Eds.) pp 338–350, Wiley, New York.
Wyssbrod, H. R., Ballardin, A., Schartz, I. L., Walter, R., Van Binst, G., Gibbons, W. A., Agosta, W. C., Field, F. H., & Cowburn, D. (1977) *J. Am. Chem. Soc. 99*, 5273–5276.
Wyssbrod, H. R., Fischman, A. J., Live, D. H., Hruby, V. J., Agarwal, N. S., & Upson, D. A. (1979) *J. Am. Chem. Soc. 101*, 4037–4043.
Zimmerman, S. S., & Scheraga, H. A. (1977) *Biopolymers 16*, 811–843.

Swine Pepsinogen Folding Intermediates Are Highly Structured, Motile Molecules†

Peter McPhie

ABSTRACT: Refolding of urea- or alkali-unfolded swine pepsinogen occurs by rapid formation of partially folded intermediates (I_s) which are slowly converted into the native protein (N). This slow reaction involves isomerization of proline residues in the protein to the configurations occurring in N. Kinetic studies on changes in absorbance or circular dichroism indicate I_s to be close to the native structure, while fluorescence and hydrogen exchange measurements show I_s to be much more open to solvent than N. Fluorescent probe binding suggests that I_s has a more hydrophobic surface than N. These contrasting results are interpreted to show that the presence of wrong proline residues does not greatly inhibit the overall folding of pepsinogen but prevents close packing of structural elements into the highly cooperative, stable, native form. I_s may be very similar to N in average structure, but is a much more fluctuating species.

Unfolding and refolding of low molecular weight proteins have recently been the subject of extensive studies, mainly by Baldwin and his group [reviewed in Kim & Baldwin (1982)]. These investigations have revealed a general mechanism (eq 1), in which the native protein (N)[1] rapidly equilibrates with

$$N \underset{K_u}{\leftrightarrow} U_f \rightleftharpoons U_s \quad (1)$$

a fast-folding unfolded form (U_f), which is in slow equilibrium with a mixture of slowing folding forms (U_s). It was originally proposed that these arose by cis–trans isomerization of proline residues in U_f (Brandts et al., 1975), but in some proteins, U_s could arise by configurational changes at peptide bonds not involving proline. Three such bonds have recently been detected in the high energy cis form in carboxypeptidase A (Rees et al., 1981). The kinetics of isomerization of all peptide bonds are very alike (Stewart & Siddall, 1970). With increasing time under unfolding conditions, the rate of refolding to N becomes limited by reversal of these slow reactions at "essential" residues, which must have the correct configurations for formation of N. Under strongly denaturing conditions, N is converted to U_f in a rapid reaction, but formation of U_s can be demonstrated by the double-jump procedure, in which protein is jumped to unfolding conditions, held for a variable delay time, and then returned to native conditions. The kinetics of refolding are found to vary with the delay time in a manner consistent with eq 1. When the native protein is only marginally unstable, the kinetics of unfolding are biphasic, corresponding to the two coupled steps in this mechanism (Brandts et al., 1975). Computer simulations by Creighton (1978) have shown that for high molecular weight proteins containing many prolines, cis–trans isomerization may drive unfolding to completion even when N is considerably more stable than U_f, because of the large number of U_s forms. Calculated unfolding curves showed slow, single phase kinetics with a rate constant which was essentially proportional to K_u, the equilibrium constant of the unfolding reaction (eq 1). These simulations also showed the dramatic effect of large numbers of proline residues in slowing down the rates of folding of such proteins. Since slow folding may be a disadvantage to a protein under physiological conditions, Creighton speculated on ways that the rate of folding of large proteins could be accelerated. He suggested intramolecular catalysis of proline isomerization during folding, division of large proteins into smaller independently folding domains, each containing fewer prolines, or rapid formation of partially folded intermediates containing incorrect proline isomers.

Swine pepsinogen is a large protein (M_r 40 000) containing a large number (i.e., 18) of prolines (Tang, 1976) which can be reversibly unfolded. Thus, it forms a good subject for the investigation of the mechanism of folding of such species. Previous studies which followed the ionization of tyrosine residues, normally buried in the native protein, and also the variation of potential pepsin activity indicated that pepsinogen unfolded very rapidly at pH 11.5 but was then slowly transformed into slowly refolding forms (U_s). The rate of this slow transformation was independent of pH and showed a temperature dependence characteristic of proline isomerization (McPhie, 1982). In contrast, unfolding by urea, at neutral pH, followed by absorbance and activity, was a slow reaction whose rate was dependent on pH, temperature, and urea concentration. The product of this reaction always refolded slowly (McPhie, 1980). The enthalpies of activation of the rates of unfolding and refolding were again consistent with the involvement of proline isomerization in the rate-limiting

† From the Laboratory of Biochemistry and Metabolism, National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20205. *Received April 14, 1982.*

[1] Abbreviations: U_s and U_f, slow- and fast-folding forms of pepsinogen, respectively; I_s and I_n, folding intermediates; N, native form of pepsinogen; ANS, 8-anilino-1-naphthalenesulfonate; Tris, tris(hydroxymethyl)aminomethane.

step. Such behavior would be anticipated from the simulations described above, as a consequence of the increased stability of the protein at neutral pH.

Refolding of U_s does not occur by direct reversal of eq 1. Detailed studies showed formation of partially folded intermediates, I_s, which are converted to N through slow reactions, including proline isomerization. For example, absorbance changes during folding of pepsinogen indicated rapid formation of I_s, in which 65–75% of the aromatic residues buried in N were already removed from solvent. However, these intermediates could not activate themselves to pepsin at pH 2 (McPhie, 1980, 1982). In the experiments reported below, refolding of pepsinogen was followed, under standard conditions, by a number of other techniques which gave complementary information on the structure of I_s, formed from both alkali-unfolded protein and also from pepsinogen which was initially in the presence of high concentrations of urea. Detailed information on the structure of these intermediates should show more clearly the effect of proline isomerization on the mechanism of folding of high molecular weight proteins.

Experimental Procedures

Materials and Methods. Pepsinogen (lot PG30J608) and hemoglobin substrate powder (lot HB39H795) were obtained from Worthington Biochemical Corp. 8-Anilino-1-naphthalenesulfonate (ANS) was purchased as the magnesium salt from Eastman Kodak. ^{3}H-Labeled water (specific activity 1 Ci/mL) was the product of New England Nuclear Co. All other reagents were analytical grade or equivalent. All experiments were performed at 25 °C. Procedures not described below were as before (McPhie, 1982).

Double-Jump Experiments. A standard procedure was used to ensure that all observations were made under the same conditions. Stock solutions of pepsinogen were made up in distilled water at concentrations of 4–50 μM. The pH values of these solutions were 6–6.5. The protein was jumped to pH 11.5 by the addition of a small aliquot of 1 M sodium hydroxide solution to 1 mL (or 0.5 mL) of solution. After a carefully measured delay time, an equal volume of 0.1 M sodium phosphate buffer, pH 7, was added to give a final pH of 7.3, and the time course of refolding was measured. The procedure was repeated for various values of the delay time. Delay times longer than 5 min were not investigated, as these had been shown to produce some irreversible denaturation. Complete reversibility of refolding at the end of these experiments was routinely checked by assay of the potential pepsin activity of the solutions.

Urea-Unfolded Pepsinogen. Stock solutions of pepsinogen (40 μM) were made up in 6 M urea and 0.1 M phosphate buffer, pH 8. Refolding was initiated by diluting 100 μL of this into 2 mL of 0.1 M phosphate buffer, pH 7.

Circular Dichroism Measurements. Studies on the kinetics of unfolding and refolding were made in a Jasco J500-C spectropolarimeter (Japan Spectroscopic Co., Easton, MD), using 1-cm path-length quartz cuvettes. Changes in ellipticity were followed as a function of time, at wavelengths in the range 210–240 nm. Circular dichroism spectra of the various forms of the protein were measured in the same instrument, using the DP-500 N data processor to increase the signal to noise ratio. The usual precautions with regard to slit width and absorbance were observed.

Fluorescence Measurements. Fluorescence changes during refolding were measured in a Farrand Mark I spectrofluorometer by using 1-cm path-length quartz cuvettes. Pepsinogen fluorescence was excited at 278 nm, and the intensity of the emission at 340 nm was recorded as a function of time on a Hewlett Packard 680 strip chart recorder.

ANS Binding to Pepsinogen. The binding of ANS to native pepsinogen was measured in the Farrand spectrofluorometer essentially as described by Wang & Edelman (1971). ANS fluorescence was excited at 340 nm.

The double-jump procedure was followed as described above to study fluorescence changes during refolding. However, for these experiments, a suitable concentration of ANS was included in the phosphate buffer used to neutralize the protein solution. The intensity of the fluorescence at 485 nm was followed as a function of time during refolding.

Hydrogen Exchange Measurements. These were performed by using the double-column method of Englander & Englander (1972). The columns used were Pharmacia disposable PD-10 columns prepacked with Sephadex G-25M (1.5 cm × 5 cm). Before each use, the columns were washed with 25 mL of buffer.

Exchange Out of Fully Labeled Pepsinogen. The protein was intially labeled by adding 1 μL of ^{3}H-labeled water to 0.5 mL of pepsinogen stock solution. The mixture was jumped to pH 11.5 and incubated for 5 min. Then 0.5 mL of sodium phosphate buffer was added to neutralize the solution. Fifteen minutes later, when refolding was complete, excess ^{3}H was removed by gel filtration. The solution was applied to the PD-10 column and washed with 2 mL of 0.1 M phosphate buffer, pH 7, and the protein was eluted with 1.5 mL of buffer. The protein came off the column within 1–2 min after being loaded. This solution was incubated at 25 °C for up to 30 h to allow exchange out. Then, 1 mL of solution was put through the column precisely as described above. The concentration of protein in the eluate was measured by the absorbance at 280 nm. ^{3}H content was measured by mixing a 100-μL aliquot of solution with 200 μL of distilled water and 5 mL of the scintillation cocktail described by Schreier (1977). After the mixture was cooled, radioactivity was counted in a Nuclear Chicago Mark I scintillation counter. The results, expressed as $^3H_{rem}$, the number of original tritium atoms still unexchanged per molecule, were calculated and corrected for the isotope effect as described by Englander & Englander (1972).

Exchange Out of Intermediates. The double-jump procedure was performed as described above. So that intermediates in folding could be labeled, pepsinogen solutions (0.5 mL) were jumped to pH 11.5, and 1-μL aliquots of ^3H_2O were added with the phosphate buffer or at carefully measured intervals after neutralization. The mixtures were incubated for 15 min, and excess tritium was then removed by gel filtrations. The remaining tritium was exchanged out for 24 h at 25 °C, and $^3H_{rem}$ was measured as described above.

Preparation of [^{3}H]Pepsin. Pepsin was prepared by a modification of the method of Rajagopalan et al. (1966). Unfolded pepsinogen or folding intermediates were labeled as described above, passed through a PD-10 column, and incubated for 24 h at 25 °C in phosphate buffer. Exchanged-out tritium was then removed by gel filtration through a PD-10 column which had been preequilibrated with 2 mM Tris-HCl buffer, pH 7.2, and eluted with the same buffer. $^3H_{rem}$ was determined for each of these solutions. The remainder of the solution was adjusted to pH 2 by the addition of 20 μL of concentrated perchloric acid. After 15 min, the pH of the solution was raised to 4.5 by the addition of 250 μL of 2 M sodium acetate buffer, pH 5. One milliliter of this solution was promptly applied to a PD-10 column, which had been repacked with SP Sephadex C-25 and was eluted with 0.1 M sodium acetate buffer, pH 4.5. This removed activation peptides and any tritium which exchanged out during this

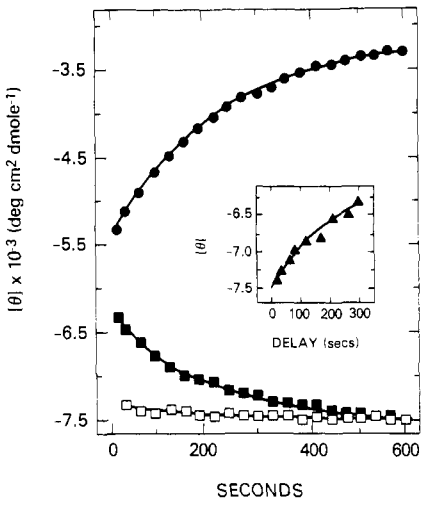


FIGURE 1: Time courses of unfolding and refolding of pepsinogen, measured by circular dichroism at 220 nm (●). The pepsinogen solution was jumped to pH 11.5 for unfolding. Solutions were held at pH 11.5 for a variable delay time, and then the pH was dropped to 7.3 for refolding. Delay times were (□) 10 s and (■) 5 min. Final pepsinogen concentration was 2 μM. Inset: Variation of the initial ellipticity in refolding with delay time at pH 11.5.

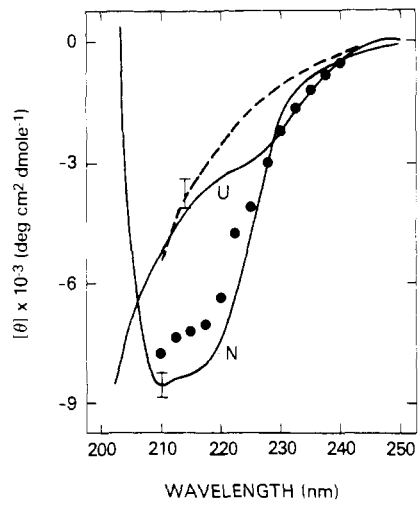


FIGURE 2: Circular dichroic spectra of refolded pepsinogen (N) at pH 7.3 and of unfolded pepsinogen (U) at pH 11.5 and in 6 M urea, pH 8 (---). The spectrum of the refolding intermediates (I_s) was measured by holding the protein at pH 11.5 for 5 min and then dropping the pH to 7.3, or by dilution from 6 M urea, pH 8, to 0.29 M urea, pH 7.3. The initial ellipticity was measured at each wavelength (●). Final pepsinogen concentration was 2 μM.

procedure. These columns were only used once. $^3H_{rem}$ was determined as described above.

Analysis of Data. Kinetic curves were fitted to exponential time courses by using the MLAB modeling system (Schrager, 1970). When more than one term was required, the "peeling back" method of analysis described by Hagerman & Baldwin (1976) was used. The size of the total change in any observable property was estimated by extrapolation of these fitted curves to zero time of refolding. The MLAB system was used to fit the results to quantitative models of the refolding process.

Results

Unfolding and Refolding Measured by Circular Dichroism. High pH unfolding was measured under standard conditions (25 °C, pH 11.5) by decreases in the intensity of circular dichroism of the peptide absorption bands. At all wavelengths, half of the expected change between the native and unfolded forms was found to occur within the mixing time (15 s), while the second half followed a slow exponential time course, k = 0.0045 ± 0.0010 s^{-1} (Figure 1). Double-jump experiments were performed to study the kinetics of refolding at neutral pH. Only small fractions of the total regain in intensity were found to occur after mixing. These changes followed an exponential time course, k = 0.006 ± 0.001 s^{-1}. The size of the resolved change was a function of the delay time after the initial jump to pH 11.5, increasing with a rate constant k = 0.0043 ± 0.0010 s^{-1} (Figure 1, inset).

The circular dichroic spectrum of refolded pepsinogen was identical with that of the native protein (N) (Figure 2). A spectrum of the unfolded protein (U) was recorded after 10 min at pH 11.5. No detectable change occurred in this spectrum during a further hour at high pH. Activity measurements showed that the extent of irreversible denaturation increased from 5% to 50% over the same time period. Pepsinogen solutions were incubated at pH 11.5 for 5 min and then neutralized to obtain a circular dichroic spectrum for I_s. Intensities were followed as a function of time at 2.5-nm intervals in the range 210–240 nm and extrapolated back to zero time of mixing as described above.

Refolding of urea-denatured pepsinogen was also found to occur in two phases. Within experimental error, the magnitudes of the slow phase were identical with those shown for I_s in Figure 2. The slow phase followed a simple exponential with time, k = 0.007 ± 0.001 s^{-1}.

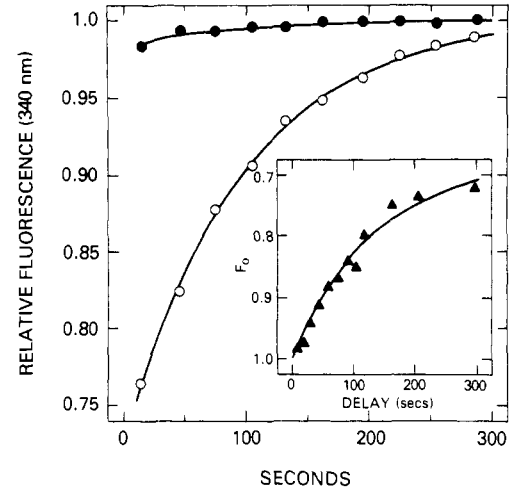


FIGURE 3: Double-jump experiments measuring the refolding of pepsinogen by the intrinsic fluorescence at pH 7.3. Protein solutions were jumped to pH 11.5 and held there for a variable delay time, and then the pH was rapidly dropped to pH 7.3. Fluorescence was excited at 278 nm and followed as a function of time at 340 nm. Delay times were (●) 10 s and (○) 5 min. Final pepsinogen concentration was 2 μM. Inset: Variation of the initial fluorescence intensity on refolding with delay time at pH 11.5.

Attempts were made to estimate the secondary structure present in the various species by analysis of the data in terms of standard spectra of the α helix, β sheet, and unordered chain, as described by White (1976). For N and I_s, the best fits were obtained by using standards derived from proteins of known structures. The analysis indicated that N contained 19 ± 1% α helix and 39 ± 2% β sheet. The values for I_s were 17 ± 1% and 30 ± 2%, respectively. No good fits could be made to the spectra of the unfolded proteins. Thus, I_s has more than 80% of the native protein's secondary structure.

Refolding Measured by Protein Fluorescence. When the double-jump procedure was used, refolding was found to be accompanied by increases in the intensity of the tryptophan

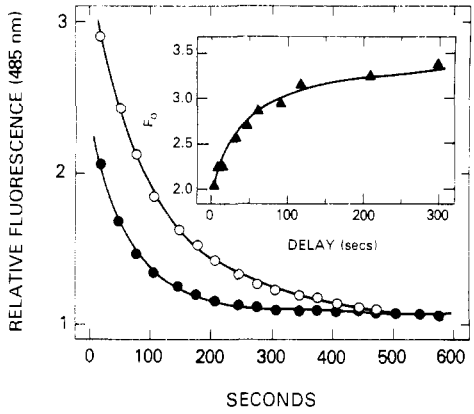


FIGURE 4: Double-jump experiments measuring the refolding of pepsinogen by the ANS fluorescence at pH 7.3. Protein solutions were jumped to pH 11.5 and held there for a variable delay time, and then the pH was rapidly dropped to pH 7.3. ANS fluorescence was excited at 350 nm and followed as a function of time at 495 nm. Delay times were (●) 10 s and (○) 5 min. Final concentrations were 13 μM pepsinogen and 45 μM ANS. Inset: Variation of the initial fluorescence intensity on refolding with delay time at pH 11.5.

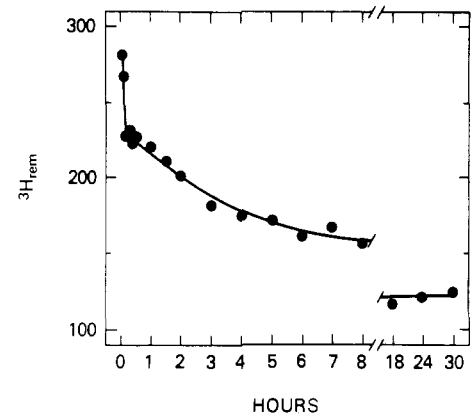


FIGURE 5: Exchange-out kinetics of refolded pepsinogen, measured by loss of tritium, in 0.1 M phosphate buffer, pH 7, 25 °C. Protein concentration during the exchange out was 17 μM.

fluorescence at 340 nm (Figure 3). The increases followed a simple exponential time course whose rate constant decreases slightly with increasing delay time, at pH 11.5, from $k = 0.02$ s^{-1} after 10 s to $k = 0.01$ s^{-1} after 5 min at pH 11.5. The total size of the increase on refolding also varied with the delay time at pH 11.5, from 2% of the total native intensity after a delay of 10 s to 28% after 5 min, with a rate constant $k = 0.07 \pm 0.001$ s^{-1}.

During refolding of urea-denatured pepsinogen, it was possible to resolve an increase of 25% of the total native intensity, which occurred with a rate constant $k = 0.005 \pm 0.001$ s^{-1}. Frattali et al. (1965) showed that urea unfolding of pepsinogen produced a 40% quenching of fluorescence, indicating that 70% of the normally buried tryptophans are still exposed in I_s.

ANS Binding to Native Pepsinogen. Titration of pepsinogen and ANS, at neutral pH (0.05 M phosphate buffer, pH 7.3), produced large increases in the fluorescence of the dye. The maximum in the uncorrected emission spectrum moved from 520 to 485 nm. Analysis of the data according to Wang & Edelman (1971) was compatible with the protein binding one molecule of ANS with a dissociation constant of 3×10^{-4} M. The bound dye had a quantum yield of 0.56 [assuming that of the free dye to be 0.004 (Li et al., 1976)].

Refolding in the Presence of ANS. Addition of excess ANS to the phosphate buffer, used in the double-jump procedure, produced dramatic changes in the fluorescence of the dye during the time when other measurements had shown pepsinogen refolding occurring (Figure 4). Immediately after neutralization, the intensity was considerably higher than the equilibrium value. After short delay times at pH 11.5, the fluorescence decayed with a rate constant $k = 0.02$ s^{-1}, but with increasing times at higher pH, a slower reaction in refolding with $k = 0.004$ s^{-1} became increasingly evident. The initial fluorescence of the ANS–pepsinogen mixture was found to be a function of the delay time at pH 11.5 (Figure 4, inset), increasing from twice the equilibrium intensity after 5 s to 3.3 times as intense after a 5-min delay. Other experiments showed that the initial fluorescent intensities on refolding were a function of protein and dye concentrations but that their variation with delay time was not. Control experiments failed to show any changes in fluorescence intensity at 485 nm when protein or dye was subjected to the double-jump procedure alone.

Hydrogen Exchange from Native Pepsinogen. The hydrogen exchange kinetics of pepsinogen have not previously been reported. Hence, it was necessary to investigate the behavior of the native protein as a reference against which to compare the behavior of folding intermediates. The exchange-out kinetics of fully tritiated pepsinogen, determined under standard conditions, pH 7 and 25 °C, are shown in Figure 5. With the assumption that we are measuring only the exchange of peptide hydrogens, they can be divided into three classes: 140 rapidly exchanging protons which are lost within 5 min, 110 slowly exchanging protons which exchange over several hours, and a plateau of 120 stable protons which show no detectable exchange between 18 and 30 h. It is this last group we are concerned with here, since they represent hydrogens in stable structures, which are protected from solvent (Englander & Englander, 1972).

^3H Labeling of Folding Intermediates. The aim of these experiments was to measure the formation of these stable structures during refolding. ^3H was added to protein solutions, folding was allowed to go to completion, and label was then exchanged out for 24 h. If the small amount of ^3H present makes no difference to the behavior of the protein, the relative value of $^3H_{rem}$ will show how much stable structure was formed after the label was added.

When ^3H was added to pepsinogen solutions before the jump to pH 11.5, the value of $^3H_{rem}$ after 24 h of exchange out was independent of the delay time at high pH from 10 s to 5 min (121 ± 3 atoms). However, if ^3H was added during or after neutralization, the results became highly dependent on the delay time (Figure 6). Labeling of protein 10 s after neutralization gave $^3H_{rem} = 27$ for a delay time of 10 s at pH 11.5 and $^3H_{rem} = 76$ for a delay time of 5 min. The extent of refolding followed in this way was a biphasic function of time with rate constants $k = 0.02$ and 0.004 s^{-1}. In a control experiment, ^3H was added with phosphate buffer to a stock solution of pepsinogen in distilled water, pH 6.5. The mixture was incubated for 15 min and exchange out measured as described above. After 24 h, a value of $^3H_{rem} = 1$ was determined.

A preliminary attempt was made to obtain information about the relative positions of ^3H incorporated into the several forms of pepsinogen. The protein was labeled under various conditions and $^3H_{rem}$ determined after 24 h as described above, except that the second Sephadex column was eluted with 2 mM Tris-HCl buffer. This modification was found to have no significant effect. These samples were activated to pepsin under standard conditions at pH 2 and adjusted to pH 4 to prevent autolysis. Activation peptides and any lost ^3H were

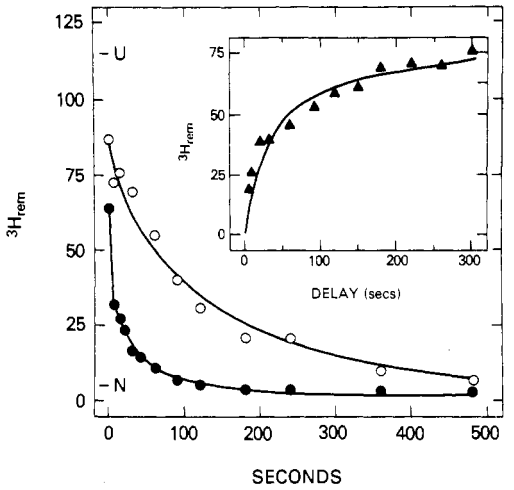


FIGURE 6: Incorporation of stable tritium into pepsinogen during folding. Pepsinogen solutions were jumped to pH 11.5 and held there for a variable delay time, and then the pH was dropped to pH 7.3. Aliquots of tritiated water were added at the times shown after the pH drop. Delay times were (●) 10 s and (○) 5 min. $^3H_{rem}$ was determined 24 h later, as described in the text. U and N are the values obtained with unfolded and native pepsinogen, respectively. Inset: $^3H_{rem}$ was determined by adding tritiated water 10 s after the pH drop, but varying the delay time at pH 11.5. The pepsinogen concentration during refolding was 25 μM.

Table I: Incorporation of Stable ^3H into Pepsinogen during Refolding and Activation[a]

labeled species	$^3H_{rem}$	% lost on activation
U	121 ± 3	13 ± 3
I_s	88 ± 3[b]	9 ± 2
	75 ± 2[c]	13 ± 2
I_n	65 ± 1[b]	15 ± 2
	28 ± 1[c]	19 ± 2
N	1 ± 1	ND[d]

[a] One microliter of ^3H_2O was mixed with 0.5 mL of pepsinogen solution and 0.5 mL of phosphate buffer as described. Either ^3H was added before the jump to pH 11.5 (U) or protein was held at pH 11.5 for 5 min (I_s) or 10 s (I_n). Native pepsinogen (N) was never exposed to pH 11.5. $^3H_{rem}$ was measured before and after activation as described in the text. Each value was measured 4–8 times and is given as the mean ± the standard deviation. On activation, pepsinogen loses 12% of its peptide groups (Tang, 1976). [b] ^3H was added with the phosphate buffer. [c] ^3H was added 10 s after the phosphate buffer. [d] ND, not determined.

removed by chromatography, and $^3H_{rem}$ values for the resultant pepsin molecules were determined (Table I). Some loss of label was found under all conditions.

Discussion

Previous studies on the unfolding of pepsinogen at pH 11.5 were made by following the changes in absorbance due to ionization of tyrosine residues in the protein (McPhie, 1982). These showed the reaction to be rapid, and on the basis of further studies on refolding, they were interpreted in terms of the mechanism of eq 1. In protein unfolding reactions, absorbance changes may not always be interpreted unambiguously. The circular dichroism results, shown in Figure 1, give a direct indication of changes in secondary structure. They can be taken to show that immediately after a jump from neutrality to pH 11.5, under the standard conditions used in these experiments, only half of the protein molecules are unfolded. However, it should be recalled that when such a solution is added to a pH 2 hemoglobin solution, no pepsin activity develops. This crucial test shows the absence of native pepsinogen and suggests the presence of an intermediate, I_n (eq 2), which has the same secondary structure as N but has

$$N \leftrightarrow I_n \leftrightarrow U_f \rightleftharpoons U_s \quad (2)$$

no potential pepsin activity. Further evidence for the existence of I_n will be discussed below.

The circular dichroism spectra of the unfolded forms of pepsinogen shown in Figure 2 differ considerably from the standard "random coil" spectra used in the analysis of the other data. However, the spectrum of pepsinogen unfolded by urea is similar to those of other proteins in the presence of high concentrations of denaturants (Cortijo et al., 1973). In the case of the alkali-unfolded form, U, the invariance of its spectrum on long incubation at pH 11.5 shows that the difference cannot arise from chemical changes accompanying irreversible denaturation. Consequently, the spectrum must be taken as showing that the unfolded protein has some residual structure, whose nature and extent could not be readily estimated. This is usually the case with proteins which are unfolded only by changes in temperature or pH (Tanford, 1968). Alkali-unfolded pepsinogen has a greatly increased molecular volume, as measured by viscosity (Frattali et al., 1965), and all of its tyrosine residues are titrated (McPhie, 1980). The present work shows all of its peptide hydrogens are readily exchangeable, suggesting that any structure is only marginally stable. An added complication is that after the maximum time for which pepsinogen could safely be exposed to high pH (5 min), the results in Figure 1 indicate that unfolding is only 90% complete. Unfolding for longer times or at higher pH values always produced some irreversible denaturation. Nevertheless, the great similarity of the refolding kinetics of U_s produced by unfolding either at high pH or in high concentrations of urea [where the protein has all the properties of a cross-linked random coil (Frattali et al., 1965; Ahmad & McPhie, 1978)] indicates that this residual structure has little effect on this reaction.

Determination of the rate constant of the slow phase in unfolding by circular dichroism gave the same value as did previous absorbance measurements, $k = 0.0045\ s^{-1}$. Double-jump experiments measuring changes in both absorbance and potential pepsin activity showed a "hidden transition" after unfolding from fast- to slow-refolding forms, $k = 0.025\ s^{-1}$ (McPhie, 1982). The present results also show changes in refolding properties with time at high pH, but their rates seem to depend on the probe used to measure refolding. After unfolding, each proline in the molecule will isomerize with a characteristic rate constant, determined principally by its neighboring residues (Grathwohl & Wuthrich, 1981). Essential prolines will manifest themselves on returning to native conditions by blocking refolding in ways which depend on their roles in the protein's structure. Probes which reflect different facets of this structure may be anticipated to show transitions which reflect the rates of isomerization of different essential residues.

For a molecule such as pepsinogen, with 18 prolines (Tang, 1976), a complete analysis of this situation may never be possible. However, as a first approximation, we may consider a model in which there are two independent groups of essential residues in the zymogen, at high pH: a rapidly isomerizing group, which in the wrong configuration inhibits activation to pepsin at low pH, and a slowly isomerizing group, which interferes with formation of the correct secondary structure on refolding (and, consequently, also prevents activation to pepsin). The model predicts the formation of three slowly folding forms of pepsinogen, U_s^1 and U_s^2, in which either of these groups has the wrong configuration, and U_s^{12}, in which

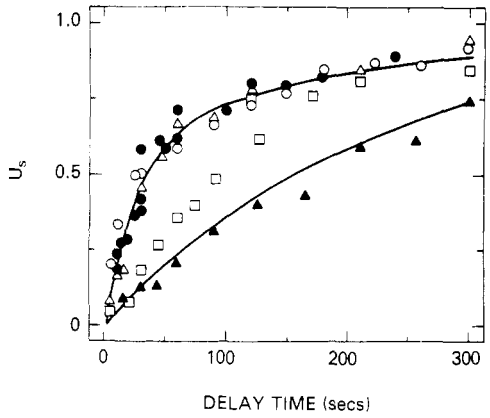
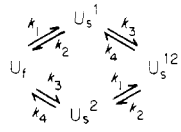


FIGURE 7: Conversion of pepsinogen into slowly refolding forms during the delay at pH 11.5, as measured by a variety of techniques: (▲) circular dichroism; (□) intrinsic fluorescence; (△) ANS fluorescence; (○) hydrogen exchange; (●) absorbance and potential pepsin activity (McPhie, 1982). The curves are derived from the model in eq 3.

both have isomerized. These are connected to U_f as shown in eq 3:



(3)

The hidden transition detected by potential pepsin activity will follow the disappearance of U_f and that detected by circular dichroism the appearance of U_s^2 and U_s^{12}. The predicted behavior is shown in Figure 7, using the values $k_1 = 0.020\ s^{-1}$, $k_2 = 0.014\ s^{-1}$, $k_3 = 0.0045\ s^{-1}$, and $k_4 = 0.000045\ s^{-1}$. [The analysis is independent of the fact that unfolding is not complete initially, if U_f is in rapid equilibrium with I_n and N (cf. eq 2).] It will be noted that the transitions measured by absorbance changes, hydrogen exchange, and ANS fluorescence follow potential pepsin activity, whereas that measuring protein fluorescence shows intermediate behavior. The model also explains why the slow phases in refolding show complex kinetics when followed by the former class of probes but not by fluorescence or circular dichroism. As discussed earlier (McPhie, 1982), it is not possible to study hidden transitions after urea unfolding of pepsinogen. The similar slow kinetics of refolding under these circumstances shown by absorbance, fluorescence, and circular dichroism indicate the absence of the rapidly isomerizing group of residues at neutral pH.

Before attempting an analysis of the structure of the intermediates formed from these unfolded forms, it is worthwhile to review very briefly the kind of information each of these probes may give. Potential pepsin activity is assumed to indicate the formation of pepsinogen's native structure. Absorbance difference spectra in proteins show the time-averaged accessibility of aromatic residues to solvent, while quenching of fluorescence is a dynamic measure of the structural fluctuations in the protein, in the nanosecond time range, which allow partial exposure to solvent (Lakowicz & Weber, 1973). Similarly, circular dichroism gives an average measure of protein secondary structure, but hydrogen exchange is taken to indicate "breathing" reactions in which regions of the protein are exposed to solvent for short periods of time. Model compound studies show the average exchange lifetime of a free amide proton to be 3 ms at pH 7.3, 25 °C (Englander & Englander, 1972). Consequently, exchange of protons which are normally protected can be taken to show that breathing reactions are occurring over a similar time scale. Finally, it is usually thought that changes in the fluorescence of bound ANS molecules reflect variations in the polarity of their environment (Stryer, 1965). However, in the present context, it is perhaps more pertinent to recall that unfolded proteins show extra binding sites for this dye (Li et al., 1976).

While the measured properties of the intermediates can be analyzed in these terms, it should be kept in mind that we are observing the average properties of a number of intermediates, arising from different slowly folding forms. Three cases will be considered below, refolding from 6 M urea and after a long (5 min) and a short (10 s) delay time at pH 11.5.

Refolding from 6 M Urea. Earlier experiments, measuring refolding under these conditions by absorbance changes, showed that over 65% of the total change expected from equilibrium measurements had occurred within the mixing time. This observation was interpreted to show the rapid formation of intermediates in refolding (I_s) which had already buried 65% of the aromatic residues normally inside the native protein (N) (McPhie, 1980). The results shown in Figure 2 indicate that 80% of the secondary structure present in N is formed at the same time, suggesting that the average I_s molecule is very similar to N. In contrast, the size of the slow phase in refolding measured by fluorescence suggested that more than 60% of the normally buried tryptophans are exposed in I_s.

Refolding after 5 min at pH 11.5. The initial unfolded state at high pH cannot be defined as clearly as that in urea (see above). Changes in absorbance and fluorescence resulting from unfolding are masked by those produced by titration of tyrosine residues. However, the close similarity of the slow phases in refolding from high pH and from high concentrations of urea shows that similar intermediates are formed in both reactions. In this case, the disparity between the conclusions drawn from absorbance and fluorescence measurements is reinforced by the other techniques. Circular dichroism measurements again indicate that I_s has 80% of the native secondary structure (Figure 2), but hydrogen exchange studies show that over 70% of the 120 protons, which are stable in the native protein, are readily labeled in I_s (Figure 6). Thus, results obtained by the two techniques which measure dynamic properties imply that I_s has much more open structures than N.

In terms of the summary given above, these contrasting results can be reconciled in a model in which I_s molecules are close to N in average structure but are much more motile. Levitt (1981) has performed conformational energy calculations on the effect of "wrong" proline isomers on the native structure of bovine pancreatic trypsin inhibitor. Forcing certain residues from the trans to cis configuration was found to produce marked destabilization of the molecule. However, the resultant deviations of the peptide backbone from the all-trans structure were small (root mean square value < 0.56 Å). Richards (1979) showed that the observed hydrogen exchange and fluorescence quenching properties of native proteins can arise from transient defects in molecular packing of similar size, produced by thermal displacements. Such defects would be greatly augmented in I_s. In the case of pepsinogen, the presence of wrong proline isomers may not greatly inhibit the overall folding of the protein but prevent the close packing of structural elements into the highly cooperative, stable, native structure. Consequently, I_s is very similar to N in average structure but shows increased hydrogen exchange and fluorescence quenching.

Refolding after 10 s at pH 11.5. The resolvable changes in absorbance, circular dichroism, and tryptophan fluorescence

during refolding were very small (Figures 1 and 3). This would be anticipated if the main product of fast refolding under these circumstances were the native zymogen. However, one other probe gave disproportionately larger changes.

Sizable increases in the fluorescence of dyes like ANS, which accompanied the denaturation of large proteins, were ascribed to the formation of large numbers of nonspecific dye binding sites in the unfolded protein (Anderson & Weber, 1966; Heitz & Brand, 1971). It seems almost certain that this phenomenon is the basis of the large changes in ANS fluorescence shown in Figure 4. Thus, I_s may be taken as having a more hydrophobic surface than N. The extent of this increase cannot readily be estimated. Binding studies at equilibrium showed that the pressure-denatured forms of lysozyme and chymotrypsinogen tightly bind one and two molecules of ANS, respectively (Li et al., 1976), while acid-dissociated alcohol dehydrogenase was claimed to bind over 20 molecules per subunit (Anderson & Weber, 1966). The transient nature of these intermediates prevents a full characterization of this binding, in the present case.

Large changes in dye fluorescence were found even after very short times under unfolding conditions (Figure 4). These might result from inhibition of folding by bound ANS. However, these measurements show the same hidden transition and biphasic kinetics of refolding as do changes in absorbance, activity, and hydrogen exchange. This argues against such inhibition and indicates that the initial folding product from U_f is identical with N in all its measured properties except for the presence of extra ANS binding sites. Unfolding studies suggested an intermediate I_n between N and U_f which was similar to N but had no potential pepsin activity. These experiments may be showing another aspect of the same species.

While this work was in progress, Privalov and co-workers reported calorimetric studies on the unfolding of pepsinogen by temperature and urea (Mateo & Privalov, 1981; Privalov et al., 1981). Their results were interpreted to show that unfolding occurred in two separate domains of the molecule. A comparison of the stabilities and relative reversibilities of unfolding of pepsinogen and pepsin led them to suggest that the region lost from pepsinogen on activation served as a nucleus for folding. There is a great deal of evidence for a dramatic reduction in stability after activation (Ahmad & McPhie, 1979). It was hoped that the hydrogen exchange studies would show a special role for the amino-terminal region as a nucleus during the formation of intermediates. An analysis of variance was applied to the total data describing the percent of label lost on activation, summarized in Table I. It was found that both the delay time at high pH and the time of labeling significantly influenced the results ($p < 0.01$ and $p < 0.05$, respectively). Furthermore, the two effects were strictly additive. Thus, different conditions of folding seem to lead to the incorporation of tritium into different regions of the zymogen. However, at the present level of resolution, these regions cannot be clearly identified [cf. Rosa & Richards (1979)]. In contrast, previous chemical modification studies on pepsinogen demonstrated increased reactivity to diethyl pyrocarbonate in the amino-terminal region, while the protein was undergoing the $I_s \rightarrow N$ reaction (McPhie & Chiu, 1981). As this region becomes protected in a late step in folding, it may stabilize the zymogen by acting as a "key stone" rather than a nucleus.

Acknowledgments

I thank Dr. Robert Blumenthal, NCI, for the use of his spectropolarimeter, Dr. David W. Alling, NIAID, for his help in the analysis of the hydrogen exchange data, Dr. Joseph van Renswoude for a critical review of the manuscript, and Dr. R. L. Baldwin for a copy of his forthcoming review.

References

Ahmad, F., & McPhie, P. (1978) *Biochemistry 17*, 241–246.
Ahmad, F., & McPhie, P. (1979) *Can. J. Biochem. 57*, 1090–1092.
Anderson, S., & Weber, G. (1966) *Arch. Biochem. Biophys. 116*, 207–223.
Brandts, J. F., Halvorson, H. R., & Brennan, M. (1975) *Biochemistry 14*, 4953–4963.
Cortijo, M., Panijpan, B., & Gratzer, W. B. (1973) *Int. J. Pept. Protein Res. 5*, 179–186.
Creighton, T. E. (1978) *J. Mol. Biol. 125*, 401–406.
Englander, S. W., & Englander, J. J. (1972) *Methods Enzymol. 26*, 406–413.
Frattali, V., Steiner, R. F., & Edelhoch, H. (1965) *J. Biol. Chem. 240*, 112–121.
Grathwohl, C., & Wuthrich, K. (1981) *Biopolymers 20*, 2623–2633.
Hagerman, P., & Baldwin, R. L. (1976) *Biochemistry 15*, 1462–1473.
Heitz, J. R., & Brand, L. (1981) *Arch. Biochem. Biophys. 144*, 286–291.
Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459–489.
Lakowicz, J. R., & Weber, G. (1973) *Biochemistry 12*, 4171–4179.
Levitt, M. (1981) *J. Mol. Biol. 145*, 251–263.
Li, T. M., Hook, J. W., Brickamer, H. G., & Weber, G. (1976) *Biochemistry* 15, 5571–5580.
Mateo, P. L., & Privalov, P. L. (1981) *FEBS Lett. 123*, 189–192.
McPhie, P. (1980) *J. Biol. Chem. 255*, 4048–4052.
McPhie, P. (1982) *J. Biol. Chem. 257*, 689–693.
McPhie, P., & Chiu, F.-T. (1981) *Arch. Biochem. Biophys. 211*, 471–477.
Privalov, P. L., Mateo, P. L., Khechinashvili, N. K., Stepanov, V. M., & Revina, L. O. (1981) *J. Mol. Biol. 152*, 445–464.
Rajagopalan, T. G., Moore, S., & Stein, W. H. (1966) *J. Biol. Chem. 241*, 4940–4950.
Rees, D. C., Lewis, M., Honzatko, R. B., Lipscomb, W. N., & Hardman, K. D. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3408–3412.
Richards, F. M. (1979) *Carlsberg Res. Commun. 44*, 47–63.
Rosa, J. J., & Richards, F. M. (1979) *J. Mol. Biol. 133*, 399–416.
Schreier, A. A. (1977) *Anal. Biochem. 83*, 178–184.
Shrager, R. J. (1970) *J. Assoc. Comput. Mach. 17*, 446–456.
Stewart, W. E., & Siddall, T. H. (1970) *Chem. Rev. 70*, 517–551.
Stryer, L. (1965) *J. Mol. Biol. 13*, 482–495.
Tanford, C. (1968) *Adv. Protein Chem. 23*, 121–182.
Tang, J. (1976) *Trends Biochem. Sci. (Pers. Ed.) 1*, 205–208.
Wang, J. L., & Edelman, G. M. (1971) *J. Biol. Chem.* 246, 1185–1191.
White, F. H. (1976) *Biochemistry 15*, 2906–2911.