of the Self-Association of Bovine Insulin*

P. D. Jeffrey† and J. H. Coates

ABSTRACT: Insulin self-associates in aqueous solution. By sedimentation equilibrium experiments the monomer molecular weight is found to be 5800 awu. At pH 2, ionic strength 0.1, monomers, dimers, tetramers, and hexamers are found to exist in dynamic equilibrium. Measurements at 15 and 25° yield k, $\Delta G°$, $\Delta H°$, and $\Delta S°$ values for the association reactions. Statistical mechanical calculations suggest that the hypothesis of limited rotational freedom of the monomers accompanied by hydrophobic bonding accounts most satisfactorily for the observed entropies of association.

Bovine insulin has long been recognized as a protein that reversibly self-associates in dilute aqueous solution. Thus it belongs to a group of proteins including α-chymotrypsin, β-lactoglobulin, and many others that are being increasingly investigated, both because of their intrinsic interest and because their study represents a relatively simple means of studying reversible protein–protein interactions.

Interacting protein systems have been studied in two main ways, by methods involving the anomalous effects occurring in transport experiments due to the superposition of interaction equilibria on the transport phenomena and by the measurement of the effects of interaction on equilibrium properties, such as average molecular weight; see, for example, Nichol *et al.* (1963).

In the study to be described the equilibrium ultracentrifuge was used to investigate the association of insulin in aqueous solution at pH 2, ionic strength (I) 0.1, with the object of determining the molecular weight of the monomer, the number and degree of polymerization of the species involved in the equilibria, and the thermodynamic parameters describing these equilibria. In previous studies, with the exception of the work of Adams (1962; Adams and Fujita, 1963), the species involved have been assumed uncharged and thermodynamically ideal. In this paper the effects of molecular charge have been estimated, and the results have been shown to be consistent with the postulated thermodynamic ideality. In order to obtain more insight into the nature of the interactions, entropies of association have been calculated for some very simple statistical mechanical models of the insulin system.

Experimental Procedures

Preparation of Insulin Solutions. Bovine crystalline zinc insulin (batch number A3) was supplied by the Australian Commonwealth Serum Laboratories. The zinc content was stated to be between 0.3 and 0.9%. Zinc is not bound by insulin at pH 2 (Cunningham *et al.*, 1955), and after dialysis the concentration in the insulin solutions was reduced to about $^1/_{2,500}$ of the original value. The buffer used was an aqueous sodium chloride–glycine–hydrochloric acid solution of ionic strength 0.1 and pH 2.00. Adjustment and measurement of pH was carried out at the temperature at which the sedimentation experiment was to be performed.

Solutions of insulin in the buffer were dialyzed at 2–6° against buffer. The pH values of diffusate and insulin solution were identical after dialysis and were always between 2.00 and 2.04.

The concentrations of the insulin stock solutions were measured in a refractometer similar to that described by Cecil and Ogston (1951). The concentration was expressed in terms of Rayleigh interference fringes in the 12-mm light path ultracentrifuge cell, j, by the use of the expression $j = \Delta nl/\lambda$ where Δn is the refractive index increment of the insulin solution, l is the optical path length (12 mm), and λ is the wavelength of the monochromatic light used in both the refractometer and the ultracentrifuge (5461 A). Insulin solutions of concentration lower than about 1 g/100 ml were prepared by diluting portions of the stock solutions with diffusate by weight.

Definition of the Macromolecular Component. Insulin has a positive charge at pH 2 and the insulin solutions were prepared as described above so that an electrically neutral macromolecular component could be defined in the way suggested by Casassa and Eisenberg (1960). The definition involves setting the concentration of diffusible ions equal inside and outside the membrane at dialysis equilibrium and allows the system to be treated formally as consisting of two components, the solvent and the macromolecular component (denoted component 2*). Casassa and Eisenberg's definition has been used throughout this work and its implications are discussed in the appropriate parts of this paper.

* From the Department of Physical and Inorganic Chemistry, University of Adelaide, Adelaide, South Australia. *Received August 16, 1965.*

† Present address: Department of Chemistry, Clark University, Worcester 10, Mass.

The concentrations of the insulin solutions determined refractometrically were consistent with the Casassa and Eisenberg definition since they were measured with diffusate as the reference solvent. The relationship between concentration in terms of Rayleigh interference fringes and in terms of grams of component 2* per liter of solution was obtained by evaporating to dryness accurately known volumes of insulin solution and of diffusate which had previously been equilibrated by dialysis. The weight of component 2* in a known volume of solution of known refractive index increment was obtained by subtracting the weight of dry residue in the diffusate from that in an equal volume of solution containing the nondiffusible component.

Measurement of Molecular Weight. The molecular weight of insulin as a function of concentration was determined by the technique of sedimentation equilibrium (Van Holde and Baldwin, 1958). The Rayleigh optical system of the ultracentrifuge was used to record the concentration distribution. Molecular weights were evaluated from the expression

$$\bar{M}_{w\,app}^{(x)} = \frac{2RT}{(1 - \bar{v}\rho)\omega^2} \frac{d \ln j^{(x)}}{d(x^2)} \qquad (1)$$

where R is the gas constant, T is the absolute temperature, j is the concentration in terms of Rayleigh interference fringes at a point in the cell x cm from the axis of rotation, $\bar{v}$ is the partial specific volume of the protein, and ρ is the density of the solution. For a thermodynamically ideal system of uncharged macromolecules in which the refractive index is a linear function of the solute concentration and in which pressure effects are negligible, the molecular weights which are evaluated from the right-hand term in eq 1 are true weight-average molecular weights. The subscript "app" in eq 1 denotes that for nonideal systems of charged molecules the molecular weight obtained is apparent, containing unknown contributions from charge and nonideality effects.

A Spinco Model E analytical ultracentrifuge was used for the determination of the apparent weight-average molecular weight of insulin as a function of concentration. Both 12- and 30-mm light path double-sector interference cells were used and were filled with a sufficient solution to give a column height of approximately 3 mm. A micrometer syringe was used to fill the cells to the same level in each experiment. The insulin solution was that obtained after dialysis and the solvent was diffusate to which 1,3-butanediol had been added to raise its refractive index to that of the solution (Richards and Schachman, 1959). A small volume of Dow Corning 555 silicone oil was placed in the solvent compartment to provide a curved bottom. Temperature control was maintained within $\pm 0.1°$ throughout the experiments by the use of the temperature control unit and refrigeration system of the ultracentrifuge. The speed was calculated from odometer and stop-watch readings taken over the final 3 hr of the experiment. Immediately after speed control was established, a schlieren photograph at 90° bar angle was taken to allow the position of the meniscus and cell bottom to be determined accurately. Photographs of the Rayleigh interference fringe patterns were taken 18 and 24 hr after the set speed had been reached. Measurements of these photographs showed that sedimentation equilibrium was attained in 18 hr in all the experiments. The photographs were taken on Kodak scientific plates, Type IIG. Photographic patterns were measured with a two-dimensional microcomparator (Optical Measuring Tools, Maidenhead, England) capable of reading to 2 μ and fitted with a projection screen. The true position of the meniscus and the cell bottom were taken as those suggested by Erlander and Babcock (1961) and were determined from the schlieren photograph. Measurements of the radial positions of interference minima were made from the meniscus to the cell bottom at increments of one fringe. The experiments were designed to give about 20 fringes from the meniscus to the cell bottom. Thus each sedimentation equilibrium experiment yielded about 20 pairs of values of concentration and distance from the axis of rotation for use in calculating molecular weights by eq 1.

Fitting of ln j vs. x^2 Data. To evaluate the apparent weight-average molecular weight as a function of concentration by eq 1, a relationship between ln j and x^2 is required. Two methods, least squares and Fourier series, were applied in obtaining such a relationship. Only the former method is considered in detail here since it was the method ultimately used. The conclusions drawn as a result of a comparison between the two methods will be given later.

LEAST-SQUARES FITTING. For all of the sedimentation equilibrium experiments performed it was found that the ln j *vs.* x^2 data could be fitted with the required degree of precision by an expression of the form

$$\ln j = a + bX + cX^2 \qquad (X = x^2)$$

The term $(d \ln j)/d(x^2)$ as a function of x was then obtained by termwise differentiation followed by substitution of the appropriate value of x. Only those x values which were actually measured on the photographic plate have been used in calculating values of $(d \ln j)/d(x^2)$ in this way. The raw ln j *vs.* x^2 data were also differentiated by the mean-value theorem in order to check that the termwise differentiation did not give anomalous results (Householder, 1953). The molecular weights obtained by mean-value theorem differentiation exhibited considerable scatter but were evenly distributed about the smooth curve obtained from least-squares fitting followed by termwise differentiation. All of the apparent weight-average molecular weight *vs.* concentration graphs presented in this paper were obtained by least-squares fitting of the ln j *vs.* x^2 data followed by termwise differentiation. If these graphs (Figure 1a and b) are examined, it is observed that, at the points where the concentration ranges covered in two experiments overlap, there is quite a consistent trend

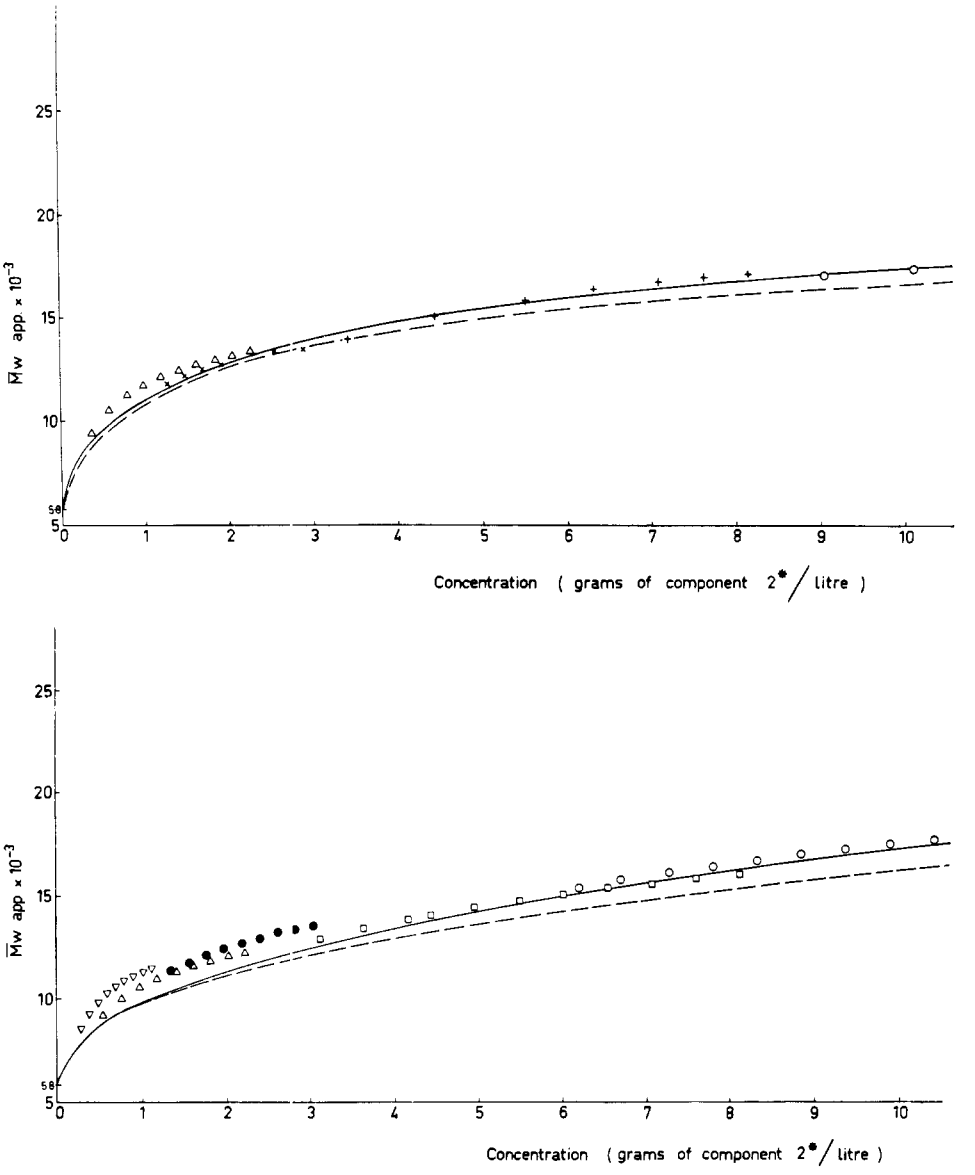

FIGURE 1: Apparent weight-average molecular weight of insulin as a function of concentration; pH 2, ionic strength 0.1. Each set of symbols refers to a particular experiment. The solid line is calculated from the equilibrium constants assuming no charge effects and ideality, the dashed line is calculated accounting for charge effects. (a top) Temperature 15°, (b bottom) 25°.

toward the formation of a cusp or step discontinuity. The effect is more marked in experiments at higher ionic strengths where the curves are steeper. This behavior is not expected on theoretical grounds since the apparent weight-average molecular weight is supposedly a function only of concentration, at fixed temperature and ionic strength, but it has been observed before with results from sedimentation equilibrium experiments in which apparent weight-average molecular weights have been evaluated by a procedure similar to that described above (Adams, 1962). In order to see whether this anomaly was connected with the way in which the ln j *vs.* x^2 data were fitted, data from two experiments were fitted by means of Fourier series also. The results, while not conclusive, indicated that the cusps or discontinuities are obtained as a consequence of the least-squares method employed in treating the experimental data. The curves obtained by least-squares fitting were found to lie within the scatter of points obtained by Fourier-series fitting, and the former method has been employed throughout this work.

THE CASASSA AND EISENBERG DEFINITION APPLIED TO AN ASSOCIATING SYSTEM. In a system consisting of a single charged macromolecular component, component 2, and a salt, component 3, the Casassa and Eisenberg definition of the macromolecular component

leads to the following expression at sedimentation equilibrium

$$M^*_2L^*_2\left[\frac{\mathrm{d}\ln m_2}{\mathrm{d}(x^2)}\right]^{-1} = 1 + m_2(\Sigma\nu^{*2}_{2i}/m_i + \beta^*_{22}) \quad (2)$$

where $L_2 = (1 - \bar{\nu}\rho)\omega^2/2RT$, the ν_{2i} are the number of moles of ion species comprising component 3 which are included in 1 mole of electrically neutral component 2, $\beta_{22} = [(\partial \ln j_2)/\partial m_2]_{P,T,m_3}$, and the concentrations are in terms of molality. Asterisks denote quantities consistent with the Casassa and Eisenberg definition and it is assumed that $\bar{\nu}$, ρ, and ν_{2i} are independent of x. The latter two assumptions will be satisfied for sedimentation equilibrium experiments at low speeds with solutions dilute in component 2 but moderately concentrated in salt. Such conditions are satisfied in the sedimentation equilibrium experiments with insulin.

The effect of the Casassa and Eisenberg definition in a system in which reversible association reactions are occurring is discussed in Appendix I. For an ideal system in which monomer, dimer, tetramer, and hexamer are in equilibrium and in which the partial specific volumes of all the constituents are the same, *i.e.*, $L^*_1 = L^*_2 = \ldots = L^*$, the following equation is obtained

$$\bar{M}^*_{w\,app} = \frac{1}{c^*_1}\Bigg[c^*_1M^*_1\left(1 - c^*_1\frac{X}{M^*_1}\right) + c^*_2M^*_2\left(1 - 2c^*_2\frac{X}{M^*_1}\right) + c^*_4M^*_4 \times \left(1 - 4c^*_4\frac{X}{M^*_1}\right) + c^*_6M^*_6\left(1 - 6c^*_6\frac{X}{M^*_1}\right)\Bigg] \quad (3)$$

where c^*_1, c^*_2, c^*_4, c^*_6 are the weight concentrations of the monomer, dimer, tetramer, and hexamer constituents, respectively, and $X = \nu^{*2}_{1+}/m_+ + \nu^{*2}_{1-}/m_-$. Thus, if we measure $(1/L^*)\ \mathrm{d}\ln c^*/\mathrm{d}(x^2)$ as a function of x in a sedimentation equilibrium experiment with an ideal associating system in which the species in equilibrium are the monomer, dimer, tetramer, and hexamer, $M^*_{w\,app}$ as a function of x and hence of c^* is given by (3). The term L^* represents $(1 - \bar{\nu}^*\rho)\omega^2/2RT$, and strictly speaking the partial specific volumes used in evaluating $M^*_{w\,app}$ should be that defined by adding 1 g of component 2* to a large volume of solution in dialysis equilibrium with salt solution. However, Casassa and Eisenberg (1961) have shown that, for solutions of bovine serum albumin in 0.1 M NaCl, the difference between the partial specific volume consistent with their definition and that as ordinarily measured is only a few tenths of a percent. The magnitude of the difference in the partial specific volumes which could be measured for insulin is unlikely to differ from that measured for bovine serum albumin; thus a value determined in the usual way (0.72 ml/g, Oncley *et al.*, 1952) has been used. The concentrations measured in the sedimentation equilibrium experiments with insulin were in terms of Rayleigh interference fringes, and it is assumed that the refractive index increment is proportional to concentration on a weight per volume scale since this is usually found to be the case for protein solutions (Casassa and Eisenberg, 1961). Equation 3 is in terms of weight/weight concentrations, but the densities of the solutions used in these experiments were so close to unity that the error incurred in equating weight/weight and weight/volume concentrations may be neglected.

It is possible to include nonideality effects and to derive an expression for the measured apparent weight-average molecular weight which includes terms expressing the dependence of the activity coefficients of the various species on concentration as well as on the effects of charge (Adams, 1962). This has not been done here because there seems at present to be no way in which the magnitude of nonideality effects can be estimated.

Evaluation of Equilibrium Constants. Steiner (1952) has shown that in an ideal, associating system of uncharged macromolecules, stepwise equilibrium constants can be evaluated from a plot of $\bar{\alpha}_w/x_1$ *vs.* x_1c/m where $\bar{\alpha}_w = \bar{M}_w/m$, m is the molecular weight of the monomer, and x_1 is the weight fraction of monomer at the concentration c. In the general case $\bar{\alpha}_w/x_1$ is related to x_1c/m by

$$\frac{\bar{\alpha}_w}{x_1} = 1 + 4k_2\left(\frac{x_1c}{m}\right) + 9k_2k_3\left(\frac{x_1c}{m}\right)^2 + 16k_2k_3k_4\left(\frac{x_1c}{m}\right)^3 + \ldots \quad (4)$$

where the k_2 is the equilibrium constant for the formation of dimer, k_3 that for trimer from a monomer and a dimer, and so on. In the present study the stepwise equilibrium constants for the insulin equilibria in acid solution have been determined by Steiner's method using a curve-fitting technique with the aid of a computer. The data in the form $\bar{\alpha}_w/x_1$ *vs.* x_1c/m were supplied to the computer which was programmed in such a way as to fit these data to polynomials of increasing order up to a predesignated limit using the method of least squares. The printout provided the values of the coefficients of the polynomials, their standard errors, and the standard errors of the polynomials. The particular polynomial which best fitted the data could thus easily be decided. Preliminary attempts by graphical methods and by the method of orthogonal polynomials (Buckingham, 1957) to evaluate equilibrium constants indicated that the polynomials required would not be in the form given by (4), and the next approach tried was based on the assumption that the species in equilibrium were the even-numbered members of the series, *i.e.*, that the monomer as such only participated in the reaction leading to the formation of dimer which subsequently acted as the effective monomer. It can easily be shown that if this is the case $\bar{\alpha}_w/x_1$ and x_1c/m are related by an expression of the form

$$\frac{\bar{\alpha}_w}{x_1} = 1 + 4k_2\left(\frac{x_1c}{m}\right) + 16k_2^2k_4\left(\frac{x_1c}{m}\right)^3 + 36k_2^3k_4k_6\left(\frac{x_1c}{m}\right)^5 + \ldots \quad (5)$$

where k_4 now refers to the formation of tetramer from two dimers and k_6 to the formation of hexamer from dimer and tetramer.

Thus the computer was programmed to fit the data to polynomials of odd order in x_1c/m. The equilibrium constants which were evaluated from the polynomial of best fit were used to calculate curves of molecular weight *vs.* concentration which were compared with the experimental curves. If the calculated and experimental molecular weights agreed within about ±500 over the entire concentration range, the fit was judged to be satisfactory, since this figure is the precision with which molecular weights could be measured.

The Effect of a Charged Monomer Species on the Evaluation of Equilibrium Constants (see also Appendix II). Equation 5 was derived for an ideal system of neutral macromolecules. If we now consider an ideal system of charged macromolecules where the molecular weight is expressed in a manner consistent with the Casassa and Eisenberg definition, the expression which can be derived when monomer, dimer, tetramer, and hexamer are in equilibrium is

$$\frac{\alpha^*}{x_1} = 1 + (4k_2 - X)\left(\frac{x_1c^*}{m^*}\right) + 16k_2^2(k_4 - X) \times \left(\frac{x_1c^*}{m^*}\right)^3 + 36k_2^3k_4k_6\left(\frac{x_1c^*}{m^*}\right)^5 - (16k_2^2k_4)^2\left(\frac{x_1c^*}{m^*}\right)^7 \times X - (36k_2^3k_4k_6)^2\left(\frac{x_1c^*}{m^*}\right)^{11} X \quad (6)$$

where $\alpha^* = \bar{M}^*_{w\,app}/m^*$, m^* is the molecular weight of the monomer consistent with the definition, c^* is the total weight per volume concentration of component 2*. The quantity x_1 in eq 5 may be evaluated from the expression

$$\ln (x_{1,c}/x_{1,0}) = \int_{c_0}^{c} \frac{(\bar{\alpha})^{-1} - 1}{c}\, dc$$

which applies to an ideal system (Steiner, 1954). In practice the expression

$$\int_{c^*_0}^{c^*} [(\alpha^*)^{-1} - 1] dc^*/c^*$$

is evaluated since it is determinable directly from experimental measurements. The expression for x_1 appropriate to eq 6 is given in Appendix II. It is found that negligible error is incurred in evaluating x_1 from

$$\int_{c^*_0}^{c^*} [(\alpha^*)^{-1} - 1] dc^*/c^*$$

On comparing eq 5 and 6 it is seen that the coefficients $4k_2$ and $16k_2^2k_4$ are replaced by $(4k_2 - X)$ and $16k_2^2(k_4 - X)$, respectively; also there are extra terms in $(x_1c^*/m^*)^7$ and $(x_1c^*/m^*)^{11}$.

Assessment of Probable Errors. The standard errors of the coefficients of the polynomials in (x_1c/m) were obtained from the computer output. These values were used to compute standard errors for the derived k values and thermodynamic parameters by standard statistical techniques (Paradine and Rivett, 1960). The quoted standard errors in derived quantities were all calculated in this way and of course do not reflect any *systematic* error that may exist in the original equilibrium data. A full discussion of the problems of error estimation in this system has been given by Jeffrey (1964).

Results

Sedimentation Equilibrium at pH 2, Ionic Strength 0.1, 25 and 15°. Sedimentation equilibrium experiments yielded values of apparent molecular weights at a number of concentrations; these are shown as points in Figures 1a and b. Empirical extrapolation of the low concentration regions of the curves to zero concentration yields a value for the molecular weight of the insulin monomer of 5800 ± 200 awu. This is consistent with the formula weight of the neutral insulin molecule obtained by amino acid analysis (Harfenist, 1953) of 5733 awu. Taking the molecular weight of the monomer to be 5800, assuming the system to be ideal and the insulin monomer to be uncharged, these data were used to compute the equilibrium constants shown in Table I. To check the validity of the procedures used the equilibrium constants were used in conjunction with arbitrarily selected values of x_1c/m to compute the solid lines shown in Figures 1a and b.

TABLE I: Equilibrium Constants for Insulin Equilibria at I = 0.1, pH 2, Assuming the Species Involved to Be Uncharged and Exhibiting Ideal Behavior.

Constant	Temp (°C)	Calcd Value of Equilibrium Constant × 10^{-3}	Standard Error
k_2	15	15.46	±0.38
	25	10.20	±0.21
k_4	15	2.04	±0.12
	25	0.785	±0.05
k_6	15	0.038	±0.026
	25	0.668	±0.051

Equilibrium constants were also evaluated taking into account the charge on the monomer unit in the interpretation of the computer-derived coefficients of the Steiner polynomials, but neglecting the effect of charge

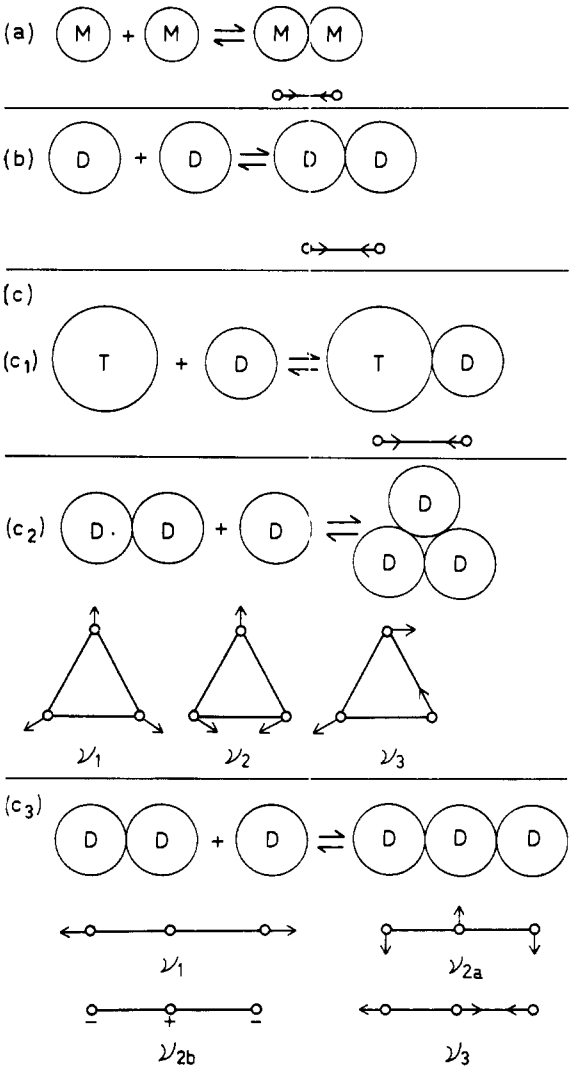

FIGURE 2: Spherical models used for the calculation of entropies of association by the Scheraga and Steinberg method. In their associated states the spheres are allowed to rotate with respect to each other subject to the restriction that from 0.1 to 0.01 of their surface areas only are available for bonding. The extra modes of vibration for each of the associated species are shown also.

on the calculation of x_1 values. The equilibrium constants obtained in this way are shown in Table II and were used in conjunction with arbitrarily selected values of x_1c/m to calculate the dashed lines shown in Figures 1a and 1b.

Calculation of Thermodynamic Quantities. The equilibrium constants (in units of liters/mole) obtained in assuming an ideal system of uncharged monomers were used to calculate $\Delta G°$, $\Delta H°$, and $\Delta S°$ together with their standard errors as listed in Table III. In such a system true thermodynamic parameters would be obtained. Here the quantities denoted "thermodynamic" should strictly be regarded as *apparent* thermodynamic quantities because of the assumptions made in their evaluation.

TABLE II: Equilibrium Constants for Insulin Equilibria at I = 0.1, pH 2, Partially Taking into Account the Effect of a Charged Monomer Species on the Evaluation of the Equilibrium Constants.

Constant	Temp (°C)	Calcd Value of Equilibrium Constant $\times 10^{-3}$	Standard Error
k_2	15	15.49	±0.38
	25	10.23	±0.21
k_4	15	2.16	±0.12
	25	0.897	±0.05
k_6	15	0.036	±0.026
	25	0.580	±0.051

TABLE III: Thermodynamic Parameters for Insulin Association Reactions Calculated Assuming the Insulin Monomer to Be Uncharged and Ideal Behavior for the Associating Species.

Parameter	Temp (°C)	Calcd Value (kcal)	Standard Error (kcal)
$\Delta G°_2$	15	−5.52	±0.014
	25	−5.47	±0.012
$\Delta G°_4$	15	−4.36	±0.036
	25	−3.95	±0.038
$\Delta G°_6$	15	−2.08	±0.40
	25	−3.85	±0.045
$\Delta H°_2$	—	−7.10	±0.24
$\Delta H°_4$	—	−16.30	±0.63
$\Delta H°_6$	—	+49.0	±5.2
		eu	eu
$\Delta S°_2$		−5.5	±0.8
$\Delta S°_4$		−41	±2
$\Delta S°_6$		+177	±18

Discussion

The computer analysis showed that at both temperatures the data are best fitted by a model involving the following equilibria. I refers to the insulin monomer of molecular weight 5800. It was at first attempted to fit

$$2I \rightleftharpoons I_2 \qquad k_2 = I_2/I^2$$
$$2I_2 \rightleftharpoons I_4 \qquad k_4 = I_4/I_2^2$$
$$I_2 + I_4 \rightleftharpoons I_6 \qquad k_6 = I_6/I_2I_4$$

h e data to a scheme including the species I_3 and I_5

in addition to those already listed. This proved abortive as the attempt gave rise to negative values for some of the polynomial coefficients. The scheme involving I, I_2, I_4, and I_6 only gave no negative coefficients and furthermore gave a smaller standard error for the over-all fit of the polynomial to the data than reaction schemes involving higher order polynomials of I_2 than I_6. In view of this the reaction scheme put forward above is regarded as most likely. Because the Steiner method involves fitting to a polynomial and the higher order equilibrium constants are computed from products of the polynomial coefficients, there is least uncertainty in k_2 and most in k_6. The uncertainties are reflected in the appropriate standard errors which are so large in the case of k_6 that the value should be regarded as of qualitative significance only. The degree of agreement between the data points and the full curves computed using k values obtained by assuming ideality and no charge is further confirmation of the reaction scheme.

Comparison of the k values shown in Tables I and II shows the magnitude of the error arising as a consequence of neglecting the charge on the monomer in the interpretation of the polynomial coefficients. When the effect of charge is included in the analysis, k_2 and k_6 increase slightly, whereas k_4 decreases slightly. The analysis assumes a charge of +5 for the insulin monomer (Tanford and Epstein, 1954) which would correspond to +30 for the hexamer. The results suggest that neglect of the effect of charge if the Casassa and Eisenberg approach is used results in errors in the k values of the same order of magnitude as the standard errors. The dashed lines shown in Figures 1a and b, calculated using the k values of Table II, do not fit the original data points as well as do the full lines calculated using the k values of Table I. We have no explanation at present for this effect. It may be that a smaller value for the charge than +5 may be more appropriate; alternatively, the good agreement between the data and the calculated curve may be due to a fortuitous cancellation of the effects of charge and nonideality when both are neglected.

Calculation of Entropy Changes for Idealized Models of the Associating System. Steinberg and Scheraga (1963) have given equations by means of which the intrinsic entropy change for the association of two similar spherical particles in solution may be calculated when relative rotation of the units in the associated species is allowed. The intrinsic entropy change is defined as the sum of the contributions from changes in translational, rotational, and vibrational entropy but excludes changes in the interaction between solvent and solutes.

The Steinberg and Scheraga methods with various modifications have been used to calculate intrinsic entropy changes for the insulin association reactions found to occur in our experiments. Their equations were used unaltered for calculations of the translational and rotational entropy changes associated with dimer and tetramer formation. Three different models were considered for hexamer formation. The Steinberg and Scheraga equations were modified for the different models shown in Figure 2 in the following ways. In calculating the contribution to the entropy from internal vibration in the dimer and tetramer, the frequency of vibration was calculated following Steinberg and Scheraga from the equation

$$\nu = \frac{1}{\pi \Delta x_{\max}} \sqrt{\frac{2\Delta U_{\max}}{n_{\text{red}}}}$$

where $\Delta U_{\max}$ is the energy needed to break the bond and $\Delta x_{\max}$ is the maximum extension of the bond before it breaks. n_{red} is the reduced mass of the particles participating in the association. However, instead of equating $\Delta U_{\max}$ with $-\Delta H/N$ as was done by Steinberg and Scheraga, we have arbitrarily assumed maximum and minimum values for $\Delta U_{\max}$ of 50 and 5 kcal mole^{-1} and for $\Delta x_{\max}$ of 3 and 1 A. The extreme values of ν obtained by this procedure were used to obtain maximum and minimum values for the vibrational contribution to the entropy of association. The mean of these values was used in estimating the intrinsic entropy change. In the case of hexamer formation the frequencies of the normal modes of vibration were calculated for cases C_2 and C_3 using the valence force approximation of Bjerrum (Herzberg, 1956). The units in the hexamer were allowed complete freedom of rotation so the force constant for the angular restoring force was put equal to zero.

Doty and Myers (1953) have also considered the problem of calculating the entropy of association for various models of the insulin molecule. Their method has been extended and applied to the end-to-end association of cylinders of the appropriate size as a model for the formation of dimer and tetramer. The change in translational entropy was calculated on the basis of the Sakur-Tetrode equation and the cylinders were considered free to rotate with respect to each other about their common axis only. The contribution to the entropy change due to vibration along the line of centers was included.

The values obtained using the Steinberg and Scheraga (1963) approach, which assumes that combined monomers may rotate freely with respect to each other in all directions but that only a fraction (between 0.1 and 0.01) of the surface area is available for bonding, are shown in Table IV. The quantities in parentheses give the maximum total variation in S° that would arise if the fraction of the surface available for bonding varies between 0.1 and 0.01 and the parameters controlling the vibration vary between the arbitrary limits, $\Delta U_{\max}$ = 50 kcal mole^{-1}, $\Delta x_{\max}$ = 1 A and $\Delta U_{\max}$ = 5 kcal mole^{-1}, $\Delta x_{\max}$ = 3 A. Three values of ΔS°_6 were calculated for the three models shown in Figure 2. Three models were considered in the latter case as it is conceivable that the particular geometrical arrangement of the three dimer molecules might have a profound effect on the entropy of association; this was found not to be the case. The experimental value for ΔS°_6 is included only for completeness; it is so highly derived from the original experimental data as to be extremely uncertain.

For ΔS°_2 the agreement between the experimental and

TABLE IV: Entropy of Association Values Obtained by Calculation[a] and Experiment.

	ΔS°_2 (eu)	ΔS°_4 (eu)	ΔS°_6 (eu)
Calcd by Steinberg and Scheraga (1963) method	(a) -10.4 (± 7)	(b) -9.9 (± 7)	(c_1) $+5.4$ (± 3.4) (c_2) -2.2 (± 9.3) (c_3) $+7.6$ (± 11.5)
Calcd by extended Doty and Myers (1953) method	-60	-66	—
Exptl values	-5.5 (± 0.8)	-41 (± 2.2)	$+177$ (± 18)

[a] The values in parentheses are the maximum total variation in S° that could occur (see text). The letters in parentheses correspond to the models in Figure 2.

Steinberg–Scheraga values can be regarded as satisfactory, this interpretation implying that if relative free rotation of the monomers is allowed, it is not necessary to include any positive contribution to the entropy of association due to the formation of electrostatic or hydrophobic bonds. On the other hand, the agreement between the Doty and Myers approach and experiment can be judged satisfactory if a positive contribution of 55 eu/mole of dimer is added. This contribution could well arise from the formation of hydrophobic or electrostatic bonds on the formation of the dimer. Némethy and Scheraga (1962) calculate a positive entropy of formation of from $+1.7$ to $+11$ eu for a pure hydrophobic bond depending on the groups involved and the extent of overlap. If the hydrophobic bond is associated with a hydrogen bond, Némethy *et al.* (1963) calculate an entropy contribution of from -3 to $+7$ eu/bond.

For ΔS°_4 the situation is a little different; the difference between the Scheraga and Steinberg value and experiment could be accounted for by a negative contribution from hydrogen bonding of 30 eu. Némethy *et al.* (1963) calculate -2 to -5 eu/side-chain hydrogen bond. The difference between the Doty and Myers value and experiment could be accounted for by a positive contribution of 25 eu from hydrophobic bonding.

It is thus apparent that the present data do not in themselves allow of a decision between the two alternative hypotheses. Hydrophobic bonds are known to be of crucial importance for the association of insulin as solvents likely to break hydrophobic bonds such as dioxane–water (Fredericq, 1957), trifluoracetic acid–ether (Crespi *et al.*, 1956), and dimethylacetamide in 0.05 M potassium thiocyanate (Rees and Singer, 1956); all lead to dissociation of insulin to the 5800 molecular weight unit. Such solvent systems would tend to enhance the strength of hydrogen and electrostatic bonds. In view of this it seems more likely that the extended Doty and Myers model for association is more applicable to insulin than the Steinberg and Scheraga model.

Appendix I

The Casassa and Eisenberg Definition in a Reversibly Associating System.

The system considered is one in which monomer, dimer, tetramer, and hexamer are in equilibrium. If ideal behavior is assumed, eq 2 may be written for the monomer constituent as

$$M^*_1 L^*_1 \left(\frac{\mathrm{d} \ln \frac{c^*_1}{M^*_1}}{\mathrm{d}(x^2)} \right)^{-1} = 1 + \frac{c^*_1}{M^*_1} \left(\frac{\nu^{*2}_{1+}}{m_+} + \frac{\nu^{*2}_{1-}}{m_-} \right) \quad \text{(a)}$$

where c^*_1 is the weight concentration of the monomer constituent. Similar equations may be written for the other constituents and, if it is assumed that the partial specific volumes of all the constituents are the same, *i.e.*, $L^*_1 = L^*_2 = \ldots = L^*$, we may write

$$L^* \left[\frac{c^*_1 M^*_1}{1 + \frac{c^*_1}{M^*_1} \left(\frac{\nu^{*2}_{1+}}{m_+} + \frac{\nu^{*2}_{1-}}{m_-} \right)} + \frac{c^*_2 M^*_2}{1 + \frac{c^*_2}{M^*_2} \left(\frac{\nu^{*2}_{2+}}{m_+} + \frac{\nu^{*2}_{2-}}{m_-} \right)} + \ldots \right] = \frac{\mathrm{d}c^*}{\mathrm{d}(x^2)} \quad \text{(b)}$$

where $\mathrm{d}c^* = \mathrm{d}c^*_1 + \mathrm{d}c^*_2 + \ldots$ If the sedimentation equilibrium experiment is performed with solutions of macromolecular component and salt which have been equilibrated with each other by dialysis, the quantity which is actually measured at each distance x from the axis of rotation is proportional to $(1/L^*)\mathrm{d}c^*/\mathrm{d}(x^2)$ which we may denote by $\bar{M}^*_{\mathrm{w\ app}}$. Thus from (b)

$$\bar{M}^*_{\mathrm{w\ app}} = \frac{1}{c^*} \left[\frac{c^*_1 M^*_1}{1 + \frac{c^*_1}{M^*_1} \left(\frac{\nu^{*2}_{1+}}{m_+} + \frac{\nu^{*2}_{1-}}{m_-} \right)} + \frac{c^*_2 M^*_2}{1 + \frac{c^*_2}{M^*_2} \left(\frac{\nu^{*2}_{2+}}{m_+} \frac{\nu^{*2}_{2-}}{m_-} \right)} + \ldots \right] \quad \text{(c)}$$

If it is assumed that $M^*_2 = 2M^*_1$ etc., *i.e.*, $\nu^*_{2+} = 2\nu^*_{1+}$ etc., and for brevity we write $X = (\nu^{*2}_{1+}/m_+$

$+ \nu^*_{1-}{}^2/m_-$) then

$$\bar{M}^*_{w\,app} = \frac{1}{c^*}\left[\frac{c^*_1 M^*_1}{1 + c^*_1\left(\frac{X}{M^*_1}\right)} + \frac{c^*_2 M^*_2}{1 + 2c^*_2\left(\frac{X}{M^*_1}\right)} + \ldots\right] \quad \text{(d)}$$

This can be simplified by making use of the binomial theorem and approximating, for example

$$\left(1 + c^*_1\frac{X}{M^*_1}\right)^{-1} \simeq 1 - c^*_1\frac{X}{M^*_1}$$

where $c^*_1(X/M^*_1) \leqslant 0.1$. Expanding all the denominators similarly, (d) becomes for the case of monomer, dimer, tetramer, and hexamer

$$\bar{M}^*_{w\,app} = \frac{1}{c^*_1}\left[c^*_1 M^*_1\left(1 - c^*_1\frac{X}{M^*_1}\right) + c^*_2 M^*_2\left(1 - 2c^*_2\frac{X}{M^*_1}\right) + c^*_4 M^*_4 \times \left(1 - 4c^*_4\frac{X}{M^*_1}\right) + c^*_6 M^*_6\left(1 - 6c^*_6\frac{X}{M^*_1}\right)\right] \quad \text{(e)}$$

Appendix II

The Effect of Charge on the Evaluation of Equilibrium Constants. The expression for x_1 appropriate to 6 is as follows.

$$\ln\frac{x_{1,c^*}}{x_{1,0}} = \int_{c^*_0}^{c^*}\frac{[\alpha^* + \mathrm{f}(X)]^{-1} - 1}{c^*}\mathrm{d}c^* \quad \text{(f)}$$

where

$$\mathrm{f}(X) = \frac{X}{c^*/m^*}\left\{\left(\frac{x_1 c^*}{m^*}\right)^2 + \left[4k_2\left(\frac{x_1 c^*}{m^*}\right)^2\right]^2 + \left[16k_2{}^2k_4\left(\frac{x_1 c^*}{m^*}\right)^4\right]^2 + \left[36k_2{}^3k_4k_6\left(\frac{x_1 c^*}{m^*}\right)^6\right]^2\right\}$$

In practice the expression

$$\int_{c^*_0}^{c^*}[(\alpha^*)^{-1} - 1]\mathrm{d}c^*/c^* \quad \text{(g)}$$

is evaluated (rather than the right-hand side of eq f) since it is determinable directly from experimental measurements. The error committed by neglecting f(X) in the integral is not large. Using k and x_1 values obtained assuming an uncharged ideal system, it was found that for 25°, ionic strength 0.1, and a concentration of 15 g/l. the x_1 value calculated using eq f differed by only 2% from that obtained using eq g. At lower concentrations the error is somewhat less. The value of X used in this calculation was obtained as follows: Eisenberg and Casassa (1960) have shown that $\nu^*_{1-} = (1 - \Gamma)z$ and $\nu^*_{1+} = -\Gamma z$ and

$$\Gamma = {}^1/_2(1 - (zm_2/4m_3) + \ldots)$$

where Γ is the membrane distribution parameter. In our case z is the charge on the insulin monomer, and m_2 and m_3 are the molal concentrations at dialysis equilibrium of nondiffusible component and salt, respectively. For the conditions referred to in the previous paragraph, and additionally assuming $z = +5$ for the insulin monomer (Tanford and Epstein, 1954), the maximum value of the term $zm_2/4m_3$ is about 0.03. Accordingly, throughout this work Γ is approximated to $^1/_2$ and the definition of macromolecular component becomes identical with that suggested by Johnson *et al.* (1954). Thus $\nu^*_{1-} = -\nu^*_{1+} = 2.5$, and $X = 125$. This value of X was used in correcting the equilibrium constants for the effect of charge.

Acknowledgments

We wish to acknowledge the help of Mr. R. A. Mills with the statistical and curve-fitting procedures in this work; also the advice and encouragement of Professor D. O. Jordan.

References

Adams, E. T., Jr. (1962), Ph.D. Thesis, University of Wisconsin.

Adams, E. T., Jr., and Fujita, H. (1963), *in* Ultracentrifugal Analysis in Theory and Experiment, Williams, J. W., Ed., New York, N. Y., Academic.

Buckingham, R. A. (1957), Numerical Methods, London, Pitman, p. 312.

Casassa, E. F., and Eisenberg, H. (1960), *J. Phys. Chem. 64*, 753.

Casassa, E. F., and Eisenberg, H. (1961), *J. Phys. Chem. 65*, 427.

Cecil, R., and Ogston, A. G. (1951), *J. Sci. Instr. 28*, 253.

Crespi, H. L., Uphaus, R. A., and Katz, J. J. (1956), *J. Phys. Chem. 60*, 1190.

Cunningham, L. W., Fischer, B. L., and Vestling, C. S. (1955), *J. Am. Chem. Soc. 77*, 5703.

Doty, P., and Myers, G. E. (1953), *Discussions Faraday Soc. 13*, 51.

Erlander, S. R., and Babcock, G. E. (1961), *Biochim. Biophys. Acta 50*, 205.

Fredericq, E. (1957), *J. Am. Chem. Soc. 79*, 599.

Harfenist, E. J. (1953), *J. Am. Chem. Soc. 75*, 5528.

Herzberg, G. (1956), Molecular Spectra and Molecular Structure, Vol. II, Princeton, N. J., Van Nostrand, Chapter III.

Householder, A. S. (1953), Principles of Numerical Analysis, New York, N. Y., McGraw-Hill, p. 226.

Jeffrey, P. D. (1964), Ph.D. Thesis, University of Adelaide.

Johnson, J. S., Kraus, K. A., and Scatchard, G. (1954), *J. Phys. Chem. 58*, 1034.
Némethy, G., and Scheraga, H. A. (1962), *J. Phys. Chem. 66*, 1773.
Némethy, G., Steinberg, I. Z., and Scheraga, H. A. (1963), *Biopolymers 1*, 43.
Nichol, L. W., Bethune, J. L., Kegeles, G., and Hess, E. L. (1963), *Proteins 2*, 305.
Oncley, J. L., Ellenbogen, E., Gitlin, D., and Gurd, F. R. N. (1952), *J. Phys. Chem. 56*, 85.
Paradine, C. G., and Rivett, B. H. P. (1960), Statistical Methods for Technologists, London, English Univ. Press, p. 161.
Rees, E. D., and Singer, S. J. (1956), *Arch. Biochem. Biophys. 63*, 144.
Richards, E. G., and Schachman, H. K. (1959), *J. Phys. Chem. 63*, 1578.
Steinberg, I. Z., and Scheraga, H. A. (1963), *J. Biol. Chem. 238*, 172.
Steiner, R. F. (1952), *Arch. Biochem. Biophys. 39*, 333.
Tanford, C., and Epstein, J. (1954), *J. Am. Chem. Soc. 76*, 2163.
Van Holde, K. E., and Baldwin, R. L. (1958), *J. Phys. Chem. 62*, 734.