

Review

## Macromolecule-Assisted *de novo* Protein Folding

Seong Il Choi <sup>1,2,\*</sup>, Ahyun Son <sup>2</sup>, Keo-Heun Lim <sup>2</sup>, Hotcherl Jeong <sup>3,\*</sup> and Baik L. Seong <sup>1,2,\*</sup>

<sup>1</sup> Translational Research Center for Protein Function Control, Yonsei University, Seoul 120-749, Korea

<sup>2</sup> Department of Biotechnology, College of Bioscience and Biotechnology, Yonsei University, Seoul 120-749, Korea; E-Mails: 50hyuny@naver.com (A.S.); keo-heun@hanmail.net (K.-H.L.)

<sup>3</sup> Vismar Co., Ltd., Ansan, Kyeonggi-do 426-791, Korea

\* Authors to whom correspondence should be addressed; E-Mails: seongilchoi@daum.net (S.I.C.); hotcherl.j@gmail.com (H.J.); blseong@yonsei.ac.kr (B.L.S.); Tel.: +82-2-393-4631 (S.I.C.).

Received: 26 July 2012; in revised form: 14 August 2012 / Accepted: 17 August 2012 /

Published: 20 August 2012

---

**Abstract:** In the processes of protein synthesis and folding, newly synthesized polypeptides are tightly connected to the macromolecules, such as ribosomes, lipid bilayers, or cotranslationally folded domains in multidomain proteins, representing a hallmark of *de novo* protein folding environments *in vivo*. Such linkage effects on the aggregation of endogenous polypeptides have been largely neglected, although all these macromolecules have been known to effectively and robustly solubilize their linked heterologous proteins in fusion or display technology. Thus, their roles in the aggregation of linked endogenous polypeptides need to be elucidated and incorporated into the mechanisms of *de novo* protein folding *in vivo*. In the classic hydrophobic interaction-based stabilizing mechanism underlying the molecular chaperone-assisted protein folding, it has been assumed that the macromolecules connected through a simple linkage without hydrophobic interactions and conformational changes would make no effect on the aggregation of their linked polypeptide chains. However, an increasing line of evidence indicates that the intrinsic properties of soluble macromolecules, especially their surface charges and excluded volume, could be important and universal factors for stabilizing their linked polypeptides against aggregation. Taken together, these macromolecules could act as folding helpers by keeping their linked nascent chains in a folding-competent state. The folding assistance provided by these macromolecules in the linkage context would give new insights into *de novo* protein folding inside the cell.

**Keywords:** folding; macromolecules; aggregation; hydrophobic interaction; molecular chaperone; surface charges; excluded volume

---

## 1. Introduction

What kinds of factors would drive newly synthesized polypeptides to efficiently fold into their native structures in crowded cellular environments? Over the last two decades, this question has been posed among other major issues in biology [1–4]. Despite much progress in our understanding of the “assisted” *de novo* protein folding *in vivo*, it still remains largely unsolved. Of note, it is extremely difficult to unambiguously uncover the underlying principles of the processes that include folding, misfolding, and aggregation due to the following reasons: They contain many complex features, such as transient and heterogeneous conformations, strict dependence of biological functions and physicochemical properties of proteins on their conformational states, and the irreversibility and multiphasic behavior of aggregation. In particular, elucidating the mechanisms of such processes *in vivo* becomes almost enigmatic when the influences of the cellular milieu are further taken into account. Along with these features, the diversity and flexibility of proteins would render the currently accepted fundamental principles fragile and biased with many exceptions and neglected aspects.

Efforts have been made to establish a single unifying principle governing protein folding. Indeed, the advent of new paradigms, such as the Anfinsen postulate and the concept of molecular chaperone, revolutionized our understanding of protein folding *in vitro* and *in vivo* [1,5]. Nonetheless, as the aforementioned complexity of the protein world implies, they appear to lack some important aspects, calling for a new viewpoint for describing *de novo* protein folding *in vivo*. In this review, we reexamine the Anfinsen postulate and the current status of molecular chaperones. We then discuss the folding assistance that is mediated by the nascent chain-linked macromolecules and their stabilizing factors against aggregation.

## 2. Uncovered Aspects of the Anfinsen Postulate and Molecular Chaperones

### 2.1. Anfinsen Postulate

Using a simple *in vitro* refolding system, Christian Anfinsen showed that all the information required for the native structure of a protein is encoded in its amino acid sequence [5]. According to his postulate, proteins can fold spontaneously, and their native structures are thermodynamically the most stable. This led us to a misconception that protein folding *in vivo* would occur in the same manner as the refolding *in vitro*. This principle turned out to be insufficient for explaining protein folding *in vivo*, based on the following considerations.

First, the Gibbs free energy of intramolecular folding conceptually provides no direct information about intermolecular association among polypeptides leading to aggregation, as pointed out previously [6]. Second, whereas single-domain proteins can fold *in vitro* in the time scale of micro-seconds to seconds, the refolding rates of larger multidomain proteins are sometimes very slow in the range of several hours to a day [7]. Such a long time scale *in vitro* appears to be physiologically irrelevant. In contrast,

cotranslational folding (at least domain-wise cotranslational folding) facilitates multidomain proteins to fold rapidly [8–11] and thus can play a decisive role in *de novo* protein folding. Moreover, translational pausing at rare codons was reported to significantly affect folding yields and even final conformations [12,13]. In addition, there exist folding catalysts, such as peptidyl prolyl isomerase and protein disulfide isomerase [14,15]. These cotranslational folding events and folding catalysts strongly indicate that protein folding can be pathway-dependent or under kinetic control. Finally, native structures are not necessarily minimum energy conformers as long as proteins can maintain their solubility by intrinsic or extrinsic factors. For example, a substantial fraction of proteins are intrinsically denatured proteins (IDPs) [16,17]. IDPs alone can maintain their solubility under physiological conditions due to relatively higher net charge [18]. The model of conformational selection and population shift suggests that the sparsely populated pre-existing native conformers in the ensemble of IDPs can recognize and bind to their partners, subsequently leading to the population shift toward the native conformers [19]. In these cases, the native conformers are thermodynamically unstable and yet able to efficiently exert their biological functions. Given the widely distributed IDPs in the proteome, a plethora of proteins are likely incompatible with the Anfinsen's thermodynamic hypothesis.

## 2.2. Molecular Chaperones

The principles of molecular chaperones have been gleaned mainly from the *in vitro* refolding experiments, using representative model proteins and simplified buffer solutions [2,3,20]. Obviously, however, there are significant differences between *in vitro* and *in vivo* environments with respect to cotranslational folding and macromolecular crowding [8–11,21]. Especially, the *in vitro* data of larger multidomain proteins regarding their folding rates and chaperone-dependent folding should be interpreted with every caution. For example, the *in vitro* refolding of firefly luciferase has been well known to be strictly dependent on the DnaK (*E. coli* hsp70 homolog) system [22,23]. However, *de novo* folding of firefly luciferase *in vivo* is independent of the DnaK system, although aggregation prevention and refolding of the denatured firefly luciferase during heat shock are dependent on this chaperone system [24].

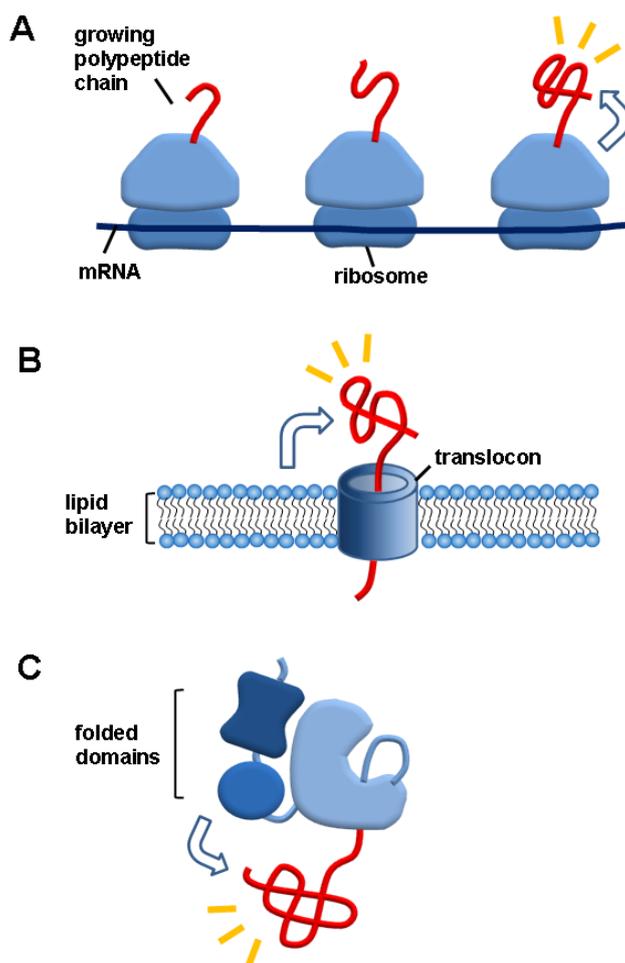
A biochemical approach using the co-immunoprecipitation (co-IP) assay showed that trigger factor (TF), DnaK, and GroEL/GroES interact with 60%, 9%–18%, and 10% of newly synthesized polypeptides, respectively [3,25,26]. These findings led to the conclusion that *de novo* folding of these proteins *in vivo* is dependent on their interacting molecular chaperones [3,26]. In fact, it is this co-IP assay that decisively established the molecular chaperones as general folding helpers *in vivo*. However, it should be noted that the hydrophobic interactions, which are considered responsible for the substrate recognition of molecular chaperones, are involved in the inhibition of protein folding. Indeed, this inhibitory ability of molecular chaperones is crucial for protein translocation [27]. In addition, molecular chaperones interact with proteins to fulfill other cellular functions, including proteolysis, translocation, and signal transduction [28–30]. These functions are conceptually different from *de novo* protein folding. Frydman articulated the limit of the co-IP approach by pointing out “*Unlike the genetic approach, these studies do not establish that folding of a given substrate requires a given chaperone but only that an interaction occurs in vivo.*” [31].

Genetic approaches are crucial for revealing the role of molecular chaperones in protein folding *in vivo*. The single deletion of either *dnaK* or *tf* gene in *E. coli* strain turns out to make no detectable effect on protein folding, whereas the deletion of both genes increases protein aggregation [25]. This aggregation can be overcome by overexpressing GroEL/GroES or SecB [32,33]. Most notably, Masters *et al.* showed that the physical depletion (knock-down) of GroEL via a tightly controlled expression system has little or no effect on the folding of newly made proteins in *E. coli* [34]. They further suggested that an *E. coli* strain without GroEL could be constructed although the folding of several essential proteins depends on this chaperonin [34]. Furthermore, GroEL is either non-essential or absent at least in some mycoplasmas [35]. In contrast, a significant or wholesale protein aggregation was observed in the *E. coli* strains harboring the GroEL conditional mutant (E461K) at the non-permissive temperature [36,37]. It has been widely believed that protein aggregation at the non-permissive temperature resulted from a loss-of-function mutation in GroEL. But Masters *et al.* pointed out that this wholesale aggregation could result from a gain-of-function mutation in GroEL [34]. Consistent with this suggestion, the GroEL mutant (E461K) does not release its bound substrate due to the loss of proper allosteric communications at non-permissive temperature [37–39]. A timely release of substrate proteins from molecular chaperones is crucial for proper folding. Otherwise they can act as folding inhibitors. Likewise, a conditional mutant of the mitochondrial chaperonin, which resulted in the global aggregation of imported proteins in the yeast mitochondria, was observed to almost completely aggregate at the non-permissive temperature [40]. It is therefore of great necessity to distinguish the protein aggregation generated by a gain-of-function mutation in chaperonin from that done by a loss-of-function mutation in chaperonin.

### 3. Macromolecule Linkage-Mediated Folding Helper Systems *in Vivo*

A hallmark of *de novo* protein folding environments *in vivo* is that, as illustrated in Figure 1, all nascent polypeptides without any exception are tightly connected to gigantic ribosomes (top), lipid bilayers in the case of secreted proteins across membranes (middle), or cotranslationally folded domains in the case of multidomain proteins (bottom). In particular, nascent chains are not released from ribosomes and lipid bilayers until the completion of translation and translocation, respectively. The potential aggregation problem of unfolded nascent chains on ribosomes or lipid bilayers has been a subject of intensive discussion in terms of *de novo* protein folding [3,41,42]. Moreover, the elucidation of the role of cotranslationally folded domains in their linked domains is crucial for understanding the *in vivo* folding of multidomain proteins. Nonetheless, the linkage effects of these macromolecules on the aggregation of their linked nascent chains have not been considered in the molecular chaperone-assisted *de novo* protein folding. By contrast, all these macromolecules have been known to solubilize their linked aggregation-prone heterologous proteins in the fusion technology or ribosome (or membrane surface) display technology [43–47]. In this review, we therefore suggest that the robust folding assistance of these macromolecules for their linked heterologous proteins could also play a pivotal role in *de novo* folding of “endogenous” proteins inside the cell.

**Figure 1.** Macromolecule linkage-mediated folding helper systems *in vivo*. A hallmark of *de novo* folding environments *in vivo* is that newly synthesized polypeptides (red tube) are tightly connected to the macromolecules, such as ribosomes (top), lipid bilayers (middle), or cotranslationally folded domains in multidomain proteins (bottom). Although the linkage effects (represented by arrows) on the folding or aggregation of their linked endogenous polypeptides remain largely unknown, these macromolecules have been known to effectively and robustly solubilize their linked heterologous proteins. Thus, this known folding assistance of these macromolecules in the linkage context could be applied to *de novo* folding of endogenous proteins *in vivo*.



### 3.1. Ribosomes

In 1986, Pelham put the following speculations on the roles of the hsc70 [48], contributing to the birth of the concept of molecular chaperone: “*What else might hsc70 do in unstressed cells? One possibility is that it recognizes nascent proteins, which by definition are “denatured,” and sorts out any aggregation problems that occur during their folding and assembly into oligomeric structures.*” The cooperative folding of domains and small proteins and the relatively slow rate of protein synthesis could potentially render the cytosol-exposed unfolded nascent chains on ribosomes as aggregation-prone for a significant timescale [3,41,42]. The degree of aggregation could be further increased by the high local effective concentration of nascent chains on the same polysomes and the

macromolecular crowding inside the cell [3,21]. Importantly, these assumptions provided a rationale for the existence of the ribosome-tethered molecular chaperones such as TF in *E. coli* [49,50]. However, the deletion of the TF gene or eukaryotic counterparts has no effect on protein folding [25,51].

Ribosomes are megadalton-sized macromolecules (~2.4 MDa in *E. coli*) with polyanionic surfaces. It is questionable if the aggregation behaviors of nascent chains alone are similar to those of the same chains that are linked to ribosomes. Of note, hydrophobic interactions do not reflect the properties of non-contact regions, such as their size, shape and surface charges. Thus, in the hydrophobic interaction-dominated stabilizing mechanism, ribosomes seem to be unnecessary for the description of the aggregation of their linked nascent chains. In contrast, the surface charges (generating electrostatic repulsions) and excluded volume (generating steric hindrance) of macromolecules, including ribosomes and molecular chaperones, could serve as important stabilizing factors against the aggregation of their linked polypeptides [6,52]. These two factors are discussed later in more detail.

The linkage effects of ribosomes could be inferred from the study on a polyhistidine tagged luciferase immobilized on chelating Sepharose beads [53]. The inhibition of aggregation and subsequent increase in the refolding yield by the bead immobilization enabled the authors to suggest that like the beads, ribosomes as massive particles could carry out similar functions for their linked polypeptides [53]. Previously, we proposed that ribosomes could keep their linked polypeptides in an aggregation-resistant and folding-competent state, due to their surface charges and excluded volume [6,52,54]. Ribosomes can efficiently promote the solubility and folding of their surface-linked aggregation-prone proteins *in vivo* that are fused to the ribosomal protein [43]. The ribosome display technology has been proven useful for the production of highly aggregation-prone domains as a functionally active form [44]. Furthermore, the folding experiments of a ribosome surface-exposed polypeptide by ribosome stalling with optical tweezers showed that the bound polypeptide remains aggregation-resistant and folding-competent [55]. Taken together, the aggregation problem of nascent chains of endogenous proteins on the ribosomes should be understood with taking these known roles of ribosomes into account. To our knowledge, there has been no report showing that the nascent chains linked to ribosomes give rise to aggregation problem.

Ribosomes and their components including 50S subunit, 23S rRNA, and the domain V of 23S rRNA have been known to function as molecular chaperones *in vitro* by reversibly interacting with proteins [56,57]. The regions including peptidyl transferase center (PTC) in the domain V are responsible for their substrate recognition. Interestingly, either neutral or positively charged amino acids on the surfaces of native structures of the tested proteins are recognized by the domain V [58]. So far, it is unknown what stabilizing factors of ribosomes and its components maintain their bound substrate in a folding-competent state. Both the charge and steric factors of ribosomes and their components might be responsible for the stabilization of their bound proteins.

### 3.2. Lipid Bilayers

As is the case of ribosomes, the aggregation problem may also persist during the translocation of newly made polypeptides across lipid bilayers of membranes of endoplasmic reticulum, mitochondria, chloroplast, and cell. The linkage effects of membranes on the aggregation of translocating polypeptides have not been considered in most studies of *de novo* protein folding. The cell surface

display technology allows protein of interest on the membranes or cell surfaces by fusing them to surface anchoring motifs [45,46,59–61]. The popular anchoring motifs for the *E. coli* outer membrane display include porins, lipoproteins, glycosyl-phosphatidylinositol anchor, and  $\beta$ -autotransporters [60]. Importantly, the display of proteins to the cell (or membrane) surfaces greatly stabilizes their linked proteins [45,46,62]. Similarly, the immobilization of proteins on the surfaces of beads is known to be a powerful method for stabilizing proteins [63,64]. The cells or membranes that display proteins were thought to be equivalent to the beads that immobilize proteins [45]. Artificial systems they might be, the accumulating evidence supports that the membrane linkage could stabilize the aqueous phase-exposed polypeptides of endogenous secreted proteins or membrane-anchored proteins against aggregation.

The membrane linkage could affect the substrate-stabilizing ability of the membrane-anchored chaperones and components, such as calnexin and translocon: these molecules can be considered to be the components of membranes. In consistent with this idea, the chaperoning activity of the membrane-anchored calnexin is more effective than that of the ER lumen-exposed domain of calnexin in solution without membrane anchor [65]. In addition, when immobilized at the surface of solid matrix, a small heat shock protein,  $\alpha$ B-crystallin, dramatically inhibit the aggregation of substrate proteins compared to  $\alpha$ B-crystallin in solution without the immobilization [66].

### 3.3. Folded Domains in Multidomain Proteins

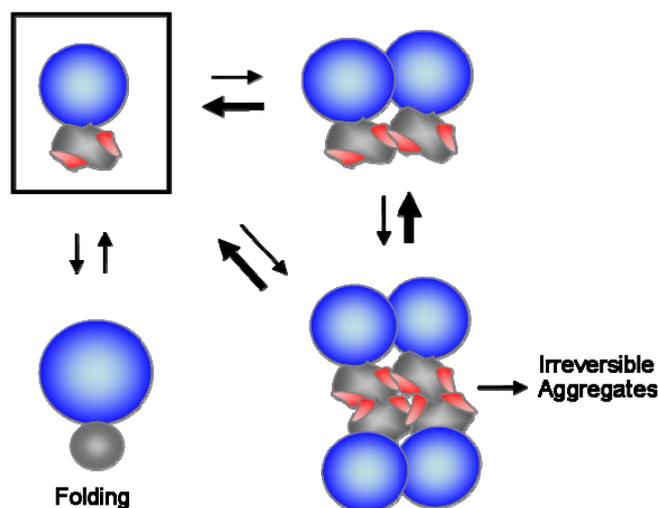
As evident in the *in vitro* refolding experiments with multidomain proteins, domains generally tend to inhibit the folding of other domains [7,67]. These *in vitro* experiments gave rise to the suggestion that multidomain proteins fold slowly and are aggregation-prone *in vivo*, thus necessitating assistance of molecular chaperones [3]. As mentioned before, the domain-wise cotranslational folding facilitates rapid folding of multidomain proteins *in vivo*, and yet the role of cotranslationally or independently folded domains remains largely unknown. In fact, demonstrating the linkage effects of folded domains, ribosomes, and membranes via the classic *in vitro* refolding method is not easy. It is partly due to the difficulties inherent to this method, where the denaturants required to unfold the target domains or linked proteins also impair or destroy the native forms of these macromolecules. In contrast, however, the *in vivo* fusion of aggregation-prone proteins to these macromolecules through genetic manipulation is relatively straightforward.

Fusion of soluble carriers to the *N*-termini of aggregation-prone proteins has been exploited as one of the efficient tools for overcoming aggregation of heterologous proteins in *E. coli* [68,69]. In essence, the artificial fusion proteins can be viewed as multidomain proteins, in which the *N*-terminal domains act as robust helpers for promoting the solubility and folding of their linked downstream domains. To demonstrate this chaperone-like function of *N*-terminal domains in natural multidomain proteins of *E. coli*, three *N*-terminal domains of *E. coli* proteins (lysyl tRNA synthetase, threonyl tRNA synthetase, and aconitase) were fused to the *N*-termini of highly aggregation-prone proteins [47]. All these domains profoundly enhanced the solubility and folding of a variety of their *C*-terminal heterologous proteins in the *E. coli* cytosol. Importantly, these findings strongly suggest that these *N*-terminal domains could also provide a similar chaperoning function for their authentic *C*-terminal domains *in vivo*, contributing to the autonomous folding of endogenous multidomain proteins. However, if the

upstream domains fail to fold properly, they would interfere with the folding of downstream domains. Therefore, depending on their folding status, domains can act as either folding enhancers or inhibitors.

The solubilizing ability of the *N*-terminal domains was shown to be closely correlated with their net charge and size. On the basis of these findings, a model was proposed to explain how soluble folded domains could increase the solubility and folding of their linked domains [47]. As shown in Figure 2, both electrostatic repulsions and steric hindrance of folded domains inhibit the intermolecular association and shift the populations toward the monomeric state, thus keeping the *C*-terminal domains in a folding-competent state. In particular, this model does not include the effects of interdomain interactions that could greatly influence the folding of multidomain proteins. Nevertheless, it explains well why soluble folded domains including soluble carriers can generally increase the solubility of a variety of their linked proteins. Importantly, this folding helper system mediated by soluble folded domains might be ubiquitous in the folding of endogenous multidomain proteins, because their charge and steric factors could stabilize their linked domains even without specific interdomain interactions.

**Figure 2.** A model for how soluble folded *N*-terminal domains can increase the solubility and folding of their linked domains. The blue, wrinkled, and gray spheres represent the folded *N*-terminal domains, incompletely folded and folded *C*-terminal domains, respectively. The red spots on wrinkled sphere indicate the aggregation-prone regions. The electrostatic repulsions and steric hindrance of the folded *N*-terminal domains inhibit intermolecular association and shift the populations from the oligomeric state to the monomeric state (boxed), thus increasing the chance for the proper folding of *C*-terminal domains. Reproduced from [47].



### 3.4. Salient Features of the Macromolecule Linkage-Mediated Folding Helper Systems

The macromolecule linkage-mediated folding helper systems described in this review have several salient features for *de novo* protein folding. First, they exert the chaperoning function in the linkage context without ATP consumption, in much the same way as an intramolecular (or autocatalytic) reaction. Second, while linked to the macromolecules, proteins can fold, leading to cotranslational folding or rapid folding. Rapid folding would allow proteins to escape from the adverse aggregation enhanced by the macromolecular crowding effects [70]. This feature is distinct from the hydrophobic

interaction-mediated mechanism by which protein folding generally stalls in the chaperone-bound state [27]. Third, soluble folded domains can persistently stabilize their linked domains against aggregation, basically independent of native interdomain interactions. Fourth, these systems can be conceptually applied to all nascent chains, as they are tightly linked to the ribosomes, lipid bilayers, or folded domains, throughout their biogenesis and folding. Lastly, nascent chains should communicate diligently with a variety of other molecules upon exposure to the cytosol. The interacting molecules include modifying enzymes, translocation systems, binding partners, and proteolytic systems [71,72]. These folding helper systems allow the nascent chains to fold in open space for the communication with their surroundings.

#### 4. Hydrophobic Interaction-Based Stabilizing Mechanisms of Molecular Chaperones

Depending on what stabilizing factors of macromolecules including molecular chaperones are important, our understanding of assisted *de novo* protein folding could be profoundly influenced. Thus, it is necessary to discuss the possible factors of macromolecules involved in protein aggregation. In the energetics of protein structure and folding and the theoretical force fields for their simulations, various components are involved, including conformational changes, hydrophobic effect, hydrogen bond, electrostatic interaction, peptide solvation, backbone conformational entropy, surface charge, and excluded volume [73–76]. These components could be applied and extended to intermolecular aggregation. Consistently, the hydrophobic interactions and hydrogen bonds are known to be important for protein aggregation [3,77,78]. Molecular chaperones prevent misfolding and aggregation by transiently binding to the exposed hydrophobic surfaces of non-native structures [2,3]. The hydrophobic interaction-mediated substrate recognition and stabilization against aggregation underlie the prevailing working mechanisms of molecular chaperones [2,3]. However, it remains largely unknown whether the hydrophobic interactions are prerequisite for the substrate recognition or stabilization against aggregation. Other stabilizing mechanisms cannot be excluded. For instance, even molecular chaperones, including GroEL, hsp70, TF, calnexin, and calreticulin, were reported to recognize their substrates mainly through the non-hydrophobic interactions, such as electrostatic interactions and glycan-binding [79–82].

GroEL assists protein folding by preventing aggregation without increasing folding rate (a passive but major role) or by increasing folding rate (an active role) [3,83–88]. Of note, the conformational changes of substrate protein by the bound chaperonin are not mandatory for the passive role of chaperonin [86]. This means that in the passive role of chaperonin, overcoming kinetic traps during folding process is entirely dependent on protein's own intrinsic ability, even though the overall folding yield is strictly dependent on chaperonin. As the representative models of the active role of chaperonin, there are iterative annealing mechanism by which the chaperonin stimulates folding by repetitive unfolding and release, and the active cage model by which the confinement and interactions with cavity wall result in smoothing of protein's energy landscape [3,85–88].

## 5. Charge and Steric Hindrance as Stabilizing Factors

### 5.1. Evidence for Charge and Steric Hindrance as Stabilizing Factors

Charge is a universal key factor determining the solubility of proteins as well as other molecules in the aqueous phase [89]. The net charge has been well established to be closely correlated with protein solubility [18,89–95]. For example, the supercharged variants of green fluorescence protein are extremely resistant to aggregation, such that they remain soluble even after boiling [94]. The aggregation rate is inversely correlated with the net charge [95]. The flanking charged residues or large soluble carriers, when covalently added, can increase the solubility of their linked polypeptides [96–99]. The net charge of soluble carriers is closely correlated with their solubilizing ability [47]. The increased electrostatic repulsions and hydration by charged groups are likely responsible for protein solubility. Taken together, accumulating evidence of the charge-mediated solubility enhancement clearly shows that the hydrophobic interaction-mediated direct contact is not a sole means for stabilizing the exposed hydrophobic regions against aggregation in the aqueous phase.

Protein aggregation is a specific process [100–102]. This means that stereospecificity is important for aggregation [103]. Thus, any steric hindrance resulting from the excluded volume of macromolecules, either covalently or non-covalently connected to proteins, is expected to profoundly inhibit aggregation. Previously, the steric hindrance of the macromolecules such as polysaccharides in *Peniophora lycii* phytase and polyethyleneglycol conjugated to human granulocyte colony stimulating factor was suggested to play a role in the stabilization of their linked proteins against aggregation [104,105]. The GroEL sequesters its substrate protein in the chamber, referred to as the Anfinsen cage, where an intermolecular interaction among substrate proteins is completely blocked [83]. In fact, this physical encapsulation can be viewed as a specific type of steric hindrance. In contrast, steric hindrance of macromolecules we here focus on can stabilize their linked polypeptides fully exposed to their surroundings.

### 5.2. Unique Features of the Charge and Steric Factors

Mechanistically, the aggregation inhibition by the charge and steric factors of macromolecules is fundamentally different from that by hydrophobic interactions, although they are not mutually exclusive. In particular, both charge and steric factors can inhibit aggregation without direct contact with aggregation-prone regions in the linkage context [6,47,52]. To highlight the unique features of these two factors of macromolecules with respect to aggregation, we here use simple illustrations, which are depicted in Figure 3. Let us take an example that a large soluble macromolecule (blue-colored) with a constant surface charge density is covalently linked to aggregation-prone polypeptide (red-colored), as shown in Figure 3A,B. In addition, neither conformational changes nor intermolecular interactions are assumed to exist between them. This situation can mimic the ribosome (or membrane surface)-linked nascent chains. In terms of conformational changes and hydrophobic interactions, the polypeptide, either alone or linked to macromolecule, can be thought to be same or similar with respect to folding and even aggregation. Since protein folding is generally an intramolecular reaction, the thermodynamic stability and folding (or unfolding) kinetics of the polypeptide can be little affected by the presence of the linked macromolecule. As for aggregation,

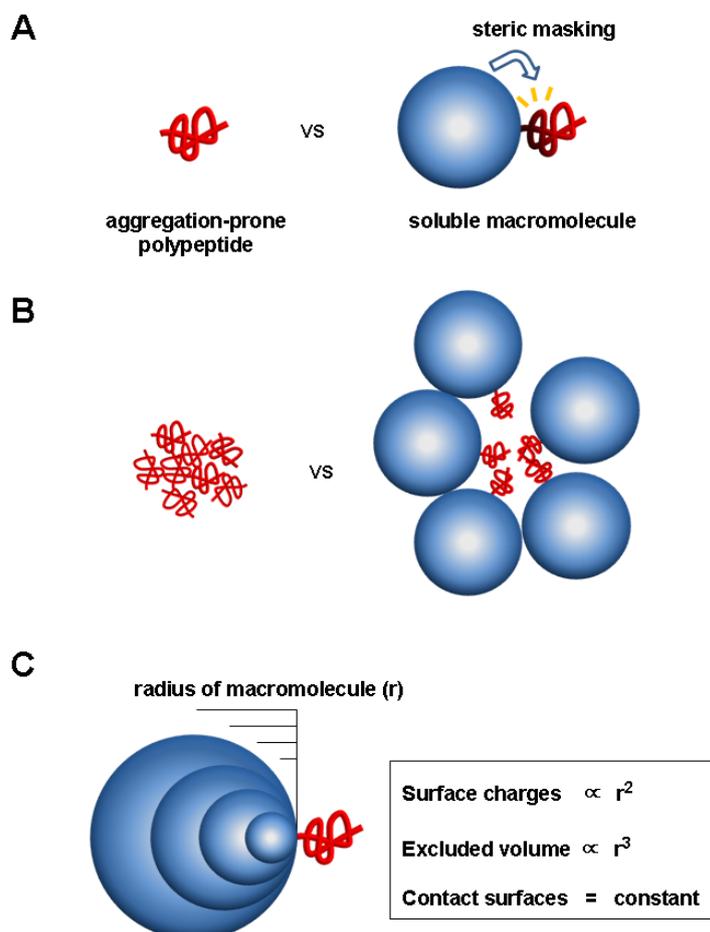
however, the local surfaces of polypeptide in close proximity to the linked macromolecule could be protected from intermolecular association, due to the direct steric masking of the macromolecule (Figure 3A), as pointed out previously [47].

Aggregation is a multimolecular reaction, neither unimolecular nor bimolecular. In a multimolecular system as shown in Figure 3B, it becomes more obvious that the total excluded volume of the macromolecule can directly or indirectly inhibit the aggregation of the whole regions of the polypeptide. Likewise, the total surface charges of the macromolecule are also expected to inhibit the aggregation of the whole regions of the polypeptide in a multimolecular reaction. Using either unimolecular or bimolecular system, however, it is difficult to explain these long-range effects of the surface charges and excluded volume of the macromolecules that can be exerted without direct interactions with aggregation-prone regions and induced conformational changes.

It is a challenging task to quantitatively determine the contribution of stabilizing factors of macromolecules. A representative chaperone DnaK binds to only tiny hydrophobic regions (e.g., NRRLLTG) of its substrates [106]. The hydrophobic interactions for the substrate recognition have been widely believed to be mainly responsible for stabilizing the bound substrate proteins against aggregation [2,3]. However, this important assumption underlying the action mechanisms of molecular chaperones is not well demonstrated experimentally. We constructed the fusion proteins in which DnaK is linked to the *N*-termini of aggregation-prone proteins to mimic the DnaK-substrate complex and expressed them in the *E. coli* cytosol [107]. In the fusion context, the residue or domain of DnaK crucial for its substrate recognition can be changed or deleted while maintaining the complex. Neither mutation of the crucial residue nor deletion of the whole substrate-binding domain has appreciable effect on the solubilizing ability of DnaK. These findings suggest that DnaK could have its intrinsic ability to stabilize its bound substrate proteins, independent of its hydrophobic interactions with substrate proteins [107].

To further elaborate on the stabilizing factors of macromolecules including DnaK, a simplified model is presented where the radius ( $r$ ) of macromolecule is changed while the macromolecule and polypeptide are non-covalently linked to each other via limited and constant hydrophobic interactions (Figure 3C). This model is modified from the previously reported [107]. It is likely that the replacement of the covalent linkage into the hydrophobic interactions between them little affects the stabilizing effects of intrinsic factors of macromolecule in the complex state. Given a radius ( $r$ ) of macromolecule, its surface charges and excluded volume are proportional to  $r^2$  and  $r^3$ , respectively, whereas the contact surfaces or intermolecular hydrophobic interactions remain constant, regardless to the change of macromolecule size. Thus, it was proposed that the surface charges and excluded volume of soluble macromolecules including molecular chaperones could serve as dominant and universal stabilizing factors for their linked proteins [6,107]. From this viewpoint, ribosomes and membranes appear to be ideal macromolecules for stabilizing their linked polypeptides. Consistent with the present model, the electrostatic and steric repulsions have been known to be mainly responsible for stabilizing colloids against aggregation [108]. In addition, the charged patches of hsp90 were reported to be important for the hsp90-mediated aggregation inhibition [109].

**Figure 3.** A simple model for assessing the stabilizing effects of the surface charges and excluded volume of macromolecule on the aggregation of its linked polypeptide. In this model, soluble macromolecule (blue sphere) with a constant surface charge density is linked to aggregation-prone polypeptide (red tube). **(A)** Unimolecular system. Assuming that macromolecule is covalently linked to the polypeptide without any conformational changes and intermolecular interactions between them, protein folding can be little affected by the macromolecule. But, as for aggregation, the local surfaces of polypeptide in close proximity to the macromolecule can be protected from intermolecular association by the direct steric masking of macromolecule. **(B)** Multimolecular system. In contrast to a unimolecular system, the total excluded volume and surface charges of the macromolecule can directly or indirectly inhibit the aggregation of the whole regions of the linked polypeptide in a multimolecular association. **(C)** Correlation of the stabilizing factors (surface charges, excluded volume, hydrophobic interactions) of macromolecule with its size. The soluble macromolecule with varying radius ( $r$ ) is linked to a polypeptide through limited and constant hydrophobic interactions. The surface charges (generating electrostatic repulsions) and excluded volume (generating steric hindrance) of the macromolecule are proportional to  $r^2$  and  $r^3$ , respectively, whereas the contact surfaces or intermolecular hydrophobic interactions remain constant, regardless to the size change of macromolecule. Thus, the surface charges and excluded volume of soluble macromolecules could be important stabilizing factors.



### 5.3. Significance of the Charge and Steric Factors of Macromolecules in Chaperoning Functions

The charge and steric factors of macromolecules would make a profound impact on our understanding of *de novo* protein folding *in vivo*. First, both factors well explains why soluble macromolecules, including ribosomes, lipid bilayers, and folded domains, can stabilize their linked polypeptides against aggregation, regardless of the nature of linkage type between them. Second, the hydrophobic interaction-mediated substrate recognition is not a prerequisite for chaperoning function. Third, the chaperoning activity mediated by both factors can be delivered to long range: direct interactions with aggregation-prone regions of substrate proteins are not mandatory for chaperoning function. Fourth, these factors were suggested to have a tendency to inhibit most, if not all, types of multimolecular assemblies, such as amorphous aggregation, native or non-native oligomers, and amyloid fibrils [52]. Finally, these factors are compatible with hydrophobic interactions.

## 6. Conclusions and Perspectives

The elucidation of the roles of the macromolecules, including ribosomes, lipid bilayers, and folded domains, in the aggregation of their linked nascent polypeptide chains is crucial for our understanding of *de novo* protein folding *in vivo*. Accumulating evidence clearly shows that all these macromolecules can promote the solubility and folding of their linked heterologous proteins. Nevertheless, little attention is paid to the roles of these macromolecules in *de novo* folding of endogenous proteins. Thus, the known chaperoning function of these macromolecules in the linkage context needs to be incorporated into the mechanisms of *de novo* folding of endogenous proteins. In addition, we suggest that the surface charges and excluded volume of soluble macromolecules could be important and universal factors for stabilizing their linked proteins against aggregation.

After the release from ribosomes or lipid bilayers, proteins interact with a variety of macromolecules. The charge and steric factors described in this review imply that these macromolecules might have the inherent ability to exert chaperoning function on their associated proteins. Thus, this new view would contribute to a better understanding of *de novo* protein folding, aggregation-associated diseases, and proteostasis.

## Acknowledgements

The work was supported by the National Strategic Research Grant from the Ministry of Knowledge Economy of the Korean Government (10031969) and the National Research Foundation of Korea Grant from the Korean Government (MEST) (2010-0001932).

## References

1. Hemmingsen, S.M.; Woolford, C.; van der Vies, S.M.; Tilly, K.; Dennis, D.T.; Georgopoulos, C.P.; Hendrix, R.W.; Ellis, R.J. Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature* **1988**, *333*, 330–334.
2. Bukau, B.; Horwich, A.L. The Hsp70 and Hsp60 chaperone machines. *Cell* **1998**, *92*, 351–366.
3. Hartl, F.U.; Hayer-Hartl, M. Molecular chaperones in the cytosol: From nascent chain to folded protein. *Science* **2002**, *295*, 1852–1858.

4. Hartl, F.U.; Bracher, A.; Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **2011**, *475*, 324–332.
5. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230.
6. Choi, S.I.; Ryu, K.; Seong, B.L. RNA-mediated chaperone type for *de novo* protein folding. *RNA Biol.* **2009**, *6*, 21–24.
7. Creighton, T.E. *Protein Folding*; W.H. Freeman and Company: New York, NY, USA, 1992; pp. 405–454.
8. Kolb, V.A.; Makeyev, E.V.; Spirin, A.S. Folding of firefly luciferase during translation in a cell-free system. *EMBO J.* **1994**, *13*, 3631–3637.
9. Fedorov, A.N.; Baldwin, T.O. Cotranslational protein folding. *J. Biol. Chem.* **1997**, *272*, 32715–32718.
10. Kramer, G.; Ramachandiran, V.; Hardesty, B. Cotranslational folding-omnia mea mecum porto? *Int. J. Biochem. Cell. Biol.* **2001**, *33*, 541–553.
11. Cabrita, L.D.; Dobson, C.M.; Christodoulou, J. Protein folding on the ribosome. *Curr. Opin. Struct. Biol.* **2010**, *20*, 33–45.
12. Zhang, G.; Hubalewska, M.; Ignatova, Z. Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.* **2009**, *16*, 274–280.
13. Kimchi-Sarfaty, C.; Oh, J.M.; Kim, I.W.; Sauna, Z.E.; Calcagno, A.M.; Ambudkar, S.V.; Gottesman, M.M. A “silent” polymorphism in the MDR1 gene changes substrate specificity. *Science* **2007**, *315*, 525–528.
14. Lilie, H.; Lang, K.; Rudolph, R.; Buchner, J. Prolyl isomerases catalyze antibody folding *in vitro*. *Protein Sci.* **1993**, *2*, 1490–1496.
15. Gruber, C.W.; Cemazar, M.; Heras, B.; Martin, J.L.; Craik, D.J. Protein disulfide isomerase: The structure of oxidative folding. *Trends Biochem. Sci.* **2006**, *31*, 455–464.
16. Dunker, A.K.; Silman, I.; Uversky, V.N.; Sussman, J.L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **2008**, *18*, 756–764.
17. Xue, B.; Williams, R.W.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N. Archaic chaos: Intrinsically disordered proteins in Archaea. *BMC Syst. Biol.* **2010**, *4*, S1.
18. Uversky, V.N.; Gillespie, J.R.; Fink, A.L. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* **2000**, *41*, 415–427.
19. Boehr, D.D.; Nussinov, R.; Wright, P.E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **2009**, *5*, 789–796.
20. Chakraborty, K.; Chatila, M.; Sinha, J.; Shi, Q.; Poschner, B.C.; Sikor, M.; Jiang, G.; Lamb, D.C.; Hartl, F.U.; Hayer-Hartl, M. Chaperonin-catalyzed rescue of kinetically trapped states in protein folding. *Cell* **2010**, *142*, 112–122.
21. Ellis, R.J. Macromolecular crowding: An important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.* **2001**, *11*, 114–119.
22. Szabo, A.; Langer, T.; Schröder, H.; Flanagan, J.; Bukau, B.; Hartl, F.U. The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70 system DnaK, DnaJ, and GrpE. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10345–10349.

23. Buchberger, A.; Schröder, H.; Hesterkamp, T.; Schönfeld, H.J.; Bukau, B. Substrate shuttling between the DnaK and GroEL systems indicates a chaperone network promoting protein folding. *J. Mol. Biol.* **1996**, *261*, 328–333.
24. Hesterkamp, T.; Bukau, B. Role of the DnaK and HscA homologs of Hsp70 chaperones in protein folding in *E. coli*. *EMBO J.* **1998**, *17*, 4818–4828.
25. Deuerling, E.; Schulze-Specking, A.; Tomoyasu, T.; Mogk, A.; Bukau, B. Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* **1999**, *400*, 693–696.
26. Houry, W.A.; Frishman, D.; Eckerskorn, C.; Lottspeich, F.; Hartl, F.U. Identification of *in vivo* substrates of the chaperonin GroEL. *Nature* **1999**, *402*, 147–154.
27. Koll, H.; Guiard, B.; Rassow, J.; Ostermann, J.; Horwich, A.L.; Neupert, W.; Hartl, F.U. Antifolding activity of hsp60 couples protein import into the mitochondrial matrix with export to the intermembrane space. *Cell* **1992**, *68*, 1163–1175.
28. Kandror, O.; Busconi, L.; Sherman, M.; Goldberg, A.L. Rapid degradation of an abnormal protein in *Escherichia coli* involves the chaperones GroEL and GroES. *J. Biol. Chem.* **1994**, *269*, 23575–23582.
29. De Los Rios, P.; Ben-Zvi, A.; Slutsky, O.; Azem, A.; Goloubinoff, P. Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 6166–6171.
30. Taipale, M.; Jarosz, D.F.; Lindquist, S. HSP90 at the hub of protein homeostasis: Emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 515–528.
31. Frydman, J. Folding of newly translated proteins *in vivo*: The role of molecular chaperones. *Annu. Rev. Biochem.* **2001**, *70*, 603–647.
32. Vorderwülbecke, S.; Kramer, G.; Merz, F.; Kurz, T.A.; Rauch, T.; Zachmann-Brand, B.; Bukau, B.; Deuerling, E. Low temperature or GroEL/ES overproduction permits growth of *Escherichia coli* cells lacking trigger factor and DnaK. *FEBS Lett.* **2004**, *559*, 181–187.
33. Ullers, R.S.; Luirink, J.; Harms, N.; Schwager, F.; Georgopoulos, C.; Genevaux, P. SecB is a bona fide generalized chaperone in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 7583–7588.
34. Masters, M.; Blakely, G.; Coulson, A.; McLennan, N.; Yerko, V.; Acord, J. Protein folding in *Escherichia coli*: The chaperonin GroE and its substrates. *Res. Microbiol.* **2009**, *160*, 267–277.
35. Wong, P.; Houry, W.A. Chaperone networks in bacteria: Analysis of protein homeostasis in minimal cells. *J. Struct. Biol.* **2004**, *146*, 79–89.
36. Horwich, A.L.; Low, K.B.; Fenton, W.A.; Hirshfield, I.N.; Furtak, K. Folding *in vivo* of bacterial cytoplasmic proteins: Role of GroEL. *Cell* **1993**, *74*, 909–917.
37. Chapman, E.; Farr, G.W.; Usaite, R.; Furtak, K.; Fenton, W.A.; Chaudhuri, T.K.; Hondorp, E.R.; Matthews, R.G.; Wolf, S.G.; Yates, J.R.; *et al.* Global aggregation of newly translated proteins in an *Escherichia coli* strain deficient of the chaperonin GroEL. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15800–15805.
38. Sewell, B.T.; Best, R.B.; Chen, S.; Roseman, A.M.; Farr, G.W.; Horwich, A.L.; Saibil, H.R. A mutant chaperonin with rearranged inter-ring electrostatic contacts and temperature-sensitive dissociation. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1128–1133.

39. Sot, B.; von Germar, F.; Mäntele, W.; Valpuesta, J.M.; Taneva, S.G.; Muga, A. Ionic interactions at both inter-ring contact sites of GroEL are involved in transmission of the allosteric signal: A time-resolved infrared difference study. *Protein Sci.* **2005**, *14*, 2267–2274.
40. Cheng, M.Y.; Hartl, F.U.; Martin, J.; Pollock, R.A.; Kalousek, F.; Neupert, W.; Hallberg, E.M.; Hallberg, R.L.; Horwich, A.L. Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. *Nature* **1989**, *337*, 620–625.
41. Ellis, R.J.; Hartl, F.U. Principles of protein folding in the cellular environment. *Curr. Opin. Struct. Biol.* **1999**, *9*, 102–110.
42. Feldman, D.E.; Frydman, J. Protein folding *in vivo*: The importance of molecular chaperones. *Curr. Opin. Struct. Biol.* **2000**, *10*, 26–33.
43. Sørensen, H.P.; Kristensen, J.E.; Sperling-Petersen, H.U.; Mortensen, K.K. Soluble expression of aggregating proteins by covalent coupling to the ribosome. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 715–719.
44. Schimmele, B.; Gräfe, N.; Plückthun, A. Ribosome display of mammalian receptor domains. *Protein Eng. Des. Sel.* **2005**, *18*, 285–294.
45. Wittrup, K.D. Protein engineering by cell-surface display. *Curr. Opin. Biotechnol.* **2001**, *12*, 395–399.
46. Ngo, K.X.; Umakoshi, H.; Shimanouchi, T.; Sugaya, H.; Kuboi, R. Chitosanase displayed on liposome can increase its activity and stability. *J. Bacteriol.* **2010**, *146*, 105–113.
47. Kim, C.W.; Han, K.S.; Ryu, K.S.; Kim, B.H.; Kim, K.H.; Choi, S.I.; Seong, B.L. N-terminal domains of native multidomain proteins have the potential to assist *de novo* folding of their downstream domains *in vivo* by acting as solubility enhancers. *Protein Sci.* **2007**, *16*, 635–643.
48. Pelham, H.R. Speculations on the functions of the major heat shock and glucose-regulated proteins. *Cell* **1986**, *46*, 959–961.
49. Wegrzyn, R.D.; Deuerling, E. Molecular guardians for newborn proteins: Ribosome-associated chaperones and their role in protein folding. *Cell. Mol. Life Sci.* **2005**, *62*, 2727–2738.
50. Craig, E.A.; Eisenman, H.C.; Hundley, H.A. Ribosome-tethered molecular chaperones: The first line of defense against protein misfolding? *Curr. Opin. Microbiol.* **2003**, *6*, 157–162.
51. Peisker, K.; Chiabudini, M.; Rospert, S. The ribosome-bound Hsp70 homolog Ssb of *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta.* **2010**, *1803*, 662–672.
52. Choi, S.I.; Lim, K.H.; Seong, B.L. Chaperoning roles of macromolecules interacting with proteins *in vivo*. *Int. J. Mol. Sci.* **2011**, *12*, 1979–1990.
53. Svetlov, M.S.; Kolb, V.A.; Spirin, A.S. Folding of the firefly luciferase polypeptide chain with immobilized C-terminus. *Mol. Biol. (Mosk)* **2007**, *41*, 96–102.
54. Choi, S.I.; Han, K.S.; Kim, C.W.; Ryu, K.S.; Kim, B.H.; Kim, K.H.; Kim, S.I.; Kang, T.H.; Shin, H.C.; Lim, K.H.; *et al.* Protein solubility and folding enhancement by interaction with RNA. *PLoS One* **2008**, *3*, e2677.
55. Kaiser, C.M.; Goldman, D.H.; Chodera, J.D.; Tinoco, I., Jr.; Bustamante, C. The ribosome modulates nascent protein folding. *Science* **2011**, *334*, 1723–1727.
56. Das, B.; Chattopadhyay, S.; Bera, A.K.; DasGupta, C. *In vitro* protein folding by ribosomes from *Escherichia coli*, wheat germ and rat liver: The role of the 50S particle and its 23S rRNA. *Eur. J. Biochem.* **1996**, *235*, 613–621.

57. Chattopadhyay, S.; Das, B.; DasGupta, C. Reactivation of denatured proteins by 23S ribosomal RNA: Role of domain V. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8284–8287.
58. Samanta, D.; Mukhopadhyay, D.; Chowdhury, S.; Ghosh, J.; Pal, S.; Basu, A.; Bhattacharya, A.; Das, A.; Das, D.; DasGupta, C. Protein folding by domain V of Escherichia coli 23S rRNA: Specificity of RNA-protein interactions. *J. Bacteriol.* **2008**, *190*, 3344–3352.
59. Chen, W.; Georgiou, G. Cell-Surface display of heterologous proteins: From high-throughput screening to environmental applications. *Biotechnol. Bioeng.* **2002**, *79*, 496–503.
60. Cornelis, P. Expressing genes in different *Escherichia coli* compartments. *Curr. Opin. Biotechnol.* **2000**, *11*, 450–454.
61. Jose, J. Autodisplay: Efficient bacterial surface display of recombinant proteins. *Appl. Microbiol. Biotechnol.* **2006**, *69*, 607–614.
62. Chen, Y.P.; Hwang, I.E.; Lin, C.J.; Wang, H.J.; Tseng C.P. Enhancing the stability of xylanase from *Cellulomonas fimi* by cell-surface display on *Escherichia coli*. *J. Appl. Microbiol.* **2012**, *112*, 455–463.
63. Polizzi, K.M.; Bommarius, A.S.; Broering, J.M.; Chaparro-Riggers, J.F. Stability of biocatalysts. *Curr. Opin. Chem. Biol.* **2007**, *11*, 220–225.
64. O’Fágáin, C. Engineering protein stability. *Methods Mol. Biol.* **2011**, *681*, 103–136.
65. Hebert, D.N.; Zhang, J.X.; Chen, W.; Foellmer, B.; Helenius, A. The number and location of glycans on influenza hemagglutinin determine folding and association with calnexin and calreticulin. *J. Cell. Biol.* **1997**, *139*, 613–623.
66. Garvey, M.; Griesser, S.S.; Griesser, H.J.; Thierry, B.; Nussio, M.R.; Shapter, J.G.; Ecroyd, H.; Giorgetti, S.; Bellotti, V.; Gerrard, J.A.; *et al.* Enhanced molecular chaperone activity of the small heat-shock protein  $\alpha$ B-cystallin following covalent immobilization onto a solid-phase support. *Biopolymers* **2011**, *95*, 376–389.
67. Netzer, W.J.; Hartl, F.U. Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* **1997**, *388*, 343–349.
68. Braun, P.; LaBaer, J. High throughput protein production for functional proteomics. *Trends Biotechnol.* **2003**, *21*, 383–388.
69. Esposito, D.; Chatterjee, D.K. Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.* **2006**, *17*, 353–358.
70. Van den Berg, B.; Ellis, R.J.; Dobson, C.M. Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.* **1999**, *18*, 6927–6933.
71. Kramer, G.; Boehringer, D.; Ban, N.; Bukau, B. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat. Struct. Mol. Biol.* **2009**, *16*, 589–597.
72. Meinnel, T.; Giglione C. Tools for analyzing and predicting N-terminal protein modifications. *Proteomics* **2008**, *8*, 626–649.
73. Kauzmann, W. Factors in interpretation of protein denaturation. *Adv. Protein Chem.* **1959**, *14*, 1–63.
74. Baldwin, R.L. Energetics of protein folding. *J. Mol. Biol.* **2007**, *10*, 283–301.
75. Bolen, D.W.; Rose, G.D. Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annu. Rev. Biochem.* **2008**, *77*, 339–362.

76. Kolinski, A. Protein modeling and structure prediction with a reduced representation. *Acta Biochim. Pol.* **2004**, *51*, 349–371.
77. Knowles, T.P.; Fitzpatrick, A.W.; Meehan, S.; Mott, H.R.; Vendruscolo, M.; Dobson, C.M.; Welland, M.E. Role of intermolecular forces in defining material properties of protein nanofibrils. *Science* **2007**, *318*, 1900–1903.
78. Fernández, A.; Crespo, A. Protein wrapping: A molecular marker for association, aggregation and drug design. *Chem. Soc. Rev.* **2008**, *37*, 2373–2382.
79. Misselwitz, B.; Staeck, O.; Rapoport, T.A. J proteins catalytically activate Hsp70 molecules to trap a wide range of peptide sequences. *Mol. Cell* **1998**, *2*, 593–603.
80. Aoki, K.; Taguchi, H.; Shindo, Y.; Yoshida, M.; Ogasahara, K.; Yutani, K.; Tanaka, N. Calorimetric observation of a GroEL-protein binding reaction with little contribution of hydrophobic interaction. *J. Biol. Chem.* **1997**, *272*, 32158–32162.
81. Martinez-Hackert, E.; Hendrickson, W.A. Promiscuous substrate recognition in folding and assembly activities of the trigger factor chaperone. *Cell* **2009**, *138*, 923–934.
82. Trombetta, E.S.; Helenius, A. Lectins as chaperones in glycoprotein folding. *Curr. Opin. Struct. Biol.* **1998**, *8*, 587–592.
83. Ellis, R.J. Protein folding: Importance of the Anfinsen cage. *Curr. Biol.* **2003**, *13*, R881–R883.
84. Apetri, A.C.; Horwich, A.L. Chaperonin chamber accelerates protein folding through passive action of preventing aggregation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17351–17355.
85. Tang, Y.C.; Chang, H.C.; Roeben, A.; Wischnewski, D.; Wischnewski, N.; Kerner, M.J.; Hartl, F.U.; Hayer-Hartl, M. Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* **2006**, *125*, 903–914.
86. Lin, Z.; Rye, H.S. GroEL-mediated protein folding: Making the impossible, possible. *Crit. Rev. Biochem. Mol. Biol.* **2006**, *41*, 211–239.
87. Todd, M.J.; Lorimer, G.H.; Thirumalai, D. Chaperonin-facilitated protein folding: Optimization of rate and yield by an iterative annealing mechanism. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4030–4035.
88. Kmiecik, S.; Kolinski, A. Simulation of chaperonin effect on protein folding: A shift from nucleation-condensation to framework mechanism. *J. Am. Chem. Soc.* **2011**, *133*, 10283–10289.
89. Shaw, B.F.; Moustakas, D.T.; Whitelegge, J.P.; Faull, K.F. Taking charge of proteins from neurodegeneration to industrial biotechnology. *Adv. Protein Chem. Struct. Biol.* **2010**, *79*, 127–164.
90. Schmittschmitt, J.P.; Scholtz, J.M. The role of protein stability, solubility, and net charge in amyloid fibril formation. *Protein Sci.* **2003**, *12*, 2374–2378.
91. Shaw, K.L.; Grimsley, G.R.; Yakovlev, G.I.; Makarov, A.A.; Pace, C.N. The effect of net charge on the solubility, activity, and stability of ribonuclease Sa. *Protein Sci.* **2001**, *10*, 1206–1215.
92. Tompa, P.; Csermely, P. The role of structural disorder in the function of RNA and protein chaperones. *FASEB J.* **2004**, *18*, 1169–1175.
93. Chiti, F.; Calamai, M.; Taddei, N.; Stefani, M.; Ramponi, G.; Dobson, C.M. Studies of the aggregation of mutant proteins *in vitro* provide insights into the genetics of amyloid diseases. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16419–16426.
94. Lawrence, M.S.; Phillips, K.J.; Liu, D.R. Supercharging proteins can impart unusual resilience. *J. Am. Chem. Soc.* **2007**, *129*, 10110–10112.

95. Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C.M. Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* **2003**, *424*, 805–808.
96. Marqusee, S.; Robbins, V.H.; Baldwin, R.L. Unusually stable helix formation in short alanine-based peptides. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5286–5290.
97. Chen, J.; Skehel, J.J.; Wiley, D.C. A polar octapeptide fused to the *N*-terminal fusion peptide solubilizes the influenza virus HA<sub>2</sub> subunit ectodomain. *Biochemistry* **1998**, *37*, 13643–13649.
98. Zhang, Y.B.; Howitt, J.; McCorkle, S.; Lawrence, P.; Springer, K.; Freimuth, P. Protein aggregation during overexpression limited by peptide extensions with large net negative charge. *Protein Expr. Purif.* **2004**, *36*, 207–216.
99. Kapust, R.B.; Waugh, D.S. *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* **1999**, *8*, 1668–1674.
100. Speed, M.A.; Wang, D.I.; King, J. Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. *Nat. Biotechnol.* **1996**, *14*, 1283–1287.
101. Rajan, R.S.; Illing, M.E.; Bence, N.F.; Kopito, R.R. Specificity in intracellular protein aggregation and inclusion body formation. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 13060–13065.
102. Wright, C.F.; Teichmann, S.A.; Clarke, J.; Dobson, C.M. The importance of sequence diversity in the aggregation and evolution of proteins. *Nature* **2005**, *438*, 878–881.
103. Fink, A.L. Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **1998**, *3*, R9–R23.
104. Høiberg-Nielsen, R.; Fuglsang, C.C.; Arleth, L.; Westh, P. Interrelationships of glycosylation and aggregation kinetics for *Peniophora lycii* phytase. *Biochemistry* **2006**, *45*, 5057–5066.
105. Rajan, R.S.; Li, T.; Aras, M.; Sloey, C.; Sutherland, W.; Arai, H.; Briddell, R.; Kinstler, O.; Lueras, A.M.; Zhang, Y.; *et al.* Modulation of protein aggregation by polyethylene glycol conjugation: GCSF as a case study. *Protein Sci.* **2006**, *15*, 1063–1075.
106. Rüdiger, S.; Germeroth, L.; Schneider-Mergener, J.; Bukau, B. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *EMBO J.* **1997**, *16*, 1501–1507.
107. Ryu, K.; Kim, C.W.; Kim, B.H.; Han, K.S.; Kim, K.H.; Choi, S.I.; Seong, B.L. Assessment of substrate-stabilizing factors for DnaK on the folding of aggregation-prone proteins. *Biochem. Biophys. Res. Commun.* **2008**, *373*, 74–79.
108. Hiemenz, P.C.; Rajagopalan, R. *In Principles of Colloid and Surface Chemistry*, 3rd ed.; CRC Press: London, UK, 1997; pp. 575–621.
109. Wayne, N.; Bolon, D.N. Charge-rich regions modulate the anti-aggregation activity of Hsp90. *J. Mol. Biol.* **2010**, *401*, 931–939.