Biochem. J. (1970) **118**, 355–365
Printed in Great Britain

Thermodynamic Studies of the Assembly *in vitro* of Native Collagen Fibrils

By ALAN COOPER*
Department of Theoretical Physics and Rheumatism Research Centre, University of Manchester, Manchester M13 9PL, U.K.

(*Received 7 January* 1970)

Measurements of the solubility of calf-skin tropocollagen in neutral phosphate buffers in the temperature range 20–37°C show that native collagen fibril formation is an endothermic process made thermodynamically favourable by a large positive entropy of precipitation associated with structural changes in the surrounding solvent. The effect of inorganic ions and small solute molecules on precipitation seems to be correlated with their structural effects on liquid water. Heterogeneity in the precipitation properties of the collagen solutions may be related to changes in the configurational entropy of the macromolecules due to intramolecular cross-linking.

The assembly *in vitro* of native collagen fibrils from tropocollagen solutions is one of the most widely studied examples of macromolecular systems in which macroscopic structures assemble spontaneously from a collection of identical molecular subunits. Self-assembly from subunits seems to be a common feature in the structural organization of many organisms (Kushner, 1969), possibly because this mechanism provides a biologically efficient way of utilizing genetic information for the construction of large intra- or extra-cellular structures (Caspar, 1963). The fundamental subunits of the collagen fibres of higher animals are the tropocollagen macromolecules. They may be extracted from young tissues by various gentle procedures (Wood, 1964; Piez, 1967), and their physical behaviour in solution is adequately described by a homogeneous collection of long thin semi-rigid rods about 2800 Å long and 13.6 Å in diameter, with a molecular weight of about 300000 (Boedtker & Doty, 1956; Rice, Casassa, Kerwin & Maser, 1964). They consist of three polypeptide chains (α-chains), of apparently equal length, with at least one differing in amino acid composition from the other two (Piez, 1967), and these lie side-by-side in a coiled-coil triple-helical arrangement. All the α-chains point in the same direction (Ramachandran, 1967) and there are short non-helical sections in the molecule, probably at the ends of the rod (Rubin *et al.* 1965; Piez, 1967).

Solutions of tropocollagen are stable at acid pH at low temperatures and ionic strength, but when the pH, ionic strength and temperature are raised to near-physiological conditions the collagen precipitates spontaneously, usually as an opaque rigid gel consisting of a matrix of fibrils indistinguishable in the electron microscope from the native fibrils seen in the tissue. The conditions required for the formation of native-type aggregates are not very strict, and for mammalian collagens gelation will occur in the pH range 5.5–8.5, with ionic strengths between 0.1 and 0.8 and temperatures 15–37°C (Gross & Kirk, 1958; Bensusan & Hoyt, 1958; Wood & Keech, 1960; Cassel, Mandelkern & Roberts, 1962). The rates of gelation are, however, very sensitive to these parameters and several studies of the kinetics of the process have been made (see review by Wood, 1964). Precipitation rates are normally observed as the increasing turbidity of the solutions as the gel develops, and Wood (1960) showed that the process may be adequately described in terms of standard crystal-growth kinetics in which there is a lag or nucleation phase followed by a rapid interface-controlled growth phase during which the nuclei develop into fibrils. Cassel *et al.* (1962) confirmed this and supported the concept that native fibril formation is a simple phase transition in which no gross change in macromolecular conformation occurs. The rate of precipitation increases rapidly with temperature up to the denaturation temperature of the collagen, and Cassel (1966) deduced from this that native aggregation is endothermic and therefore an entropy-driven process. Previous workers (Bensusan & Hoyt, 1958; Wood, 1960;

* Present address: Laboratory of Molecular Biophysics, Department of Zoology, Parks Road, Oxford OX1 3PS, U.K.

Bensusan & Scanu, 1960) had assumed that aggregation of the macromolecules involved an activation step and that the rate of precipitation at the surface of a fibril was controlled by the activation energy, which is not necessarily a measure of the total enthalpy change during precipitation. In principle the kinetics of aggregation may be controlled by both an activation step and by the final strength of the intermolecular interaction, and it is not possible to obtain from kinetics alone the equilibrium thermodynamic parameters of the assembly process.

Since macromolecular systems are subject to exactly the same physical laws as systems containing only small molecules their equilibrium properties must obey the general principles of classical thermodynamics (Tanford, 1961). Any understanding of the interactions between macromolecules in solution should be based on experimental knowledge of the thermodynamics of the system, and it is the purpose of the present paper to present some measurements of the free energy, enthalpy and entropy of the native collagen assembly process. The free energies of fibril formation may be calculated from measurements of the solubility of tropocollagen in aqueous buffer solutions by using a simple modification of the standard thermodynamic treatment of solutions. The method can only be used with confidence for very dilute solutions, but has been used extensively in the past to estimate the thermodynamic parameters of small poorly soluble molecules in water (see, for example, Némethy & Scheraga, 1962), and also to calculate the heats of solution of crystalline proteins (Cohn & Edsall, 1943).

Before presentation and discussion of the results for collagen it will be necessary to present the relevant theory and a discussion of some of the problems of protein solubility in general, though the relevance of some of the points will only become apparent later.

THEORY

The condition for thermodynamic equilibrium between several phases in contact is that the chemical potential, μ_i, of any component, i, must be the same in all the phases, i.e. $\mu_i^{(1)} = \mu_i^{(2)}$ if the phases 1 and 2 are in equilibrium. Thus in the specific case of a condensed phase in equilibrium with a saturated solution the chemical potential of the solute is the same whether it be in the solution or in the precipitate. The chemical potential of a component is defined as the change in free energy/mol of component involved in transferring an infinitesimal amount of the component from a standard state into the system, without changing any other components in the phase. The choice of a standard reference state for the potentials is important, though it need not correspond to any physically attainable state.

A collagen solution might ideally be represented as a two-component system of n_1 solvent molecules and n_2 rigid-rod-like macromolecules. We take a standard state in which the n_2 rods are distributed at fixed points in space (far enough apart to be non-interacting) with fixed orientations. The solvent molecules are in contact with the rods but free to move, and the overall concentration is the same as in the real solution. This hypothetical reference state is equivalent to a real solution in which all the macromolecules have been frozen in their instantaneous positions, leaving the solvent in its original state.

Defined in this way the chemical potential, μ_2', of the condensed fibrillar phase is a physically significant quantity representing the free-energy change/mol in assembling fibrils from a collection of widely separated macromolecules. The free-energy change/mol of transfer of n_2 macromolecules from the standard state to the real saturated solution containing n_1 solvent molecules is the solute potential μ_2, and this must equal μ'_2 because of the condition of thermodynamic equilibrium. If the solutions are sufficiently dilute that interactions between macromolecules may be neglected μ_2 consists only of the entropy changes involved in the transfer from the standard state to the real solution. This may be separated into two contributions: first, the increase in entropy due to mixing the macromolecules with the solvent, and, secondly, the entropy increase due to the freedom of the rods to take up any orientation, i.e.:

$$\mu_2 = \mu_2^{\text{mixing}} + \mu_2^{\text{orientation}}$$

The free energy of a concentrated mixture of rigid rods in monomeric solvents was studied by Flory (1956). He showed that such solutions were always isotropic (i.e., completely disordered) provided that:

$$v_2 < \frac{8}{x}\left(1 - \frac{2}{x}\right)$$

where v_2 is the volume fraction of the rods and x is their axial ratio. For tropocollagen $x \simeq 200$, giving isotropic mixtures for $v_2 < 4\%$. In practice collagen concentrations rarely exceed 0.1 g/100 ml, and the condition for complete disorder is well satisfied. In addition, the solutions are sufficiently dilute (10 μM or less) to be treated by ideal solution theory.

For ideal solutions the free energy of mixing of n_1 and n_2 molecules is:

$$F = -RT.\ln\left(\frac{(n_1+n_2)!}{n_1!\,n_2!}\right)$$

which, on using Stirling's approximation, gives:

$$\mu_2^{\text{mixing}} = \frac{\partial F}{\partial n_2} = RT.\ln \alpha_2$$

where $\alpha_2 = n_2/(n_1+n_2)$ = mole fraction of macromolecules in solution.

[This treatment effectively assumes that the solvent and solute molecules are the same size. This is obviously incorrect in the case of macromolecular solutions, but the correction due to excluded volume effects is negligible at the collagen concentrations met in practice (Cooper, 1969).]

The orientational part of the chemical potential is independent of the concentration and may be estimated

as follows. With one end of the rigid rod fixed the other end is free to take up any position on the surface of a sphere of radius L (where L is the length of the macromolecule). The additional free energy comes from the entropy of mixing of the end of one rod with the solvent molecules in a surface area $4\pi L^2$. If the mean cross-section area of a solvent molecule is a then:

$$\mu_2^{\text{orientation}} = -RT.\ln(4\pi L^2/a)$$

For water $a \simeq 10\,\text{Å}^2$, and for collagen $L \simeq 2800\,\text{Å}$. Allowing one order of magnitude either way for this rather crude estimate of the number of possible orientations:

$$\frac{-\mu_2^{\text{orientation}}}{RT} = 16 \pm 2$$

So the total chemical potential for a dilute solution of rigid-rod-like macromolecules at a mole fraction α_2 is:

$$\mu_2 = RT(\ln \alpha_2 - \beta)$$

where β is a constant dependent only on the molecular dimensions, and for collagen $\beta = 16 \pm 2$. The potential of the standard state has been arbitrarily defined as zero.

For a saturated solution of macromolecules in equilibrium with a precipitated phase α_2 is equal to the solubility α_s and μ_2 is the free-energy change ΔF/mol of rods when the rods condense from a widely separated array in the presence of solvent to a closely packed quasi-crystalline aggregate. This free energy will contain, not only those direct interactions between the macromolecules, but also any contributions due to the change in solvent–macromolecule interactions during precipitation.

Protein solutions are unlikely to be completely homogeneous in practice and the solubility α_s of the precipitating component must be obtained from a solubility diagram constructed from a series of measurements over a range of total protein concentrations (Cohn & Edsall, 1943). If the initial protein concentration is C then the supernatant concentration α' after precipitation to equilibrium is:

$$\alpha' = \alpha_s + bC \quad \text{(above saturation)}$$
$$\alpha' = C \quad \text{(below saturation)}$$

where b is the fraction of the total protein that remains completely soluble under the experimental conditions. A theoretical derivation of this behaviour is given in the Appendix, where it is shown to be true for the general case of precipitation from inhomogeneous solutions in which the quasi-crystalline phase may well consist of a co-precipitate of a mixture of the macromolecules from the solution. In this case the size of the non-precipitating fraction, b, and the composition of the co-precipitate may depend on experimental conditions.

For rigid-rod-like molecules the mean free energy of interaction during precipitation is given by:

$$\Delta F = RT(\ln \alpha_s - \beta)$$

Provided that the composition of the precipitate does not change with temperature the entropy and enthalpy of the process may be obtained:

$$\Delta S = -\frac{\partial}{\partial T}\Delta F$$

$$\Delta H = -T^2.\frac{\partial}{\partial T}\left(\frac{\Delta F}{T}\right)$$

An additional problem arises with collagen in that precipitated gels seem to become more insoluble with time (Gross, 1958, 1964), probably by the slow spontaneous formation of covalent intermolecular cross-links between adjacent macromolecules (Piez, 1968). This makes it difficult to demonstrate the reversibility of fibril formation since a large portion of the heat-precipitated gel will not redissolve on cooling, and solution equilibrium cannot be approached from both directions. It has been shown, however, that at least part of the gel will redissolve on cooling (Fessler, 1960*a*; Wood, 1962; Bannister, 1969). Necessary conditions for true thermodynamic equilibrium are that precipitation should obey the phase (Cohn & Edsall, 1943) and that protein solubility should be independent on the length of incubation of solution with precipitate (after the initial precipitation is complete). In the experiments reported below both these conditions seemed to be satisfied, indicating that although the overall process of fibril formation is made irreversible by slow cross-linking of the gel, the initial assembly process may be considered reversible involving only non-covalent interactions.

MATERIALS AND METHODS

Collagen solutions. The two preparations of tropocollagen used in this work were obtained by cold acetic acid-extraction of the dermis of a 1-week-old calf (Gross & Kirk, 1958; Bensusan & Hoyt, 1958; Rice *et al.* 1964; Jackson & Cleary, 1967). Exhaustive extraction of the tissue with neutral M-NaCl containing thiomersal (Merthiolate; 0.01 g/100 ml) as antibacterial agent continued for about a week, and was followed by four successive 2-day extractions with 0.2 M-acetic acid. Pooled acid-extracted fractions, salted out with solid NaCl to 20% (w/v), were redissolved in 1% (v/v) acetic acid. Dialysis against 1% acetic acid was followed by centrifugation at $80000 g_{av.}$ for 1 h. The resulting solutions (preparation A) had a reduced viscosity of 10.8 dl/g (measured at 25°C in a Cannon–Fenske capillary viscometer). The residual dermis was drained and stored in the deep-freeze.

Collagen preparation B was obtained by further acid extraction of the dermis after storage for several months in the deep-freeze. The thawed dermis was briefly macerated and extracted with two changes of cold 0.2 M-acetic acid. The pooled extracts, salted out and redissolved as before, were neutralized with M-NaOH and precipitated with precooled ethanol added to a final concentration of 14% (v/v). This last step removes contaminating glycoprotein (J. Gross, unpublished work reported by Jackson & Cleary, 1967). The precipitate, redissolved and dialysed against 1% acetic acid, was centrifuged at $80000 g_{av.}$ for 1 h before use.

Both preparations sedimented as single hypersharp boundaries in the analytical ultracentrifuge.

Absolute protein concentrations were estimated by micro-Kjeldahl determination of nitrogen, the nitrogen content of soluble collagen being taken as 17.7% (Bowes, Elliott & Moss, 1955). The hydroxyproline content of solutions was measured by the overnight procedure of Bergman & Loxley (1963) after hydrolysis in 6 M-HCl (sealed glass tubes, 110°C, 24–36 h). Preparations A and B contained 12.1 and 13.2 g of hydroxyproline/100 g of protein respectively.

Measurement of collagen solubility. Solubility diagrams for tropocollagen in neutral phosphate buffers in the temperature range 20–37°C were constructed as follows. Batches of the stock solution, at concentrations of 800–1000 μg/ml in 1% acetic acid, dialysed against cold water until the acid concentration was less than 0.01%, were diluted with samples of the final dialysis medium to give a series of five different protein concentrations in the range 50–800 μg/ml. Then 3 ml samples of each concentration were pipetted into stoppered test tubes, followed by 1 ml of concentrated buffer solution made up so as to give the desired final concentration, and the contents of the tubes were mixed immediately by brief agitation. All manipulation was performed in the cold-room. The stoppered tubes, transferred to water baths at temperatures in the range 20–37°C, were normally incubated for about 36 h, though experiments with incubation times between 10 and 72 h gave similar results, within experimental error, except for occasional samples at low temperature which required at least 24 h for complete precipitation. Collagen solutions precipitated as rigid gels, or in the more dilute solutions as fibrous suspensions. The gels were broken up with a glass rod, to free any trapped air bubbles and to disperse the gel, and the precipitates were spun off at approx. 2000*g*. After centrifugation 1 ml samples of the supernatants (two samples of each) were transferred to separate tubes for protein determination by the method of Lowry, Rosebrough, Farr & Randall (1951), as modified by Oyama & Eagle (1956). This colorimetric method was calibrated against collagen solutions of known protein concentration, obtained by micro-Kjeldahl nitrogen determination, and could detect concentrations as low as 5 μg/ml. Precipitation could be examined in up to four different buffer systems simultaneously, each at four (occasionally five) different temperatures, and one of the four buffers was always a control. Concentrated buffer solutions were made up fresh for each experiment and the control buffer throughout was KH_2PO_4–Na_2HPO_4, made up to give a final ionic strength 0.2 and pH 7.3 ($\pm$0.05). Reagents were of A.R. grade where available. Representative samples of most of the collagen precipitates were taken for examination by electron microscopy.

RESULTS

Examples of the solubility diagrams obtained are shown in Fig. 1. Data were fitted by least-squares analysis to linear solubility plots. (Attempts at fitting the data to quadratic curves gave no significant improvement. Occasional experiments at the lowest temperature gave only two points on the upper linear portion of the diagram. Extrapolations from such plots were given very low weight in the free energy curves.) The results are as expected from the phase rule for a single precipitating protein in the presence of a completely soluble non-precipitating fraction (Cohn & Edsall, 1943), though the magnitude of this fraction varies with temperature and with the buffer. Extrapolated solubilities lie in the range 0–300 μg/ml, depending on buffer and temperature, with standard deviations usually less than 3 μg/ml. The solubility decreases rapidly as the temperature is raised. Presentation in detail as free-energy curves (Figs. 2–6) shows that ΔF becomes more negative and the native fibril becomes more stable with increasing temperature. A characteristic feature is the curvature of the plots of ΔF versus T. For the detailed data of Figs. 2–4 this curvature is statistically significant, and the solid curves in these figures were obtained from weighted least-squares fits of the data from several independent experiments to a quadratic equation. The entropy and enthalpy of the process were computed from the parameters of this equation.

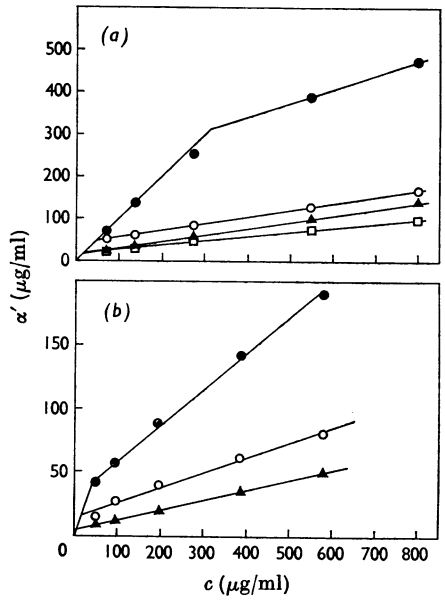

Fig. 1. Examples of solubility plots obtained for tropocollagen in KH_2PO_4–Na_2HPO_4 buffer, pH 7.3 and I 0.2. (*a*) Preparation A: ●, 25°C; ○, 30°C; □, 33°C; ▲, 37°C. (*b*) Preparation B: ●, 25°C; ○, 29°C; ▲, 37°C.

The size of the non-precipitating fraction (i.e. the slope, b, of the solubility diagrams) also varies characteristically with temperature (Figs. 2*a*, 3*a* and 4*a*), initially decreasing as T is raised but, after passing through a minimum, rising as T approaches 37°C. This behaviour indicates that the fibrils formed are in fact a co-precipitate of a mixture of slightly differing macromolecules, and that the composition of this co-precipitate varies slightly with temperature. Under these conditions the breakdown of ΔF into ΔH and ΔS components is not strictly valid, though the errors involved are probably quite small. This problem is discussed in more detail in the Appendix.

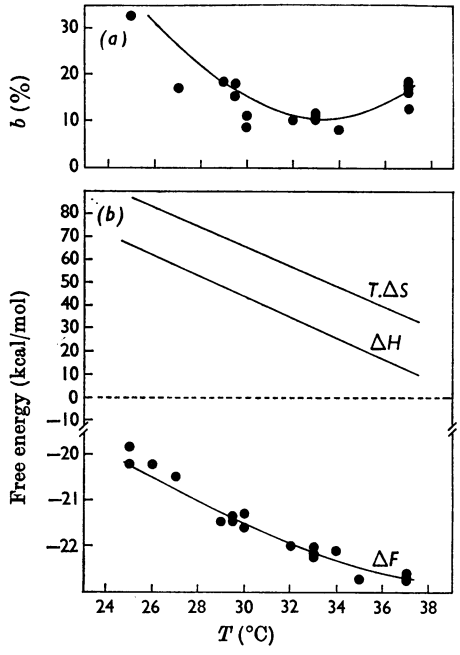

Fig. 2. Collagen preparation A in KH_2PO_4–Na_2HPO_4 buffer, pH 7.3 and I 0.2. (*a*) Variation with temperature of the non-precipitating fraction *b*. (*b*) Free energy of precipitation, with the enthalpy and entropy computed from the weighted least-squares equation:

$$\Delta F = 722.5 - 4.96\,T + 0.00738\,T^2 \text{ kcal/mol}$$

(±0.2 kcal/mol, with 95% confidence; T in °K).

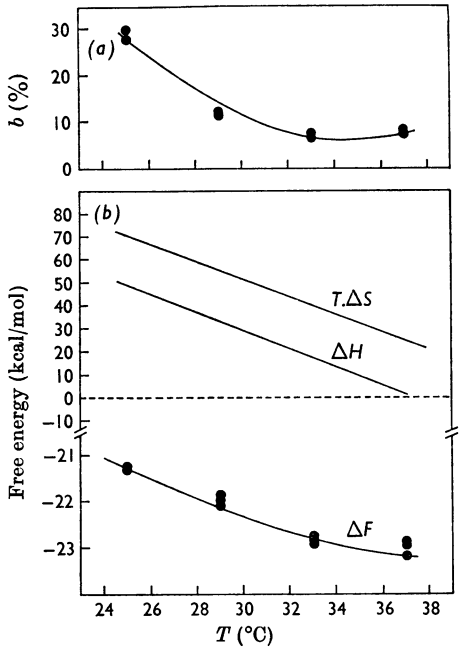

Fig. 3. Collagen preparation B in KH_2PO_4–Na_2HPO_4 buffer, pH 7.3 and I 0.2. (*a*) Variation with temperature of the non-precipitating fraction *b*. (*b*) Free energy of precipitation, with the enthalpy and entropy computed from the weighted least-squares equation:

$$\Delta F = 636.6 - 4.09\,T + 0.00647\,T^2 \text{ kcal/mol}$$

(±0.2 kcal/mol, with 95% confidence; T in °K).

Neither the solubility nor the size of the non-precipitated fraction varied significantly with incubation times between 9 and 72 h, except for occasional samples at low temperature, which required at least 24 h for complete precipitation. This indicates that reasonable equilibrium was achieved and that bacterial attack during incubation was negligible.

For collagen preparation A the free energy of fibril formation, in phosphate buffer (KH_2PO_4–Na_2HPO_4) of ionic strength 0.2 and pH 7.3, lies in the range −20 to −23 kcal/mol (Fig. 2*b*) and the computed entropy and enthalpy of the interaction are both positive, though they decrease with temperature. The free-energy curve of sample B under identical conditions parallels that of sample A, but is displaced downwards by about 0.6 kcal/mol; the ΔS and ΔH values are also decreased slightly, but show the same temperature variation (Fig. 3*b*).

Fibril formation is very sensitive to the concentration of the buffer salts and the solubility rises rapidly with increasing phosphate ionic strength. At low ionic strengths (I 0.05 and 0.1) the extrapolated solubilities were too low to be determined with confidence and the resultant free energies can only be crudely indicated by the shaded regions of Fig. 6. The gels formed at these low ionic strengths were weak and almost water-clear, and electron micrographs showed them to consist of loose amorphous precipitates with only slight indications of band–interband periodicities, in contrast with the typical native fibrils observed under other conditions. In general ΔF, ΔH and ΔS were all increased with increasing ionic strength of the phosphate buffer up to I 0.25.

Precipitation was relatively insensitive to pH in the range 6.7–7.8 (Fig. 5), ΔF decreasing slightly as the phosphate buffer pH increases at constant ionic strength (0.2).

Replacing most of the phosphate by chloride ions, while keeping the total ionic strength and pH constant, resulted in a significant increase in fibril stability (i.e. ΔF more negative), with a simultaneous decrease in ΔH and ΔS (Fig. 4*b*). Comparison of a series of anions at constant ionic strength showed that the free energy was sensitive

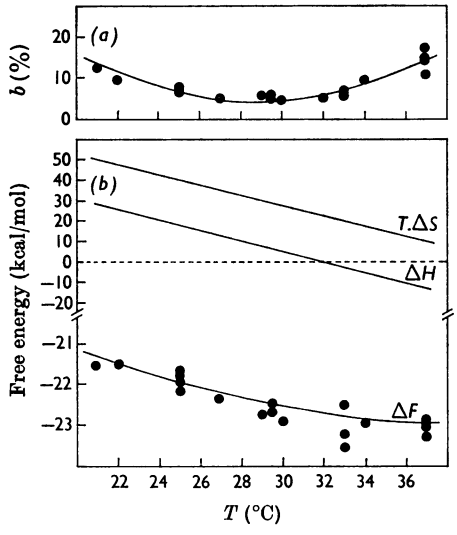

Fig. 4. Collagen preparation A in KH_2PO_4–Na_2HPO_4–NaCl buffer, pH 7.3 and I 0.05 (phosphate)+0.15 (NaCl). (*a*) Variation with temperature of the non-precipitating fraction *b*. (*b*) Free energy of precipitation, with the enthalpy and entropy computed from the weighted least-squares equation:

$$\Delta F = 399.3 - 2.69T + 0.00429T^2 \text{ kcal/mol}$$

(±0.4 kcal/mol, with 95% confidence; T in °K).

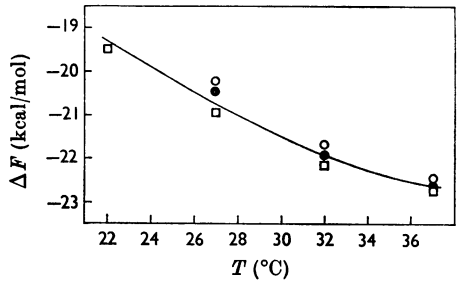

Fig. 5. Free energy of precipitation of collagen preparation A in KH_2PO_4–Na_2HPO_4 buffer, I 0.2; ○, pH 6.7; ●, pH 7.3; □, pH 7.8. The solid curve is the least-squares fit of the accumulated data of Fig. 2 for pH 7.3.

to the type of anion present in the buffer and that the anions could be ranked in a series according to their effectiveness in increasing fibril stability (i.e. decreasing ΔF): $Br^- > Cl^-$, $F^- > SO_4^{2-}$ > phosphate (at pH 7.3 and I 0.2, with Na^+ as common cation) (see Table 1). In all these cases a decrease in ΔF was coupled with a simultaneous decrease in both ΔH and ΔS, though they still remained positive over most of the temperature range. Not enough

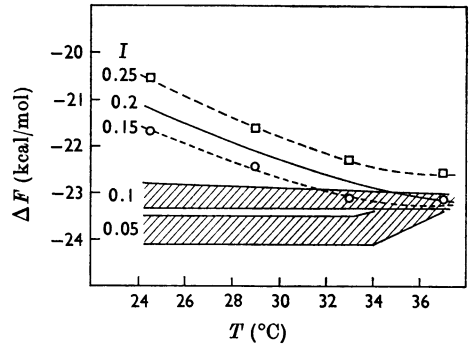

Fig. 6. Effect of phosphate ionic strength on the free energy of fibril formation of preparation B in KH_2PO_4–Na_2HPO_4 buffer, pH 7.3. The shaded areas indicate the probable ranges of the free energies at low ionic strengths, where the solubilities are too low to be estimated with confidence. The solid curve is the least-squares fit of the accumulated data of Fig. 3 for I 0.2.

Table 1. *Effect of different anions on the free energies of collagen preparation B at constant ionic strength*

The buffer in each case was KH_2PO_4–Na_2HPO_4, pH 7.3 and I 0.05, made up to I 0.2 with the respective sodium salt (except in the case of phosphate alone in which KH_2PO_4–Na_2HPO_4 buffer, pH 7.3 and I 0.2, was used).

Anion	ΔF at 25°C (kcal/mol)	ΔF at 37°C (kcal/mol)
Br^-	<−23.7	−23.59 (±0.10)
Cl^-	−22.71 (±0.08)	−23.30 (±0.05)
F^-	−22.73 (±0.08)	−23.40 (±0.11)
SO_4^{2-}	−21.63 (±0.06)	−23.20 (±0.06)
Phosphate	−21.26 (±0.06)	−22.93 (±0.08)

results were collected to allow meaningful estimates of the magnitudes of these parameters to be made. A comparison of a series of cations could not be made because of technical difficulties associated with the Folin–Lowry reaction (for details see Cooper, 1969); however, replacement of Na^+ ions by K^+ ions in the buffer produced no significant change in the free-energy curve at pH 7.3 and I 0.2.

The addition of 0.2M-methanol to the neutral phosphate buffer at I 0.2 produced no change in ΔF, though 0.2M-propan-2-ol did increase the free energy slightly and induced a small rise in the entropy and enthalpy values.

It was noted throughout these experiments that those factors that increased the stability of the fibrils also increased their apparent rate of formation with respect to the controls.

DISCUSSION

The solubility behaviour of tropocollagen confirms that fibril formation in neutral phosphate buffers is generally an endothermic process (ΔH positive), which is made thermodynamically favourable by a large positive entropy contribution, presumably arising from structural reorganization of the solvent during precipitation. The observed enthalpy changes are of similar magnitude to the activation energies obtained from kinetic data (Bensusan & Hoyt, 1958; Wood, 1960), and the effect of different anions on ΔH is the same as that seen by Bensusan & Hoyt (1958) for the activation energy. This sensitivity of the thermodynamic parameters to the type of anion is striking and parallels the effect seen in the precipitation-rate studies of Gross & Kirk (1958) and Bensusan & Hoyt (1958), who noted that the order of effectiveness of various ions resembled the Hofmeister series in reverse. A similar parallel occurs with the effect of ionic strength, where increasing the concentrations of inorganic salts increases the solubility and retards precipitation rates. The relative insensitivity of the free energy to pH near neutrality seems reasonable in view of the isoelectric plateau for collagen in this region (Bowes & Kenten, 1948), and the various, and often contradictory, reports of the effects of pH on precipitation kinetics are more readily explained in terms of the simultaneous variations in relative and absolute concentrations of the different buffer ions as the pH of the buffer system is changed (Cooper, 1969).

The importance of solvent structure in molecular interactions has been recognized for some time (Lumry & Biltonen, 1969), and interactions involving both polar and non-polar molecules in water can be stabilized by the large positive entropy contributions coming from structural changes in the surrounding liquid (Kauzmann, 1959). For collagen the plots of ΔF versus T, and the associated curves for ΔH and ΔS, are qualitatively similar to those obtained from the solubility of small non-polar molecules in water (Némethy & Scheraga, 1962), indicating that the dominant contribution in the fibril formation process comes from hydrophobic interactions between non-polar regions of adjacent molecules. This interpretation depends on the slight curvature of the free-energy plots, which may be suspect because of possible changes in fibril composition with temperature (see the Appendix); however, hydrophobic interactions have been favoured in the past as the dominant mechanism (Fessler & Hodge, 1962; Cassel, 1966; Cassel & Christensen, 1967), and the thermodynamic evidence now supports this hypothesis. Bello & Bello (1963) showed that rabbit-skin tropocollagen was precipitated more readily in phosphate buffers made up in deuterium oxide, consistent with the conclusions of Kreshek, Schneider & Scheraga (1965) that hydrophobic bonds between the side chains of non-polar amino acids are stronger in deuterium oxide than in water. Similarly, urea is known to weaken hydrophobic bonds (Bruning & Holtzer, 1961; Steinberg & Scheraga, 1962) and also to reversibly inhibit collagen precipitation (Gross & Kirk, 1958).

In spite of the apparent dominance of non-polar interactions the tropocollagen molecule is by no means completely apolar, and its polar characteristics must make some contribution to the overall interaction, though it is not possible to extract this from the data. The increase in solubility with increasing ionic strength is reminiscent of the salting-in phenomenon of dipolar molecules in aqueous salt solutions; the order of effectiveness of the various anions for 'salting-in' collagen is, however, the reverse of that usually observed for small dipolar ions (Cohn & Edsall, 1943). Moreover, the conventional theories of salting-in (Edsall & Wyman, 1958) predict that a salting-in term would make negative contributions to both the entropy and enthalpy of precipitation and that ΔH and ΔS should decrease with increasing ionic strength. This is the reverse of the effect observed here with increasing phosphate ionic strength. Nor can this be explained in terms of non-polar effects. The addition of simple inorganic ions to aqueous solutions of small non-polar molecules invariably results in salting-out, even for dilute salt solutions (Long & McDevit, 1952), and the ions may be ranked in the familiar Hofmeister series in order of salting-out efficiencies.

Thus none of the conventional concepts seems to fit the observed effects of dilute salts on collagen precipitation and, though a complete explanation must await a better understanding of biomolecular forces, there is a phenomenological approach that seems worthy of discussion and that may point the way to a more complete understanding. Many of the interactions between molecules in water depend on the unique structural properties of this liquid, and it seems a reasonable hypothesis to suppose that the presence of other molecules in solution might only affect these interactions indirectly through their effect on the structure of the water as a whole. Empirically the effects of adding various solutes to water have been described in terms of 'structure-making' or 'structure-breaking', depending on how the solute molecules affect the apparent structural organization of the water (for review see Kavanau, 1964). Bernal & Fowler (1933) introduced the concept of the 'structural temperature' of an aqueous solution, being the temperature at which pure water would have the same internal

structure as the solution. The notion is not entirely valid, since the structural changes introduced by solutes are not necessarily the same as those due to change in temperature; however, the concept will be useful. Thus structure-making solutes are said to decrease the structural temperature whereas structure-breakers increase it. The hypothesis is that changing the structural temperature of a buffer solution should have a similar effect on collagen precipitation as an equivalent change in the thermodynamic temperature, i.e. the addition of structure-breaking salts raises the structural temperature and thereby increases fibril stability (ΔF more negative) while decreasing both ΔH and ΔS in accord with the observed temperature effect, and the addition of structure-makers should have the reverse effect. These expectations seem to be borne out. Phosphate and sulphate ions are strong structure-makers (Krestov, 1962*a*,*b*), with sulphate slightly the less efficient, and the Na^+ and K^+ ions are weak structure-makers, though K^+ ions are sometimes classified as weak structure-breakers (Kavanau, 1964). Data from many sources, reviewed by Kavanau (1964), show that most univalent anions are structure-breakers, and the sequence of ions in order of increasing the structural temperature (and, I propose, fibril stability) closely resembles the Hofmeister series but in reverse. A similar correlation exists for aliphatic alcohols in water, which are structure-makers at low concentrations but become structure-breaking as the concentration is increased (Frank & Evans, 1945). Bensusan (1960) found that collagen precipitation rates were initially diminished by the addition of aliphatic alcohols, but after passing through a minimum the rates rose rapidly with increasing alcohol concentration, and the initial degree of inhibition increased with increasing hydrocarbon chain length of the alcohol, in parallel with the greater structure-making properties of the longer aliphatic chains. This is supported by the observation here of the slight destabilizing effect of 0.2M-propan-2-ol coupled with a small increase in ΔH and ΔS. The reversible inhibition of precipitation by dilute urea (Gross & Kirk, 1958) is consistent with the observation that urea is a structure-making solute (Abu-Hamidyyah, 1965).

Thus there seems strong circumstantial evidence that native collagen fibrils are stabilized by structural reorganization of the surrounding solvent during precipitation. The usual salting-in and salting-out effects are apparently suppressed, and the presence of other solutes seems to influence the interaction only indirectly via their effect on the structural properties of the bulk solvent.

The nature of the inhomogeneities that give rise to the variable non-precipitating fraction can only be guessed. Fessler (1960*a*,*b*) showed that tropocollagen solutions could be inhomogeneous in their precipitating properties, and later work (Wood, 1962; Bannister, 1969) indicates that this probably arises from differences in intramolecular cross-linking in the macromolecules. The differences in interaction properties may be related to the chemical modifications introduced during the process of cross-linking, or possibly to the structural changes involved when the covalent links are formed in the non-helical regions of the macromolecule (Piez, 1968). The increase in b at higher temperatures may be due to partial denaturation of some of the collagen, since the denaturation temperature of calf-skin tropocollagen is about 40°C. Partial unfolding of the macromolecule can be detected at temperatures well below denaturation (Piez, 1967), and Fessler (1960*a*,*b*) described a tropocollagen fraction that would not form a precipitate at 37°C, having a decreased specific optical rotation and that apparently contained no intramolecular cross-links (Wood, 1962).

One possible way in which intramolecular cross-links might affect fibril formation is by their effect on the configurational entropy of the macromolecule in solution. Up to this point it has been assumed that the molecules behave as perfectly rigid rods, but in reality the rods will have a degree of flexibility that will contribute to their entropy in solution. In particular there are regions of the α-chains at each end of the macromolecule that cannot fit into the triple-helical polyproline-like helix characteristic of the rest of the molecule (Piez, 1967), which are susceptible to attack by proteolytic enzymes and which seem to take up random-coil-like configurations (since the rest of the molecule is only susceptible to those enzymes after denaturation). During precipitation the conformational freedom of these non-helical regions would be diminished by the topological constraints of the surrounding fibril, and there will be a negative (destabilizing) contribution to the net entropy of fibril formation. The introduction of a covalent cross-link between non-helical regions of adjacent α-chains (Piez, 1968) would decrease the initial configurational entropy of the molecule and thus decrease the entropy lost during precipitation, the net effect being a further stabilization of the fibril. The studies by Wood (1962) and D. W. Bannister (personal communication) do show that the more cross-linked fractions of tropocollagen precipitate more rapidly and more completely than the less cross-linked molecules, and it seems possible that intramolecular cross-linking *in vivo* might act as a subtle control mechanism in the growth and stability of developing fibrils simply by decreasing the conformational entropy of the macromolecules in solution.

It is a pleasure to thank Dr J. A. Chapman and Dr A. Herzenberg for their help and guidance during this work, Miss J. Weiss for biochemical advice, Dr D. W. Bannister for discussion of his work before publication, and Mrs W. Staley for technical assistance. I am grateful to Professor S. F. Edwards and Professor J. H. Kellgren for the hospitality in their departments and to the Science Research Council for financial support.

REFERENCES

Abu-Hamidiyyah, M. (1965). *J. Phys. Chem., Ithaca*, **69**, 2720.
Bannister, D. W. (1969). *Biochem. J.* **113**, 419.
Bensusan, H. B. (1960). *J. Am. chem. Soc.* **82**, 4995.
Bensusan, H. B. & Hoyt, B. L. (1958). *J. Am. chem. Soc.* **80**, 719.
Bensusan, H. B. & Scanu, A. (1960). *J. Am. chem. Soc.* **82**, 4990.
Bello, J. & Bello, H. R. (1963). *Nature, Lond.*, **197**, 77.
Bergman, I. & Loxley, R. (1963). *Analyt. Chem.* **35**, 1961.
Bernal, J. D. & Fowler, R. H. (1933). *J. chem. Phys.* **1**, 515.
Boedtker, H. & Doty, P. (1956). *J. Am. chem. Soc.* **78**, 4267.
Bowes, J. H., Elliott, R. G. & Moss, J. A. (1955). *Biochem. J.* **61**, 143.
Bowes, J. H. & Kenten, R. H. (1948). *Biochem. J.* **43**, 358.
Bruning, W. & Holtzer, A. (1961). *J. Am. chem. Soc.* **83**, 4865.
Caspar, D. L. D. (1963). *Adv. Protein Chem.* **18**, 37.
Cassel, J. M. (1966). *Biopolymers*, **4**, 989.
Cassel, J. M. & Christensen, R. G. (1967). *Biopolymers*, **5**, 431.
Cassel, J. M., Mandelkern, L. & Roberts, D. E. (1962). *J. Am. Leath. Chem. Ass.* **57**, 556.
Cohn, E. J. & Edsall, J. T. (1943). *Proteins, Amino Acids and Peptides*, pp. 576–584. New York: Reinhold Publishing Corp.
Cooper, A. (1969). Ph.D. Thesis: University of Manchester.
Edsall, J. T. & Wyman, J. (1958). *Biophysical Chemistry*, vol. 1, pp. 296–320. New York: Academic Press Inc.
Fessler, J. H. (1960*a*). *Biochem. J.* **76**, 452.
Fessler, J. H. (1960*b*). *Biochem. J.* **76**, 463.
Fessler, J. H. & Hodge, A. J. (1962). *J. molec. Biol.* **5**, 446.
Flory, P. J. (1956). *Proc. R. Soc.* A, **234**, 73.
Frank, H. S. & Evans, M. W. (1945). *J. chem. Phys.* **13**, 507.
Gross, J. (1958). *Nature, Lond.*, **181**, 556.
Gross, J. (1964). *Science, N.Y.*, **143**, 960.
Gross, J. & Kirk, D. (1958). *J. biol. Chem.* **233**, 355.
Jackson, D. S. & Cleary, E. G. (1967). In *Methods of Biochemical Analysis*, vol. 15, p. 25. Ed. by Glick, D. New York: Academic Press Inc.
Kauzmann, W. (1959). *Adv. Protein Chem.* **14**, 1.
Kavanau, J. L. (1964). *Water and Solute–Water Interactions*, pp. 52–66. San Francisco: Holden–Day Inc.
Kreshek, G. C., Schneider, H. & Scheraga, H. A. (1965). *J. phys. Chem., Ithaca*, **69**, 3132.
Krestov, G. A. (1962*a*). *Zh. strukt. Khim.* **2**, 402.
Krestov, G. A. (1962*b*). *Zh. strukt. Khim.* **3**, 137.
Kushner, D. J. (1969). *Bact. Rev.* **33**, 302.
Long, F. A. & McDevit, W. F. (1952). *Chem. Rev.* **51**, 119.
Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951). *J. biol. Chem.* **193**, 265.
Lumry, R. & Biltonen, R. (1969). In *Biological Macromolecules*, vol. 2, p. 65. Ed. by Timasheff, S. N. & Fasman, G. D. New York: Marcel Dekker Inc.
Némethy, G. & Scheraga, H. A. (1962). *J. chem. Phys.* **36**, 3401.
Oyama, V. I. & Eagle, H. (1956). *Proc. Soc. exp. Biol. Med.* **91**, 305.
Piez, K. A. (1967). In *Treatise on Collagen*, vol. 1, p. 207. Ed. by Ramachandran, G. N. London: Academic Press (Inc.) Ltd.
Piez, K. A. (1968). *A. Rev. Biochem.* **37**, 547.
Ramachandran, G. N. (1967). In *Treatise on Collagen*, vol. 1, p. 103. Ed. by Ramachandran, G. N. London: Academic Press (Inc.) Ltd.
Rice, R. V., Casassa, E. F., Kerwin, R. E. & Maser, M. D. (1964). *Archs Biochem. Biophys.* **105**, 409.
Rubin, A. L., Drake, M. P., Davison, P. F., Pfahl, D., Speakman, P. T. & Schmitt, F. O. (1965). *Biochemistry, Easton*, **4**, 181.
Steinberg, I. Z. & Scheraga, H. A. (1962). *J. Am. chem. Soc.* **84**, 2890.
Tanford, C. (1961). *Physical Chemistry of Macromolecules*, p. 180. New York: John Wiley and Sons Inc.
Wood, G. C. (1960). *Biochem. J.* **75**, 598.
Wood, G. C. (1962). *Biochem. J.* **84**, 429.
Wood, G. C. (1964). *Int. Rev. conn. Tissue Res.* **2**, 1.
Wood, G. C. & Keech, M. K. (1960). *Biochem. J.* **75**, 588.

APPENDIX

Solubility of Co-precipitating Mixtures

Consider an isotropic solution consisting of a mixture of almost identical rigid-rod-like macromolecules, N_i moles of molecular species i ($i = 1 \ldots m$), and N_0 solvent molecules. The free energy of this system, with respect to the standard state defined previously, is:

$$-\boldsymbol{RT}.\ln\left(\frac{\left(\sum_{i=0}^{m} N_i\right)!}{\prod_{i=0}^{m} N_i!}\right)-\beta\boldsymbol{RT}.\sum_{i\neq 0} N_i$$

The first term is the entropy of mixing of all the molecules, and the second the entropy of disorientation of the macromolecules.

Under suitable conditions this solution will be unstable and a precipitate will form to give thermodynamic equilibrium between the quasi-crystalline precipitate and the supernatant solution. Many, if not all, of the individual species may be too dilute to form a separate phase of their own, but if the macromolecular species are sufficiently alike they may combine to form a uniform co-precipitate of a

mixture of macromolecules. The composition of the co-precipitate will be such as to minimize the free energy of the whole system and will be independent of the total concentration of macromolecules (since free energy is an intensive quantity in a uniform system). Assuming one such co-precipitate forms leaving a supernatant containing a mixture n_i mol of each species i, the free energy of the whole is now:

$$\sum(N_i-n_i)\mu_i - \boldsymbol{R}T.\ln\left(\frac{[\sum(N_i-n_i)]!}{\Pi(N_i-n_i)!}\right)$$
$$-\boldsymbol{R}T.\ln\left(\frac{(\sum n_i)!}{\Pi n_i!}\right)-\beta\boldsymbol{R}T.\sum_{i\neq 0} n_i$$

where μ_i is the free energy/mol of species i of transfer from the standard state to the co-precipitate, and the sums and products are over all values $i = 0, 1, \ldots m$, unless otherwise specified. The second term is the entropy of mixing of the species in the co-precipitate and the last terms are the entropy of the residual supernatant.

Combining these two expressions, and using Stirling's approximation, gives for the total free energy change during precipitation of the original solution:

$$\begin{aligned}\Delta F^* = &-\beta\boldsymbol{R}T.\sum_{i\neq 0}(N_i-n_i)-\sum(N_i-n_i)\mu_i\\ &-\boldsymbol{R}T.\{[\sum N_i].\ln[\sum N_i]-\sum N_i.\ln N_i\\ &\quad-[\sum(N_i-n_i)].\ln[\sum(N_i-n_i)]\\ &\quad+\sum(N_i-n_i).\ln(N_i-n_i)\\ &\quad-[\sum n_i].\ln[\sum n_i]+\sum n_i.\ln n_i\}\end{aligned}$$

For thermodynamic equilibrium: $\dfrac{\partial}{\partial n_i}\Delta F^* = 0$ for all i. This gives:

$$n_i = \frac{(N_i-n_i).\sum n_j}{\sum(N_j-n_j)}.\exp(\beta+\mu_i/\boldsymbol{R}T)\ (i\neq 0) \quad (1)$$

For very dilute solutions:

$$\sum n_j \simeq N_0$$

Defining:

Total number of solute macromolecules $N = \sum_{i\neq 0}^{m} N_i$

Fraction of component i in original mixture:

$$\gamma_i = N_i/N$$

Mole fraction of i in supernatant $\alpha = n_i/N_0$

Molar concentration of original solution $C = N/N_0$

Eqn. (1) now becomes:

$$\alpha_i = \frac{\gamma_i C}{1+C.\exp[-(\beta+\mu_i/\boldsymbol{R}T)]}\ (i\neq 0)$$

This expression can be simplified further under certain conditions. First, for the components forming the co-precipitate:

$$\ln C > \beta+\mu_i/\boldsymbol{R}T$$

i.e.

$$C.\exp[-(\beta+\mu_i/\boldsymbol{R}T)] \gg 1$$

and

$$\alpha_i \simeq \gamma_i.\exp(\beta+\mu_i/\boldsymbol{R}T)$$

Secondly, if species i is excluded from the co-precipitate:

$$C.\exp[-(\beta+\mu_i/\boldsymbol{R}T)] \ll 1$$

and

$$\alpha_i \simeq \gamma_i C$$

The quantity we measure is the total supernatant concentration $\alpha' = \sum_{i\neq 0}^{m}\alpha_i$, and if under the conditions of precipitation the first p species $i = 1 \ldots p$ precipitate together, whereas species $i = (p+1) \ldots m$ do not, then:

$$\alpha' = \sum_{i=1}^{p}\gamma_i\exp(\beta+\mu_i/\boldsymbol{R}T) + C.\sum_{i=p+1}^{m}\gamma_i$$

This is the familiar form of a solubility graph:

$$\alpha' = \alpha_s + bC$$

where the solubility

$$\alpha_s = \sum_{i=1}^{p}\gamma_i\exp(\beta+\mu_i/\boldsymbol{R}T)$$

and the non-precipitating fraction

$$b = \sum_{i=p+1}^{m}\gamma_i$$

Both α_s and b will vary with experimental conditions as the number of precipitating components, p, changes.

The experimentally determined quantity is:

$$\begin{aligned}\Delta F &= \boldsymbol{R}T.\ln\alpha_s - \beta\boldsymbol{R}T\\ &= \boldsymbol{R}T.\ln\left[\sum_{i=1}^{p}\gamma_i\exp(\mu_i/\boldsymbol{R}T)\right]\end{aligned}$$

Writing $\mu_i = \mu + \epsilon_i$, where μ is some average of the chemical potentials, gives:

$$\Delta F = \mu + \boldsymbol{R}T.\ln\left[\sum_{i=1}^{p}\gamma_i.\exp(\epsilon_i/\boldsymbol{R}T)\right]$$

Thus the free energy ΔF obtained from the extrapolated solubility is equal to the mean free energy of the interaction plus a correction term. The sign and magnitude of this term will depend on the composition of the original mixture of macromolecules and on the size of the free-energy devia-

tions, ϵ_i, of the species. It is difficult to estimate how large these deviations might be, but they are unlikely to be very big because of the apparently continuous variation of the parameter b.

Since

$$\sum_{i=1}^{p} \gamma_i = 1 - \sum_{i=p+1}^{m} \gamma_i = 1 - b$$

the correction will be of order $RT \cdot \ln(1-b)$ provided the ϵ_i values are not significantly greater than RT. This correction is small and probably within the experimental error of the results, and unless the ϵ_i values are large and/or very temperature-dependent the basic behaviour of the ΔF curve should not differ greatly from the mean μ.