



Review

# Is Protein Folding a Thermodynamically Unfavorable, Active, Energy-Dependent Process?

Irina Sorokina <sup>1,\*</sup>, Arcady R. Mushegian <sup>2,3</sup> and Eugene V. Koonin <sup>4,\*</sup>

<sup>1</sup> Strenic LLC, McLean, VA 22102, USA

<sup>2</sup> Division of Molecular and Cellular Biosciences, National Science Foundation, Alexandria, VA 22314, USA; mushegian2@gmail.com

<sup>3</sup> Clare Hall College, University of Cambridge, Cambridge CB3 9AL, UK

<sup>4</sup> National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

\* Correspondence: irinsoro@gmail.com (I.S.); koonin@ncbi.nlm.nih.gov (E.V.K.)

**Abstract:** The prevailing current view of protein folding is the thermodynamic hypothesis, under which the native folded conformation of a protein corresponds to the global minimum of Gibbs free energy  $G$ . We question this concept and show that the empirical evidence behind the thermodynamic hypothesis of folding is far from strong. Furthermore, physical theory-based approaches to the prediction of protein folds and their folding pathways so far have invariably failed except for some very small proteins, despite decades of intensive theory development and the enormous increase of computer power. The recent spectacular successes in protein structure prediction owe to evolutionary modeling of amino acid sequence substitutions enhanced by deep learning methods, but even these breakthroughs provide no information on the protein folding mechanisms and pathways. We discuss an alternative view of protein folding, under which the native state of most proteins does not occupy the global free energy minimum, but rather, a local minimum on a fluctuating free energy landscape. We further argue that  $\Delta G$  of folding is likely to be positive for the majority of proteins, which therefore fold into their native conformations only through interactions with the energy-dependent molecular machinery of living cells, in particular, the translation system and chaperones. Accordingly, protein folding should be modeled as it occurs in vivo, that is, as a non-equilibrium, active, energy-dependent process.

**Keywords:** protein folding; entropy; free energy; free energy landscape; energy-dependent protein folding; co-translational protein folding; molecular chaperones; physical model of protein folding



**Citation:** Sorokina, I.; Mushegian, A.R.; Koonin, E.V. Is Protein Folding a Thermodynamically Unfavorable, Active, Energy-Dependent Process? *Int. J. Mol. Sci.* **2022**, *23*, 521. <https://doi.org/10.3390/ijms23010521>

Academic Editors: Istvan Simon and Csaba Magyar

Received: 6 December 2021

Accepted: 31 December 2021

Published: 4 January 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

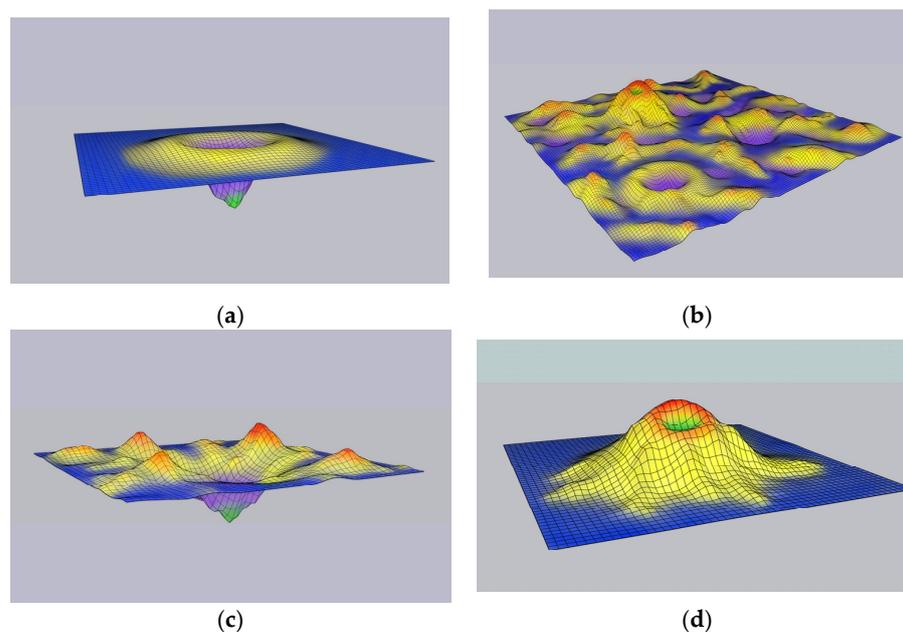
For the last six decades, the general understanding in the protein folding field has been that proteins fold into their native conformations driven by decrease in Gibbs free energy (negative  $\Delta G$ ). This thermodynamic hypothesis of protein folding stems from the iconic experiments of Anfinsen on in vitro folding of RNase A. Based on the successful refolding of this enzyme into the active, native conformation, Haber and Anfinsen concluded in a seminal 1962 paper that “the unique secondary and tertiary structure of RNase is, thermodynamically, the most stable configuration” [1]. Codified in Anfinsen’s 1973 Nobel lecture-based review [2], the thermodynamic hypothesis has become the default physical description of protein folding.

The thermodynamic hypothesis of folding, and in particular, the idea that the native state is the most stable one, that is, the global  $G$  minimum, is indeed highly attractive and appears natural. Furthermore, this view drastically simplifies theory development and modeling by effectively avoiding the need to explain how and why a protein reaches the unique native conformation: indeed, the global minimum is unique by definition. Assuming that the native conformation occupies a local rather than the global minimum of

$G$  immediately complicates the problem because this demands an explanation of how this particular minimum is selected among the many other local minima.

In the last two decades, the protein folding problem has been addressed primarily in terms of a free energy landscape that is usually represented as containing a funnel, the bottom of which corresponds to the global  $G$  minimum, that is, the native conformation; many different shapes of this hypothetical funnel have been considered [3–11].

Volcano-shaped landscapes have been also proposed, where during folding, the protein initially has to overcome a barrier of positive  $\Delta G$  due to entropy decrease, but the native structure still occupies the global  $G$  minimum [12,13] (Figure 1a). However, there is effectively no information on the actual structure of the landscape, and the possibility that the native conformation represents a local minimum in a rugged landscape that is generated and continuously affected by dynamic interactions within the cell environment, rather than the global minimum (Figure 1b), has not been systematically addressed. The distinction between the two classes of models can be formulated, in general terms, as thermodynamic vs. kinetic control of protein folding. Indeed, the early work of Wetlauffer and others (reviewed in [14]) emphasized that the native conformation would be the one with the minimum  $G$  among the kinetically accessible structures. However, this approach to the study of protein folding has not received much attention or further development, arguably, because it dramatically complicates modeling compared to the straightforward thermodynamic approach.



**Figure 1.** Protein folding energy landscapes in vitro and in vivo. Blue areas are occupied by “perfectly unfolded” conformations with no stable interactions between non-contiguous residues. Yellow and purple areas are populated by more compact protein conformations. Red zones are thought to arise mostly as the result of interactions between the protein and cellular components in a crowded environment. Green zones correspond to proteins in native conformation. (a) Canonical funnel-shaped energy landscape that most likely applies only to folding of small, thermodynamically stable proteins as it occurs spontaneously, in vitro, in isolation from all cellular compounds. (b) Folding energy landscape for a protein that folds in vivo is poorly understood, but most likely, is complex, rugged, dynamic, and shaped by interactions of the folding polypeptide with multiple cellular components. (c) Folding energy landscape of the same small protein as in (a) is most likely substantially different and far more complex when folding occurs in a crowded cellular environment. (d) Native conformations of most proteins are likely to occupy local thermodynamic minima with higher Gibbs free energy than their unfolded conformations (positive  $\Delta G$  of folding). Such native conformation can only arise as a result of active, energy dependent folding process.

Although the funnel landscape concept dominates the protein folding field, it is not without its critics. As argued in detail by Ben-Naim, the funnel folding landscape is effectively a metaphor that lacks substantial support [15]. Furthermore, as pointed out by Shakhnovich [16], in simulations, the shape of the landscape in low-dimensional spaces is sensitive to the procedure used for dimensionality reduction, and the procedures that yield the funnel landscapes tend to be physically unrealistic.

Despite the attractive simplicity of the notion that the native conformation of a protein occupies the global minimum of  $G$ , it remains a hypothesis. Several lines of evidence can be and often are construed as supporting this thermodynamic hypothesis, including direct measurements of the  $\Delta G$  of folding for multiple proteins, refolding of numerous proteins after denaturation, and spontaneous folding of proteins that were produced by complete chemical synthesis. Crucially, however, all this data pertains to a small number of small, highly stable proteins that have been studied *in vitro*, in isolation. Even apart from problems with the quantity and quality of this data, the question remains how generalizable these results are and how relevant are they for protein folding under native conditions, that is, in the crowded cell environment (compare panels (a) and (c) in Figure 1).

In this article, we critically assess the empirical data behind the thermodynamic hypothesis of protein folding and discuss an alternative, non-equilibrium folding hypothesis.

## 2. Review of Protein Folding

### 2.1. Experimental Data on Free Energy of Protein Folding

The  $\Delta G$  of protein folding can be determined from denaturation-renaturation experiments under the basic assumptions that proteins are completely denatured in the well-controlled experimental conditions and that such denaturation is fully reversible. A careful examination of the methodology of these experiments, however, reveals a complicated picture, with each class of methods employed for assessing the degree of denaturation rife with its own assumptions and biases (see, for example, [17] on chemical denaturation methods assayed by spectrophotometry, [18,19] for updates on urea- and guanidinium chloride-mediated denaturation methods, and [20,21] for thermal denaturation methods and microscanning calorimetry assays). A recent discussion of the biases, sensitivity issues, and other concerns in the analysis of denaturation-renaturation data can be found in [22]. All told, the results of such experiments, that have been reported for only a handful of proteins, have led to the general consensus that  $\Delta G$  of folding is a small negative value, that is, proteins (at least, single domain ones) fold spontaneously, but are only marginally stable reviewed in [23–25].

There are several reasons why, in our view, the experimental evidence in support of the thermodynamic hypothesis of folding is far less compelling than it is usually perceived to be. In particular, only in very few folding experiments, the completeness of protein unfolding at the start of the experiment has been convincingly demonstrated. Although it is often claimed that proteins in such experiments were completely denatured, a closer examination shows that typically this is an assumption rather than an experimentally validated observation. In early work (1950s–1970s), the extent of denaturation was typically assessed using indirect methods, such as circular dichroism (CD), which yields a general measure of the proportion of secondary structure in a protein, or fluorescence, which assesses the exposure of individual aromatic residues to the solute, or other, similarly indirect, approaches. However, reanalysis of a subset of cases with more advanced, direct methods has shown that proteins that have been initially characterized as completely denatured often turn out to be only partially unfolded [26]. For example, an NMR analysis of staphylococcal nuclease, the second enzyme extensively studied by Anfinsen after the seminal experiments with RNase A, has demonstrated persistence of native-like structure in the protein that was denatured in 8 M urea [27]. Subsequent NMR analysis of multiple, diverse proteins has similarly revealed preservation of extensive structure in 10 M urea [28]. Strikingly, for the paradigmatic case for the thermodynamic hypothesis, RNase A itself, advanced methods, such as small-angle X-ray scattering (SAXS) and time-resolved fluorescence energy transfer,

have demonstrated that compact regions survive many thermal and chemical denaturation regimes [29]. Furthermore, it has been shown that both the degree and the character of unfolding of the same protein can substantially differ depending on the denaturation protocol (e.g., [30,31]). A computational study of protein conformers, in which backbone torsion angles were randomly varied for only 8% of the residues, while the remaining 92% of the residues remained fixed in their native conformations, has shown that the vast majority of these ensembles had end-to-end distances and mean radii of gyration that were within the range of the random-coil expectations. Therefore, it has been concluded that observation of random-coil statistics for denatured proteins cannot be taken as evidence of the absence of residual structure [32].

Measurements of  $\Delta G$  of folding have been collected in protein thermodynamics databases, of which the most comprehensive one is ProTherm/ProThermDB. The latest release of this database [33] contains more than 30,000 entries. It would be important to determine how many records in the database are informative for estimating ranges of folding  $\Delta G$ . We analyzed the slightly smaller, 2017 release of ProTherm (available at [34]) that contained 26,045 records representing ~700 unique wild-type proteins, of which less than 500 were annotated as reversibly denatured. Strikingly, only for 18 of these proteins, denaturation was monitored using a rigorous method, such as NMR, and for three of those the actual values of  $\Delta G$  for the entire molecule were not reported (Supplementary Files S1 and S2). Thus, in actuality, the “vast body of curated literature” does not amount to much. Nearly all experimental data that are cited in support of the hypothesis of spontaneous refolding was obtained on a limited set of compact, stable, globular proteins. Most of these are small, single-domain monomers that are marginally stable ( $\Delta G$  of folding between  $-3.5$  and  $7$  kcal/mol) and have been shown to fold rapidly- typically, on the millisecond time scale [35]. Furthermore, the most thoroughly studied set of spontaneous refolders is enriched in extracellular proteins, often containing disulfide bonds, which are likely to dominate the fold stabilization mechanisms (see more on this below where we discuss total protein synthesis).

With all these caveats in mind, the reported  $\Delta G$  values are within the range of  $-1$  to  $-20$  kcal/mol, with a Poisson-like distribution peaking around  $-5$  kcal/mol [36]. The common range that is pervasively quoted in the literature is  $-5$ – $15$  kcal/mol, which is typically interpreted qualitatively as “proteins are marginally stable”, or in other words, the folding funnel is thought to be extremely shallow (e.g., [37–40]).

New opportunities to study the thermodynamics of protein folding/unfolding could be provided by single-molecule methods; for an overview of these methods as applied to protein folding, see [41,42], and for  $\Delta G$  measurement using these methods, see [43–45]. However, these studies face the same major problem as bulk denaturation experiments discussed above: most proteins do not unfold completely in single-molecule manipulations, such as atomic force microscopy or optical tweezers. Almost always, there is an uncertainty about the amount of residual structure, as indicated by the fact that the stretched form of a protein is often measured to be shorter, or occasionally, paradoxically longer than the theoretically predicted length (e.g., [46–48]). There also indications that the theoretical length of a polypeptide chain can be sequence- and structure-dependent [48]. Taken together, these data suggest that single-molecule methods are not yet sufficiently reliable for a confident determination of the state of protein unfolding.

Overall, the survey of the experimental study of protein folding/unfolding shows that  $\Delta G$  has been measured only for a highly biased set of a few small, compact, single domain proteins, and even for most of these, the obtained values cannot be considered reliable due to the lack of evidence of complete unfolding or, worse, presence of evidence of persisting, residual secondary structure. Even for those few proteins, for which reliable experimental data have been obtained, the negative  $\Delta G$  values were low, in many cases, not far above the level of thermal fluctuations.

## 2.2. Chemically Synthesized Proteins Folding into Native Conformations

Another major argument in support of the thermodynamic hypothesis of folding is thought to come from experiments on proteins obtained by complete chemical synthesis. By and large,  $\Delta G$  of folding for these proteins has not been measured directly, but is strongly believed to be negative because these proteins were produced by ligation of individual amino acids or peptides in a chemical reaction, in the complete absence of ribosomes, chaperones or other cell components, and then folded into native conformations in solution. Such spontaneous folding from the denatured form is commonly seen as direct, highly convincing evidence in support of the thermodynamic hypothesis (notably, the Nobel Prize has been awarded to Anfinsen only after the appearance of papers from Hirschmann and Merrifield groups on the complete synthesis of RnaseA [49,50]).

To assess the evidence obtained from this type of experiments, we performed a literature search on the proteins that, in the last 50 years or so, were produced by complete chemical synthesis and refolded to the biologically active form. In the set of about 60 unique proteins (not counting mutants and variants) studied in these experiments, only two were longer than 200 aa; the mean protein length in this group was 94 amino acids, which is at least 3–5-fold less than the proteome-wide mean lengths in Archaea, Bacteria, and Eukarya (Table 1). The proportions of secreted proteins and proteins containing disulfide bonds (DSB) in this dataset was several-fold higher than in the complete proteomes of various organisms (cysteine preference in these sequences is built in because modern methods of complete chemical synthesis assemble proteins from peptides, which usually requires internal cysteine residues for conjugation chemistry). Thus, recapitulating the properties of spontaneously refolded proteins discussed in the preceding section, the set of chemically synthesized proteins is heavily biased and not at all representative of real proteomes. Furthermore, there is no reliable data on those targets that were synthesized but could not be refolded. Even apart from these limitations, successful folding of chemically synthesized proteins requires non-physiological renaturation times in the hours' range. The yields of the native conformations are often omitted from the reports, but vary widely when reported (5–48%; Supplementary File S3); the folding protocols are complex and are developed on a case-by-case basis. Thus, even for this collection of privileged proteins, folding to the active forms in vitro is not straightforward and likely proceeds differently than in vivo.

**Table 1.** Properties of 59 proteins produced by total chemical synthesis and refolded to their active forms, as compared to the properties of whole proteomes.

	Total Chemical Synthesis <sup>1</sup>	Archaea	Bacteria	Eukarya	Data Sources for Archaea, Bacteria and Eukarya
mean protein length, amino acids	94	283	320	472	[51]
% secreted	62	6–19	18–30	13 (humans)	[52–54]
% with DSB in the known 3-D structures	57	15	11	30	[55]

<sup>1</sup> For the full data compilation from the literature, see Supplementary File S3.

Then, there is an even deeper problem with the folding of chemically synthesized proteins as the ultimate argument for the thermodynamics hypothesis. Through the course of the chemical synthesis, these proteins remain attached to solid phase, with limited degrees of freedom for the main chain rearrangement. The structure of the Gibbs energy landscapes (or other landscapes) for such proteins and their precursor peptides, before or after they are released from the solid phase into the solution, is completely unknown. These landscapes would be important to study, not only to resolve this open question as

such, but also because this might yield clues both to the mechanisms of protein folding inside cells and to the folding of primordial peptides during the early evolution of life (see discussion towards the end of this paper).

### *2.3. Refolding of Insoluble Overexpressed Proteins from Denatured Bacterial Inclusion Bodies into Soluble Active Proteins in Native Conformations*

A widely used approach to protein production is based on the fact that, when a recombinant protein is overexpressed in bacterial cells, it often forms insoluble inclusion bodies that are easy to isolate from other cellular components. Such aggregates of overexpressed proteins can be collected, further purified, denatured *in vitro* and often can be successfully refolded into soluble, active proteins. Because of the numerous industrial applications, protein purification and refolding from bacterial inclusion bodies have been extensively studied (for overview, see [56–59]).

If indeed the proteins that are purified from inclusion bodies were shown to unfold completely and then routinely refold to the native, active conformation, this would comprise strong evidence that spontaneous protein folding is common, in accord with the thermodynamic hypothesis. However, experiments with overexpressed proteins that form inclusion bodies resulted in a key observation that suggests quite different conclusions. Typically, proteins within the aggregates that form the inclusion bodies are neither disordered nor unfolded, but have specific secondary and tertiary structures, which are substantially ordered and are often enriched in in-register beta-sheets [60–63]. Detailed analysis of the refolding process shows that some of the ordered structure is preserved throughout the purification stages ([64–66], reviewed in [67]). Moreover, harsh denaturing conditions tend to be detrimental for protein refolding to the native conformation, so that new protocols for unfolding–refolding under mild conditions are constantly proposed in attempts to improve the yield of functional proteins (e.g., [68–70]).

Thus, experiments on refolding of overexpressed proteins demonstrate the key role of the residual secondary and tertiary structures, which are generated in the first place by the cell during protein expression *in vivo* and apparently have to be retained by the protein for efficient refolding. Furthermore, even when refolding occurs, it barely resembles the folding processes that occur in living cells. Indeed, purification and refolding of nearly every protein requires extensive protocol development, which often includes solutions and treatments that are far from physiological conditions and refolding times that are typically much longer than the biologically relevant time scale. All these efforts notwithstanding, the yields of the refolded native proteins vary widely [57,59]. Therefore, in general, refolding of proteins from inclusion bodies cannot be counted as a showcase for spontaneous refolding of completely denatured proteins and hence does not provide clear support for the thermodynamic hypothesis.

### *2.4. Scarcity of Data on $\Delta G$ of Protein Folding Reflects Pervasive Non-Refoldability and Instability of Proteomes*

A general conclusion from all of the above is that the evidence for the negative  $\Delta G$  of folding is quite thin, at best, and that the data on the ability of proteins to refold from a completely denatured state is fragmentary and comes from small, heavily biased datasets. What causes this scarcity of data, especially for larger proteins? The principal cause appears to be quite simple: most proteins actually cannot refold once completely unfolded, but the negative results of this type, that is, failed attempts to refold proteins, are almost never published. To our knowledge, no representative samples of proteins have been studied under this angle until very recently. However, in a recent proteome-wide study, protease-resistance assay was used in combination with quantitative mass-spectrometry to show that about 50% of the proteins in *E. coli* cell lysates could not refold into their native states following chemical denaturation, even when the conditions were optimized for refolding [71]. These findings indicate that non-refoldability *in vitro* is a general characteristic of at least this bacterial proteome, especially, taking into consideration that the completeness of unfolding was not monitored in these experiments.

Another widely observed but often misinterpreted phenomenon is the general proteome instability, that is, pervasive spontaneous loss of native structure and activity in proteins that have been originally properly folded in the living cells. This loss of native conformation and functional activity is commonly observed both in vitro, in preparations of isolated proteins, and in vivo. Indeed, it is well known that all cells maintain elaborate proteostasis machineries that functions to repair or destroy any proteins that have irreversibly lost their active conformations [72,73].

Spontaneous protein unfolding (denaturation) in vitro is an extremely common observation which suggests that at equilibrium many if not most proteins are in unfolded conformations. Unfortunately, to our knowledge, there is no published comprehensive statistics on protein (in)stability in vitro. Studies on protein stability and approaches to stabilization are a major expense in the pharma/biotech/industrial enzymology industry, and apparently, much of the results comprise intellectual property of these companies. The physical and chemical processes that are associated with instability have been thoroughly studied for a relatively small number of proteins [74–76]. The principal take-home message from these experiments is that even correctly folded proteins are often not intrinsically stable, as it would have been expected if they were in a deep free energy minimum, either global or local; instead, many lose native conformation easily.

Protein engineering experiments indicate that many proteins are easily destabilized with small sequence changes. However, despite years of research, predicting the effect of mutations on protein stability remains challenging. Nevertheless, the general conclusion is that most proteins are only barely stable, such that there are many destabilizing mutations (reviewed in [77,78]). Poorly understood tradeoffs seem to exist between protein stability and solubility such that a mutation on the exterior of a protein that increases its solubility is often destabilizing [79–82]. Similarly, there are tradeoffs between protein activity and stability such that mutations that enhance enzyme activity often destabilize the protein, and vice versa, stabilizing mutations often decrease activity [83].

Over the decades, many ad hoc explanations have been given for the spontaneous unfolding, denaturation, and destabilization that proteins typically undergo. Mostly, some irreversible events are postulated to occur during unfolding, such as protein oxidation, other chemical modifications, and/or aggregation, and such secondary effects are claimed to shift the equilibrium towards the unfolded state, preventing thermodynamically driven folding [74,76,84,85]. However, few targeted studies of protein denaturation mechanisms have been published. Usually, the loss of the native conformation and consequently activity by an isolated protein is perceived as a (often major) nuisance and is rarely seen as an opportunity to study the mechanisms of irreversible denaturation, and apparently, for this reason, not much systematic research has been done in this field. A notable exception are experiments of Klibanov and colleagues on the mechanisms of amylase denaturation. In these studies, the processes involved in thermal inactivation of this enzyme were dissected, showing that denaturation (unfolding), chemical modification, and aggregation are all distinct processes separated in time, and irreversible denaturation of this enzyme precedes chemical modifications and aggregation [86,87]. Several studies on other enzymes have also demonstrated that denaturation by irreversible chain unfolding is a process distinct from protein aggregation [88–90].

There is a call in the literature to apply modern approaches, such as new methods of spectroscopy and mass-spectrometry, for the proteome scale analysis of protein stability [91–94], but the actual experiments of this type remain to be performed.

A rough estimate of the failure rate of attempts on isolation of proteins in the native conformations can be extracted from large-scale structural genomics projects, which publish some statistics of protein production and purification. For example, Page et al. [95] reported that of more than 1800 proteins encoded in the genome of the hyperthermophilic bacterium *Thermotoga maritima* and cloned into expression plasmids, only 539 (~29%) could be purified in the form suitable for crystallization. In the Northeast Structural Genomics Consortium project [96], 6493 proteins could be purified out of the total 16,992 expressed (34%). The

New York Structural Genomics consortium has not reported consolidated statistics, but ~30% purified-to-expressed ratio seems to be a general trend across many participating projects [97]. It should be noted that in all these efforts, except for the *Thermotoga* case, the set of targets is strongly biased by pre-selection for predicted solubility, globularity, and evolutionary conservation. Even in these privileged sets of proteins, two-thirds could not be purified in the native conformation.

Generally, for the vast protein space, the thermodynamic parameters of protein folding and unfolding remain effectively unknown. There is no strong evidence that negative  $\Delta G$  of folding is a general property of many proteins. On the contrary, a wealth of data seem to present evidence against this possibility, showing instead that unfolding of most proteins occurs spontaneously, whereas folding does not. Thus, protein folding appears to be a non-equilibrium process that is accompanied by free energy increase.

### 2.5. Special Features of Protein Folding In Vivo

Complex proteostasis systems operate in every cell, and malfunction of these systems leads to (often lethal) accumulation of unfolded and misfolded proteins in the cell [72,73]. Proteomics shows that more than two thirds of all proteins in yeast specifically interact with one, or often more than one, of the proteostasis systems, and more specifically, with molecular chaperones [98]. Hundreds of proteins in *E. coli* interact with the chaperone GroEL-GroES alone [99]. The most abundant proteins in eukaryotes (actin, tubulin) do not fold in vitro at all, and to fold in vivo, they require, in addition to general chaperone systems, also the specialized co-chaperone prefoldin [100,101]. Apparently, proteostasis mechanisms consume a substantial fraction of the cellular energy supply; although the estimates do not seem to achieve high precision, the fraction of the energy budget dedicated to these processes is thought to be greater than 10% [102].

Chaperone clients are classified based on how closely they interact with the chaperones (for example, obligately-dependent vs. partially-dependent clients, based on the occupancy of the client-chaperone complexes [99]) and what, specifically, do they need chaperones for some proteins aggregate in the absence of chaperones, others stay soluble but are inactive, yet others need chaperones only under stress [103,104].

How do chaperones facilitate protein folding? The dominant view is that they help client proteins to quickly reach the minimum of free energy, that is, the chaperones create conditions for the thermodynamically driven folding of a substrate protein molecule into the native conformation. Some specific mechanisms include: (1) holding the client in isolation so that it does not aggregate with other proteins and folds correctly by itself, a mechanism known as “Anfinsen’s cage” in the case of GroES/GroEL [105–107], (2) preventing client proteins from getting stuck in kinetic traps during folding, conceivably, via partial unfolding [108–110], (3) unfolding misfolded or aggregated substrates before proceeding with mechanisms (1) or (2) [111–114], (4) reshaping the folding landscape in ways different from mechanism 2, known as “kinetic assistance”, but typically not specified further [115–118].

Some chaperones, known as foldases, are ATPases, whereas others, dubbed holdases, are not [119], but the distinction does not appear to map well onto the mechanisms listed above. Indeed, some chaperones from each functional class seem to exercise mechanisms 1–4 (see, e.g., [120]), whereas some appear to combine properties of foldases and holdases, as argued for the ATP-independent chaperone trigger factor [121] as well as the ATP-dependent HSP70 [122,123].

Crucially, all proposed chaperone mechanisms are predicated on the thermodynamic hypothesis, and to our knowledge, the relevance of these mechanisms has not been tested against the alternatives. A different view of the chaperone mechanisms will be discussed below.

### 2.6. Is Protein Folding In Vivo an Active, Energy-Dependent Process?

In our view, the above discussion shows that there is very little experimental evidence that  $\Delta G$  of folding is negative for most proteins. Conversely, a massive amount of experi-

mental observations indicates that native conformations of proteins are only metastable. Taken together, these lines of evidence compel us, in the least, to seriously consider the possibility that, for the majority of proteins,  $\Delta G$  of folding is positive (Figure 1d). The key implication of this hypothesis is that protein folding *in vivo* does not occur spontaneously, but rather, is an active, energy-dependent process.

This conceptual shift in our understanding of protein folding further implies that the ribosome itself is likely to act as a giant chaperone and the most important part of the protein folding machinery [124–126]. Clearly, this possibility is fully compatible with the numerous observations indicating that folding of most if not all proteins occurs co-translationally [127–151].

The most obvious way the ribosome could cause the increase in the Gibbs free energy that seems to accompany protein folding is by lowering the entropy of the protein by reducing the number of possible conformations of the peptide backbone. Some strained conformations with elevated enthalpy appear possible, too.

A common assumption in modeling of protein folding and in theoretical discussions is that the protein backbone can be well approximated by a freely jointed chain (FJC), so that all energy that could be applied to it would rapidly dissipate because of unrestricted rotation around each psi and phi bond. However, theoretical argument against this assumption, based on the available data on excluded volume effects and steric hindrances, has been brought up (e.g., [152]). Recently, our all-atom molecular dynamics simulations have revealed the situations when the backbone indeed is not FJC. When rotational force is applied to the protein backbone during the simulation, diverse helical peptides, despite their purported freedom to rotate about the psi and phi bonds, rapidly fold into the native structure, which remains stable [153].

It is important to recall that translation is coupled to the hydrolysis of massive amounts of GTP, but there is no clarity as to what this energy is actually expended on [154]. If  $\Delta G$  of protein folding is positive, it appears likely that at least some fraction of the energy of GTP hydrolysis contributes to active, non-equilibrium, co-translational folding. Apart from the ribosome, other molecular players are likely to be involved in active co-translational (and “co-translocational”) folding as well, in particular, the signal recognition particle (SRP) that contains two GTPases of its own, while the role of GTP hydrolysis is no better understood than it is in the case of the ribosome [155].

The energy-coupling machine framework has been suggested also for chaperone mechanisms as an alternative to the Anfinsen’s cage. Once again, it is unclear what the energy of ATP hydrolysis by ATP-dependent chaperones is actually spent on. Most studies link the ATPase activity with rearrangements of the chaperone itself [156,157]. However, the energy balance of these reactions remains unknown, and the possibility of coupling between ATP hydrolysis and the client protein rearrangements is typically not even considered because folding is assumed to be spontaneous. In contrast, a series of studies pioneered by Lorimer, De Los Rios and Goloubinoff argue that ATP-dependent chaperones, such as HSP60, HSP70 or HSP90, might expend at least part of the energy of ATP hydrolysis to manipulate the substrate directly (“non-equilibrium activation”) although the mechanistic details remain unclear [158–163].

Apart from the empirical evidence and thermodynamic considerations, the notion of active, non-equilibrium protein folding also appears to be better compatible with the evolutionary history of the relevant cellular components than the thermodynamic hypothesis of spontaneous folding. Indeed, the ribosome, translation factors with GTPase activity, and the SRP are universal to all cellular life, and several key chaperones also are among the most highly conserved proteins. All these molecular machines are likely to antedate the Last Universal Cellular Ancestor [164–166]. During all the 4 billion years or so of their existence, natural selection (including purifying selection for most of this time) would have acted primarily on the foldability of proteins on these machines, rather than their ability to fold/re-fold spontaneously. Perhaps, spontaneous folding could be a factor only

occasionally, in particular, for secreted proteins that have little access to chaperones once outside the cell.

### 2.7. Towards a Realistic Physical Model of Active Protein Folding

If the thermodynamic concept of protein folding generally fails, a new physical model of protein folding as an active, energy-dependent process is needed. Where to start? To begin with, a better definition of a “perfectly unfolded” protein conformation is essential. In such a fully unfolded conformation, there are no stable contacts between any two amino acids that are not adjacent in the polypeptide chain. It is currently unclear whether a perfectly unfolded conformation actually exists *in vitro* or *in vivo* for any particular protein, but this definition will be an appropriate starting point for building a physical model and recreating the folding process *in silico*.

Depending on the length of the polypeptide chain, there are theoretically on the order of  $10^{100}$  perfectly unfolded conformations for each protein [167,168]; more recently, suggestions have been made for a more conservative upper bound which, however, remains astronomically high [169]. From this vast set of perfectly unfolded conformations, one can build up and arrive at all kinds of structures: active intrinsically disordered (“natively unfolded”) conformations, stable misfolded conformations, conformations that only emerge through interaction with other proteins, classic globular native conformations with hydrophobic cores, and more. What do we know about “perfectly unfolded” conformations? If we measure (calculate) Gibbs free energy for these conformations, we should observe approximately the same value for each of them because these are random conformations with no interactions other than with the solvent. Even a single contact that forms within the polypeptide chain, whether it is a short or long distance one, makes the polypeptide more compact and increases the Gibbs free energy both due to the entropy reduction resulting from limiting the degrees of freedom and to changing enthalpy if, for example, the contact is hydrophobic or ionic.

Considering that most of protein folding *in vivo* takes place co-translationally, while the polypeptide chain is built up one amino acid at a time, simultaneously exploring the shifting folding landscape while interacting with multiple other molecules in a crowded environment, the task of incorporating all known cellular biochemistry and structural biology into the physical model of non-equilibrium protein folding as it occurs *in vivo* seems daunting. Nevertheless, this goal no longer appears to be out of reach. Advanced methods for quantitative measurement of various energy inputs, molecular motions, heat transfers and other relevant quantities should provide the values, or at least the bounds, of many parameters that determine protein folding as *in vivo*. Such work has already started although it is notable that many crucial parameters of the relevant processes, even some basic ones, such as the translation rate, rely on estimates obtained decades ago [170,171] and refined only very recently [172].

A complementary class of approaches involves building, both *in silico* and *in vitro*, simplified artificial protein folding machines that apply various forces to the folding polypeptide in an attempt to directly manipulate the peptide backbone into the desired conformations, imposing various kinds of physical constraints on the folding process, and thus, causing shifts and introducing kinetic barriers into the folding landscape. Work in this direction has already started as well. In the next section, we provide a brief overview of several advanced techniques and some recent observations, which suggest a more sophisticated understanding of the mechanisms of protein folding than what was provided by the canonical models of spontaneous protein folding *in vitro*.

### 2.8. Non-Equilibrium Protein Folding: New Approaches and Recent Results

In recent years, a variety of novel experimental techniques have been applied to study co-translational and chaperone-assisted protein folding. Particularly informative are methods that can manipulate a defined single molecule using a specific force probe, such as atomic force microscope, optical tweezers, or magnetic tweezers; these methods are

sometimes collectively referred to as single-molecule force spectroscopy methods (SMFS; reviewed in [173]). The SMFS methods have been recently applied to the study of co-translational folding of nascent protein chains, using “life-like” *in vitro* translation systems (reviewed in [149]). Although SMFS approaches have not yet reached the precision required to infer the thermodynamic parameters of protein folding (see Section 2.1 above), these approaches are well-suited to address questions about the effects of specific treatments and interactions on the folding process. Recent observations made using such methods include, for example, detection of co-translational folding intermediates, suggesting a defined folding pathway for a small domain that had been thought to fold in a two-step fashion *in vitro* [174], and observation of a direct accelerating and stabilizing effect of the ribosomal tunnel on the co-translational folding of another small domain [151]. Although often discussed within the conventional framework of thermodynamically-driven folding, these and similar results can be productively exploited to develop the non-equilibrium protein folding model. Furthermore, with regard to chaperone-assisted protein refolding, SMFS methods have revealed that the chaperones of the Hsp90 family use the energy of ATP hydrolysis to perform mechanical work, which is applied to compact unfolded chains against the counteracting denaturing forces [175], in an apparent contradiction to the traditionally envisaged, Anfinsen chamber-like mechanisms of chaperone action.

Another group of powerful methods are modern structural biology approaches, including cryo-electron microscopy, solid state nuclear magnetic resonance, SAXS and others, which reveal the structure of nascent polypeptide chains during protein synthesis. Many of these methods are focused on the kinetics and regulation of protein synthesis [176,177] and on the functions of nascent chain, such as sensing the state of regulatory metabolites in the environment and communicating the results to the peptidyltransferase center of the translating ribosome [178,179]. These studies also highly informative for the study of co-translational protein folding, and have already illuminated defined secondary and tertiary structures adopted by nascent peptides in the ribosome tunnel and exit vestibule [180–182]. In the forthcoming years, we expect to see more explicit investigation of the interactions of the nascent peptide with the peptidyl transferase center, ribosome exit tunnel, and other components of the protein folding machinery.

Computational modeling of protein folding also is taking a new direction towards a closer mimicking of the folding environment encountered by proteins *in vivo*. We recently reported the results of all-atom molecular dynamics simulations, in which the standard force field was augmented by the application of a mechanical force that rotated a single N-terminal amino acid of peptides, while simultaneously restricting the movements of a distal amino acid. Such directional rotation changed the peptide backbone behavior, facilitating rapid formation of native structures in several diverse alpha-helical peptides [153]. Apparently, steric clashes arising due to the forced directional rotation resulted in the behavior of the peptide backbone that no longer resembled an FJC. Further studies are needed to determine whether such an effect can be observed in single-molecule experiments *in vitro* as well. Other attempts to build simplified folding machines to model aspects of co-translational peptide folding *in vivo* include the molecular-dynamics studies of folding in a tubular chamber representing the ribosome exit tunnel, either with uncharged elastic walls or with charged walls [183–186]. Finally, sophisticated methods of visualization and analysis of the massive dynamic data on protein folding, unfolding, and refolding are also undergoing active development (see [187] for a recent review). Such methods should greatly aid our understanding of the complex mechanisms of protein folding *in vivo*.

### 3. Conclusions

The cornerstone assumption in the field of protein folding is that proteins spontaneously fold into their native conformations driven by negative  $\Delta G$ . Furthermore, it is generally assumed that the native conformation of a protein is the global minimum of Gibbs free energy. However, a survey of the available data on spontaneous protein folding and refolding, in particular, for chemically synthesized and over-expressed proteins, presents

little evidence in support of this thermodynamic hypothesis of folding. On the contrary, the majority of proteins appear not to be spontaneously foldable and are only marginally stable, at best. The totality of these observations along with thermodynamic considerations suggest that across the protein world, there is a wide variety of rugged, dynamic landscapes of folding free energy, resulting in a broad range of thermodynamic and kinetic stability, and refoldability of proteins. For different proteins,  $\Delta G$  of folding can be either negative or positive, conceivably, for the majority of the proteins. Even regardless of the specific value of  $\Delta G$ , folding of most proteins is likely to be an active, non-equilibrium, energy-dependent process. This conceptual shift in our understanding of protein folding appears to be best compatible with the extensive molecular data on the universal translation and proteostasis machineries that operate in all cells, and with the evolutionary history of these molecular machines that is traced to the earliest stages of life evolution. We believe that this change in perspective on protein folding can and should stimulate a dedicated program of theoretical, modeling, and experimental studies.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijms23010521/s1>.

**Author Contributions:** Conceptualization, I.S., A.R.M. and E.V.K.; methodology, I.S. and A.R.M.; investigation, I.S., A.R.M. and E.V.K.; resources, I.S.; data curation, I.S., A.R.M.; writing—original draft, I.S. and A.R.M.; writing—review and editing, I.S., A.R.M. and E.V.K.; visualization, I.S.; funding acquisition, I.S., A.R.M. and E.V.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** A.R.M.'s stay at Clare Hall College was partially supported by the U.S. National Science Foundation Long-Term Professional Development program; E.V.K. is supported by the Intramural Research Program of the National Institutes of Health of the USA (National Library of Medicine).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data supporting reported results have been submitted as Supplementary Files to this manuscript, except the ProTherm data that have been converted by Protabit LLC with permission from ProTherm and deposited at <https://github.com/protabit/protherm-conversion> (accessed on 6 December 2021).

**Acknowledgments:** We are grateful to Connie Wang (Protabit, LLC) for the advice on the converted ProTherm dataset, to Natalia Mushegian for help with data analysis, and to Sophia Belkin for the artwork.

**Conflicts of Interest:** A.R.M. is a program director at the National Science Foundation (NSF), an agency of the U.S. Government, but the statements and opinions expressed herein are made in the personal capacity and do not constitute the endorsement by NSF or the government of the United States. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Haber, E.; Anfinsen, C.B. Side-Chain Interactions Governing the Pairing of Half-Cystine Residues in Ribonuclease. *J. Biol. Chem.* **1962**, *237*, 1839–1844. [[CrossRef](#)]
2. Anfinsen, C.B. Principles That Govern the Folding of Protein Chains. *Science* **1973**, *181*, 223–230. [[CrossRef](#)] [[PubMed](#)]
3. Bryngelson, J.D.; Wolynes, P.G. Intermediates and Barrier Crossing in a Random Energy Model (with Applications to Protein Folding). *J. Phys. Chem.* **1989**, *93*, 6902–6915. [[CrossRef](#)]
4. Zwanzig, R.; Szabo, A.; Bagchi, B. Levinthal's Paradox. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 20–22. [[CrossRef](#)]
5. Leopold, P.E.; Montal, M.; Onuchic, J.N. Protein Folding Funnels: A Kinetic Approach to the Sequence-Structure Relationship. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8721–8725. [[CrossRef](#)]
6. Bryngelson, J.D.; Onuchic, J.N.; Socci, N.D.; Wolynes, P.G. Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. *Proteins* **1995**, *21*, 167–195. [[CrossRef](#)] [[PubMed](#)]
7. Wolynes, P.G. Energy Landscapes and Solved Protein-Folding Problems. *Philos. Trans. A Math. Phys. Eng. Sci.* **2005**, *363*, 453–464; discussion 464–467. [[CrossRef](#)] [[PubMed](#)]

8. Dill, K.A.; Ozkan, S.B.; Shell, M.S.; Weikl, T.R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37*, 289–316. [[CrossRef](#)] [[PubMed](#)]
9. Dill, K.A.; MacCallum, J.L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338*, 1042–1046. [[CrossRef](#)] [[PubMed](#)]
10. Schafer, N.P.; Kim, B.L.; Zheng, W.; Wolynes, P.G. Learning To Fold Proteins Using Energy Landscape Theory. *Isr. J. Chem.* **2014**, *54*, 1311–1337. [[CrossRef](#)] [[PubMed](#)]
11. Nassar, R.; Dignon, G.L.; Razban, R.M.; Dill, K.A. The Protein Folding Problem: The Role of Theory. *J. Mol. Biol.* **2021**, *433*, 167126. [[CrossRef](#)]
12. Rollins, G.C.; Dill, K.A. General Mechanism of Two-State Protein Folding Kinetics. *J. Am. Chem. Soc.* **2014**, *136*, 11420–11427. [[CrossRef](#)]
13. Finkelstein, A.V.; Badretdin, A.J.; Galzitskaya, O.V.; Ivankov, D.N.; Bogatyreva, N.S.; Garbuzynskiy, S.O. There and Back Again: Two Views on the Protein Folding Puzzle. *Phys. Life Rev.* **2017**, *21*, 56–71. [[CrossRef](#)]
14. Wetlaufer, D.B.; Ristow, S. Acquisition of Three-Dimensional Structure of Proteins. *Annu. Rev. Biochem.* **1973**, *42*, 135–158. [[CrossRef](#)]
15. Ben-Naim, A. *Myths and Verities in Protein Folding Theories*; World Scientific: Singapore, 2015; ISBN 978-981-4725-98-9.
16. Shakhnovich, E. Protein Folding Thermodynamics and Dynamics: Where Physics, Chemistry, and Biology Meet. *Chem. Rev.* **2006**, *106*, 1559–1588. [[CrossRef](#)] [[PubMed](#)]
17. Santoro, M.M.; Bolen, D.W. Unfolding Free Energy Changes Determined by the Linear Extrapolation Method. 1. Unfolding of Phenylmethanesulfonyl Alpha-Chymotrypsin Using Different Denaturants. *Biochemistry* **1988**, *27*, 8063–8068. [[CrossRef](#)] [[PubMed](#)]
18. Grimsley, G.R.; Huyghues-Despointes, B.M.P.; Pace, C.N.; Scholtz, J.M. Determining a Urea or Guanidinium Chloride Unfolding Curve. *CSH Protoc.* **2006**, *2006*, pdb-prot4242. [[CrossRef](#)]
19. Shaw, K.L.; Scholtz, J.M.; Pace, C.N.; Grimsley, G.R. Determining the Conformational Stability of a Protein Using Urea Denaturation Curves. *Methods Mol. Biol.* **2009**, *490*, 41–55. [[CrossRef](#)] [[PubMed](#)]
20. Privalov, P.L. Microcalorimetry of Proteins and Their Complexes. *Methods Mol. Biol.* **2009**, *490*, 1–39. [[CrossRef](#)]
21. Ibarra-Molero, B.; Naganathan, A.N.; Sanchez-Ruiz, J.M.; Muñoz, V. Modern Analysis of Protein Folding by Differential Scanning Calorimetry. *Methods Enzymol.* **2016**, *567*, 281–318. [[CrossRef](#)]
22. Naganathan, A.N.; Perez-Jimenez, R.; Muñoz, V.; Sanchez-Ruiz, J.M. Estimation of Protein Folding Free Energy Barriers from Calorimetric Data by Multi-Model Bayesian Analysis. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17064–17076. [[CrossRef](#)]
23. Makhatazde, G.I.; Privalov, P.L. Energetics of Protein Structure. *Adv. Protein Chem.* **1995**, *47*, 307–425. [[PubMed](#)]
24. Baldwin, R.L. Energetics of Protein Folding. *J. Mol. Biol.* **2007**, *371*, 283–301. [[CrossRef](#)] [[PubMed](#)]
25. Bedouelle, H. Principles and Equations for Measuring and Interpreting Protein Stability: From Monomer to Tetramer. *Biochimie* **2016**, *121*, 29–37. [[CrossRef](#)] [[PubMed](#)]
26. Basharov, M.A. Residual Ordered Structure in Denatured Proteins and the Problem of Protein Folding. *Indian J. Biochem. Biophys.* **2012**, *49*, 7–17.
27. Shortle, D.; Ackerman, M.S. Persistence of Native-like Topology in a Denatured Protein in 8 M Urea. *Science* **2001**, *293*, 487–489. [[CrossRef](#)]
28. Vendruscolo, M.; Paci, E.; Karplus, M.; Dobson, C.M. Structures and Relative Free Energies of Partially Folded States of Proteins. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 14817–14821. [[CrossRef](#)]
29. Sosnick, T.R.; Trewheella, J. Denatured States of Ribonuclease A Have Compact Dimensions and Residual Secondary Structure. *Biochemistry* **1992**, *31*, 8329–8335. [[CrossRef](#)]
30. Lim, W.K.; Rösgen, J.; Englander, S.W. Urea, but Not Guanidinium, Destabilizes Proteins by Forming Hydrogen Bonds to the Peptide Group. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 2595–2600. [[CrossRef](#)]
31. Lapidus, L.J. Protein Unfolding Mechanisms and Their Effects on Folding Experiments. *F1000Research* **2017**, *6*, 1723. [[CrossRef](#)]
32. Fitzkee, N.C.; Rose, G.D. Reassessing Random-Coil Statistics in Unfolded Proteins. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12497–12502. [[CrossRef](#)] [[PubMed](#)]
33. Nikam, R.; Kulandaisamy, A.; Harini, K.; Sharma, D.; Gromiha, M.M. ProThermDB: Thermodynamic Database for Proteins and Mutants Revisited after 15 Years. *Nucleic Acids. Res.* **2021**, *49*, D420–D424. [[CrossRef](#)] [[PubMed](#)]
34. ProTherm Conversion. 2017. Available online: <https://github.com/protabit/protherm-conversion> (accessed on 6 December 2021).
35. Braselmann, E.; Chaney, J.L.; Clark, P.L. Folding the Proteome. *Trends Biochem. Sci.* **2013**, *38*, 337–344. [[CrossRef](#)] [[PubMed](#)]
36. Zeldovich, K.B.; Chen, P.; Shakhnovich, E.I. Protein Stability Imposes Limits on Organism Complexity and Speed of Molecular Evolution. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16152–16157. [[CrossRef](#)]
37. Taverna, D.M.; Goldstein, R.A. Why Are Proteins Marginally Stable? *Proteins* **2002**, *46*, 105–109. [[CrossRef](#)] [[PubMed](#)]
38. Godoy-Ruiz, R.; Perez-Jimenez, R.; Ibarra-Molero, B.; Sanchez-Ruiz, J.M. Relation between Protein Stability, Evolution and Structure, as Probed by Carboxylic Acid Mutations. *J. Mol. Biol.* **2004**, *336*, 313–318. [[CrossRef](#)]
39. Williams, P.D.; Pollock, D.D.; Goldstein, R.A. Functionality and the Evolution of Marginal Stability in Proteins: Inferences from Lattice Simulations. *Evol. Bioinform. Online* **2007**, *2*, 91–101. [[CrossRef](#)]
40. Wilson, A.E.; Kosater, W.M.; Liberles, D.A. Evolutionary Processes and Biophysical Mechanisms: Revisiting Why Evolved Proteins Are Marginally Stable. *J. Mol. Evol.* **2020**, *88*, 415–417. [[CrossRef](#)]

41. Borgia, A.; Williams, P.M.; Clarke, J. Single-Molecule Studies of Protein Folding. *Annu. Rev. Biochem.* **2008**, *77*, 101–125. [[CrossRef](#)]
42. Bustamante, C.; Alexander, L.; Maciuba, K.; Kaiser, C.M. Single-Molecule Studies of Protein Folding with Optical Tweezers. *Annu. Rev. Biochem.* **2020**, *89*, 443–470. [[CrossRef](#)] [[PubMed](#)]
43. Tinoco, I., Jr.; Bustamante, C. The Effect of Force on Thermodynamics and Kinetics of Single Molecule Reactions. *Biophys. Chem.* **2002**, *101–102*, 513–533. [[CrossRef](#)]
44. Junier, I.; Mossa, A.; Manosas, M.; Ritort, F. Recovery of Free Energy Branches in Single Molecule Experiments. *Phys. Rev. Lett.* **2009**, *102*, 070602. [[CrossRef](#)]
45. Wang, J.; Ferguson, A.L. Nonlinear Reconstruction of Single-Molecule Free-Energy Surfaces from Univariate Time Series. *Phys. Rev. E* **2016**, *93*, 032412. [[CrossRef](#)] [[PubMed](#)]
46. Yang, G.; Cecconi, C.; Baase, W.A.; Vetter, I.R.; Breyer, W.A.; Haack, J.A.; Matthews, B.W.; Dahlquist, F.W.; Bustamante, C. Solid-State Synthesis and Mechanical Unfolding of Polymers of T4 Lysozyme. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 139–144. [[CrossRef](#)]
47. Dietz, H.; Rief, M. Exploring the Energy Landscape of GFP by Single-Molecule Mechanical Experiments. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16192–16197. [[CrossRef](#)] [[PubMed](#)]
48. Ainarapu, S.R.K.; Brujic, J.; Huang, H.H.; Wiita, A.P.; Lu, H.; Li, L.; Walther, K.A.; Carrion-Vazquez, M.; Li, H.; Fernandez, J.M. Contour Length and Refolding Rate of a Small Protein Controlled by Engineered Disulfide Bonds. *Biophys. J.* **2007**, *92*, 225–233. [[CrossRef](#)]
49. Hirschmann, R.; Nutt, R.F.; Veber, D.F.; Vitali, R.A.; Varga, S.L.; Jacob, T.A.; Holly, F.W.; Denkwalter, R.G. Studies on the Total Synthesis of an Enzyme. V. The Preparation of Enzymatically Active Material. *J. Am. Chem. Soc.* **1969**, *91*, 507–508. [[CrossRef](#)] [[PubMed](#)]
50. Gutte, B.; Merrifield, R.B. The Synthesis of Ribonuclease A. *J. Biol. Chem.* **1971**, *246*, 1922–1941. [[CrossRef](#)]
51. Tiessen, A.; Pérez-Rodríguez, P.; Delaye-Arredondo, L.J. Mathematical Modeling and Comparison of Protein Size Distribution in Different Plant, Animal, Fungal and Microbial Species Reveals a Negative Correlation between Protein Size and Protein Number, Thus Providing Insight into the Evolution of Proteomes. *BMC Res. Notes* **2012**, *5*, 85. [[CrossRef](#)]
52. Saleh, M.T.; Fillon, M.; Brennan, P.J.; Belisle, J.T. Identification of Putative Exported/Secreted Proteins in Prokaryotic Proteomes. *Gene* **2001**, *269*, 195–204. [[CrossRef](#)]
53. Saleh, M.; Song, C.; Nasserulla, S.; Leduc, L.G. Indicators from Archaeal Secretomes. *Microbiol. Res.* **2010**, *165*, 1–10. [[CrossRef](#)] [[PubMed](#)]
54. Uhlén, M.; Karlsson, M.J.; Hober, A.; Svensson, A.-S.; Scheffel, J.; Kotol, D.; Zhong, W.; Tebani, A.; Strandberg, L.; Edfors, F.; et al. The Human Secretome. *Sci. Signal.* **2019**, *12*, eaaz0274. [[CrossRef](#)] [[PubMed](#)]
55. Bošnjak, I.; Bojović, V.; Šegvić-Bubić, T.; Bielen, A. Occurrence of Protein Disulfide Bonds in Different Domains of Life: A Comparison of Proteins from the Protein Data Bank. *Protein Eng. Des. Sel.* **2014**, *27*, 65–72. [[CrossRef](#)]
56. Guise, A.D.; West, S.M.; Chaudhuri, J.B. Protein Folding in Vivo and Renaturation of Recombinant Proteins from Inclusion Bodies. *Mol. Biotechnol.* **1996**, *6*, 53–64. [[CrossRef](#)] [[PubMed](#)]
57. Panda, A.K. Bioprocessing of Therapeutic Proteins from the Inclusion Bodies of Escherichia Coli. *Adv. Biochem. Eng. Biotechnol.* **2003**, *85*, 43–93. [[PubMed](#)]
58. Cabrita, L.D.; Bottomley, S.P. Protein Expression and Refolding—A Practical Guide to Getting the Most out of Inclusion Bodies. *Biotechnol. Annu. Rev.* **2004**, *10*, 31–50.
59. Vera, A.; González-Montalbán, N.; Arís, A.; Villaverde, A. The Conformational Quality of Insoluble Recombinant Proteins Is Enhanced at Low Growth Temperatures. *Biotechnol. Bioeng.* **2007**, *96*, 1101–1106. [[CrossRef](#)]
60. Georgiou, G.; Valax, P.; Ostermeier, M.; Horowitz, P.M. Folding and Aggregation of TEM Beta-Lactamase: Analogies with the Formation of Inclusion Bodies in Escherichia Coli. *Protein Sci.* **1994**, *3*, 1953–1960. [[CrossRef](#)]
61. Przybycien, T.M.; Dunn, J.P.; Valax, P.; Georgiou, G. Secondary Structure Characterization of Beta-Lactamase Inclusion Bodies. *Protein Eng.* **1994**, *7*, 131–136. [[CrossRef](#)]
62. de Groot, N.S.; Sabate, R.; Ventura, S. Amyloids in Bacterial Inclusion Bodies. *Trends Biochem. Sci.* **2009**, *34*, 408–416. [[CrossRef](#)]
63. Ramón, A.; Señorale-Pose, M.; Marín, M. Inclusion Bodies: Not That Bad. *Front. Microbiol.* **2014**, *5*, 56. [[CrossRef](#)]
64. Bowden, G.A.; Paredes, A.M.; Georgiou, G. Structure and Morphology of Protein Inclusion Bodies in Escherichia Coli. *Biotechnology* **1991**, *9*, 725–730. [[CrossRef](#)]
65. Chaffotte, A.F.; Guillou, Y.; Goldberg, M.E. Inclusion Bodies of the Thermophilic Endoglucanase D from Clostridium Thermocellum Are Made of Native Enzyme That Resists 8 M Urea. *Eur. J. Biochem.* **1992**, *205*, 369–373. [[CrossRef](#)]
66. Vandebroek, K.; Martens, E.; D’Andrea, S.; Billiau, A. Refolding and Single-Step Purification of Porcine Interferon-Gamma from Escherichia Coli Inclusion Bodies. Conditions for Reconstitution of Dimeric IFN-Gamma. *Eur. J. Biochem.* **1993**, *215*, 481–486. [[CrossRef](#)] [[PubMed](#)]
67. Doglia, S.M.; Ami, D.; Natalello, A.; Gatti-Lafranconi, P.; Lotti, M. Fourier Transform Infrared Spectroscopy Analysis of the Conformational Quality of Recombinant Proteins within Inclusion Bodies. *Biotechnol. J.* **2008**, *3*, 193–201. [[CrossRef](#)]
68. Kudou, M.; Yumioka, R.; Ejima, D.; Arakawa, T.; Tsumoto, K. A Novel Protein Refolding System Using Lauroyl-l-Glutamate as a Solubilizing Detergent and Arginine as a Folding Assisting Agent. *Protein Expr. Purif.* **2011**, *75*, 46–54. [[CrossRef](#)] [[PubMed](#)]
69. Singh, S.M.; Sharma, A.; Upadhyay, A.K.; Singh, A.; Garg, L.C.; Panda, A.K. Solubilization of Inclusion Body Proteins Using n-Propanol and Its Refolding into Bioactive Form. *Protein Expr. Purif.* **2012**, *81*, 75–82. [[CrossRef](#)]

70. Singh, A.; Upadhyay, V.; Upadhyay, A.K.; Singh, S.M.; Panda, A.K. Protein Recovery from Inclusion Bodies of Escherichia Coli Using Mild Solubilization Process. *Microb. Cell Fact.* **2015**, *14*, 41. [CrossRef] [PubMed]
71. To, P.; Whitehead, B.; Tarbox, H.E.; Fried, S.D. Nonrefoldability Is Pervasive Across the E. Coli Proteome. *J. Am. Chem. Soc.* **2021**, *143*, 11435–11448. [CrossRef]
72. Hartl, F.U.; Bracher, A.; Hayer-Hartl, M. Molecular Chaperones in Protein Folding and Proteostasis. *Nature* **2011**, *475*, 324–332. [CrossRef]
73. Saibil, H. Chaperone Machines for Protein Folding, Unfolding and Disaggregation. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 630–642. [CrossRef] [PubMed]
74. Manning, M.C.; Patel, K.; Borchardt, R.T. Stability of Protein Pharmaceuticals. *Pharm. Res.* **1989**, *6*, 903–918. [CrossRef] [PubMed]
75. Manning, M.C.; Chou, D.K.; Murphy, B.M.; Payne, R.W.; Katayama, D.S. Stability of Protein Pharmaceuticals: An Update. *Pharm. Res.* **2010**, *27*, 544–575. [CrossRef] [PubMed]
76. Wang, W. Advanced Protein Formulations. *Protein Sci.* **2015**, *24*, 1031–1039. [CrossRef]
77. Magliery, T.J.; Lavinder, J.J.; Sullivan, B.J. Protein Stability by Number: High-Throughput and Statistical Approaches to One of Protein Science's Most Difficult Problems. *Curr. Opin. Chem. Biol.* **2011**, *15*, 443–451. [CrossRef]
78. Magliery, T.J. Protein Stability: Computation, Sequence Statistics, and New Experimental Methods. *Curr. Opin. Struct. Biol.* **2015**, *33*, 161–168. [CrossRef]
79. Klesmith, J.R.; Bacik, J.-P.; Wrenbeck, E.E.; Michalczyk, R.; Whitehead, T.A. Trade-Offs between Enzyme Fitness and Solubility Illuminated by Deep Mutational Scanning. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2265–2270. [CrossRef]
80. García-Fruitós, E.; Arís, A.; Villaverde, A. Localization of Functional Polypeptides in Bacterial Inclusion Bodies. *Appl. Environ. Microbiol.* **2007**, *73*, 289–294. [CrossRef] [PubMed]
81. Broom, A.; Jacobi, Z.; Trainor, K.; Meiering, E.M. Computational Tools Help Improve Protein Stability but with a Solubility Tradeoff. *J. Biol. Chem.* **2017**, *292*, 14349–14361. [CrossRef] [PubMed]
82. Broom, A.; Trainor, K.; Jacobi, Z.; Meiering, E.M. Computational Modeling of Protein Stability: Quantitative Analysis Reveals Solutions to Pervasive Problems. *Structure* **2020**, *28*, 717–726.e3. [CrossRef] [PubMed]
83. Siddiqui, K.S. Defying the Activity-Stability Trade-off in Enzymes: Taking Advantage of Entropy to Enhance Activity and Thermostability. *Crit. Rev. Biotechnol.* **2017**, *37*, 309–322. [CrossRef]
84. Manning, M.; Colón, W. Structural Basis of Protein Kinetic Stability: Resistance to Sodium Dodecyl Sulfate Suggests a Central Role for Rigidity and a Bias Toward  $\beta$ -Sheet Structure. *Biochemistry* **2004**, *43*, 11248–11254. [CrossRef]
85. Kazlauskas, R. Engineering More Stable Proteins. *Chem. Soc. Rev.* **2018**, *47*, 9026–9045. [CrossRef]
86. Ahern, T.J.; Klibanov, A.M. Analysis of Processes Causing Thermal Inactivation of Enzymes. *Methods Biochem. Anal.* **1988**, *33*, 91–127.
87. Tomazic, S.J.; Klibanov, A.M. Mechanisms of Irreversible Thermal Inactivation of Bacillus Alpha-Amylases. *J. Biol. Chem.* **1988**, *263*, 3086–3091. [CrossRef]
88. Nury, S.; Meunier, J.C. Molecular Mechanisms of the Irreversible Thermal Denaturation of Guinea-Pig Liver Transglutaminase. *Biochem. J.* **1990**, *266*, 487–490. [CrossRef]
89. Blaber, S.I.; Culajay, J.F.; Khurana, A.; Blaber, M. Reversible Thermal Denaturation of Human FGF-1 Induced by Low Concentrations of Guanidine Hydrochloride. *Biophys. J.* **1999**, *77*, 470–477. [CrossRef]
90. Jahromi, R.R.F.; Morris, P.; Martinez-Torres, R.J.; Dalby, P.A. Structural Stability of E. Coli Transketolase to Temperature and PH Denaturation. *J. Biotechnol.* **2011**, *155*, 209–216. [CrossRef]
91. Leurs, U.; Mistarz, U.H.; Rand, K.D. Getting to the Core of Protein Pharmaceuticals—Comprehensive Structure Analysis by Mass Spectrometry. *Eur. J. Pharm. Biopharm.* **2015**, *93*, 95–109. [CrossRef]
92. Gan, J.; Ben-Nissan, G.; Arkind, G.; Tarnavsky, M.; Trudeau, D.; Noda Garcia, L.; Tawfik, D.S.; Sharon, M. Native Mass Spectrometry of Recombinant Proteins from Crude Cell Lysates. *Anal. Chem.* **2017**, *89*, 4398–4404. [CrossRef]
93. Kaur, U.; Meng, H.; Lui, F.; Ma, R.; Ogburn, R.N.; Johnson, J.H.R.; Fitzgerald, M.C.; Jones, L.M. Proteome-Wide Structural Biology: An Emerging Field for the Structural Analysis of Proteins on the Proteomic Scale. *J. Proteome Res.* **2018**, *17*, 3614–3627. [CrossRef]
94. Atsavapranee, B.; Stark, C.D.; Sunden, F.; Thompson, S.; Fordyce, P.M. Fundamentals to Function: Quantitative and Scalable Approaches for Measuring Protein Stability. *Cell Syst.* **2021**, *12*, 547–560. [CrossRef] [PubMed]
95. Page, R.; Grzechnik, S.K.; Canaves, J.M.; Spraggon, G.; Kreuzsch, A.; Kuhn, P.; Stevens, R.C.; Lesley, S.A. Shotgun Crystallization Strategy for Structural Genomics: An Optimized Two-Tiered Crystallization Screen against the Thermotoga Maritima Proteome. *Acta Crystallogr. Biol. Crystallogr.* **2003**, *59*, 1028–1037. [CrossRef] [PubMed]
96. Northeast Structural Genomics Consortium Statistics. 2021. Available online: [https://www.nesg.org/statistics\\_00.html](https://www.nesg.org/statistics_00.html) (accessed on 6 December 2021).
97. New York Structural Genomics Research Consortium. 2021. Available online: [http://www.nysgsrc.org/psi3/progress\\_statistics.html](http://www.nysgsrc.org/psi3/progress_statistics.html) (accessed on 6 December 2021).
98. Gong, Y.; Kakihara, Y.; Krogan, N.; Greenblatt, J.; Emili, A.; Zhang, Z.; Houry, W.A. An Atlas of Chaperone-Protein Interactions in Saccharomyces Cerevisiae: Implications to Protein Folding Pathways in the Cell. *Mol. Syst. Biol.* **2009**, *5*, 275. [CrossRef] [PubMed]

99. Kerner, M.J.; Naylor, D.J.; Ishihama, Y.; Maier, T.; Chang, H.-C.; Stines, A.P.; Georgopoulos, C.; Frishman, D.; Hayer-Hartl, M.; Mann, M.; et al. Proteome-Wide Analysis of Chaperonin-Dependent Protein Folding in *Escherichia Coli*. *Cell* **2005**, *122*, 209–220. [[CrossRef](#)]
100. Lopez-Fanarraga, M.; Avila, J.; Guasch, A.; Coll, M.; Zabala, J.C. Review: Postchaperonin Tubulin Folding Cofactors and Their Role in Microtubule Dynamics. *J. Struct. Biol.* **2001**, *135*, 219–229. [[CrossRef](#)] [[PubMed](#)]
101. Povarova, O.I.; Uversky, V.N.; Kuznetsova, I.M.; Turoverov, K.K. Actinous Enigma or Enigmatic Actin: Folding, Structure, and Functions of the Most Abundant Eukaryotic Protein. *Intrinsically Disord. Proteins* **2014**, *2*, e34500. [[CrossRef](#)]
102. Finka, A.; Goloubinoff, P. Proteomic Data from Human Cell Cultures Refine Mechanisms of Chaperone-Mediated Protein Homeostasis. *Cell Stress Chaperones* **2013**, *18*, 591–605. [[CrossRef](#)]
103. Fujiwara, K.; Ishihama, Y.; Nakahigashi, K.; Soga, T.; Taguchi, H. A Systematic Survey of in Vivo Obligate Chaperonin-Dependent Substrates. *EMBO J.* **2010**, *29*, 1552–1564. [[CrossRef](#)]
104. Azia, A.; Unger, R.; Horovitz, A. What Distinguishes GroEL Substrates from Other *Escherichia Coli* Proteins? *FEBS J.* **2012**, *279*, 543–550. [[CrossRef](#)]
105. Hartl, F.U.; Hayer-Hartl, M. Molecular Chaperones in the Cytosol: From Nascent Chain to Folded Protein. *Science* **2002**, *295*, 1852–1858. [[CrossRef](#)]
106. Gupta, A.J.; Haldar, S.; Miličić, G.; Hartl, F.U.; Hayer-Hartl, M. Active Cage Mechanism of Chaperonin-Assisted Protein Folding Demonstrated at Single-Molecule Level. *J. Mol. Biol.* **2014**, *426*, 2739–2754. [[CrossRef](#)]
107. Singhal, K.; Vreede, J.; Mashaghi, A.; Tans, S.J.; Bolhuis, P.G. The Trigger Factor Chaperone Encapsulates and Stabilizes Partial Folds of Substrate Proteins. *PLoS Comput. Biol.* **2015**, *11*, e1004444. [[CrossRef](#)]
108. Grantcharova, V.; Alm, E.J.; Baker, D.; Horwich, A.L. Mechanisms of Protein Folding. *Curr. Opin. Struct. Biol.* **2001**, *11*, 70–82. [[CrossRef](#)]
109. Balchin, D.; Hayer-Hartl, M.; Hartl, F.U. In Vivo Aspects of Protein Folding and Quality Control. *Science* **2016**, *353*, aac4354. [[CrossRef](#)] [[PubMed](#)]
110. Balchin, D.; Miličić, G.; Strauss, M.; Hayer-Hartl, M.; Hartl, F.U. Pathway of Actin Folding Directed by the Eukaryotic Chaperonin TRiC. *Cell* **2018**, *174*, 1507–1521. [[CrossRef](#)]
111. Shtilerman, M.; Lorimer, G.H.; Englander, S.W. Chaperonin Function: Folding by Forced Unfolding. *Science* **1999**, *284*, 822–825. [[CrossRef](#)]
112. Sousa, R. Structural Mechanisms of Chaperone Mediated Protein Disaggregation. *Front. Mol. Biosci.* **2014**, *1*, 12. [[CrossRef](#)] [[PubMed](#)]
113. Nillegoda, N.B.; Bukau, B. Metazoan Hsp70-Based Protein Disaggregases: Emergence and Mechanisms. *Front. Mol. Biosci.* **2015**, *2*, 57. [[CrossRef](#)] [[PubMed](#)]
114. Deville, C.; Carroni, M.; Franke, K.B.; Topf, M.; Bukau, B.; Mogk, A.; Saibil, H.R. Structural Pathway of Regulated Substrate Transfer and Threading through an Hsp100 Disaggregase. *Sci. Adv.* **2017**, *3*, e1701726. [[CrossRef](#)] [[PubMed](#)]
115. Bukau, B.; Horwich, A.L. The Hsp70 and Hsp60 Chaperone Machines. *Cell* **1998**, *92*, 351–366. [[CrossRef](#)]
116. Zhang, X.; Kelly, J.W. Chaperonins Resculpt Folding Free Energy Landscapes to Avoid Kinetic Traps and Accelerate Protein Folding. *J. Mol. Biol.* **2014**, *426*, 2736–2738. [[CrossRef](#)]
117. Sekhar, A.; Rosenzweig, R.; Bouvignies, G.; Kay, L.E. Hsp70 Biases the Folding Pathways of Client Proteins. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E2794–E2801. [[CrossRef](#)]
118. Bagdany, M.; Veit, G.; Fukuda, R.; Avramescu, R.G.; Okiyoneda, T.; Baaklini, I.; Singh, J.; Sovak, G.; Xu, H.; Apaja, P.M.; et al. Chaperones Rescue the Energetic Landscape of Mutant CFTR at Single Molecule and in Cell. *Nat. Commun.* **2017**, *8*, 444. [[CrossRef](#)] [[PubMed](#)]
119. Çetinbaş, M.; Shakhnovich, E.I. Is Catalytic Activity of Chaperones a Selectable Trait for the Emergence of Heat Shock Response? *Biophys. J.* **2015**, *108*, 438–448. [[CrossRef](#)] [[PubMed](#)]
120. Suss, O.; Reichmann, D. Protein Plasticity Underlines Activation and Function of ATP-Independent Chaperones. *Front. Mol. Biosci.* **2015**, *2*, 43. [[CrossRef](#)]
121. Hebert, D.N.; Chandrasekhar, K.D.; Gierasch, L.M. You Got to Know When to Hold (or Unfold) ‘Em. *Mol. Cell* **2012**, *48*, 3–4. [[CrossRef](#)]
122. Guisbert, E.; Yura, T.; Rhodius, V.A.; Gross, C.A. Convergence of Molecular, Modeling, and Systems Approaches for an Understanding of the *Escherichia Coli* Heat Shock Response. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 545–554. [[CrossRef](#)]
123. Fernandez-Funez, P.; Sanchez-Garcia, J.; de Mena, L.; Zhang, Y.; Levites, Y.; Khare, S.; Golde, T.E.; Rincon-Limas, D.E. Holdase Activity of Secreted Hsp70 Masks Amyloid-B42 Neurotoxicity in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E5212–E5221. [[CrossRef](#)]
124. Sorokina, I.; Mushegian, A. The Role of the Backbone Torsion in Protein Folding. *Biol. Direct* **2016**, *11*, 64. [[CrossRef](#)]
125. Sorokina, I.; Mushegian, A. Rotational Restriction of Nascent Peptides as an Essential Element of Co-Translational Protein Folding: Possible Molecular Players and Structural Consequences. *Biol. Direct* **2017**, *12*, 14. [[CrossRef](#)]
126. Sorokina, I.; Mushegian, A. Modeling Protein Folding in Vivo. *Biol. Direct* **2018**, *13*, 13. [[CrossRef](#)]
127. Netzer, W.J.; Hartl, F.U. Recombination of Protein Domains Facilitated by Co-Translational Folding in Eukaryotes. *Nature* **1997**, *388*, 343–349. [[CrossRef](#)]
128. Basharov, M.A. Cotranslational Folding of Proteins. *Biochemistry* **2000**, *65*, 1380–1384.

129. Lorimer, G.H. A Personal Account of Chaperonin History. *Plant Physiol.* **2001**, *125*, 38–41. [[CrossRef](#)]
130. Bashan, A.; Yonath, A. Ribosome Crystallography: Catalysis and Evolution of Peptide-Bond Formation, Nascent Chain Elongation and Its Co-Translational Folding. *Biochem. Soc. Trans.* **2005**, *33*, 488–492. [[CrossRef](#)]
131. Lim, V.I.; Curran, J.F.; Garber, M.B. Ribosomal Elongation Cycle: Energetic, Kinetic and Stereochemical Aspects. *J. Mol. Biol.* **2005**, *351*, 470–480. [[CrossRef](#)] [[PubMed](#)]
132. Ziv, G.; Haran, G.; Thirumalai, D. Ribosome Exit Tunnel Can Entropically Stabilize Alpha-Helices. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18956–18961. [[CrossRef](#)] [[PubMed](#)]
133. Krishna, M.M.G.; Englander, S.W. A Unified Mechanism for Protein Folding: Predetermined Pathways with Optional Errors. *Protein Sci.* **2007**, *16*, 449–464. [[CrossRef](#)] [[PubMed](#)]
134. Steitz, T.A. A Structural Understanding of the Dynamic Ribosome Machine. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 242–253. [[CrossRef](#)]
135. 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. [[CrossRef](#)]
136. Kaiser, C.M.; Goldman, D.H.; Chodera, J.D.; Tinoco, I.; Bustamante, C. The Ribosome Modulates Nascent Protein Folding. *Science* **2011**, *334*, 1723–1727. [[CrossRef](#)]
137. Wilson, D.N.; Beckmann, R. The Ribosomal Tunnel as a Functional Environment for Nascent Polypeptide Folding and Translational Stalling. *Curr. Opin. Struct. Biol.* **2011**, *21*, 274–282. [[CrossRef](#)]
138. Choi, S.I.; Kwon, S.; Son, A.; Jeong, H.; Kim, K.-H.; Seong, B.L. Protein Folding in Vivo Revisited. *Curr. Protein Pept. Sci.* **2013**, *14*, 721–733. [[CrossRef](#)]
139. Krobath, H.; Shakhnovich, E.I.; Faísca, P.F.N. Structural and Energetic Determinants of Co-Translational Folding. *J. Chem. Phys.* **2013**, *138*, 215101. [[CrossRef](#)]
140. Gloge, F.; Becker, A.H.; Kramer, G.; Bukau, B. Co-Translational Mechanisms of Protein Maturation. *Curr. Opin. Struct. Biol.* **2014**, *24*, 24–33. [[CrossRef](#)] [[PubMed](#)]
141. Javed, A.; Christodoulou, J.; Cabrita, L.D.; Orlova, E.V. The Ribosome and Its Role in Protein Folding: Looking through a Magnifying Glass. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 509–521. [[CrossRef](#)]
142. Thommen, M.; Holtkamp, W.; Rodnina, M.V. Co-Translational Protein Folding: Progress and Methods. *Curr. Opin. Struct. Biol.* **2017**, *42*, 83–89. [[CrossRef](#)] [[PubMed](#)]
143. Sharma, A.K.; O'Brien, E.P. Non-Equilibrium Coupling of Protein Structure and Function to Translation-Elongation Kinetics. *Curr. Opin. Struct. Biol.* **2018**, *49*, 94–103. [[CrossRef](#)] [[PubMed](#)]
144. Wruck, F.; Avellaneda, M.J.; Koers, E.J.; Minde, D.P.; Mayer, M.P.; Kramer, G.; Mashaghi, A.; Tans, S.J. Protein Folding Mediated by Trigger Factor and Hsp70: New Insights from Single-Molecule Approaches. *J. Mol. Biol.* **2018**, *430*, 438–449. [[CrossRef](#)]
145. Alexander, L.M.; Goldman, D.H.; Wee, L.M.; Bustamante, C. Non-Equilibrium Dynamics of a Nascent Polypeptide during Translation Suppress Its Misfolding. *Nat. Commun.* **2019**, *10*, 2709. [[CrossRef](#)] [[PubMed](#)]
146. Waudby, C.A.; Dobson, C.M.; Christodoulou, J. Nature and Regulation of Protein Folding on the Ribosome. *Trends Biochem. Sci.* **2019**, *6*, 8. [[CrossRef](#)] [[PubMed](#)]
147. Cassaignau, A.M.E.; Cabrita, L.D.; Christodoulou, J. How Does the Ribosome Fold the Proteome? *Annu. Rev. Biochem.* **2020**, *89*, 389–415. [[CrossRef](#)]
148. Cassaignau, A.M.E.; Włodarski, T.; Chan, S.H.S.; Woodburn, L.F.; Bukvin, I.V.; Streit, J.O.; Cabrita, L.D.; Waudby, C.A.; Christodoulou, J. Interactions between Nascent Proteins and the Ribosome Surface Inhibit Co-Translational Folding. *Nat. Chem.* **2021**, *21*, 796. [[CrossRef](#)] [[PubMed](#)]
149. Maciuba, K.; Rajasekaran, N.; Chen, X.; Kaiser, C.M. Co-Translational Folding of Nascent Polypeptides: Multi-Layered Mechanisms for the Efficient Biogenesis of Functional Proteins. *Bioessays* **2021**, *43*, e2100042. [[CrossRef](#)]
150. Plessa, E.; Chu, L.P.; Chan, S.H.S.; Thomas, O.L.; Cassaignau, A.M.E.; Waudby, C.A.; Christodoulou, J.; Cabrita, L.D. Nascent Chains Can Form Co-Translational Folding Intermediates That Promote Post-Translational Folding Outcomes in a Disease-Causing Protein. *Nat. Commun.* **2021**, *12*, 6447. [[CrossRef](#)] [[PubMed](#)]
151. Wruck, F.; Tian, P.; Kudva, R.; Best, R.B.; von Heijne, G.; Tans, S.J.; Katranidis, A. The Ribosome Modulates Folding inside the Ribosomal Exit Tunnel. *Commun. Biol.* **2021**, *4*, 523. [[CrossRef](#)] [[PubMed](#)]
152. Fleming, P.J.; Rose, G.D. Conformational Properties of Unfolded Proteins. In *Protein Science Encyclopedia*; Fersht, A.R., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2008; p. sf06. ISBN 978-3-527-61075-4.
153. Sahakyan, H.; Nazaryan, K.; Mushegian, A.; Sorokina, I. Energy-Dependent Protein Folding: Modeling How a Protein Folding Machine May Work. *F1000Res* **2021**, *10*, 3. [[CrossRef](#)]
154. Spirin, A.S. *Ribosomes*; Kluwer Academic/Plenum Publishers: New York, NY, USA, 1999.
155. Doudna, J.A.; Batey, R.T. Structural Insights into the Signal Recognition Particle. *Annu. Rev. Biochem.* **2004**, *73*, 539–557. [[CrossRef](#)]
156. Lavery, L.A.; Partridge, J.R.; Ramelot, T.A.; Elnatan, D.; Kennedy, M.A.; Agard, D.A. Structural Asymmetry in the Closed State of Mitochondrial Hsp90 (TRAP1) Supports a Two-Step ATP Hydrolysis Mechanism. *Mol. Cell* **2014**, *53*, 330–343. [[CrossRef](#)]
157. Clerico, E.M.; Meng, W.; Pozhidaeva, A.; Bhasne, K.; Petridis, C.; Gierasch, L.M. Hsp70 Molecular Chaperones: Multifunctional Allosteric Holding and Unfolding Machines. *Biochem. J.* **2019**, *476*, 1653–1677. [[CrossRef](#)] [[PubMed](#)]
158. De Los Rios, P.; Barducci, A. Hsp70 Chaperones Are Non-Equilibrium Machines That Achieve Ultra-Affinity by Energy Consumption. *Elife* **2014**, *3*, e02218. [[CrossRef](#)] [[PubMed](#)]

159. Barducci, A.; De Los Rios, P. Non-Equilibrium Conformational Dynamics in the Function of Molecular Chaperones. *Curr. Opin. Struct. Biol.* **2015**, *30*, 161–169. [[CrossRef](#)] [[PubMed](#)]
160. Goloubinoff, P.; Sassi, A.; Fauvet, B.; Barducci, A.; De Los Rios, P. Molecular Chaperones Inject Energy from ATP Hydrolysis into the Nonequilibrium Stabilisation of Native Proteins. *Nat. Chem. Biol.* **2018**, *14*, 388–395. [[CrossRef](#)]
161. Assenza, S.; Sassi, A.S.; Kellner, R.; Schuler, B.; De Los Rios, P.; Barducci, A. Efficient Conversion of Chemical Energy into Mechanical Work by Hsp70 Chaperones. *Elife* **2019**, *8*, 8491. [[CrossRef](#)]
162. Chakrabarti, S.; Hyeon, C.; Ye, X.; Lorimer, G.H.; Thirumalai, D. Molecular Chaperones Maximize the Native State Yield on Biological Times by Driving Substrates out of Equilibrium. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E10919–E10927. [[CrossRef](#)]
163. Xu, H. ATP-Driven Nonequilibrium Activation of Kinase Clients by the Molecular Chaperone Hsp90. *Biophys. J.* **2020**, *119*, 1538–1549. [[CrossRef](#)]
164. Mirkin, B.G.; Fenner, T.I.; Galperin, M.Y.; Koonin, E.V. Algorithms for Computing Parsimonious Evolutionary Scenarios for Genome Evolution, the Last Universal Common Ancestor and Dominance of Horizontal Gene Transfer in the Evolution of Prokaryotes. *BMC Evol. Biol.* **2003**, *3*, 2. [[CrossRef](#)]
165. Mushegian, A. Gene Content of LUCA, the Last Universal Common Ancestor. *Front. Biosci.* **2008**, *13*, 4657–4666. [[CrossRef](#)]
166. Kannan, L.; Li, H.; Rubinstein, B.; Mushegian, A. Models of Gene Gain and Gene Loss for Probabilistic Reconstruction of Gene Content in the Last Universal Common Ancestor of Life. *Biol. Direct* **2013**, *8*, 32. [[CrossRef](#)]
167. Levinthal, C. Are There Pathways for Protein Folding? *J. Chim. Phys. Phys. Chim. Biol.* **1968**, *65*, 44–45. [[CrossRef](#)]
168. Levinthal, C. How to Fold Graciously. In *Mössbauer Spectroscopy in Biological Systems Proceedings*. University of Illinois Bulletin; Forgotten Books: London, UK, 1969; Volume 67, pp. 22–24.
169. Grosberg, A. A Few Disconnected Notes Related to Levinthal Paradox. *J. Biomol. Struct. Dyn.* **2002**, *20*, 317–321. [[CrossRef](#)]
170. Talkad, V.; Schneider, E.; Kennell, D. Evidence for Variable Rates of Ribosome Movement in Escherichia Coli. *J. Mol. Biol.* **1976**, *104*, 299–303. [[CrossRef](#)]
171. Olofsson, S.-O.; Boström, K.; Carlsson, P.; Borén, J.; Wettsten, M.; Bjursell, G.; Wiklund, O.; Bondjers, G. Structure and Biosynthesis of Apolipoprotein B. *Am. Heart J.* **1987**, *113*, 446–452. [[CrossRef](#)]
172. Li, G.-W.; Burkhardt, D.; Gross, C.; Weissman, J.S. Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. *Cell* **2014**, *157*, 624–635. [[CrossRef](#)]
173. Petrosyan, R.; Narayan, A.; Woodside, M.T. Single-Molecule Force Spectroscopy of Protein Folding. *J. Mol. Biol.* **2021**, *433*, 167207. [[CrossRef](#)]
174. Liutkute, M.; Maiti, M.; Samatova, E.; Enderlein, J.; Rodnina, M.V. Gradual Compaction of the Nascent Peptide during Cotranslational Folding on the Ribosome. *Elife* **2020**, *9*, e60895. [[CrossRef](#)] [[PubMed](#)]
175. Mashaghi, A.; Moayed, F.; Koers, E.J.; Kramer, G.; Mayer, M.P.; Tans, S.J. Direct Observation of Hsp90-Induced Compaction in a Protein Chain. Available online: <https://www.biorxiv.org/content/10.1101/2021.08.08.455546v1#page> (accessed on 6 December 2021).
176. Julián, P.; Milon, P.; Agirrezabala, X.; Lasso, G.; Gil, D.; Rodnina, M.V.; Valle, M. The Cryo-EM Structure of a Complete 30S Translation Initiation Complex from *Escherichia coli*. *PLoS Biol.* **2011**, *9*, 1001095. [[CrossRef](#)] [[PubMed](#)]
177. Bögeholz, L.A.K.; Mercier, E.; Wintermeyer, W.; Rodnina, M.V. Kinetic Control of Nascent Protein Biogenesis by Peptide Deformylase. *Sci. Rep.* **2021**, *11*, 24457. [[CrossRef](#)] [[PubMed](#)]
178. Herrero Del Valle, A.; Seip, B.; Cervera-Marzal, I.; Sacheau, G.; Seefeldt, A.C.; Innis, C.A. Ornithine Capture by a Translating Ribosome Controls Bacterial Polyamine Synthesis. *Nat. Microbiol.* **2020**, *5*, 554–561. [[CrossRef](#)]
179. van der Stel, A.X.; Gordon, E.R.; Sengupta, A.; Martínez, A.K.; Klepacki, D.; Perry, T.N.; Herrero Del Valle, A.; Vázquez-Laslop, N.; Sachs, M.S.; Cruz-Vera, L.R.; et al. Structural Basis for the Tryptophan Sensitivity of TnaC-mediated Ribosome Stalling. *Nat. Commun.* **2021**, *12*, 5340. [[CrossRef](#)]
180. Nilsson, O.B.; Hedman, R.; Marino, J.; Wickles, S.; Bischoff, L.; Johansson, M.; Müller-Lucks, A.; Trovato, F.; Puglisi, J.D.; O'Brien, E.P.; et al. Cotranslational Protein Folding inside the Ribosome Exit Tunnel. *Cell Rep.* **2015**, *12*, 1533–1540. [[CrossRef](#)]
181. Bañó-Polo, M.; Baeza-Delgado, C.; Tamborero, S.; Hazel, A.; Grau, B.; Nilsson, I.; Whitley, P.; Gumbart, J.C.; von Heijne, G.; Mingarro, I. Transmembrane but not Soluble Helices Fold inside the Ribosome Tunnel. *Nat. Commun.* **2018**, *9*, 5246. [[CrossRef](#)] [[PubMed](#)]
182. Schulte, L.; Mao, J.; Reitz, J.; Sreeramulu, S.; Kudlinzki, D.; Hodirnau, V.V.; Meier-Credo, J.; Saxena, K.; Buhr, F.; Langer, J.D.; et al. Cysteine Oxidation and Disulfide Formation in the Ribosomal Exit Tunnel. *Nat. Commun.* **2020**, *11*, 5569. [[CrossRef](#)] [[PubMed](#)]
183. Bui, P.T.; Hoang, T.X. Protein Escape at the Ribosomal Exit Tunnel: Effects of Native Interactions, Tunnel Length, and Macromolecular Crowding. *J. Chem. Phys.* **2018**, *149*, 045102. [[CrossRef](#)] [[PubMed](#)]
184. Bui, P.T.; Hoang, T.X. Protein Escape at the Ribosomal Exit Tunnel: Effect of the Tunnel Shape. *J. Chem. Phys.* **2020**, *153*, 045105. [[CrossRef](#)]
185. Bui, P.T.; Hoang, T.X. Hydrophobic and Electrostatic Interactions Modulate Protein Escape at the Ribosomal Exit Tunnel. *Biophys. J.* **2021**, *9*, 27. [[CrossRef](#)]
186. Joiret, M.; Rapino, F.; Close, P.; Geris, L. Ribosome Exit Tunnel Electrostatics. Available online: <https://www.biorxiv.org/content/10.1101/2020.10.20.346684v1.full> (accessed on 6 December 2021).
187. Ferina, J.; Daggett, V. Visualizing Protein Folding and Unfolding. *J. Mol. Biol.* **2019**, *431*, 1540–1564. [[CrossRef](#)]