Biotechnol. Bioprocess Eng. **2001**, 6: 237-243

Protein Folding, Misfolding, and Refolding of Therapeutic Proteins

Hang-Cheol Shin

Department of Bioinformatics and Life Science, Soongsil University, Seoul 156-743, Korea
and Bioproducts Research Center, Yonsei University, Seoul 120-749, Korea

Abstract Substantial progress has been made towards understanding the folding mechanisms of proteins in vitro and in vivo even though the general rules governing such folding events remain unknown. This paper reviews current folding models along with experimental approaches used to elucidate the folding pathways. Protein misfolding is discussed in relation to disease states, such as amyloidosis, and the recent findings on the mechanism of converting normally soluble proteins into amyloid fibrils through the formation of intermediates provide an insight into understanding the pathogenesis of amyloid formation and possible clues for the development of therapeutic treatments. Finally, some commonly adopted refolding strategies developed over the past decade are summarized.

Keywords: protein folding, misfolding, amyloidosis, *E. coli* expression, refolding of inclusion bodies

INTRODUCTION

Understanding how a protein molecule finds its unique functional three-dimensional structure (protein folding) following its synthesis on the ribosome remains an unsolved problem in molecular biology. Protein folding is usually referred to as the second half of the genetic code. As such, a solution to the folding problem will not only complete our knowledge on the cellular events that link a gene sequence and the three-dimensional structure of a protein, but also play a key role in the rational design of novel protein molecules with a predetermined function. In addition, the understanding of protein folding has a variety of implications in biology in fields such as molecular recognition, diseases such as amyloidosis (disease state caused by amyloid formation), cell invasion, membrane fusion, and protein translocation [1].

Protein folding is the process by which an unstructured protein alters its conformation to reach its native structure. The protein folding problem consists of determining how and why this happens, that is, characterizing the typical structural changes that occur during folding and the interactions governing these changes. An important goal in the study of folding is to develop a model that can make quantitative predictions about the effects of certain factors, such as the amino acid sequence, chain topology, pH, salt concentration, and temperature, on the kinetics and thermodynamics of the folding process. The development of such a model requires detailed knowledge of the structural and thermodynamic properties of the different species involved in folding, as well as the kinetics that relate them.

Historically, protein folding studies began as denaturation/renaturation (unfolding/refolding) experiments shortly after Sanger [2] determined the sequence of insulin, thereby proving that the basic covalent structure of a protein is a linear peptide chain and that the amino acid sequence of the chain is unique for a given protein. By showing that an unfolded (denatured) polypeptide chain could spontaneously refold to form a native protein with full biological activity, Anfinsen [3] concluded that this sequence, by itself, contains all the information necessary to define the three-dimensional structure of the protein and, thus, its biological function. Yet how and why the linear sequence of a protein adopts a specific native conformation remains unsolved and poses a challenging intellectual puzzle for modern science.

The field of folding has expanded in recent years to embrace both biophysical studies of folding mechanisms, and the more complex, yet biologically important, issues of protein folding *in vivo* and the relationship between protein misfolding, aggregation, and disease. This review summarizes the above topics along with the refolding strategies currently used in the production of biologically active proteins in *E. coli*.

Levinthal Paradox and Folding Models

Consider a chain of 100 amino acids, and assume that each amino acid exists either in an extended, helical, or loop conformation. There are 3^{100} (or about 10^{48}) possible ways to arrange this chain. If changes of state take place at a rate of 10^{14} s^{-1} (estimated rate of bond rotation in small molecules), a random search for the right conformation would take of the order of 10^{34} s or 10^{26} years, which is orders of magnitude greater than the age

* **Corresponding author**
Tel: +82-2-820-0451 Fax: +82-2-825-1795
e-mail: hcshin@saint.ssu.ac.kr

of the universe [4]. Real proteins fold over a time scale of 10^{-1} to 10^{8} seconds, and for a small protein of 100 residues, folding in less than 1 second is not unusual. This is called Levinthal's paradox and suggests the necessity of folding pathways and probable folding intermediates.

Three mechanisms have emerged to provide these routes. These mechanisms simplify the search for the folded state by uncoupling the formation of the secondary structure from the tertiary structure, thereby removing the stringent requirement of the simultaneous formation of the secondary and tertiary structure. The framework model (or diffusion-collision-adhesion model) proposes that local elements of the native local secondary structure form independently of the tertiary structure [5,6]. These elements diffuse until they collide, successfully adhere, and coalesce to give the tertiary structure. The nucleation/growth model postulates that certain neighboring residues in the sequence form the native secondary structure that then acts as a nucleus from which the native structure propagates, in a stepwise manner [7]. Thus, the tertiary structure forms as a necessary consequence of the secondary structure. The hydrophobic collapse model hypothesizes that a protein collapses rapidly around its hydrophobic side chains and then rearranges from the restricted conformational space occupied by the intermediate [8,9]. Here, the secondary structure is directed by native-like tertiary interactions. In contrast, the jigsaw puzzle model proposes that each molecule of a protein folds by a different distinct path [10]. These models are illustrated in Fig. 1. The hydrophobic collapse mechanism and framework model imply the existence of folding intermediates, whereas nucleation does not. Due to the subsequent finding of many apparent folding intermediates, it is now assumed that the presence of intermediates on pathways is an essential requirement for folding.

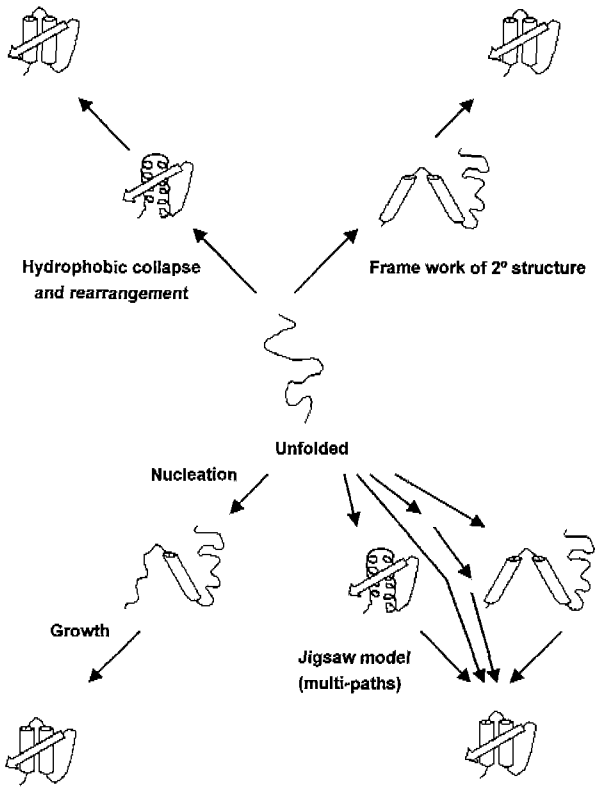

Fig. 1. Classical models for mechanisms of protein folding.

Protein Folding *in vitro* and *in vivo*

Protein folding in vitro is generally a spontaneous process leading to a native three-dimensional structure depending on the amino acid sequence of the protein [3]. However, in many cases, folding does not proceed efficiently because of intermolecular aggregation and slow folding reactions, such as disulfide formation or proline isomerization. The folding process can be particularly slow in the case of cysteine-containing proteins that require the formation of disulfide bonds [11]. Disulfide formation in vivo is catalyzed by specialized enzymes belonging to a thioredoxin superfamily, such as PDI (protein disulfide isomerase) in eukaryotes and Dsb (disulfide bond) proteins in prokaryotes.

PDI is a 55 kDa protein found in the endoplasmic reticulum (ER) where it catalyzes disulfide-bond formation and rearrangement during the maturation of most secretory and membrane-bound proteins [12,13]. PDI is not a particularly effective enzyme, consistent with its abundance in the ER, and high concentrations are often required to observe the catalysis of protein folding both *in vivo* and *in vitro* [14].

Generally, in the absence of catalysts such as PDI, the oxidative folding of cysteine-containing proteins that requires the formation of native disulfide bonds is considerably slower than the folding of proteins without disulfide bonds or the folding of the same protein with disulfide bonds intact [15,16]. This is mainly due to the formation of non-native disulfide bonds or kinetically trapped species with native disulfide bonds [17-19]. In both cases, the disulfides have to be rearranged by a thiol/disulfide exchange to reach the native state, and this rearrangement frequently constitutes the rate-determining step. In the folding pathways of ribonuclease A (RNase A), a well-known model protein for folding studies, one or more species in the ensemble of non-native three-disulfide intermediates undergo a disulfide rearrangement to form two native-like intermediates in the rate-determining steps. The regeneration process of RNase A is significantly more complex than the regeneration of the bovine pancreatic trypsin inhibitor (BPTI), another well-known protein that has been used in folding studies. During the regeneration of BPTI, only five well-populated disulfide-bonded intermediates accumulate [19]. However, in the case of RNase A, most of the theoretically possible intermediates have been found during regeneration [20,21]. To understand the protein folding mechanism in vivo, the effect of PDI on the kinetics and distribution of intermediates during the re-

generation of RNase A was studied in detail at 25°C [22]. According to this study, there is no PDI-altered folding mechanism; PDI simply appears to accelerate all steps of the oxidation, reduction, and rearrangement.

The rate of formation of a native protein at 37°C increases 62-fold in the presence of PDI [23], yet there is no noticeable change in the distribution, indicating that the folding mechanism is similar at 37°C in both the absence and presence of PDI. The cellular conditions of a high protein concentration, temperature, and ionic strength lead to aggregation, thereby disfavoring proper folding reactions, in contrast to well-optimized in vitro conditions. These results indicate that the significantly increased folding rate at a higher temperature in the presence of PDI can be an efficient means of solving unfavorable cellular conditions for protein folding and provide a clue to understand the folding mechanism in vivo, where folding takes place at 37°C in the presence of auxiliary factors, such as PDI.

Protein Misfolding and Disease

The failure of a protein to fold correctly leads to a functional deficit, which can have serious consequences. Cystic fibrosis is an example of a genetic disease where a variant protein is unable to fold correctly to a stable state in the endoplasmic reticulum, thereby failing to reach the cell membrane and be secreted in the quantity required for proper function. Other diseases, including some types of emphysema, result from mutations leading to the improper trafficking of proteins to the sites where they are needed. Recently, disease states, such as amyloidosis, which is caused by amyloid formation, have attracted a lot of attention. A list of some human diseases for which altered protein folding may be the molecular defect is presented in Table 1.

Amyloid formation is associated with some 20 sporadic, genetic, or infectious diseases, such as Alzheimer's, spongiform encephalopathy, and Huntington's disease [24]. Each of these diseases is associated with a particular protein and the aggregates of these proteins are thought to be the direct or indirect origin of the pathological conditions associated with the disease in question. Interestingly, these amyloid fibrils have similar morphologies, despite their origin from unrelated polypeptides, indicating that similar cellular processes are involved in the fibril deposition [25]. One of these diseases, Alzheimer's, has been estimated to be the most expensive medical problem in the Western world. Also, bovine spongiform encephalopathy (BSE), commonly known as mad cow disease, has recently caused a lot of concern around the world and come to the attention of the scientific community. BSE is associated with a transmissible agent and the nature of the BSE agent is still a matter of debate. According to the prion theory, the agent is largely composed, if not en-

Table 1. Some diseases caused by protein misfolding*

Disease	Mutant protein/protein involved	Molecular phenotype
Toxic fold		
Scrapie/Creutzfeldt-Jacob/Familial insomnia	Prion protein	Aggregation
Alzheimer's disease	β-Amyloid/α-synuclein fragment	Aggregation
Parkinson's disease	α-Synuclein	Aggregation
Huntington's diseae	Polyglutamine in huntingtin	Aggregation
Type-II diabetes	Amylin	Aggregation
Familial amyloidosis	Transthyretin/lysozyme	Aggregation
Cataracts	Crystallins	Aggregation
Inability to fold		
Cystic fibrosis	cystic fibrosis transmembrane conductance regulator (CFTR)	Misfolding/altered Hsp70 and clnexin interactions
Marfan syndrome	Fibrillin	Misfolding
Amyotrophic lateral sclerosis	Superoxide dismutase	Misfolding
Scurvy	Collagen	Misfolding
Maple syrup urine disease	α-Ketoacid dehydrogenase complex	Misassembly/misfolding
Cancer	p53	Misfolding/altered Hsp70 interaction
Osteogenesis imperfecta	Type I procollagen pro α	Misassembly/altered BiP Expression
Mislocalization owing to misfolding		
Familial hypercholesterolemia	LDL receptor	Improper trafficking
Familial emphysema	α_1-Antitrypsin	Improper trafficking
Tay-Sachs disease	β-Hexosaminidase	Improper trafficking
Retinitis pigmentosa	Rhodopsin	Improper trafficking
Leprechaunism	Insulin receptor	Improper trafficking

* Data mostly adaptedfrom [55].

tirely, of a self-replicating protein, referred to as a prion. Despite the association of prions with spongiform encephalopathies, no recombinant prion protein has yet been shown to be able to induce infectivity in animals and hence provide the ultimate proof of the 'protein only' hypothesis for prion diseases [26,27]. However, recently, the formation of fibrils with the characteristic morphology of the amyloid from a large recombinant fragment of the human prion protein has been reported [28]. This was achieved by exposing the soluble form of the protein to rather extreme conditions, including the reduction of the single disulfide bond that stabilizes the native structure. Interestingly, this conversion was associated with the presence of a soluble intermediate with a predominantly β-sheet structure, instead of the largely helical nature of the native protein.

Studies of the mechanism of the conversion of normally soluble proteins into amyloid fibrils have benefited from the fact that, in many cases, the structural transitions of disease-associated molecules can be reproduced under laboratory conditions [29]. In order to achieve this, a common procedure has been developed to expose the folded proteins to mildly denaturing conditions, such as low pH or elevated temperatures. Those transitions that have been studied have generally revealed the existence of intermediates prior to the formation of well-defined aggregates, and some studies have shown a clear kinetic lag phase before the rapid development of fibrils occurs. As studies on the mechanism of the conversion of soluble proteins into insoluble aggregates progress, a more generalized view on the pathogenesis of amyloid formation will become available.

To devise therapeutic treatments, several different approaches are feasible, including the suppression of the production of an amyloid precursor, prevention of amyloid formation, and stimulation of amyloid degradation. This important area of medical application will significantly benefit from a better understanding of protein folding/misfolding mechanisms.

Refolding of Therapeutic Proteins Produced as Inclusion Bodies

Major advances in recombinant DNA technology have resulted in the development of bacterial expression systems, particularly in *Escherichia coli*, capable of producing large amounts of proteins from cloned genes [30,31]. However, with some exceptions, these recombinant proteins give low yields of active material following high-level expression in *E. coli*, with the majority of the protein being deposited as inclusion bodies [32,33]. Therefore, to obtain biologically active recombinant proteins from inclusion bodies, it is necessary to develop a simple and efficient procedure for the renaturation of these proteins [34].

The formation of inclusion bodies offers several advantages in the production of recombinant proteins. If the target proteins are unstable in the cytoplasm of *E. coli* due to proteolysis or toxic to the host cell in the native conformation, it is preferable to produce the target proteins as inclusion bodies. Under appropriate conditions, the level of recombinant protein deposited in inclusion bodies can amount up to 50% of the total cellular protein. Because inclusion bodies have a relatively high density, they can be isolated from the cellular proteins by simple centrifugation, and the purity of the resulting preparation can reach 80-90% under optimal conditions. Therefore, the production of target proteins as inclusion bodies provides a cost-effective downstream process. However, the problem lies in the refolding process of the inclusion bodies since there is no known generalized procedure for efficient refolding, which can be applied to all inclusion body proteins. So far, refolding requires a trial-and-error type of approach even though several common features affecting refolding efficiency have been identified. Some of the commonly adopted refolding protocols are summarized in this section.

Inclusion bodies isolated by centrifugation are normally solubilized by strong denaturants, such as 6 M Gdn-HCl or 8 M urea, in combination with reducing reagents in the case of cysteine-containing proteins. Various experimental protocols used for the solubilization of inclusion bodies have been compared by Fischer *et al.* [35].

To obtain correctly folded proteins after solubilization of inclusion bodies, the excess denaturant and reducing agent have to be removed, and the reduced proteins transferred to oxidizing conditions. The renaturation of solubilized inclusion bodies is initiated by either dilution or dialysis to remove the denaturant. The efficiency of the renaturation depends on the competition between intra-molecular folding and inter-molecular aggregation. To slow down the aggregation process, refolding is usually performed at low protein concentration, within a range of 10-100 μg/mL. Furthermore, the renaturation conditions must be carefully optimized with regard to external parameters, such as temperature, pH, and ionic strength, for each individual protein [36].

Most secretory proteins contain disulfide bonds in their native state. If a target protein contains disulfide bonds, the renaturation buffer has to be supplemented with a redox system. The addition of a mixture of the reduced and oxidized forms of low molecular weight thiol reagents, such as glutathione, cysteine, and dithiothreitol, provides the appropriate redox potential to facilitate the formation and reshuffling of disulfides [36,37]. These systems increase both the rate and yield of renaturation/reoxidation by facilitating a rapid reshuffling of incorrect disulfide bonds [38]. In order to prevent the fortuitous oxidation of thiols by dissolved oxygen, catalyzed by trace amounts of metal ions such as Cu^{2+}, EDTA should be added to the buffer solutions. Some of the optimal redox conditions for refolding of disulfide-containing proteins from inclusion bodies are listed in Table 2.

In addition to controlling the temperature, pH, or redox conditions, the presence of low molecular weight compounds in the renaturation buffer can also have a

Table 2. Optimal conditions for refolding of disulfide-containing proteins from inclusion bodies by the glutathione system

Recombinant Protein	Number of disulfide bonds	Reduced/Oxidized glutathione (mM)	pH	Temp. (°C)	Time (h)	Ref.
Fab-fragment	5	5 / 0.5	8	10	150	41
ht-PA	17	0.5 / 0.3	8. 75	15	24	56
truncated ht-PA	9	2 / 0.2	8.6	20	24	57
truncated hM-CSF	9	0.5 / 0.1	8.5	4	48	58
hIL-2	1	10 / 1	8	RT	16	59
hIL-4	3	2 / 0.2	8	RT	16	60
hIL-6	2	0.01 / 0.02	8.5	22	4	61
hTIMP-1	6	2 / 0.2	8	4	16	62
truncated hTIMP-2	3	0. 78 / 0.44	9.75	25	2	63

marked effect on the renaturation yield [34,39]. The most popular additive is L-arginine [34]. In the case of human t-PA, the renaturation yield is markedly increased in the presence of 0.5 M L-arginine, whereas in its absence almost no reactivity is observed [38]. The positive effect of L-arginine on renaturation efficiency has also been confirmed for various other proteins, such as single-chain immunotoxins [40] and antibody Fab fragments [41]. The presence of polyethylene glycol (PEG) has been reported to significantly increase the recovery of active proteins by reducing aggregation, as shown in the cases of carbonic anhydrase B, deoxyribonuclease, t-PA (tissue-type plasminogen activator), and IFN-γ (interferon-gamma) [42,43]. Likewise, an increased solubilization of the folding intermediates can explain the positive effect of detergents on the refolding yield. The Use of lauryl-maltoside, CHAPS (3-[3-chloramidopropyl]dimethylammonia-1-propane sulfonate) or some other detergents during renaturation, improves the yield of a renatured protein [44].

Other approaches, such as production of target proteins as a fusion protein with an appropriate fusion partner that improves solubility and facilitates simple purification [45-50 *and references cited therein*], or secretion into periplasmic space that provides favorable condition for disulfide formation [51-53 *and references cited therein*], also show promise in devising an efficient way of producing therapeutically important proteins in *E. coli*, even though these areas of research are not the focus in this review. Many additional details and references can be found in Baneyx's excellent review [54].

As the human genome project (HGP) approaches completion, vast amounts of DNA sequence information are becoming available. In most cases, a gene sequence alone reveals little about the protein function or disease relevance. Accordingly, the true value of genome sequence information will only be realized after a function has been assigned to all the encoded proteins. Structural and functional analyses of proteins, especially those for therapeutic or industrial application, require large amounts of recombinant proteins. The E. coli system for the production of recombinant proteins as inclusion bodies, together with a suitable renaturation procedure, may provide an efficient avenue for meeting these requirements.

CONCLUSION

Proteins are engaged in promoting or controlling virtually every event on which our lives depend, as such, attaining a biologically active structure through the folding process is an essential step prior to the involvement of proteins in such events. This paper reviewed the various research areas related to protein folding and misfolding, and discussed the different experimental approaches.

A better understanding of amyloid fibrils and the way they form should result in improved knowledge of the pathological conditions that lead to many of the diseases that are currently emerging in the aging population of the modern world, thereby providing opportunities for their prevention and treatment. In addition, new ideas on the nature of protein sequences are anticipated from the various genome projects.

In the production of recombinant proteins using E. coli as the host cell, understanding the folding mechanism will provide important information in devising a downstream process, and the development of a generalized folding scheme will greatly facilitate the production of thousands of potential protein drug targets that will become available from the human genome sequence.

REFERENCES

[1] Dobson, C. M. and R. J. Ellis (1998) Protein folding and misfolding inside and outside the cell. *EMBO J.* 17: 5251-5254.

[2] Sanger, F. (1952) The arrangement of amino acids in protein. *Adv. Protein Chem.* 7: 1-67.

[3] Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* 181: 223-230.

[4] Levinthal, C. (1968) Are there pathways for protein folding? *J. Chim. Phys.* 65: 44-45.

[5] Kim, P. S. and R. L. Baldwin (1990) Intermediates in the folding reactions of small proteins. *Ann. Rev. Biochem.* 59: 631-660.

[6] Karplus, M. and D. L. Weaver (1994) Protein folding dynamics: the diffusion-collision model and experimental data. *Protein Sci.* 3: 650-668.

[7] Wetlaufer, D. B. (1990) Nucleation in protein folding--confusion of structure and process. *Trends Biochem. Sci.* 15: 414-415.
[8] Ptitsyn, O. B. (1995) How the molten globule became. *Trends Biochem. Sci.* 20: 376-379.
[9] Kuwajima, K. (1989) The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins: Struct. Funct. Genet.* 6: 87-103.
[10] Harrison, S. C. and R. Durbin (1985) Is there a single pathway for the folding of a polypeptide chain? *Proc. Natl. Acad. Sci. USA* 82: 4028-4030.
[11] Huppa, J. B. and H. L. Ploegh (1998) The eS-Sence of –SH in the ER. *Cell* 92: 145-148.
[12] Freedman, R. B., T.R. Hirst, and M.F. Tuite (1994) Protein disulfide isomerase: Building bridges in protein folding. *Trends Biochem. Sci.* 19: 331-336.
[13] Gilbert, H. F. (1998) Protein disulfide isomerase. *Methods Enzymol.* 290: 26-50.
[14] Lyles, M. M. and H. F. Gilbert (1991) Catalysis of the oxidative folding of ribonuclease A by protein disulfide isomerase: dependence of the rate on the composition of the redox buffer. *Biochemistry* 30: 613-619.
[15] Creighton, T. E. (1986) Disulfide bonds as probes of protein folding pathways. *Methods Enzymol.* 131: 83-106.
[16] Baldwin, R. L. (1989) How does protein folding get started? *Trends Biochem. Sci.* 163: 291-294.
[17] Hantgan, R. R., G. G. Hammes, and H. A. Scheraga (1974) Pathways of folding of reduced bovine pancreatic ribonuclease. *Biochemistry* 13: 3421-3431.
[18] Creighton, T. E. (1977) Conformational restrictions on the pathway of folding and unfolding of the pancreatic trypsin inhibitor. *J. Mol. Biol.* 113: 275-293.
[19] Weissman, J. S. and P. S. Kim (1991) Reexamination of the folding of BPTI – predominance of native intermediates. *Science* 253: 1386-1393.
[20] Xu, X. B., D. M. Rothwarf, and H. A. Scheraga (1996) Non-random distribution of the one-disulfide intermediates in the regeneration of ribonuclease A. *Biochemistry* 35: 6406-6417.
[21] Volles, M. J., X. B. Xu, and H. A. Scheraga (1999) Distribution of disulfide bonds in the two-disulfide intermediates in the regeneration of bovine pancreatic ribonuclease A: Further insights into the folding process. *Biochemistry* 38: 7284-7293.
[22] Shin, H.-C. and H. A. Scheraga (2000) Catalysis of the oxidative folding of bovine pancreatic ribonuclease A by protein disulfide isomerase. *J. Mol. Biol.* 300: 995-1003.
[23] Shin, H.-C. and H. A. Scheraga (1999) Effect of protein disulfide isomerase on the regeneration of bovine ribonuclease A with dithiothreitol. *FEBS Lett.* 456: 143-145.
[24] Dobson, C. M. (1999) Protein misfolding, evolution and disease. *Trends Biochem Sci.* 24: 329-32.
[25] Sunde, M. and C. C. F. Blake (1997) The structure of amyloid fibrils by electron microscopy and X-ray diffraction. *Adv. Protein Chem.* 50: 123-159.
[26] Prusiner, S. B. (1997) Prion diseases and the BSE crisis. *Science* 278: 245-251.
[27] Weissman, C. (1999) Molecular genetics of transmissible spongiform encephalopathies. *J. Biol. Chem.* 274: 3-6.
[28] Jackson, G. S., I. L. Hosszu, A. Power, A. F. Hill, J. Kenney, H. Saibil, C. J. Craven, J. P. Waltho, A. R. Clarke, and J. Collinge (1999) Reversible conversion of monomeric human prion protein between native and fibrilogenic conformations. *Science* 283: 1935-1937.
[29] Kelly, J. W. (1998) The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* 8: 101-106.
[30] Hartley, D. L. and J. F. Kane (1988) Properties of inclusion bodies from recombinant *Escherichia coli*. *Biochem. Soc. Trans.* 16: 101-102.
[31] Marston, F. A. (1986) The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. *Biochem. J.* 240: 1-12.
[32] Kane, J. F. and D. L. Hartley (1991) Properties of recombinant protein-containing inclusion bodies in *Escherichia coli*. *Bioprocess Technol.* 12: 121-145.
[33] Schein, C. (1990) Solubility as a function of protein structure and solvent components. *Bio/Technol.* 8: 308-317.
[34] Rudolph, R. and H. Lilie (1996) In vitro folding of inclusion body proteins. *FASEB J.* 10: 49-56.
[35] Fischer, B., I. Sumner, and P. Goodenough (1993) Isolation and renaturation of bio-active proteins expressed in *Escherichia coli* as inclusion bodies. *Biotechnol. Bioeng.* 41: 3-13.
[36] Lilie, H., E. Schwarz, and R. Rudolph (1998) Advances in refolding of proteins produced in *E. coli*. *Curr. Opin. Biotechnol.* 9: 497-501.
[37] Wetlaufer, D. B., P. A. Branca, and G.-X. Chen (1987) The oxidative folding of proteins by disulfide plus thiol does not correlate with redox potential. *Protein Eng.* 2: 141-146.
[38] Rudolph, R. (1990) In: H. Tschesche (ed.). *Modern Methods in Protein and Nucleic Acid Research*. pp.149-172, Walter de Gruyter, NY, USA.
[39] Hofmann, A., M. Tai, W. Wong, and C. G. Glabe (1995) A sparse matrix screen to establish initial conditions for protein renaturation. *Anal. Biochem.* 230: 8-15.
[40] Brinkmann, U., J. Buchner, and I. Pastan (1992) Independent domain folding of Pseudomonas exotoxin and single-chain immunotoxins: influence of interdomain connections. *Proc. Natl. Acad. Sci. USA* 89: 3075-3079.
[41] Buchner, J. and R. Rudolph (1991) Renaturation, purification and characterization of recombinant Fab-fragments produced in *Escherichia coli*. *Bio/Technol.* 9: 157-162.
[42] Cleland, J. L., C. Hedgepeth, and D. I. C. Wang (1992) Polyethylene glycol enhanced refolding of bovine carbonic anhydrase B. Reaction stoichiometry and refolding model. *J. Biol. Chem.* 267: 13327-13334.
[43] Cleland, J. L., S. E. Builder, J. R. Swartz, M. Winkler, J. Y. Chang, and D. I. C. Wang (1992) Polyethylene glycol enhanced protein refolding. *Bio/Technol.* 10: 1013-1019.
[44] Tandon, S. and P. M. Horwitz (1987) Detergent-assisted refolding of guanidinium chloride-denatured rhodanese. The effects of the concentration and type of detergent. *J. Biol. Chem.* 262: 4486-4491.
[45] Smith, D. B. and K. S. Johnson (1988) Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* 67: 31-40.
[46] Lavallie, E., E. A. DiBlasio, S. Kovacic, K. L. Grant, P. F.

Schendel, and J. M. McCoy (1993) A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. R. *Bio/Technol.* 11: 187-193.
[47] Shin, N.-K., D.-Y. Kim, C.-S. Shin, M.-S. Hong, J. Lee, and H.-C. Shin (1998) High level production of recombinant human growth hormone in *Escherichia coli* by a simple recombinant process. *J. Biotechnol.* 62: 143-151.
[48] Zhang, Y., D. R. Olsen, K. B. Nguyen, P. S. Olson, E. T. Rhodes, and D. Mascarenhas (1998) Expression of eukaryotic proteins in soluble form in *Escherichia coli*. *Protein Expr. Purif.* 12: 159-165.
[49] Forrer, P. and R. Jaussi (1998) High-level expression of soluble heterologous proteins in the cytoplasm of *Escherichia coli* by fusion to the bacteriophage lambda head protein D. *Gene* 224: 45-52.
[50] Meyer, D. E. and A. Chilkoti (1999) Purification of recombinant proteins by fusion with thermally-responsive polypeptides. *Nature Biotechnol.* 17: 1112-1115.
[51] Nilsson, B., E. Holmgren, S. Josephson, S. Gatenbeck, L. Philipson, and M. Uhlen (1985) Efficient secretion and purification of human insulin-like growth factor I with a gene fusion vector in Staphylococci. *Nucleic Acids Res.* 13: 1151-1162.
[52] Bedouelle, H. and P. Duplay (1988) Production in Escherichia coli and one-step purification of bifunctional hybrid proteins which bind maltose. Export of the Klenow polymerase into the periplasmic space. *Eur. J. Biochem.* 171: 541-549.
[53] Jeong, K. J. and S. Y. Lee (2000) Secretory production of human leptin in *Escherichia coli*. *Biotechnol. Bioeng.* 67: 398-407.
[54] Baneyx, F. (1999) Recombinant protein expression in *Escherichia coli*. *Curr. Opin. Biotechnol.* 10: 411-421.
[55] Thomas, P. J., B.-H. Qu, and P. L. Pedersen (1995) Defective protein folding as a basis of human disease. *Trends Biochem. Sci.* 20: 456-459.
[56] Sarmientos, P., M. Duchesne, P. Denefle, J. Boiziau, N. Fromage, N. Delporte, F. Parker, Y. Lelievre, J.-F. Mayaux, and T. Cartwright (1989) Synthesis and purification of active human tissue plasminogen activator from *Escherichia coli*. *Bio/Technol.* 7: 495-501.
[57] Kohnert, U., R. Rudolph, J. H. Verheijen, E. J. D. Weening-Verhoeff, A. Stern, U. Opitz, U. Martin, H. Lill, H. Printz, M. Lechner, G.-B. Kresse, P. Buckel, and S. Fischer (1992) Biochemical properties of the kringle 2 and protease domains are maintained in the refolded t-PA deletion variant BM 06.022. *Protein Eng.* 5: 93-100.
[58] Yamanishi, K., M. Takahashi, T. Nishida, Y. Ohtomo, M. Takano, S. Nakai, and Y. Hirai (1991) Renaturation, purification, and characterization of human truncated macrophage colony-stimulating factor expressed in *Escherichia coli*. *J. Biochem.* 109: 404-409.
[59] Tsuji, T., R. Nakagawa, N. Sugimoto, and K. Fukuhara (1987) Characterization of disulfide bonds in recombinant proteins: Reduced human interleukin 2 in inclusion bodies and its oxidative refolding. *Biochemistry* 26: 3129-3134.
[60] van Kimmenade, A., M. W. Bond, J. H. Schumacher, C. Laquoi, and R. A. Kastelein (1988) Expression, renaturation and purification of recombinant human interleukin-4 from *Escherichia coli*. *Eur. J. Biochem.* 173: 109-114.
[61] Ejima, D., M. Watanabe, Y. Sato, M. Date, N. Yamada, and Y. Takahara (1999) High yield refolding and purification process for recombinant human interleukin-6 expressed in *Escherichia coli*. *Biotechnol. Bioeng.* 62: 301-310.
[62] Hodges, D. J., D. G. Reid, A. D. Rowan, I. M. Clark, and T. E. Cawston (1998) Preparation of recombinant tissue inhibitor of metalloproteinases-1 (TIMP-1) in high yield and identification of a hydrophobic surface feature. *Eur. J. Biochem.* 257: 562-569.
[63] Williamson, R. A., D. Natalia, C. K. Gee, G. Murphy, M. D. Carr, and R. B. Freedman (1996) Chemically and conformationally authentic active domain of human tissue inhibitor of metalloproteinases-2 refolded from bacterial inclusion bodies. *Eur. J. Biochem.* 241: 476-488.

[Received May 22, 2001; accepted July 21, 2001]