Biochemistry

Volume 11, Number 22 October 24, 1972

Conformation of Proinsulin. A Comparison of Insulin and Proinsulin Self-Association at Neutral pH†

Allen H. Pekar and Bruce H. Frank*

ABSTRACT: The self-association of porcine insulin and proinsulin at pH 7.00 and 24.5° has been studied using sedimentation equilibrium techniques. We found that the apparent weight-average molecular weight of insulin in solution was less than 11,550 at a protein concentration of about 0.17 mg/ml. This finding indicates that the aggregates of insulin are in equilibrium with the 5775 molecular weight monomer of insulin. The molecular weight data can be fit with a model involving monomers, dimers, hexamers, and polymers of the hexamer units. Calculations based on the equilibrium constant for dimer formation show that the mole ratio of monomer to dimer would be about 800,000 to 1 at physiological concentrations (about 0.1 ng/ml of serum). We found the self-association of proinsulin to be essentially identical to that of insulin. This suggests that the same sites of association are involved for both proinsulin and insulin, and that the connecting peptide does not interfere with the process of self-association. These data are also consistent with the earlier proposal [Frank, B. H., and Veros, A. J. (1968), *Biochem. Biophys. Res. Commun. 32*, 155] that the insulin moiety of proinsulin exists in the same conformation as does insulin itself.

Studies comparing the physical properties of insulin and proinsulin are of interest in terms of relating the conformations of the two proteins. Jeffery and Coates (1966a,b) examined the self-association of insulin at pH 2.0 and suggested that the species present at equilibrium were monomers, dimers, tetramers, and hexamers. Our laboratory previously reported (Frank and Veros, 1968) that proinsulin also self-associates at pH 2.0, and that its self-association can be described in terms of the same set of aggregates proposed for insulin by Jeffery and Coates. Further, the free energies of the association steps for proinsulin are similar to those reported for insulin (Jeffery and Coates, 1966a). Based on the similarity in association properties and the results of circular dichroic studies (Frank and Veros, 1968), we proposed that the insulin moiety of proinsulin is in the same conformation as that of the insulin molecule.

At pH 7.0 proinsulin and insulin also exhibit similar properties of aggregation in the presence of zinc (Grant *et al.*, 1972). Both proteins form a hexamer complex (two zinc atoms and six protein molecules) when the molar ratio of zinc to protein is between 0.33 and 1.0 and the protein concentration is above 0.5 mg/ml (Frank and Veros, 1970; Grant *et al.*, 1972). Little is known of the aggregation behavior of insulin and proinsulin under conditions more closely resembling the physiological situation, *i.e.*, neutral pH and little or no zinc. In fact, the molecular size of the physiologically active form of insulin is still open to question (Blundell *et al.*, 1971).

Materials and Methods

Single component porcine insulin and proinsulin were kindly supplied to us by Dr. R. E. Chance of the Lilly Research Laboratories. The buffer, 0.1 M NaCl–0.1 M Tris–0.001 M EDTA (pH 7.00 at 24.5°), was prepared from analytical grade reagents. EDTA was used as a precaution against the effects of accidental metal ion contamination. Protein concentrations were determined on a Cary 14R spectrophotometer at 276 nm using extinction coefficients of 1.05 and 0.667 ODU/cm (mg/ml) for insulin and proinsulin, respectively (Frank and Veros, 1968).

All sedimentation experiments were done on a Spinco Model E analytical ultracentrifuge using either interference optics or the photoelectric scanner. The double-sector centerpieces were made of either aluminum- or charcoal-filled epon. Overspeeding techniques (Richards *et al.*, 1968) were used to reach equilibrium in less than 24 hr. The Kodak II-G plates were analyzed on a Nikon 6C microcomparator. Synthetic boundary experiments were used to relate fringes to concentration. In the scanner runs, the chart paper was calibrated at 275 nm by means of tyrosine solutions of known optical

† From the Lilly Research Laboratories, Indianapolis, Indiana 46206.
Received June 19, 1972.

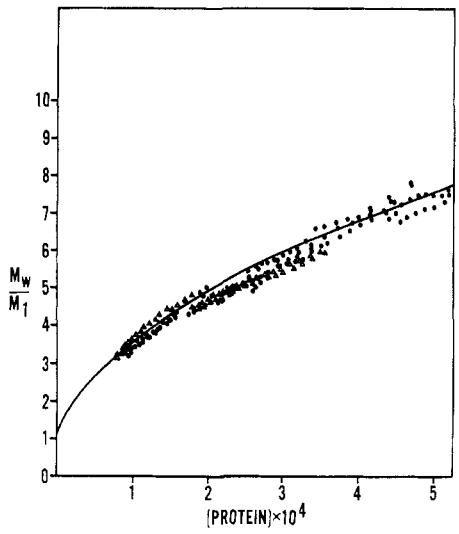

FIGURE 1: Comparison of insulin and proinsulin association at pH 7.00, 24.5°. (●) Insulin and (△) proinsulin. The data were obtained using the interference optical system of the Model E ultracentrifuge.

density. When the 12-mm centerpieces were used, a six-hole AN-G rotor containing the protein and four calibration cells was employed. At the lowest concentration, experiments were done using the two-hole AN-E rotor (30-mm centerpiece). The chart paper was calibrated afterward by a series of runs with calibration solutions in the cell.

Since the maximum protein concentration used was about 4 mg/ml and the pH of the study was very close to the isoelectric point of both proteins (Chance *et al.*, 1968), the solutions were assumed to be ideal. The weight-average molecular weights were calculated from

$$M_w = \frac{RT}{(1 - \bar{v}\rho)\omega^2} \frac{1}{r} \frac{1}{C} \frac{dC}{dr} \quad (1)$$

where T is the absolute temperature, $\bar{v}$ is the partial specific volume, ρ is the density, R is the gas constant, C is the concentration, r is the distance from the axis of rotation, and ω is the rotor speed. The value of 0.73 ml/g was used for the partial specific volume of both proteins (Frank and Veros, 1968).

Data points were taken at equally spaced radial intervals so that least-squares methods could be used for smoothing and differentiating the data (Savitsky and Golay, 1964).

The fringe values in the interference experiments were calculated using the equation for the conservation of mass in sector-shaped cells

$$J_m = J_0 - \int_{r_m}^{r_b} (J_r - J_m) dr^2/(r_b^2 - r_m^2) \quad (2)$$

where J_m is the meniscus fringe value, J_0 is the initial fringe value, J_r is the fringe value at radius r, r_b is the radius at the cell bottom, and r_m is the radius at the meniscus.

We have assumed a model for the insulin association at neutral pH which involves monomers, dimers formed from two monomers, hexamers formed from three dimers, and multiple aggregates of the hexamer. The reasons for this choice of aggregate forms are given in the Discussion section. Assuming that only these types of aggregates exist in solutions of insulin and proinsulin, we have been able to fit the experimental data shown in Figures 2 and 3. Denoting the monomer by I, the appropriate equations describing the model are:

$$2I \rightleftharpoons I_2 \qquad K_2 = [I_2]/[I]^2$$
$$3I_2 \rightleftharpoons I_6 \qquad K_6 = [I_6]/[I_2]^3$$
$$I_6 + I_6 \rightleftharpoons I_{12} \qquad K = [I_{12}]/[I_6]^2$$
$$I_{12} + I_6 \rightleftharpoons I_{18} \qquad K = [I_{18}]/[I_6][I_{12}]$$
$$I_{18} + I_6 \rightleftharpoons I_{24} \qquad K = [I_{24}]/[I_6][I_{18}]$$
$$\vdots \qquad \vdots$$
$$I_{6n} + I_6 \rightleftharpoons I_6(n+1) \qquad K = [I_6(n+1)]/[I_6][I_{6n}]$$

In these equations, an intrinsic equilibrium constant is used to describe the hexamer aggregation process. In processing the data, concentrations were converted from units of molarity to milligrams per milliliter. Upon defining new equilibrium constants $k_2 = K_2/M_1$, $k_6 = K_6/M_1^2$, and $k = K/M_1$, where M_1 is the monomer molecular weight, the expressions for the concentrations of aggregates are:

$$C_2 = 2k_2C_1^2$$
$$C_6 = 6k_6k_2^3C_1^6$$
$$C_{12} = 12kk_6^2k_2^6C_1^{12}$$
$$C_{18} = 18k^2k_6^3k_2^9C_1^{18}$$
$$C_{24} = 24k^3k_6^4k_2^{12}C_1^{24}$$
$$\vdots$$
$$C_{6n} = 6nk^{n-1}k_6^nk_2^{3n}C_1^{6n}$$

The total concentration is given by

$$C_T = C_1 + C_2 + \sum_{i=1}^{n} C_{6i}$$

while the reduced molecular weight is given by

$$\frac{M_w}{M_1} = \frac{C_1 + 2C_2 + \sum_{i=1}^{n} 6i \cdot C_{6i}}{C_T}$$

By assuming trial values for the upper limit of the summations and the three equilibrium constants, theoretical curves can be generated and the best fit to the experimental molecular weight data obtained.

Results

Figures 1, 2, and 3 summarize the molecular weight variation as a function of concentration for both proteins. As eq 1 shows, the accuracy of the molecular weights depends on the accuracy with which the concentrations can be determined. In the scanner runs, the precision of the data is limited by the small size of the chart paper and the noise in the pen trace. The use of the scanner "stair steps" to calibrate the chart paper can lead to serious errors and was avoided in the present study (Pekar *et al.*, 1971). Instead, multiple standard optical density tyrosine solutions were run simultaneously with the protein solution. In the interference runs, the plots of concentration *vs.* radius are smoother than in the scanner runs;

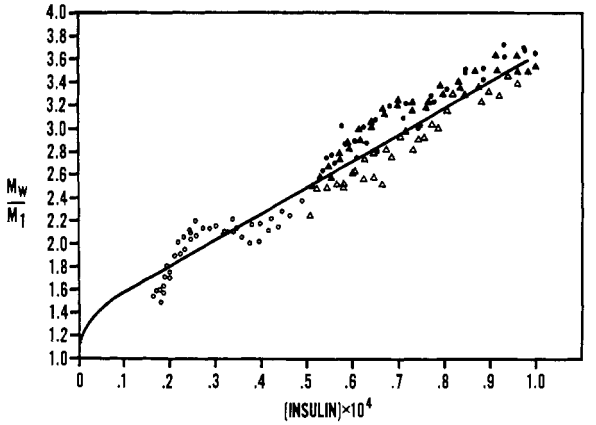

FIGURE 2: Reduced molecular weight of insulin as a function of concentration at pH 7.00, 24.5°. The data were obtained using the scanner unit of the Model E ultracentrifuge. Each symbol represents the data from a given experiment.

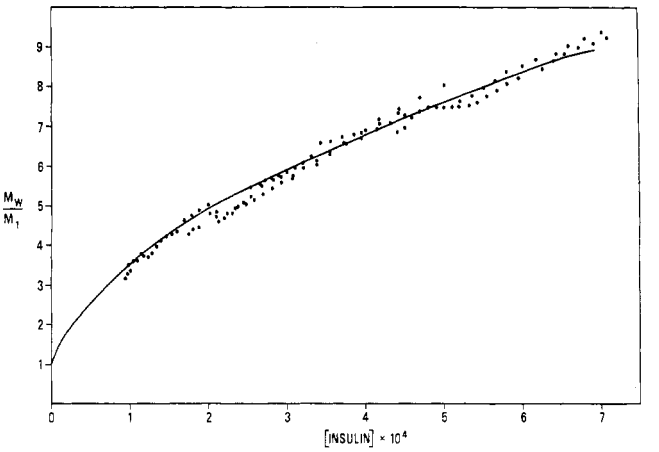

FIGURE 3: Reduced molecular weight of insulin as a function of concentration at pH 7.00, 24.5°.

this in turn is reflected in the lower degree of scatter in the molecular weights calculated using the interference data (Figure 1) than in the molecular weights derived from the scanner data (Figure 2). In the calculation of concentrations, the initial concentration fringe value J_0 in eq 2 must be determined. Figure 4 shows the relationship between fringes as determined from synthetic boundary cell runs and initial concentration. The spectroscopically determined concentrations are quite accurate, and the deviations of points from the line in large part reflect errors in the synthetic boundary runs. These errors can be very significant; their effect is to cause discontinuities of "cusps" in the plot of M_w *vs.* C_T. In our molecular weight calculations, J_0 was therefore determined from the optical density of the protein solution and the least squares slope from Figure 4.

The data in Figure 2 show that below about 0.17 mg/ml, the experimentally determined molecular weight of insulin becomes less than 11,550. Thus the insulin aggregates are in equilibrium with a monomer of mol wt 5775.

Since the monomer molecular weights of insulin and proinsulin are different, plots of M_w *vs.* C_T cannot be used to compare the aggregation of the two proteins. However, plots of M_w/M_1 *vs.* C_T/M_1 will coincide if the two materials aggregate identically.[1] These two quantities were calculated using values of M_1 = 5775 for insulin and 9080 for proinsulin (Chance *et al.*, 1968) and are shown in Figure 3. Clearly, the two proteins associate identically to within experimental error over the concentration range which was studied.

Since insulin was available in much larger quantities than proinsulin, its aggregation was studied more extensively and considered to be representative of both proteins. The data are summarized in Figures 1 and 2. Attempts to fit the experimental data to the monomer, dimer, hexamer, polymer model were fairly successful. Aggregates through "54-mer" were assumed to exist and the most satisfactory fit to the data is represented by the solid line. The values of the equilibrium constants are k_2 = 25 l./g, k_6 = 12 l.²/g², and k = 0.9 l./g

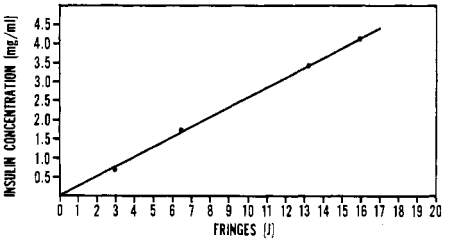

FIGURE 4: Relationship of insulin concentration to fringes determined in the synthetic boundary cell.

which correspond to values of K_2 = 1.4 × 10⁵ l./mole, K_6 = 4 × 10⁸ l./mole, and K = 5 × 10³ l./mole. Using the relationship $\Delta G^\circ = 2.303\ R \cdot T \cdot \log K$, the respective standard state free energies are −7, −12, and −5 kcal per mole.

Discussion

The results of the present study show that insulin and proinsulin self-associate essentially indentically at neutral pH. Frank and Veros (1968) made a similar observation under acid pH conditions; however, we find the degree of association to be much greater at neutral pH than at pH 2.0. For example, the highest M_w observed at pH 2.0 was 20,000 compared to 60,000 at pH 7.0 in the proinsulin solutions. The similarity of self-association of the two proteins suggests that the sites of association are also similar and that these sites reside solely with the insulin moiety. Further, the position of the connecting peptide[2] in proinsulin must be such that no marked interference with the sites occurs. The present data do not exclude the possibility that the connecting peptide may interfere with the formation of very large aggregates. This type of effect has been observed in zinc–proinsulin solutions. Though the presence of the connecting peptide apparently does not interfere with the formation of the zinc–proinsulin hexamer complex, it does prevent the formation of the large insoluble aggreagates produced under the same conditions in insulin solutions (Grant *et al.*, 1972).

The model chosen for fitting our experimental data involves

[1] The expressions for M_w and C_T are given by $C_T = \Sigma C_i = \Sigma M_i[P_i] = M_1 \Sigma i[P_i]$ and $M_w = \Sigma C_i M_i / C_T = M_1^2 \Sigma i^2[P_i]/C_T$. $[P_i]$ represents an aggregate with i monomer units. Both of the expressions become independent of monomer molecular weight when divided by M_1. Thus, $C_T/M_1 = \Sigma i[P_i]$ and $M_w/M_1 = \Sigma i^2[P_i]/\Sigma i[P_i]$. Clearly, if two materials of different monomer molecular weights aggregate identically, then plots of the latter two quantities against each other will coincide.

[2] The connecting peptide is the portion of the proinsulin chain connecting residues B30 and A1.

TABLE I: Aggregation State of Insulin at Physiological Concentrations.

Insulin (ng/ml)	[Monomer], M	[Dimer], M	[Monomer]/[Dimer]
1000	1.7×10^{-7}	2.2×10^{-9}	80
3[a]	5.2×10^{-10}	1.9×10^{-14}	26,000
0.1[b]	1.7×10^{-11}	2.2×10^{-17}	800,000

[a] Serum level under glucose challenge (Cahill, 1971). [b] Fasting serum level (Berson and Yalow, 1966).

monomers, dimers, hexamers, and aggregates of the hexamer. However, other models could probably be used to fit the experimental data since the accuracy of the data is not sufficient to allow one to decide between the closeness of fit of different models to the experimental data. We limited our choice of model by considering only those aggregate forms for which there was evidence of their existence in insulin solutions under various conditions (see discussion above for references). The use of an intrinsic equilibrium constant was based on the assumption that the energies of association (and the association constants) of possible multiple association sites would not differ greatly.

Our model fits the data (Figures 1 and 2) well throughout the concentration range used in these studies. The association constant for dimer formation derived from this curve is 1.4×10^5 l./mole, which corresponds to a ΔG° of -7 kcal/mole. Jeffery and Coates (1966a) obtained a ΔG° for insulin dimerization at pH 2.0 of -5.5 kcal/mole and Goldman (1971), working at pH 8.0 calculated a ΔG° of -7.4 kcal/mole. Goldman's value at pH 8.0 and ours at pH 7.0 are in excellent agreement. This agreement has further significance since Goldman's model involved monomers, dimers, tetramers, and hexamers and was therefore a different model than the one used in our work. The fact that ΔG° does not vary greatly with pH (*e.g.*, between 2 and 8) is in agreement with available knowledge of the interactions involved in dimer formation. The crystallographic studies of insulin structure (Blundell *et al.*, 1971) indicate that mainly nonpolar residues come into contact upon dimer formation. These contacts (van der Waal) involve the B12 valines and the B24 and B25 phenylalanines from both monomer units. In addition, a series of hydrogen bonds form a region of pleated sheet in the monomer interface region. Neither type of interaction would be expected to be particularly sensitive to pH changes over the pH 2–8 range.

The data in Figure 2 indicate that insulin solutions are in equilibrium with the 5775 molecular weight monomer. Additional evidence for the existence of the monomer–dimer association at neutral pH was obtained in earlier studies using ultraviolet difference spectroscopy (Frank and Veros, 1970). The amount of monomer and dimer at physiological concentrations of insulin can be calculated using the value of K_2 obtained in these studies (see Table I). Over the range of 0.1–3 ng of insulin/ml, the ratio of monomer to dimer varies from 800,000 to 1 to 26,000 to 1 and the actual concentration of dimer changes from 2×10^{-17} to 1.9×10^{-14} mole/l. The possible effect of zinc can be neglected since the available data (Grant *et al.*, 1972) indicate that the binding of zinc to insulin would be insignificant at these hormone concentrations. Although the present work cannot rule out the possibility that the active form of insulin is a dimer, we believe that these very low concentrations of dimer provide strong circumstantial evidence against this possibility. Therefore, we suggest that the functional role of self-association as well as zinc binding may be to ensure the efficient storage of this protein in the β cell.

Finally, the strong parallel in the self-association properties of insulin and proinsulin supports our earlier proposal that the conformation of the insulin moiety in proinsulin is essentially the same as that of insulin itself.

Acknowledgments

We thank Mr. Max M. Marsh and Dr. William W. Bromer for helpful discussions and comments on the manuscript.

References

Berson, S. A., and Yalow, R. S. (1966), *Amer. J. Med. 40*, 676.
Blundell, T. L., Dodson, G., Dodson, E., Hodgkin, D. C., and Vijayan, M. (1971), *Progr. Hormone Res. 27*, 1.
Cahill, G. F., Jr. (1971), *Diabetes 20*, 785.
Chance, R. E., Bromer, W. W., and Ellis, R. M. (1968), *Science 161*, 3837.
Frank, B. H., and Veros, A. J. (1968), *Biochem. Biophys. Res. Commun. 32*, 155.
Frank, B. H., and Veros, A. J. (1970), *Biochem. Biophys. Res. Commun. 38*, 284.
Goldman, J. (1971), Ph.D. Thesis, University of California, Berkeley.
Grant, P. T., Coombs, T. L., and Frank, B. H. (1972), *Biochem. J. 126*, 433.
Jeffery, P. D., and Coates, J. H. (1966a), *Biochemistry 5*, 489.
Jeffery, P. D., and Coates, J. H. (1966b), *Biochemistry 5*, 3821.
Pekar, A. H., Byers, R. A., Weller, R. E., and Frank, B. H. (1971), *Anal. Biochem. 42*, 516.
Richards, E. G., Teller, D. C., and Schachman, H. K. (1968), *Biochemistry 7*, 1054.
Savitsky, A., and Golay, M. J. E. (1964), *Anal. Chem. 36*, 1627.