Biochemistry

 Volume 27, Number 6 March 22, 1988

*Perspectives in Biochemistry*

# Folding of the Nascent Peptide Chain into a Biologically Active Protein†

Chen-Lu Tsou

*Laboratory of Molecular Enzymology, Institute of Biophysics, Academia Sinica, Beijing, 100080 China*

*Received October 7, 1987; Revised Manuscript Received December 23, 1987*

ABSTRACT: The refolding of denatured proteins with complete sequences may not be fast enough to account for the in vivo folding of growing peptide chains during biosynthesis. As some peptide fragments have secondary structures not unlike those of the corresponding segments in the intact molecules and native disulfide bonds of some proteins can form cotranslationally, it is suggested that the folding of the nascent chain begins early during synthesis. However, further adjustments may be necessary during chain elongation and after posttranslational modifications of the completed peptide chain to generate the native conformation of a biologically active protein.

Since the original proposal of the one gene–one enzyme theory by Beadle and Tatum (1941), the steps leading to the biosynthesis of peptide chains from mRNA templates have been largely elucidated. However, the biological activities of proteins depend not only on their amino acid sequences but also on the discrete conformations of the proteins concerned, and slight disturbances to the conformational integrity of a protein destroy its activity (Tsou, 1986). Consequently, to understand fully the entire process of gene expression, it is equally important to understand the process for the folding of the peptide chain into a biologically active protein as that for the synthesis of the primary sequence. On the basis of the recovery of ribonuclease A activity by oxidation of the fully reduced and unfolded molecule, Anfinsen (1973) proposed that the conformation of a protein is determined solely from its amino acid sequence, and the folding of the peptide chain to form the native three-dimensional structure of a protein is spontaneous. Although this proposition is now generally accepted, little is known about how a nascent peptide chain folds into its native conformation so as to generate a functional protein.

Does the nascent peptide chain of a protein fold while it is being synthesized (Figure 1a) or does the folding into its native conformation begin after the synthesis of the peptide chain has been completed (Figure 1b)? In other words, is the folding of the nascent peptide chain during the synthesis of a protein cotranslational or posttranslational? Previous discussions of folding of peptide chains into biologically active proteins during biosynthesis have concentrated almost entirely on models based on the refolding of completed peptide chains after being unfolded by denaturants (Creighton, 1984; Goldberg, 1985). However, the folding of a completed sequence would certainly be different from that of a growing peptide chain, as has been pointed out by Bergman and Kuehl (1979), whereas the above approach is based on the supposition that folding is entirely posttranslational. It is the purpose of the present paper to discuss the folding of a growing peptide chain during biosynthesis on whatever evidence now available. It is suggested that the folding of the growing peptide chain begins early in the synthesis, undergoes constant adjustments during chain elongation, and ends after the completion of the synthesis and the necessary posttranslational covalent modifications of the peptide chain.

*Refolding of the Unfolded Molecule Is Not a Valid Model.* Although it is not precisely known how fast a nascent peptide chain folds into its native conformation, the events leading finally to the formation of the completed and fully active protein molecules are known to be fast processes. The appearance of β-galactosidase activity takes place only a few minutes after the addition of the inducer to a growing culture of *Escherichia coli*, and the time for the synthesis of the native enzyme molecule from its mRNA template has been estimated to be 80–90 s (Kepes & Beguin, 1966). Approximately the same time has been reported for the synthesis of a number of proteins (Conconi & Bank, 1966; Palmiter, 1975). It should also be pointed out that if the folding is entirely a posttranslational event, the above estimation would have to include

†Supported in part by Research Grant RO1 DK 34035 from the National Institutes of Health.

0006-2960/88/0427-1809$01.50/0 

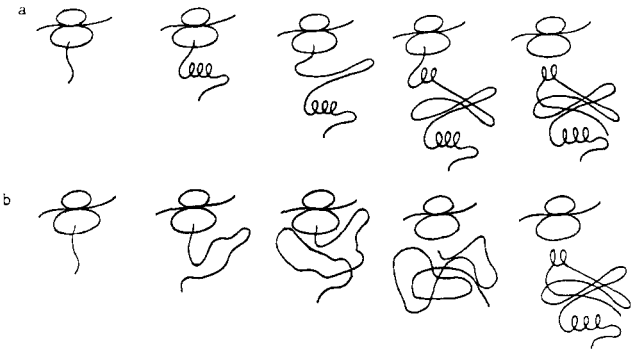

FIGURE 1: Co- and posttranslational folding of the nascent peptide chain. (a) Cotranslational folding. Each peptide segment folds into its final conformation during synthesis. (b) Posttranslational folding. Folding takes place after the completion of the entire peptide chain.

both the time required for the synthesis of the peptide chains and the time of folding of the completed chains into native conformations required for the biological activities of the proteins concerned. The folding time alone for the nascent chains must be much less and has been estimated to be 30 s for rat serum albumin (Peters & Davidson, 1982)). This is not surprising as it is vital for the organism that the completed peptide chains should fold up rapidly so as not to hold up life's processes of the living cell. The half-life for a rapidly metabolized protein has been reported to be only 11 min (Dice & Goldberg, 1975), and the generation times of some cells can be as short as 10–20 min.

Small protein molecules like ribonuclease A or lysozyme fully denaturated in guanidine or urea are known to refold within a few seconds during in vitro experiments (Garel & Baldwin, 1973; Kuwajima et al., 1985). However, these are not valid models, as in the cases studied the unfolded proteins still have the disulfide linkages intact, whereas during the biosynthetic process the peptide chains are synthesized with the Cys residues in the reduced form. Furthermore, with all the native disulfide bonds intact, the molecules may not be completely unfolded even in concentrated urea or guanidine. Little is known about the in vivo formation of disulfide bonds during the synthesis of the peptide chain except that for certain proteins some of their disulfide linkages are formed cotranslationally (Bergman & Kuehl, 1979; Peters & Davidson, 1982). From the extensive studies now available for a few small proteins like ribonuclease A or bovine pancreatic trypsin inhibitor, the in vitro processes of the formation of the native proteins are relatively slow, starting from either the fully reduced and unfolded chains or the scrambled molecules, even when assisted by the presence of protein disulfide isomerase (Creighton et al., 1980).

In general, except for a few smaller proteins (Esterhout & Nall, 1985; Johnson & Price, 1987), the refolding of completely denatured and unfolded proteins takes a much longer time, usually many minutes or even hours. For instance, recovery of about 85% of creatine kinase activity after dilution of the guanidine or urea denatured enzyme takes some 30 min or so, and full activity recovery takes many hours (Zhou & Tsou, 1986). Closely similar results were reported for the refolding of D-glyceraldehyde-3-phosphate dehydrogenase (Jaenicke et al., 1980). The refolding of denatured phosphoglycerate kinase (Betton et al., 1985) and a number of other enzymes (Garel & Dautry-Varsat, 1980; Hermann et al., 1983; Price & Stevens, 1983) is also a very slow process. Even for a small protein containing no disulfide bonds, like cytochrome *c*, the refolding of the denatured molecule takes several minutes (Esterhout & Nall, 1985), which may not be fast enough to account for its folding during biosynthesis.

It is known that the cis–trans isomerization of proline imide peptide bonds is a slow reaction that could be responsible for the slow refolding of unfolded protein molecules (Brandts et al., 1975), and recently Lang et al. (1987) showed that for some proteins the refolding process can be considerably accelerated by the enzyme peptidyl-prolyl cis–trans isomerase. However, it is not certain whether all the slow steps during the refolding processes can be attributed to the cis–trans isomerization of proline residues; other slow conformational changes may have been responsible (Lin & Brandts, 1983, 1987). Moreover, it is not known whether, during the elongation of the nascent peptide chain, proline residues are incorporated exclusively as the correct isomer required at a particular position or as a mixture of both forms, followed by isomerization to give the correct isomer. This isomerization, probably catalyzed by the isomerase, can in turn take place with the growing peptide chain or after its completion. In other words, it can also be either co- or posttranslational.

In any case, there is no a priori reason to suppose that the folding of a nascent peptide chain is entirely posttranslational so as to warrant the use of the refolding of the fully denatured complete peptide sequence as a valid model for the study of protein folding during biosynthesis.

*Peptides Have Secondary Structures.* The folding of chain fragments located in different parts of the complete sequences, especially the N-terminal sequences of different lengths, could provide valuable clues for the elucidation of the folding of the nascent peptide chains. It is now known that short peptides are not without their own conformations even though the conformation of a peptide is flexible and very much dependent on the environment. This is not at all surprising as secondary structures of proteins depend largely on the interaction through hydrogen bonds of neighboring residues. That short peptides do have recognizable ordered secondary structures has now been extensively studied both by experimental methods (Hruby, 1985) and by theoretical calculations based on the principle of energy minimization (Rapaport & Scheraga, 1981). Like the proteins, the conformational aspects of peptide structure are also important for their biological activities such as the binding of peptide hormones to their respective receptors.

It is particularly important in this connection that isolated N-terminal fragments of some proteins do have significant amounts of secondary structures. The insulin B chain, which is N-terminal in the proinsulin molecule, has been shown to contain a significant amount of ordered secondary structure, which increases upon pairing of the B chain with the A chain (Hua et al., 1985). Removal of a pentapeptide and especially an octapeptide fragment from the C-terminal of the B chain results in considerable loosening of its ordered secondary structure (Wang & Tsou, 1986). The above would seem to suggest that folding begins when the peptide chain of the insulin precursor molecule reaches a certain length during synthesis.

Furthermore, the conformation of an isolated peptide fragment has been shown to be not unlike that of the same peptide segment within the native protein molecule (Wetlaufer, 1981). It has been recently shown that some peptide fragments of proteins can induce the formation of antibodies that can bind, sometimes very tightly, to the intact protein from which the peptide fragments are derived, and these peptide fragments compete with the native protein for binding to the antibodies. This is highly suggestive that these peptide fragments can very probably adopt such conformations that are similar to those of the same segments in the intact protein molecules. Although

the conformations of the peptide fragments are known to be flexible, it is probably not coincidental that the corresponding segments in the native proteins are also located in relatively flexible regions of the molecules concerned (Attassi, 1984; Tainer et al., 1984).

In some cases, there is evidence to show that the conformation of an isolated N-terminal peptide fragment resembles that of the corresponding segment in the intact protein. The peptide segment 1–15 in pancreatic ribonuclease A with a chain length of 124 residues is known to be helical. It has recently been shown by Mitchinson and Baldwin (1986) that the isolated N-terminal peptide fragment 1–15 is also helical when the $\alpha$-amino group of the N-terminal lysine residue is substituted, as would be the case during biosynthesis of this protein. Similarly, the isolated N-terminal fragment 1–45 of adenylate kinase with a chain length of 194 residues has been shown to be able to bind MgATP tightly like the native enzyme (Fry et al., 1985). Presumably it has a conformation closely similar to the N-terminal segment 1–45 in the intact molecule. The formation of part of the disulfide linkages in some proteins is known to be cotranslational, such as the Cys-35–Cys-100 loop of the nascent immunoglobulin light chain; this must depend on the correct juxtaposition of the two Cys residues in this region of the peptide fragment, which would in turn depend on the folding of this part of the molecule during biosynthesis into a conformation resembling, at least to some extent, that of the native protein. Indeed, on the basis of this observation, Bergman and Kuehl (1979) suggested that the folding of proteins in vivo is initiated sequentially from the N-terminal.

*Adjustments during and after Completion of Synthesis.* The suggestion that folding begins with the lengthening of the peptide chain does not necessarily mean that it ends with the completion of the synthesis of the entire peptide chain. Further growth of the length of the peptide chain could have important effects on the conformation of the peptide fragment already synthesized. Moreover, conformational adjustment of the whole molecule may be necessary when termination of peptide chain synthesis occurs, such as formation of disulfide bonds and other covalent modifications of the nascent chain for the generation of the properly folded protein molecule required for its biological activity.

Studies on the CD spectra of the separated and mixed insulin A and B chains show that although both chains contain significant amounts of secondary structure, interaction of the chains leads to an increase in ordered secondary structure, as shown by comparing the CD spectrum of the mixed A and B chains with the summation of the spectra of the individual chains separately, which indicates conformational adjustment upon pairing of the chains. Oxidation of the mixed reduced chains under suitable conditions favorable for the formation of a significant amount of the native hormone (Du et al., 1965) leads to a further increase in ordered secondary structure (Hua et al., 1985), especially the $\alpha$-helix content, indicating further folding with the formation of the disulfide bonds.

In this connection the careful study by Taniuchi (Taniuchi, 1970; Andria & Taniuchi, 1978) on the folding of pancreatic ribonuclease A with six residues removed at its C-terminal (fragment 1–118) is particularly significant. Under conditions where the intact reduced molecule can refold into its native conformation and recover its activity completely, reoxidation of the reduced and truncated protein leads to inactive products that differ significantly in conformation as well as in the joining of the disulfide bonds from the corresponding portion of the native enzyme. In a similar way, in the presence of the disulfide exchange enzyme the fragment 1–118 with native disulfide bonds loses both its native conformation and its ability to be complemented by the fragment 105–124 to generate enzymatic activity, indicating that its stable conformation is different from that of the corresponding portion of the native enzyme. The scrambled and inactive fragment 1–118 can be reactivated if reshuffling of the disulfide bonds by protein disulfide isomerase is carried out in presence of the C-terminal fragment 105–124, indicating that the folding of the fragment 1–118 is strongly affected by the C-terminal portion of the enzyme. From all the above, it seems clear the folding of ribonuclease A is probably completed when the elongation of the chain is at or near its end, but this does not necessarily exclude the possibility that folding has occurred as the N-terminal portion of the molecule is synthesized.

*Folding Is both Co- and Posttranslational.* Earlier authors had proposed cotranslational folding (DeCoen, 1970; Phillips, 1967), but this was later considered unlikely by a number of authors (Anfinsen & Scheraga, 1975; Baldwin & Creighton, 1980; Creighton, 1984). This was not surprising as short peptides were hitherto considered structureless and the C-terminal portions of some proteins were shown to be extremely important for the folding as well as for the activity of the proteins concerned. As new experimental evidence accumulates to show that peptides could have ordered secondary structures which may resemble the corresponding peptide segments in native protein molecules and that some of the disulfide bonds of certain proteins from cotranslationally (Bergman & Kuehl, 1979; Peters & Davidson, 1982), it now appears likely that the nascent peptide chain folds while it is being synthesized, as recently proposed by some authors (Bergman & Kuehl, 1979; Purvis et al., 1987).

Levinthal (1968) pointed out that had the folding of the peptide chain taken a random course, it would take a protein of 100 amino acid residues $10^{50}$ years to find a particular conformation. Clearly, the amino acid sequence of a peptide contains enough information to direct its folding. It seems that the peptide chain not only folds while being elongated but the folding of the segment already synthesized should affect the subsequent folding of the rest of the molecule. This is also consistent with the recent findings that the process of refolding of denatured proteins with a nucleation process from the folding of some peptide segments (Scheraga, 1980; Goldberg, 1985; Baldwin, 1986).

On the other hand, during the elongation of the peptide chain, the conformation of that part of the molecule already synthesized should be continuously affected by the newly synthesized segments and undergo constant adjustments. The folding is not entirely cotranslational in the sense that folding does not terminate at the same time with the growth of the peptide chain. On the contrary, results from the study on ribonuclease A (Taniuchi, 1970; Andria & Taniuchi, 1978) seem to suggest that, before the synthesis of the entire sequence is completed, the partially finished translation product does not have quite the same conformation as the native protein, and some final adjustments are necessary during the final stages of the synthesis of the peptide chain for the molecule to assume the conformation required for its activity. In other words, for ribonuclease A the C-terminal portion would in turn affect the folding of that part of the molecule already synthesized. However, this may not be the case for other proteins. The unfolded sequence of bacteriorhodopsin with 17 residues removed from its C-terminal folds under suitable conditions into the native conformation with full biological activity recovery (Liao & Khorana, 1984). Thus different proteins may

FIGURE 2: Proposed scheme for the folding of the nascent peptide chain during synthesis. Peptide segments fold while being synthesized, but adjustments are necessary both during and after completion of the entire peptide chain.

behave differently in their folding processes in vivo. Nevertheless, it is suggested that the folding of the growing peptide chain, generally speaking, begins early during synthesis, undergoes constant conformational adjustments with the elongation of the chain, and ends after the completion of the synthesis of the entire sequence. Folding is therefore both cotranslational and posttranslational as illustrated in Figure 2.

The above discussion on the folding of the nascent peptide chain is unavoidably sketchy because of lack of sufficient direct experimental evidence. Some important factors have been mentioned only briefly, such as the formation of disulfide linkages during biosynthesis of a protein, or neglected altogether, such as the effect of the local environment in the cell on the folding of the nascent peptide chain. This is particularly important for those proteins synthesized by polysomes on the endoplasmic reticulum and transported through membrane systems. Comparison of peptide fragments starting from the N-terminal end and of different chain lengths should provide a better model than the study of the refolding of denatured protein molecules containing the entire amino acid sequence for the process of folding in vivo.

ACKNOWLEDGMENTS

I thank Drs. C. C. Wang and J. H. Wang for their helpful comments.

REFERENCES

Andria, G., & Taniuchi, H. (1978) *J. Biol. Chem. 253*, 2262–2270.

Anfinsen, C. B. (1973) *Science (Washington, D.C.) 181*, 223–230.

Anfinsen, C. B., & Scheraga, H. (1975) *Adv. Protein Chem. 29*, 205–300.

Atassi, M. Z. (1984) *Eur. J. Biochem. 145*, 1–20.

Baldwin, R. L. (1986) *Trends Biochem. Sci. (Pers. Ed.) 11*, 6–9.

Baldwin, R. L., & Creighton, T. E. (1980) in *Protein Folding* (Jaenicke, R., Ed.) pp 217–260, Elsevier, Amsterdam.

Beadle, G. W., & Tatum, E. L. (1941) *Proc. Natl. Acad. Sci. U.S.A. 27*, 499–506.

Bergman, L. W., & Kuehl, W. M. (1979) *J. Biol. Chem. 254*, 8869–8876.

Betton, J.-M., Desmadril, M., Mitraki, A., & Yon, J. M. (1985) *Biochemistry 24*, 4570–4577.

Brandts, J. F., Halvorson, H. R., & Bremman, M. (1975) *Biochemistry 14*, 4953–4963.

Conconi, F. M., & Bank, P. A (1966) *J. Mol. Biol. 19*, 525–540.

Creighton, T. E. (1984) *Adv. Biophys. 18*, 1–20.

Creighton, T. E., Hillson, D. A., & Baldwin, R. B. (1980) *J. Mol. Biol. 142*, 43–62.

DeCoen, J. L. (1970) *J. Mol. Biol. 49*, 405–414.

Dice, J. F., & Goldberg, A. L. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3893–3897.

Du, Y. C., Jiang, R. Q., & Tsou, C. L. (1965) *Sci. Sin. (Engl. Ed.) 14*, 229–236.

Fry, D. C., Kuby, S. A., & Mildvan, A. S. (1985) *Biochemistry 24*, 4680–4684.

Garel, J.-R., & Baldwin, R. L. (1973) *Proc. Natl. Acad. Sci. U.S.A. 70*, 3347–3351.

Garel, J.-R., & Dautry-Varsat, A. (1980) in *Protein Folding* (Jaenicke, R., Ed.) pp 485–498, Elsevier, Amsterdam.

Goldberg, M. E. (1985) *Trends Biochem. Sci. (Pers. Ed.) 10*, 388–391.

Hermann, R., Rudolph, R., Jaenicke, R., Price, N. C., & Scobbie, A. (1983) *J. Biol. Chem. 258*, 11014–11019.

Hruby, V. J. (1985) *Pept., Anal., Synth., Biol. 7*, 1–14.

Hua, Q. X., Qian, Y. Q., & Tsou, C. L. (1985) *Sci. Sin. (Engl. Ed.) 28B*, 854–862.

Jaenicke, R., Krebs, H., Rudolph, R., & Woenckhaus, C. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1966–1969.

Johnson, C. M., & Price, N. C. (1987) *Biochem. J. 245*, 525–530.

Kepes, A., & Beguin, S. (1966) *Biochim. Biophys. Acta 123*, 546–560.

Kuwajima, K., Hiraoka, Y., Ikeguchi, M., & Sugai, S. (1985) *Biochemistry 24*, 874–881.

Lang, K., Schmid, F. X., & Fischer, G. (1987) *Nature (London) 329*, 268–270.

Levinthal, C. (1968) *J. Chim. Phys. 65*, 44–45.

Liao, M.-J., & Khorana, H. G. (1984) *J. Biol. Chem. 259*, 4194–4199.

Lin, L.-N., & Brandts, J. F. (1983) *Biochemistry 22*, 559–563.

Lin, L.-N., & Brandts, J. F. (1987) *Biochemistry 26*, 3537–3543.

Mitchinson, C., & Baldwin, R. L. (1986) *Proteins: Struct. Funct. Genet. 1*, 23–33.

Osterhout, J. J., Jr., & Nall, B. T. (1985) *Biochemistry 24*, 7999–8005.

Palmiter, R. D. (1975) *Cell (Cambridge, Mass.) 4*, 189–197.

Peters, T., Jr., & Davidson, K. (1982) *J. Biol. Chem. 257*, 8847–8853.

Phillips, D. C. (1967) *Proc. Natl. Acad. Sci. U.S.A. 57*, 484–495.

Price, N. C., & Stevens, E. (1983) *Biochem. J. 209*, 763–770.

Purvis, I. J., Bettany, A. J. E., Santigo, T. C., Coggins, J. R., Duncan, K., Eason, R., & Brown, A. J. P. (1987) *J. Mol. Biol. 193*, 413–417.

Rapaport, D. C., & Scheraga, H. A. (1981) *Macromolecules 14*, 1238–1246.

Scheraga, H. A. (1980) in *Protein Folding* (Jaenicke, R., Ed.) pp 261–288, Elsevier, Amsterdam.

Tainer, J. A., Getzoff, E. D., Alexander, H., Houghton, R. A., Olson, A. J., Lerner, R. A., & Hendrickson, W. A. (1984) *Nature (London) 312*, 127–134.

Taniuchi, H. (1970) *J. Biol. Chem. 245*, 5459–5468.

Tsou, C. L. (1986) *Trends Biochem. Sci. (Pers. Ed.) 11*, 427–429.

Wang, C. C., & Tsou, C. L. (1986) *Biochemistry 25*, 5336–5340.

Wetlaufer, D. B. (1981) *Adv. Protein Chem. 34*, 61–92.

Zhou, H. M., & Tsou, C. L. (1986) *Biochim. Biophys. Acta 869*, 69–74.