

## Protein folding and aggregation in bacteria

Raimon Sabate · Natalia S. de Groot ·  
Salvador Ventura

Received: 18 January 2010/Revised: 19 February 2010/Accepted: 5 March 2010/Published online: 1 April 2010  
© Springer Basel AG 2010

**Abstract** Proteins might experience many conformational changes and interactions during their lifetimes, from their synthesis at ribosomes to their controlled degradation. Because, in most cases, only folded proteins are functional, protein folding in bacteria is tightly controlled genetically, transcriptionally, and at the protein sequence level. In addition, important cellular machinery assists the folding of polypeptides to avoid misfolding and ensure the attainment of functional structures. When these redundant protective strategies are overcome, misfolded polypeptides are recruited into insoluble inclusion bodies. The protein embedded in these intracellular deposits might display different conformations including functional and  $\beta$ -sheet-rich structures. The latter assemblies are similar to the amyloid fibrils characteristic of several human neurodegenerative diseases. Interestingly, bacteria exploit the same structural principles for functional properties such as adhesion or cytotoxicity. Overall, this review illustrates how prokaryotic organisms might provide the bedrock on which to understand the complexity of protein folding and aggregation in the cell.

**Keywords** Protein folding · Protein aggregation · Inclusion bodies · Amyloid · Bacteria · Protein synthesis · Chaperones

### Introduction

Proteins are comprised of repetitive amino acid units that differ in the chemical nature of their side chains. After synthesis at the ribosome in the highly crowded cellular environment, the polypeptide side chains are exposed to manifold non-specific interactions; nevertheless, the proteins usually manage to selectively form the restricted set of intramolecular contacts that funnel the folding of the protein towards the unique, thermodynamically stable and functional native structure [1]. At times, however, partially folded intermediates accumulate during the folding reaction or already folded proteins lose some of their functional intramolecular contacts. Both of these cases result in a population of non-native protein conformations [2]. Under these particular conditions, folding reactions are efficiently competed by anomalous intermolecular self-assembly of polypeptide chains, ultimately leading to the formation of macromolecular aggregated structures. Interestingly, the interactions leading to the formation of such aberrant protein conformations also involve specific recognitions between amino acid side chains [2].

Protein misfolding and aggregation are becoming central issues in biology, medicine, and biotechnology. On the one hand, the formation of insoluble protein deposits in human tissues triggers the onset of dozens of devastating human diseases like Alzheimer's or Parkinson's [2]. These protein aggregates are formed by fibrillar structures known as amyloids and have a common cross- $\beta$  supramolecular conformation [2]. Since Rudolf Virchow was looking for cellulose-like substances in the human body using iodine staining [3, 4] the term “amyloid” has been increasingly used. Nevertheless, the “amyloid” designation refers to different concepts, depending on the

---

R. Sabate and N. S. de Groot contributed equally to this work.

---

R. Sabate · N. S. de Groot · S. Ventura (✉)  
Departament de Bioquímica i Biologia Molecular,  
Institut de Biotecnologia i de Biomedicina, Universitat  
Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain  
e-mail: salvador.ventura@uab.es

research area. By definition, amyloid deposits are *in vivo* formed intracellular or extracellular deposits that contain the main protein fibrils and usually other components like glycosaminoglycans, apolipoprotein E, and serum amyloid P-component [5]. These deposits could be distinguished from non-amyloid aggregates by histological staining assays, principally the binding of Congo Red dye that results in green birefringence. They also present a fibrillar appearance under electron microscopy and a characteristic X-ray diffraction pattern with reflections at 4.7 and 10 Å [5, 6]. This definition excludes the fibrils formed *in vitro* by synthetic proteins and peptides despite that they possess many of the above-described hallmarks [5, 7]. Nevertheless, the study of synthetic fibrils formed by proteins linked or even unrelated to human disease has provided the benchmark to understand protein aggregation *in vivo*. They have been suggested to be termed as amyloids (amyloid homologs) [8].

Protein misfolding and aggregation into insoluble inclusion bodies (IBs) are commonly observed during recombinant expression in prokaryotic systems, thus limiting the application of microbial cell factories for the industrial production of protein-based drugs [9]. Studies on bacteria have provided the basis on which to understand the molecular machinery and the mechanism of protein synthesis [10] and protein folding [11] in higher organisms. However, protein aggregation in prokaryotes has long been seen as a non-specific process unrelated to amyloid formation in eukaryotes, and little interest has been expressed regarding this process in prokaryotes, except perhaps regarding a potential role in protein production. However, an increasing body of evidence suggests that the formation and structure of bacterial aggregates share a number of common features with the highly ordered and pathogenic amyloid fibrils linked to human diseases [12, 13]. This similarity is especially important because it appears that although proteins have co-evolved with their cellular environments to fold accurately and resist aggregation, they have little margin of safety to respond to genetic and environmental factors that decrease their folding ability or increase their aggregation propensity [14]. Because the genetic, biochemical, and environmental manipulation of bacteria is straightforward, bacteria can provide a valuable physiologically relevant benchmark to understand the generic pathways that a polypeptide follows inside the cell after its synthesis. The mechanisms employed by bacterial cells to deal with protein misfolding and aggregation, the structural and functional properties of prokaryotic aggregates, and the use of the information gained in bacterial systems to understand and tackle disease-linked protein deposition are the focus of this review.

## Protein folding in bacterial cells

Protein folding control during transcription and translation

### *Protein synthesis at the ribosome*

The ribosome is the factory for cellular proteins. The bacterial 70S ribosome is composed of three RNA molecules (that represent about two-thirds of the ribosome mass) and more than 50 usually highly conserved proteins [15, 16]. The ribosome comprises two subunits: a large 50S subunit and a small 30S subunit. The 30S subunit consists of a 16S rRNA (ribosomal RNA) plus 20 proteins. The 50S subunit contains a 23S and a 5S rRNA and over 30 proteins [17]. The rRNA molecules fold into defined spatial conformations, e.g., the 23S rRNA contains six secondary structural domains (I–VI) formed by over 130 RNA helices [17]. In recent years, we have witnessed the deciphering of the mechanism by which this highly complex molecular machine performs protein synthesis and assists protein-folding reactions [15].

Translation initiation is the rate-limiting step in protein synthesis. This process involves a series of sequential and finely regulated events. In bacteria, the correct mRNA translation start site is selected upon recognition of the initiation codon at the peptidyl site of the 30S ribosomal subunit by the fMet-tRNA anticodon. This precise matching is assisted by several initiation factors, resulting in the formation of the 30S initiation complex. Subsequently, this complex joins the 50S ribosomal subunit and releases the initiation factors, rendering the 70S initiation complex [18, 19]. The high-resolution crystal X-ray structure of the two subunits of the large ribosome has revealed how it executes the crucial peptidyl transferase reaction, which consists of the successive joining of amino acids through amide linkages. This reaction requires specific selection of the amino acids to be added to the growing polypeptide chain by reading successive mRNA codons [15, 20]. This remarkable task involves wrapping about 30 mRNA nucleotides in the groove that encircles the neck of the 30S subunit where about eight of these nucleotides are exposed almost exclusively to 16S rRNA and centered on the junction between the amino-acyl-tRNA and peptidyl-tRNA binding sites for their translation [21]. At this interface region, the peptidyl transferase center of the large subunit initiates the formation of the new peptide bonds [17]. Importantly, rRNA molecules are not inert framework organizing catalytic proteins, but are the structural units that help to organize key ribozyme (catalytic RNA) elements in forming a specific and active structure [20, 22, 23].

Newly synthesized proteins emerge from the ribosome in an environment crowded with a highly diverse collection of molecules and polypeptides. In these conditions, nascent proteins are especially vulnerable to misfolding and aggregation, which are two side-reactions that impede the acquisition of a correct native conformation in approximately 30% of new proteins [24]. Protein aggregation has harmful consequences for the cell, such as reduction of amino acid recycling, recruitment and blockage of molecular chaperones and proteases, non-specific kidnapping of unrelated functional proteins, formation of toxic polypeptide species or simply the loss of function of the misfolded protein [25–28]. Accordingly, nature has evolved a battery of strategies aimed at preventing the anomalous self-assembly of newly synthesized polypeptide chains. As we will see in the following sections, the cell controls protein misfolding and self-assembly before the protein can even fold by regulating the level of protein expression or the rate of protein translation or providing the ribosome with folding assistance activity.

#### *Correlation between protein solubility and gene expression levels*

The number of mRNAs in the bacterial cytosol encoding for a given protein might vary from 1 to  $10^6$  copies. Despite variable turnover at the protein and mRNA levels, gene expression levels tend to be proportional, on average, to protein abundance. This fact implies that the cell needs to cope with the misfolding of proteins that are present in a large dynamic range of protein concentrations. Protein aggregation usually follows a nucleation–polymerization mechanism, in which the build up of the initial self-assembled nuclei constitutes the rate-limiting step of the reaction. Nucleation is a second-order reaction and, therefore, the rate of protein aggregation is strongly dependent on the initial protein concentration. This correlation implies that, in principle, avoiding the self-assembly of highly translated sequences would be more challenging for the cell than controlling the solubility of low copy proteins. In support of this view, it has been recently shown that both in bacteria and in higher organisms, mRNA expression levels are highly correlated with the solubility of the encoded proteins [14, 29]. Highly translated proteins tend to be more soluble than proteins with low expression rates. This decreased aggregation propensity would work to prevent their wrong misassembly even if they become concentrated at specific subcytosolic locations, as might occur immediately after protein synthesis. In addition, because of the abundance of proteins with high rates of translation, their high solubility would make important contributions to an overall decrease in cell aggregation propensity [195]. The correlation between protein

expression levels and protein solubility in bacteria is such that one can use the expression levels of a bacterial protein to predict its solubility with good accuracy or, conversely, derive its maximal mRNA expression levels from its intrinsic aggregation propensity [29]. This tight relationship suggests an evolutionary control of protein expression aimed at ensuring protein levels that suffice to execute properly the biological functions but that also guarantee solubility, at least in normal physiological environments. It follows that, as suggested by Tartaglia and co-workers [14], in terms of their solubility, proteins are maintained at the edge of aggregation. This situation explains why deregulation of protein homeostasis usually results in the formation of toxic protein aggregates.

#### *The rate of protein translation*

As we will see in the next chapter, protein folding might start before the end of protein translation [30–32]. Indeed, because protein folding and aggregation compete kinetically inside the cell, a fast-folding reaction constitutes perhaps the best strategy to prevent protein aggregation. Co-translational folding is modulated by the mRNA translation rate that defines, at each particular instant, both the size of the nascent protein and the time that a polypeptide stretch resides inside the ribosome [33, 34]. The 20 proteinogenic amino acids are encoded by 61 sense codons, with usually more than one synonymous codon encoding one amino acid. The protein elongation rate depends on the codons encoding each particular amino acid in the sequence and the availability of the corresponding tRNAs, an availability that defines the frequency of codon usage [35–37]. Accordingly, the presence of the most frequently used codons results in faster translation rates. The availability of frequently used codons is especially important in the case of highly expressed genes [36, 37].

An increasing number of studies suggests that co-translational folding is a sequential event in which the presence of rare codons establishes transcriptional pauses that provide enough time to the nascent protein to acquire a correct conformation [31, 36, 38, 39]. This view is supported by the preferential location of rare codons resulting in slow-translating regions down-stream of domain boundaries [39, 40] and by the observation that rare codon clusters define the limits of polypeptide chain regions with identical secondary structures [38]. The presence of continuous regions rich in rare codons is not a common feature of mRNAs since the presence of these regions might lead to frameshifts [41] or premature termination of translation [40, 42, 43]. Therefore, the conservation of these regions suggests that they play a specific function. This idea leads to the provocative hypothesis that codon usage might be a general tool used to increase the fidelity of co-translational

folding and to contain structural information for the encoded polypeptide [31].

The elongation of nascent chains is, generally, a faster reaction than their folding. Therefore, it is not strange that the presence of infrequently used codons has been found to be associated with the formation of large structural units [40]. The presence of slowly translated sequence stretches would slow down the elongation reaction, allowing for synchronization with the usually slow speed of folding of large protein subunits. In this way, the fidelity of co-translational folding would be increased.

The exchange of natural codons for codons with higher usage frequency does not impede the acquisition of a correct protein conformation. As recurrently observed during the expression of heterologous codon-optimized proteins in bacteria, however, this exchange results in an uncoupling of the natural folding and translation rates and usually increases the amount of misfolded and aggregated protein in the cell, suggesting a co-evolution of the relative velocities of both processes [44].

#### *The ribosome-borne protein-folding activity*

The folding of a polypeptide might start before the chain completely leaves the ribosomal interior, suggesting that the ribosome itself might assist the folding process. In fact, the presence of ribosomes of different organisms in *in vitro* refolding mixtures has been shown to improve and accelerate the attainment of a native conformation for different, unrelated and initially denatured proteins [45–47]. This result suggests that the ribosome-borne protein-folding activity might represent a universal folding strategy to help proteins in the attainment of their functional states [45–48]. Initially, this activity was attributed to different chaperones able to associate with the ribosome; however, experiments in chaperone-deficient cellular backgrounds indicate that the ribosome itself has the capability to assist protein folding [46, 47].

The ribosome inner channel from the peptide synthesis site to the exit aperture is about 100 Å long and 20 Å wide [49]. Several studies indicate that the size of the cavity is compatible with the presence of protein secondary structure elements [49–51]. Using a fluorescence-resonance energy-transfer approach, it has been observed that the energy transfer efficiency measured when a transmembrane sequence is located in the ribosome exit tunnel is similar to that exhibited by the polypeptide in a lipid bilayer, indicating that the nascent transmembrane sequence assembles inside the ribosome cavity into an  $\alpha$ -helix structure analogous to that it adopts when is incorporated into the membrane [51].

The folding capability of the different ribosomal subunits has been analyzed separately to determine which

particular ribosomal component assists this reaction. It has been found that the activity is mainly located in the large ribosomal subunit and specifically in the domain V of *E. coli* 23S rRNA where the peptidyl transferase activity is also situated [52]. In agreement with these data, it has been observed that antibiotics able to bind the peptidyl transferase center and stop translation are also able to inhibit protein folding *in vitro* [17, 46]. Overall, these results suggest that the ribosome peptidyl transferase center performs the dual role of protein synthesis and folding [46, 53, 54].

#### Protein folding control in the bacterial cytosol

##### *A generic genetic response against misfolded proteins*

Despite the described fine regulation of protein levels, the ribosomal protection of the nascent polypeptide chains and the synchronization of translation and folding rates, in many cases misfolding and aggregation of new synthesized polypeptides cannot be completely avoided [24]. On these occasions, a general genetic transcriptional response against protein misfolding is triggered in bacterial cells [55, 56]. This response involves an increase in the levels of heat-shock proteins and chaperones as well as changes in the levels of ribosome-associated proteins, suggesting a translational and post-translational regulation [56]. The heat-shock sigma factor  $\sigma^{32}$ , which regulates the expression of many proteins with chaperoning activity like ClpB, DnaJ, DnaK, GroEL, GroES, GrpE, IbpA, and IbpB [57], appears to be the central component that controls the cellular response against protein misfolding [56]. The synthesis, stability and activity of  $\sigma^{32}$  are modulated by the balance between the amount of chaperones and unfolded proteins present in the cellular cytoplasm [56]. In this way, the multi-chaperone DnaK/DnaJ/GrpE complex binds  $\sigma^{32}$  and impedes its translational activity in a negative feedback loop [58, 59]. Misfolded proteins block these chaperones and compete with their binding to  $\sigma^{32}$ , providing the precise  $\sigma^{32}$  transcriptional activity required for the concentration of misfolded proteins present at each instant [56].

Because  $\sigma^{32}$  controls the expression levels of the series of genes up-regulated in the presence of insoluble proteins, it is an attractive target in strategies aimed at preventing protein aggregation in bacteria, a topic of special interest in biotechnology. In addition, because these genes act as effective reporters of the levels of misfolded protein in the cell at any point in time, their promoters can be fused to easily detectable gene reporters, like  $\beta$ -galactosidase, and subsequently the activity of the reporter can be used to monitor and optimize protein solubility at the different stages of the protein production process [55].

### Chaperone activity in the bacterial cytosol

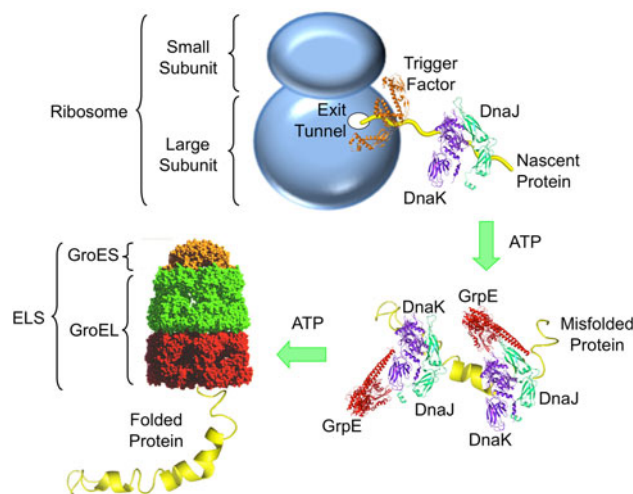
The molecular chaperone machinery comprises a specific cellular mechanism for preventing misfolding and aggregation of newly synthesized proteins. The chaperones are proteins with the capacity to recognize non-native states of other proteins and assist their folding and/or prevent their aggregation by controlled binding and release mechanisms. The best-characterized type of chaperone proteins are the so-called “heat-shock proteins” (Hsps). As described above, their expression is induced under unfavorable cellular conditions resulting in an increase of misfolded proteins. However, various families of Hsps are also involved in diverse processes under normal physiological conditions such as the folding of newly synthesized polypeptides [11]. In the bacterial cytosol, the folding of new proteins is assisted by three major molecular chaperone complexes: (1) Trigger factor (TF), (2) DnaK–DnaJ–GrpE and (3) GroEL–GroES (ELS) complexes [60, 61].

The first chaperone to interact with nascent polypeptide chains coming from the bacterial ribosome is a 48-kDa protein named trigger factor (TF) [11]. TF folds into an elongated structure comprising three domains: an N-terminal ribosome-binding domain (RBD), a central domain with peptidyl-prolyl *cis/trans* isomerase activity (PPIase), and a C-terminal domain that forms two arms between RBD and PPIase domains, for which no function has been clearly defined yet [11, 60, 62]. The activity of PPIase domain is modulated by the presence of the two other domains, indicating cooperativity [60]. TF combines its PPIase activity with a chaperone-like function [60, 63]. TF binds to the ribosome as a monomer near the nascent peptide exit tunnel. In contrast, the high cytosolic concentration of TF favors its dimerization in solution [60, 62, 63]. All three domains are implicated in the interaction with nascent polypeptide chains at the ribosome during the translation process while also being involved in the dimer interaction surface of free TF, suggesting a possible competition between both processes [62]. The N- and C-ends of TF seem to be the main regions involved in the binding of the substrate and/or modulating its accessibility to the binding pocket [60]. TF polypeptide affinity is low in comparison with the vast majority of chaperones and is ATP independent, which results in a fast binding and release cycle that can be essential for its folding capacity, a quality that is coupled to the elongation of polypeptide chains [64]. Interestingly, as happens with most chaperone systems, there is redundancy in the activity exerted by the TF, and this action partially overlaps with that of the DnaK–DnaJ–GrpE chaperone complex. Accordingly, the deletion of the gene encoding for TF increases significantly the binding of DnaK to nascent polypeptides. This result is consistent with the observation that TF and DnaK can cooperate in nascent polypeptide

folding [60, 65, 66]. Deletion of TF alone does not result in any observable growth disadvantage, while the TF/DnaK double knockout exhibits synthetic lethality, indicating that the presence of at least one of these folding activities is essential in normal physiological conditions. However, this requirement can be overcome under specific controlled-growth conditions in which major folding defects accumulate, but the presence of additional, yet unidentified, chaperones might assist the folding of newly synthesized proteins and allow cell survival [60].

The DnaK–DnaJ–GrpE and ELS complexes are the best characterized molecular chaperone systems in the cytoplasm of *E. coli*. Additional proteins such as Clps (i.e., ClpA, ClpB), HtpG and IbpA/B, act as molecular chaperones in vitro while their function in cellular protein folding remains unclear [67].

The DnaK–DnaJ–GrpE complex forms a cellular chaperone system capable of repairing protein damage [68]. DnaK binds to solvent-exposed hydrophobic regions in unfolded polypeptide chains, assisting the folding and preventing misfolding and/or aggregation in an ATP-driven process that is regulated by the co-chaperone DnaJ and the nucleotide-release cofactor GrpE (see Fig. 1) [69]. DnaK is composed of two main domains: a 44-kDa N-terminal nucleotide-binding domain (NBD) that contains the ATPase activity, and a 25-kDa substrate-binding domain (SBD) that harbors the substrate-binding site. Both protein moieties are linked by a flexible region that can rotate  $\sim 35^\circ$ , favoring allosteric communication between the domains [70, 71]. When DnaK is not complexed with other chaperones and/or substrate proteins, it is mostly found in an ATP-bound state that displays low substrate affinity. The



**Fig. 1** Vectorial-assisted folding of a newly synthesized polypeptide involving the three major chaperone systems in the bacterial cytosol. ELS complex reproduced with permission from Ref. [89]

binding and release of the initially misfolded polypeptide by DnaK has been described as an ATP hydrolysis-dependent cycle that involves four steps: (1) the unfolded polypeptide initially binds to DnaJ; (2) substrate-bound DnaJ interacts with DnaK, activating its ATP hydrolytic activity and stabilizing an ADP state that has a high binding affinity for the substrate; (3) GrpE releases ADP from DnaK; and (4) ATP binding to DnaK stabilizes the low affinity state and triggers the release of substrate protein, thus completing the reaction cycle [69]. DnaK assistance might involve consecutive binding-release cycles until the folding processes have been successfully completed [11] or the released polypeptide chain is transferred to the ELS complex where the protein can be accommodated in its central cavity in an ATP-dependent process [72, 73]. DnaK can interact with two other chaperones, namely CbpA (an analogue of DnaJ with an overall amino acid sequence identity of 39% [74]) and DjIA, a 30-kDa type III membrane protein with a cytosolic putative J-domain in the C-terminal extreme that is fully compatible with J-domain of DnaJ [60, 75]. The physiological role of such interactions is still poorly understood.

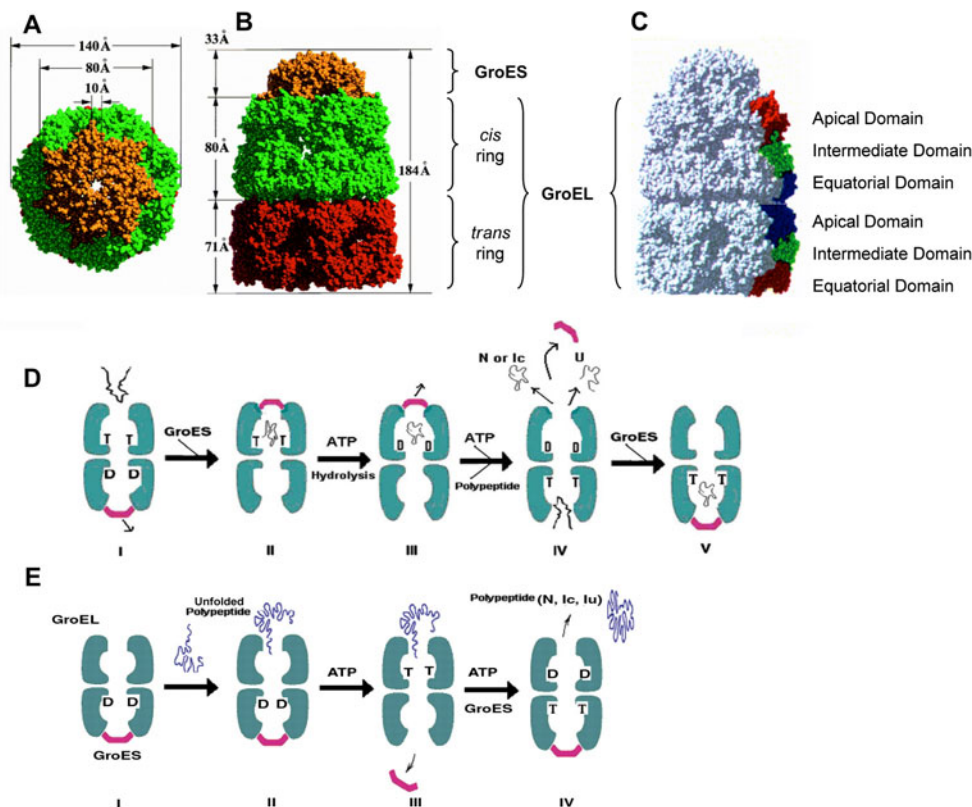
GroEL is a homotetradecameric complex that consists of 14 identical 57-kDa subunits arranged in two heptameric rings stacked back to back [76, 77]. Each single subunit is composed of three domains: (1) a flexible apical domain, responsible for binding both substrate as well as the co-chaperone GroES; (2) an equatorial domain, responsible for most of the intra-ring and all of the inter-ring contacts, wherein polypeptide possesses its own ATP-binding site that works in a cooperative manner following a concerted mechanism within the same ring [78, 79] and in a negative manner between the rings [80]; and (3) an intermediate domain that acts as a hinge and connects the other two domains covalently [81]. The folding activity of GroEL requires the cooperation of the co-chaperone GroES to form the ELS complex in the presence of ATP (see Fig. 2a–c) [81, 82]. GroES is a disk-shaped heptamer composed of seven identical 10-kDa units with a GroEL-binding loop per unit that form a dome-shaped structure on top of the GroEL rings [81]. The region of substrate and GroES binding at the apical domain of GroEL largely overlaps, but the mechanism by which GroES competes with the substrate in the GroEL-substrate complex is still unclear. Recent studies suggest the existence of an intermediate GroEL-substrate-GroES complex in which the substrate and GroES bind to GroEL by sharing seven common binding sites [83].

Although many bacterial proteins, ranging from small-single domains to large multi-domains, need assistance from the ELS complex during their lifetimes, this process is not completely understood and the precise mechanism of folding by ELS is still under debate. However, diverse

studies have recently clarified crucial points of the process, suggesting possible mechanisms of action [81]. ELS can assist the protein folding by two different pathways: (1) a *cis* folding action, initially considered the unique ELS mechanism until the work of Inbar and Horovitz [84] or (2) a *trans* folding mechanism [81, 85]. Independently of the mechanism, for many proteins, multiple binding-release cycles are necessary in order to obtain correct folding [81].

The *cis*-folding mechanism is usually used to assist the folding of small proteins (<70 kDa). In this case, the misfolded polypeptide chain rests encapsulated in the *cis*-ring of GroEL, in contact with GroES (see Fig. 2d) [81, 86]. In the free GroEL complex, the two rings of the chaperone are identical. The formation of GroEL-substrate-GroES complex causes large asymmetric conformational changes in one of the rings [82]; the conformational changes in the *cis*-ring (the GroEL ring in contact with GroES) create a large hydrophilic cavity, named “*cis*-cavity” inside the complex. On one hand, the cavity can play a passive role as an “Anfinsen cage” [87] by simply removing misfolded species and encapsulating the substrate, protecting it from the external media and so it can be correctly folded without aggregation. [82, 88, 89]. In addition, the entropic penalty upon folding is reduced in a confined environment, thermodynamically favoring the reaction. On the other hand, the cavity might play an active role in removing energy traps by rearranging the bonds, unfolding the protein to some extent or some other direct action. It is possible that the substrate protein itself influences the GroEL structure and reaction cycle [81, 90–92]. The *cis* cycle is governed by an allosteric communication process that can be summarized in three steps ( $T \rightarrow R \rightarrow R''$ ): (1) in the T state, the allosteric cycle starts with a high substrate affinity caused by the binding of ATP in the *cis*-ring subunits of GroEL that induces a cooperative aperture in the apical domains of all subunits of the ring, allowing capture of the substrate in the cylindrical cavity by structural changes at the binding site; (2) in the R state, GroES binds to the GroEL-substrate complex which after ATP hydrolysis assists the substrate unfolding-refolding reactions, propagating the conformational changes from the substrate-binding site to other distal zones of the structure by a cooperative effect; and (3) in the  $R''$  state, the positive initial cooperative process is finished upon release of ADP from *cis*-ring and binding of ATP to the *trans*-ring (the GroEL ring distal to GroES), promoting the release of GroES and the substrate in a negatively cooperative process. A new functional state is then attained that is stabilized by a redistribution of intra- or/and inter-molecular interactions as a result of the collective conformational rearrangement [93–96].

The *trans* folding mechanism is used to assist the folding of large polypeptides (>70 kDa) or is used in the



**Fig. 2** Structure and function of the GroEL–GroES (ELS) complex. **a** Top view of the complex looking down from the GroES-binding (*cis*) side. **b** Side view, *trans* GroEL ring is shown in red, *cis* GroEL ring in green and GroES in gold. **c** Each subunit in each GroEL ring is composed of three domains: the apical, equatorial, and intermediate domains. Misfolded polypeptides bind to the apical domain (in red), the equatorial domain (in blue) binds and hydrolyzes ATP and the intermediate domain (in green) acts as connector. GroES binding to GroEL-ATP causes the apical domain to move upward and rotate 120°. This large conformational change increases the size of central cavity where the polypeptide resides. Accordingly, the *cis* and *trans* rings display different volumes [81]. **d** *cis* mechanism of ELS complex. (I) The GroEL apical domains in the open ring capture misfolded polypeptides. (II) Formation of the GroEL-ATP-GroES complex promotes conformational changes in the apical domain resulting in substrate release inside the *cis* cavity. (III) Folding continues until ATP hydrolysis results in debilitation of the

GroEL–GroES interaction. (IV) ATP binding to the *trans*-ring favors the release of substrate from the *cis*-ring. Consecutive cycles can be necessary until a correct protein structure is attained. (V) Binding of GroES to GroEL allows the rings to alternate between binding-active and folding-active states. **e** *trans* mechanism of ELS. (I) The apical domains in the open ring of GroEL–GroES-ADP complex can capture misfolded polypeptides. (II) Binding occurs through hydrophobic interactions. (III) Binding of ATP to the *cis*-equatorial domain of GroEL promotes the release of GroES and ADP from the *trans* ring. (IV) ATP and GroES rebound to the *trans* GroEL ring, promoting a conformational change on the *cis* GroEL ring and liberating the polypeptide into the solution. The committed form of the polypeptide is not perfectly folded but will become fully active as it reaches the solution. Protein conformers unable to attain the native state can rebound to GroEL and go through subsequent cycles. **a–c** adapted with permission from Ref. [89]; **d** adapted with permission from Ref. [81]; and **e** adapted with permission from Ref. [97]

degradation pathway of damaged proteins that cannot achieve the native conformation. The polypeptide and GroES bind to the opposite rings of GroEL since the substrate is too large to be encapsulated by GroES. The apical domains of the ring occupied with substrate remain in a state similar to the free GroEL conformation (see Fig. 2e) [81, 85, 86, 97]. In the *trans* mechanism, binding of GroES and ATP in the *trans* ring is mandatory for the release of GroEL-bound substrates. In this case, the chaperone promotes a passive kinetic partitioning in which the newly synthesized polypeptides, which are susceptible to misfolding, may avoid aggregation through non-covalent association with GroEL, increasing the molecular flux

through productive folding pathways [98]. Once the aggregation phenomenon has been avoided, the released polypeptides might reach the folded state by themselves. Like in the *cis* mechanism, the folding of substrates in *trans* might require multiple rounds of binding and release.

Finally, GroEL has been shown to associate with the ribosome translation complex. In this configuration, GroEL holds the nascent polypeptide chain on the ribosome in a length-dependent manner and the protein is post-translationally encapsulated by the binding of the GroES cap to initiate the above-described chaperonin-assisted folding process [99], a process that again links protein synthesis to protein folding.

## Protein aggregation in bacterial cells

### Inclusion bodies formation and processing

#### *Misfolded proteins aggregate in bacteria into inclusion bodies*

The large number of usually redundant mechanisms and the large amount of energy the cell employs to assist proper protein folding reflects how important it is to prevent misfolding and aggregation in the cellular background. At any given time, the concentration of a protein in its native state results from a balance between the rate of protein synthesis, the rate of de novo folding, and the stability of the protein conformation. Situations of stress that affect this equilibrium, for example decreasing the stability of the native fold, would increase the number of misfolded molecules in the cell and promote aggregation. Heat shock has traditionally been the most studied stress factor, but recombinant protein expression can also dramatically modify the cell balance. Recombinant proteins are usually expressed in bacteria under the control of heterologous strong gene promoters. This results in a continuous and high translation rate, so that as much as 50% of the cellular proteins can correspond to recombinant polypeptide, usually saturating and/or de-coordinating the mechanisms to assist protein folding described in the previous chapter, which ultimately leads to the aggregation of the polypeptide of interest into inclusion bodies (IBs) [100]. These intracellular aggregates are dense, porous, hydrated, apparently amorphous, and refractile particles of nearly 1  $\mu\text{m}$  in diameter and are commonly localized at the cell poles [101, 102]. IBs can be formed both in the cytoplasm and periplasm, although the aggregates formed in these compartments exhibit morphological, compositional, and conformational differences [9, 103]. During its existence, a bacterium could contain one or two of these aggregates, the composition of which varies along the cell life span [102, 104]. At early stages of aggregation, host polypeptides could represent up to 50% of the proteins in IBs. However, mature IBs usually contain little cellular protein; many times, the heterologous chains represent more than 90% of the protein in the aggregate [102, 105–107] while the rest of the components in purified IBs are proteolytic fragments of the recombinant protein, ribosomal components, traces of membrane proteins, phospholipids and nucleic acids [108]. Whether these cellular elements are integral components of IBs or they are co-purified with the aggregate after cell lysis is not known. Importantly, it has been discovered that several elements of the cellular protein quality control, like small heat-shock proteins (IbpA and IbpB) or chaperones (DnaK and GroEL) [109], are systematically associated to IBs. This result reveals a relationship between

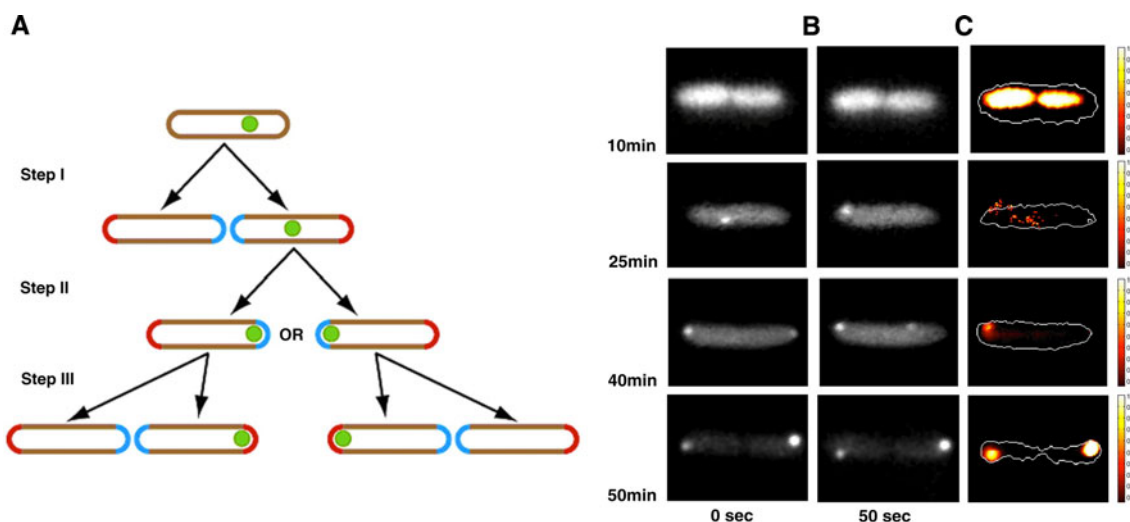
the folding machinery and the formation of proteinaceous deposits and suggests remodeling of IBs as an alternative cellular strategy to refolding or eliminating aggregated protein [110].

#### *Asymmetric segregation of inclusion bodies*

The formation of protein deposits in *E. coli* has been recently associated with a reduction in growth rate, a decrease in reproductive ability, and an increase in the chance of death, all phenotypes typically associated with aging [27]. These results suggest that bacteria could be an excellent model organism to study the effect of protein aggregation on life span. The formation of large protein deposits at one or both poles of the cell has been shown to be an energy-dependent process that comprises two sequential steps: the first stage is the formation of multiple small aggregates with a random cellular distribution and their further transport and concentration to the poles, where they remain immobilized as IBs (see Fig. 3). This process suggests that as in the case of the eukaryotic aggresome, IBs could in fact be aggregates of aggregates [26]. The second stage is dependent on two energy sources: the proton motive force and ATP. ATP is required for the activity of DnaK, which together with DnaJ is specifically needed for the formation of polar aggregates, justifying the association of these chaperones with bacterial aggregates. The maintenance of the cellular rod-shape is also an essential feature needed to concentrate the IBs specifically at the poles. Despite their apparent simplicity and reduced size, bacteria manage to maintain a spatial asymmetric distribution of their structural components, like those involved in cell division or chemotaxis, which are essential for correct cell function. This distribution could also be the case of insoluble aggregates.

*E. coli* reproduction consists of a symmetric division and an asymmetric segregation of the intracellular content, allowing a differential distribution of aged and young components between parent and offspring [25–27]. The old pole cell corresponds to the aging parent and the new one to the progeny. As a consequence of this asymmetry, any component localized preferentially at the poles, like aggregated protein in IBs, will remain longer in the old cell pole than in the young one [25, 26, 111, 112]. Because the presence of IBs diminishes the fitness of bacterial cells, the segregation of these deposits could have evolved as an effective strategy to avoid damage in the progeny and ensure the perpetuation of the population [25, 26]. Moreover, this process makes possible the concentration of misfolded protein species in a concrete space, facilitating subsequent action of the disaggregating machinery [26]. Only some of the daughter cells are designated for processing of the aggregates, whereas the rest of the





**Fig. 3** Localization of IBs to the cellular poles. **a** Schematic diagram illustrating IB segregation within a bacterial population. Independently of the initial position in the mother cell, at polar, mid-cell, or quarter-cell positions, after a reduced number of cellular divisions, IBs would become invariably located and consistently inherited by the old-pole cell (in red) whereas new poles (in blue) and daughter cells become free of aggregates. **b, c** Migration of protein aggregates

to the bacterial poles. **b** Images of a time course experiment indicating the location of a recombinant aggregation-prone protein fused to GFP after induction of its expression. **c** Next to each frame is shown the frequency of GFP fluorescence at each particular point over time according to the scale on the right, where brighter colors indicate higher GFP permanence. **a** Adapted with permission from Ref. [25]; **b, c** adapted with permission from Ref. [26]

population becomes free of IBs. These observations suggest that the cell controls the protein quality through multi-step processes, controlled both spatially and temporally, in which initially misfolded proteins can be refolded or aggregated. In this latter case, misfolded proteins are confined at the poles for posterior processing [26].

#### *Processing of aggregated protein in bacteria*

In bacteria, a network of chaperones and proteases carries out the processing of aggregated protein. Two protein types constitute the key elements in this network: (1) Hsp70 and (2) Clp ATPases belonging to the AAA+ superfamily [113]. Both chaperone and protease activities are essential for the cell and contribute to protein aggregate disintegration. Accordingly, over-expression of chaperones and proteases in recombinant systems has been shown to be an excellent strategy for the disaggregation of IBs and a concomitant formation of soluble species [108]. Aggregated and misfolded proteins usually expose hydrophobic clusters that constitute interaction sites for both chaperones and proteases. This process implies that although these activities are usually exerted in an independent manner, they can, at least in part, substitute for one another [113].

As discussed above, the Hsp70 family, of which DnaK is the main chaperone, is associated with prevention of protein aggregation by assisting protein refolding and aggregate localization. However, the Hsp70 family can also participate in aggregate processing [113, 114]. DnaK actively assists the gradual disassembly of oligomers and aggregates by

refolding misfolded polypeptides accessible at the aggregate surface. Interestingly, the specific activity of the chaperone depends on the aggregate size. Individual DnaK molecules are efficient in the fragmentation of large aggregates into medium-sized aggregates, but the subsequent fragmentation becomes gradually less efficient with decreasing aggregate size. Thus, the cooperation of several DnaK molecules is required in order to assist the folding of proteins in the smallest aggregate species [115, 116]. Although speculative, it is interesting to propose that this differential activity is related to the conformations of the polypeptide in the aggregate. If, as proposed recently, bacterial protein deposits are aggregates of aggregates, the initial action of DnaK could consist in reverting the large aggregate, in which the contacts between compacted aggregates would be superficial, into the original, smallest, assemblies in which most of the protein would be involved in the network of contacts that sustain the aggregate and block its disaggregation by DnaK. Accordingly, in the absence of DnaK, intracellular bacterial aggregates tend to be larger than in wild-type cells [117]. Apart from DnaK, the other main bacterial chaperone GroEL could be important for aggregate formation and/or processing since its absence results in a population of numerous, but very small, inclusion bodies in the cell [117]. Although other chaperones like IbpA and IbpB might participate in the process and are usually found associated to IBs, the impact of their null mutations on aggregate formation is rather moderate [118].

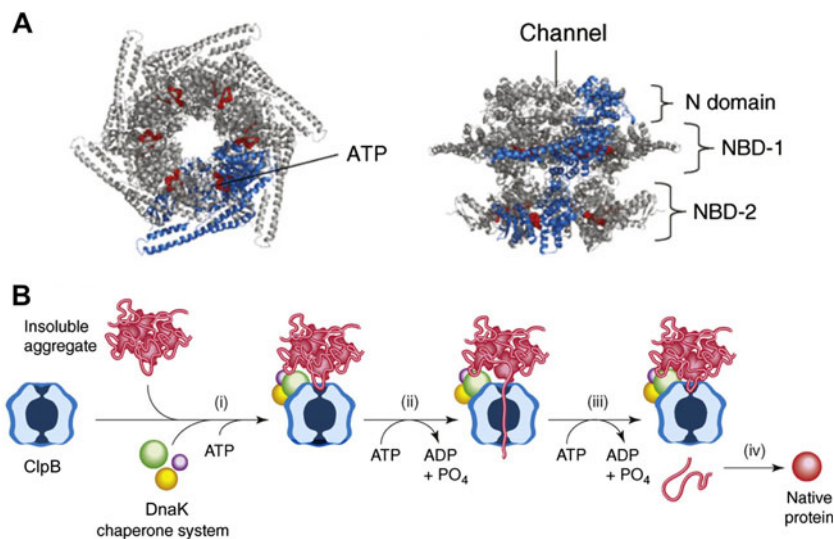
In general, Clp ATPases are formed by a protease core and oligomeric ATPase rings and are involved in the

degradation of irreversibly damaged proteins. They act as proteolytic machines (i.e., ClpA or ClpX) in a process similar to that of the proteasome. The function of the ATPase rings seems to be the energy-dependent unfolding of polypeptides by threading them through the center of the oligomeric ring. Unfolded substrates are then directly transferred into the associated protease for their irreversible degradation [113]. ClpB activity constitutes an important exception to this generic proteolytic action, since it is specifically involved in protein aggregate disintegration [113]. Remarkably, this activity is exerted in cooperation with the DnaK–DnaJ–GrpE system [67, 119, 120]. This DnaK system action occurs up-stream of ClpB, although it can also act down-stream of ClpB by refolding the solubilized proteins. Initially, DnaJ associates with the aggregate and drives DnaK to the aggregate surface; later on, DnaK mediates the binding of ClpB to the aggregate. Once at the proper location, the bichaperone network can work jointly in the solubilization of polypeptide molecules for their subsequent folding [121–123]. Each ClpB subunit consists of two ATPase domains, NBD1 and NBD2, and a long coiled-coil domain that is likely involved in substrate interaction [124]. The functional unit of ClpB is a dynamic homohexameric ring in which domains and subunits cooperate to execute the chaperone and ATPase activities [119, 120]. In an analogous manner to proteolytic members of the AAA+ family, the probable mechanism of disaggregation–refolding by ClpB involves the extraction of polypeptides from aggregates by forced unfolding,

translocation through ClpB central pore and release into cellular milieu for spontaneous or chaperone-mediated refolding (see Fig. 4) [125, 126]. The mechanical energy for this process is provided by ATP hydrolysis at the two nucleotide-binding domains of each monomer.

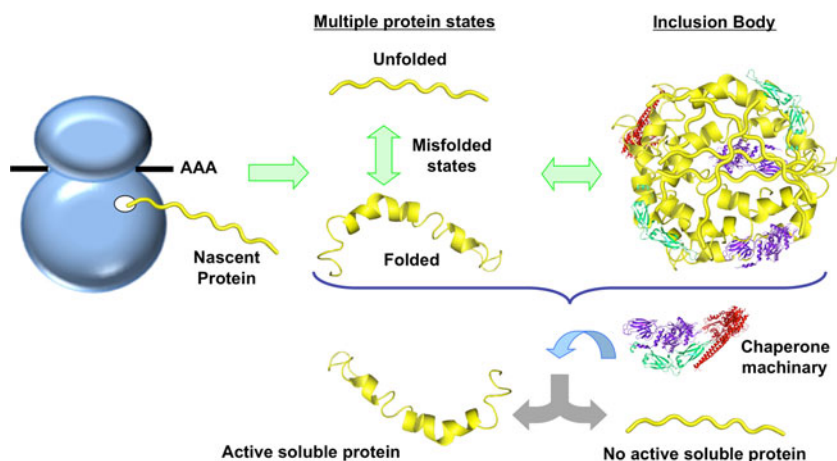
#### *A dynamic equilibrium between soluble and aggregated compartments in bacteria*

On the one hand, in a living cell, the activity of the quality control machinery operates in the disaggregation, unfolding and activation of the proteins embedded in intracellular deposits to promote protein solubility [103, 113, 127–131]. On the other hand, ongoing protein translation provides a continuous supply of aggregation-prone, partially folded intermediates, promoting protein deposition. These two opposite effects result in the establishment of a dynamic equilibrium between protein deposition and solubilization in the cell, with continuous polypeptide traffic between the soluble and insoluble compartments [9, 102, 106, 132–136] (see Fig. 5). Any intrinsic or extrinsic factor affecting this balance would modify the ratio between soluble and insoluble intracellular protein. In this way, in the absence of protein translation, IBs are almost totally disaggregated in a few hours, likely because the population of molecular chaperones is no longer limiting for processing the aggregated protein [133]. In an analogous manner, increasing the amount of available chaperones by over-expression usually promotes and increases the soluble to insoluble protein



**Fig. 4** Remodeling of bacterial protein aggregates. **a** Structural model of ClpB hexamer with bound ATP. An individual subunit is shown in blue and ATP in red. **b** Disaggregation of bacterial aggregates by ClpB in combination with the DnaK system. The DnaK chaperone system interacts with ClpB, acting before or together with ClpB. DnaK is likely involved in the presentation of substrates to ClpB (i) and/or in the hydrolysis of ATP. ClpB acts as the main

disaggregase. It extracts proteins from the aggregate by translocating them through their central channel (ii) and releasing them in an unfolded form into the solution (iii) where they can refold spontaneously or with assistance of other chaperones including the DnaK system (iv). **a** Reproduced from [125] and **b** adapted with permission from [125]



**Fig. 5** Conformations adopted by a recombinant protein inside bacteria. After protein synthesis, the polypeptide can fold, either spontaneously or with the assistance of chaperones, or remain totally or partially unfolded. These initially soluble structures can aggregate, preferably through selective interactions, to form a supramolecular IB in which multiple conformations can coexist. The protein quality machinery acts on both the soluble and insoluble protein conformers,

promoting a kinetic equilibrium between both cellular fractions. These actions result in functional and inactive protein structures coexisting in both the soluble and insoluble compartments. Accordingly, solubility alone does not present an accurate measure of protein quality during recombinant expression and IBs cannot be excluded as a source of functional protein

ratio [137, 138]. Conversely, the depletion of molecular chaperones [139] and/or the use of strong promoters resulting in very high translation rates [137, 138], especially during recombinant expression, almost invariably increases the proportion of aggregated species. In this dynamic context, one of the functions of IBs might be to act as a protein reservoir [103, 106, 133] where polypeptides can be potentially stored and extracted according to cellular conditions [132, 140, 141].

As a result of the continuous protein exchange between the soluble and insoluble compartments, a large variety of protein conformations co-exist in both fractions. Accordingly, in addition to properly folded protein, the soluble fraction usually contains misfolded inactive molecules or even soluble aggregates [139, 142, 143]. Moreover, the presence of active protein is not an exclusive feature of the soluble compartment, since functional conformations have been demonstrated to coexist with aggregated structures in many IBs [139, 142–144]. A relevant biotechnological implication of this conformational heterogeneity is that protein solubility by itself is not an accurate indicator of protein quality, since the presence of functional insoluble and misfolded soluble species might result in the proteins in the two fractions having similar effective specific activity [139, 142, 143].

Importantly, the protein quality-control machinery is not only involved in the cellular distribution of protein solubility but it also determines the structural features of the protein aggregates. In this way, in the absence of chaperones, intracellular aggregates are structurally more homogeneous, more compact and contain less intramolecularly bound

functional conformations [139]. Also, because the constant disintegrating-refolding activity of chaperones is sterically restricted to the aggregate surface and preferentially targeted to misfolded conformations [114, 145], the action of chaperones causes the structural composition and activity of the proteins at the surface to be different from those residing in the core of the aggregate [146].

#### Structure of inclusion bodies

##### *Sequential and structural specificity of inclusion bodies formation*

The sequential and structural analysis of proteins forming IBs in bacteria does not show any pattern or similarity, leading to the view that the formation of such aggregates depends on non-specific and general protein features, such as the exposure of hydrophobic clusters in partially folded intermediates during the folding process. However, an increasing body of evidence points to bacterial aggregation being a rather selective process modulated by the protein sequence and conformation. This selectivity was first observed during the *in vitro* refolding of a mixture of folding intermediates from P22 coat and tailspike proteins, two polypeptides that form IBs under overexpression but display segregated self-association in the test tube [147]. Besides, it has been demonstrated that purified bacterial intracellular aggregates are able to promote the aggregation of homologous soluble proteins in a dose-dependent manner but do not affect the aggregation of heterologous proteins, supporting the establishment of specific

sequential interactions between the soluble and aggregated species. Apparently, the aggregation determinants responsible for such contacts are exposed in partially folded intermediates but protected in the native structure, since the aggregates are not able to capture homologous but compactly folded polypeptides from the solution [148].

The sequential specificity of IB formation has also been addressed *in vivo* by simultaneously co-expressing two different heterologous proteins in the same bacterial cell. The simultaneous coexpression of *Vitreoscilla* hemoglobin (VHb) and a beta-lactamase precursor resulted in the formation of two types of cytoplasmic aggregates with different morphology and composition [149]. In a related approach, our group has co-expressed two aggregation-prone polypeptides fused to different fluorescent tags to facilitate their subcellular localization and to determine their proximity in the aggregate using fluorescence resonance energy transfer (FRET) [12]. As a control, we co-expressed two fluorescent versions of one of these proteins. These proteins generate deposits displaying high FRET efficiency, indicating that both fluorophores co-localize inside the IB. In contrast, the associated fluorescence signals do not significantly co-localize and exhibit a low transfer of energy in the IBs formed by the deposition of the two different polypeptide chains; this result indicates that they are distant in the space and likely do not interact significantly. Interestingly, one protein was principally situated in the inner core whereas the other one was found at the outer face. This particular distribution suggests that inside the cell, the preferential interaction between homologous sequences results in different aggregation rates that contribute to the kinetic and spatial partitioning in the aggregation process [12].

The aggregational effect of a large number of polypeptide mutants have been studied, both *in vitro* and inside bacteria, to understand how the amino acid sequence modulates protein deposition. Overall, the comparison between the different sequential substitutions and the resultant depositional changes shows a significant correlation between the protein aggregation propensity and the intrinsic physicochemical of the amino acid substitution, such as the charge, the propensity to acquire certain secondary structure or the hydrophobicity [150, 151]. However, the generality of this rule apparently clashes with the recurrent observation that the introduction of the same amino acid substitution in different proteins or even at different sequence positions in the same protein has dissimilar effects in the aggregation propensity of the mutated polypeptide, suggesting that the impact of the mutation depends critically on its location in the protein sequence. This apparent discrepancy can be reconciled if one considers that, as demonstrated for amyloid formation, the protein sequence might contain short specific regions able

to disrupt or assist the protein aggregation process [152, 153]. Regions promoting the deposition process have been designated hot spots (HS), and substitutions in these regions almost invariably result in changes in protein solubility, whereas mutations outside of these regions have a more moderate or even neutral effect [151, 154–157]. HS regions are enriched in aliphatic and aromatic residues (Val, Leu, Ile, Phe, Tyr, Trp) [152, 158]. The presence of charges of the same sign is avoided but not the existence of opposing charges, which favors specific interactions between polypeptide chains [152]. These aggregation-prone regions are commonly buried in the inner hydrophobic core of globular proteins or involved in the network of contacts that stabilizes the monomeric or multimeric native states of polypeptides [159, 160].

Even if they impact negatively protein solubility, HSs are widely present in biological polypeptides because they are obligatory elements for the establishment of  $\beta$ -sheet structures and molecular interactions [160]. Because these dangerous stretches are common elements in nature, globular proteins have evolved to block them by intramolecular contacts or hiding them at the hydrophobic core of the native structure, explaining why IBs are unable to specifically recognize properly folded polypeptides. This result implies that the aggregation of globular proteins occurs from at least partially unfolded states that expose HS regions to the solvent and facilitate the self-assembly process. Importantly, it has been demonstrated that local conformational changes during regular thermal fluctuation suffice to trigger aggregation under physiological conditions, which means that not only intermediates in the folding pathway can aggregate but also that initially folded conformations are at danger of becoming aggregation precursors. Structural fluctuations are largely determined by the conformational stability of a protein. A high thermodynamic stability will, in principle, reduce the population of locally unfolded structures and, therefore, of assembly-competent structures [161]. In agreement with this hypothesis, we have shown that, in bacteria, the intracellular aggregation of globular proteins tightly correlates with relative conformational stability of these proteins [162].

The composition analysis of the regions that flank the aggregation-prone regions in *E. coli* proteins has shown that Pro, Arg and Lys frequently surround these regions. These residues act as gatekeepers and prevent the self-assembly of HS even if they are exposed after synthesis or by local fluctuations. The Pro-constrained conformation is incompatible with  $\beta$ -structures, whereas the positively charged residues promote repulsion between adjacent chains and, being long and flexible, make the self-assembly process unfavorable from an entropic point of view [152, 158]. In addition, when associated to hydrophobic

stretches, these residues act as reporters for the cellular quality control machinery to facilitate HS identification and blockage by the molecular chaperones [152]. Overall, the cell employs different and redundant strategies to deal with sequences containing short stretches of aggregation-prone residues that are otherwise necessary for attaining properly folded structures in most globular proteins.

#### *Amyloid protein structure in bacterial inclusion bodies*

In humans, protein aggregation correlates with the development of several deleterious disorders such as Alzheimer's disease, Parkinson's disease, prion-associated transmissible spongiform encephalopathies, and type II diabetes [2]. The proteins associated with these pathogeneses are structurally and sequentially unrelated, but all aggregate into amyloid fibrils that share a common structure in which  $\beta$ -strands of polypeptide chains are stacked perpendicularly to the fibril axis in a cross- $\beta$  architecture [2]. The process of amyloid self-assembly is sequence-specific, sensitive to protein mutations, requires total or local protein unfolding, is nucleation driven, regulated by the chaperone activity, and results in the formation of highly homogeneous protein aggregates [2, 163]. As described, these features are also characteristic of intracellular aggregate formation. This fact, together with the recurrent observation that, independently of the presence of globular or unfolded conformations, the main structural component of IBs consists of a network of tightly packed extended intermolecular  $\beta$ -sheets with spectroscopic properties close to those found in pathogenic fibrils, suggested that the biogenesis and structure of these two type of aggregates could be similar. However, the detailed structural characterization of bacterial aggregates is extremely challenging and the extent to which this similitude applies has remained essentially unknown. Only recently, two independent research groups have converged to demonstrate unequivocally the existence of amyloid structures inside bacterial IBs [12, 13]. Riek and co-workers analyzed the conformational properties of the inclusion bodies formed by three disease-unrelated polypeptides representative of the three major protein folds:  $\alpha$ -helix,  $\alpha$ -helix/ $\beta$ -sheet and  $\beta$ -sheet [13]. The CD spectra and, more specifically, the X-ray diffraction pattern of the aggregates demonstrated the presence of characteristic cross- $\beta$  structural conformations that are less strongly aligned than in amyloid fibrils. Accordingly, these protein deposits contain fibrillar structure observable by transmission electron microscopy (TEM) and bind to specific amyloid dyes with similar affinity to that of pathogenic fibrils. Moreover, using quenched hydrogen/deuterium exchange in nuclear magnetic resonance, they could demonstrate that, like in amyloids, specific short sequence stretches with high aggregation propensity are selectively

protected in the core of the aggregate and are likely the main residues contributing to the specific interactions that promote formation of the aggregate [13]. In a complementary study, we used the Alzheimer's-related A $\beta$ 42 peptide as a model with the aim of deciphering to what extent a polypeptide embedded in IBs and the same molecule polymerized into amyloids are structurally related. We used an extensive battery of analytical techniques and demonstrated the existence of an enrichment in cross- $\beta$ -sheet structure in A $\beta$ 42 IBs by Fourier transform infrared spectrometry (FTIR), the presence of fibrillar material by TEM and atomic force microscopy (AFM), binding to amyloid dyes, the specific capability of accelerating fibril formation only in the presence of homologous peptides and, as for amyloids, the presence of a compact, proteolytically protected core enriched in amyloid-like conformations [12]. Taken together, these results strongly support the establishment of amyloid-like interactions as an omnipresent mechanism driving the formation of protein aggregates across all realms of life [164].

#### *Inclusion bodies as infectious particles*

Prions represent a particular subclass of amyloids for which the aggregation process becomes self-perpetuating in vivo and thus infectious. Several proteins with prionic behavior have been shown to be deposited inside IBs upon heterologous expression, raising the possibility that they might acquire in bacteria a fibrillar structure similar to that of the infectious form and therefore become transmissible themselves. Recently, Meier and co-workers have analyzed the conformational and infectivity properties of the IBs formed by the C-terminal domain of the *Podospora anserina* HET-s prion. This region, spanning residues 218–289, is the most essential domain for amyloid formation and prion propagation in this filamentous fungus. HET-s IBs were used to infect prion-free *P. anserina* strains using a protein transfection method. While insoluble extracts from a control strain containing only the empty vector did not induce the development of the prion form, strains transfected with HET-s(218–289) IBs acquired the prion phenotype at a high frequency, comparable to that promoted by HET-s(218–289) infectious amyloid fibrils [165]. This result indicates that the HET-s(218–289) acquires an infectious prion fold in *E. coli*. The biophysical characterization of the HET-s IBs formed by the full-length prion and the isolated amyloidogenic domain and the solid-state NMR analysis of these aggregates converge to indicate that the presence of a fibrillar cross-beta conformation at the C-terminus domain of HET-s when embedded in IBs is responsible for the high infectivity of these bacterial aggregates [165, 166]. This result demonstrates that, at least in certain cases, the processes of amyloid fibril and IB

formation are related even at the fine molecular structural level.

The possibility of reproducing the aggregation of mammalian, fungus, or yeast prion proteins using bacteria as a consistent cellular model in which the effect of autologous or heterologous protein quality machineries and/or anti-aggregational and anti-prionic drugs can be further studied opens an exciting avenue for understanding the correlation between prion strain phenotypes, “species barriers” and prion particle conformation and infectiveness. Since the *E. coli* cytoplasm apparently contains all the molecular elements to support the formation of prion aggregates, the existence of as yet unidentified prionic bacterial proteins cannot be disregarded. The discovery of prionic proteins in bacteria would have important implications for human health.

### Function and toxicity of bacterial amyloids

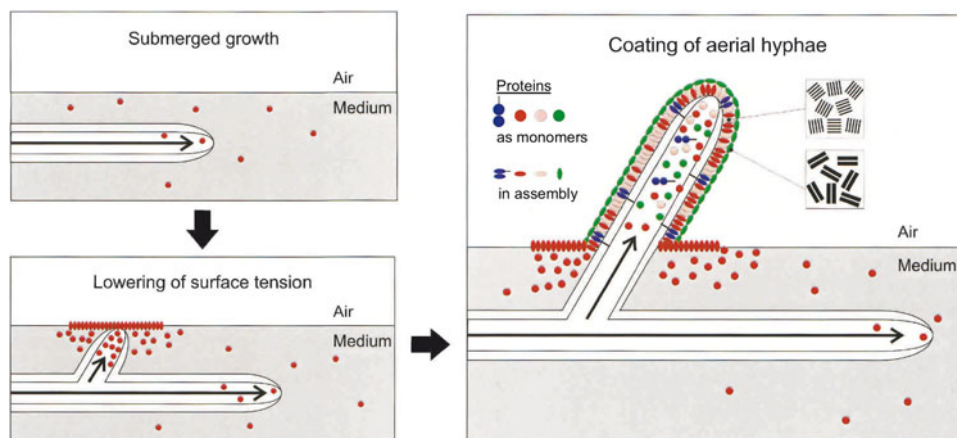
#### Functional amyloids in bacteria

The bacterial ability to construct amyloid structures is not restricted to IB assemblies; in fact, an increasing number of bacterial polypeptides able to form functional amyloid fibrils are being discovered [167]. As detailed below, these structures can be formed by structurally and sequentially unrelated proteins and can serve diverse biological functions such as surface adhesion, host invasion, colonization or cytotoxicity [168–171].

*E. coli*, *Salmonella enterica*, and other *Enterobacteriaceae* exploit the rigidity of the amyloid structure to form a fibrillar extracellular aggregate, named curli or tafi (thin aggregative fimbriae), able to perform functions related with bacterial adhesion, biofilm development and host

invasion [172, 173]. The main component of curli is the protein subunit CsgA, a 131-residue polypeptide rich in glycine repeats that rapidly self-assembles amyloid fibrils in vitro [174]. In vivo, the formation of fibrillar conformations depends on the presence of specific cellular machinery. In this way, the CsgG subunit facilitates the transport of CsgA to the cell surface [175] where the protein CsgB, helped by CsgF, acts as a nucleator for the specific assembly of CsgA [176]. The interaction of these fibrils with bacterial cellulose fibers is a crucial step in allowing attachment to different surfaces and in bacterial pathogenesis [168, 170, 177].

Chaplin amyloids have been found on the surface of the Gram-positive actinomycetes species *Streptomyces* [178]; however, there is genetic evidence that suggests the presence of chaplin amyloids in other actinomycetes [179]. These bacteria are characterized by a complex morphological differentiation composed by a vegetative mycelium state that could extend into the air and form chains of exospores. The developed spores could then be dispersed to allow the expansion of a new colonizing mycelium. Chaplins self-assemble to construct an amphipathic membrane [178], implicated in fixing the cell on hydrophobic surfaces [180], and they facilitate hyphae release from the water environment by reducing the water surface tension (see Fig. 6) [178]. Chaplin amyloid formation is controlled spatially and temporally [178, 179]. In this sense, the bldN developmental sigma factor regulates chaplin expression and guarantees that the amyloid fibrils are developed at the right time [179]. *Streptomyces coelicolor* possesses eight chaplins, five short (ChpD–H, up to 63aa) and three long (ChpA–C,  $\approx 225$ aa) forms that each contains an extra cell-wall-anchoring domain. The ChpD–H monomers self-assemble at the water–air interface and form the rigid



**Fig. 6** Formation of functional amyloid-like structures allows streptomycetes to invade the air. The water surface tension at the substrate-air interface is reduced by proteins that self-assemble into an amphipathic rigid membrane. This membrane is composed of

amyloid-like fibrils formed by the supramolecular organization of chaplins. Aerial hyphae continue to secrete these proteins that assemble at the hyphal surface, making them hydrophobic. Figure adapted with permission from Ref. [178]

amyloid fibrils that reduce water surface tension [178]. On the contrary, over a hydrophobic surface, ChpD–H monomers are soluble and adopt an alpha-helical conformation [181] that has been suggested to be a possible intermediate in the assembly process.

Bacteria also profit from the cytotoxic properties of amyloid assemblies, as in the case of harpins. These proteins are generated by plant pathogens and injected in the extracellular space of the host through a bacterial molecular syringe (a type III secretion system). Inside the plant, harpins induce a hypersensitive response, which is an early defence response that restricts the growth of plant pathogens by causing cell death in a mechanism similar to the apoptosis promoted in animal cells by different amyloids [182]. Accordingly, mutated harpins unable to self-assemble into amyloid structures do not elicit the plant programmed death response [182].

Bacterial amyloid structures might also serve to inhibit the growth of adjacent cells. This strategy reduces the competition with other bacterial species and improves pathogenicity. In this way, *Klebsiella pneumoniae* kills enterobacteria using the small peptide E492 (Mcc), a bacteriocin that causes the development of pores in the cytoplasmic membrane [183]. Surprisingly, the in vivo maturation of Mcc into amyloid fibrils structures results in the inhibition of the antibacterial function [184] suggesting that, while protofibrillar pore-promoting assemblies act as cytotoxic forms, mature amyloids might be inert protein deposits of aggregated Mcc that stop the lethal activity when it is no longer necessary. Interestingly, Mcc cytotoxic pores share morphological similarities with the small annular protofibrils formed by A $\beta$  and  $\alpha$ -synuclein, pointing to a common protein conformation able to generate toxic membrane permeabilization serving different functions [185, 186].

The endospore is a highly resistant structure generated in harsh environments by *Bacillus* and *Clostridium* species. Structural analysis of the spore coat has shown that it is principally composed of fibrillar structures with a morphology compatible with amyloids [187].

*Mycobacterium tuberculosis* produces pili that are assumed to be important for adherence, colonization and infection [188]. These pili display fibrillar structures detectable by electron microscopy that share structural and tinctorial properties with amyloid fibrils.

Overall, the formation of amyloid structures in bacteria cannot be simply considered an anomalous side-effect of producing non-bacterial proteins in prokaryotes, but is a more general phenomenon that responds to evolutionary strategies by exploiting the mechanical, cytotoxic and protective properties of these macromolecular assemblies. It is important to note that harboring polypeptides with amyloidogenic potential implies an inherent risk for the

cell. Accordingly, nature has developed strategies to avoid dangerous uncontrolled aggregation of these polypeptides by using auxiliary proteins that prevent or assist their assembly or separating the implicated polypeptides in different cellular compartments [171].

#### *Are bacterial amyloids potentially dangerous for humans?*

An important concern derived from the ability of bacterial and non-bacterial proteins to form amyloid structures in prokaryotic cells is the possibility that these aggregates might directly or indirectly be dangerous for humans. On one hand, the formation of infective conformations in bacteria might be a generic property of prion proteins; therefore, recombinant human, sheep or cow prions might also have the ability to form transmissible aggregated fibrillar structures and their infectivity for humans should not be completely discarded. On the other hand, the presence of natural amyloid structures in bacterial species known to infect humans might result in an anomalous interaction between the bacterial and host amyloidogenic proteins in the human body. Recent data suggest that these molecular interactions might have pathogenic consequences, and bacterial amyloids might act as seeds that could induce the onset of amyloidogenic diseases [170]. For example, the injection of curli fibrils or intact *E. coli* cells into mice accelerates the aggregation of amyloid protein AA, suggesting that exposure to bacteria by ingestion or inhalation may be an important risk factor in patients suffering this amyloidogenic disease [189]. Interestingly, the exposure of neuronal cells to the spirochete *Borrelia burgdorferi* and Mycolata species (*Nocardia* and *Mycobacterium*) generate phenotypes similar to those developed in Alzheimer's and Parkinson's disease, respectively. Although the precise mechanisms underlying such cellular changes are still not known, *Borrelia* generate cysts and Mycolata generate amyloid-like fibrillar spores and pili that might nucleate and accelerate the respective aggregation processes [190–192].

#### **Future perspectives**

In humans, failures in the control of protein folding and aggregation result in the development of degenerative and usually highly debilitating disorders. However, amyloids also can provide biological functionality as shown in bacteria (i.e. curli, chaplins), fungi (i.e., hydrophobins, heterokaryon incompatibility) and metazoa (i.e., spidroins, chorion). Importantly, two recent reports demonstrate that the formation of functional amyloids can contribute to normal cell and tissue physiology in humans. In this way,

the amyloid derived from the protein Pmel17 within melanosomes is involved in the generation of the melanin pigment [193]. In addition, peptide and protein hormones in secretory granules of the endocrine system have been shown to be stored in amyloid-like conformations [194]. Although *in vitro* studies have provided a wealth of information on the structural and sequential determinants that control the folding and polymerization of proteins, less is known about how these intrinsic propensities are modulated in the complex cellular environment. This lack of knowledge complicates the development of effective therapeutic approaches against these diseases. The information contained in the present review illustrates how bacterial organisms provide the bedrock on which to understand the complexity of protein folding and aggregation *in vivo*.

Folding and, especially, misfolding and aggregation in eukaryotic and prokaryotic environments have been long thought to be regulated by different mechanisms. Accordingly, the important data obtained in bacterial backgrounds has little influence on our present view of disease-related protein misfolding and aggregation in higher organisms. In addition, in contrast to other research fields, bioinformatics efforts aimed at gaining a global view of the problem have been rather scarce. As a consequence, a theoretical, integrated model of protein folding and aggregation inside the cell is still missing. The confirmation that the mechanism underlying these processes in bacteria, as well as the structural and functional properties of the final aggregated species, are related to those occurring in higher eukaryotic cells during disease should push scientists to develop integrative systems biology approaches using mathematical modeling that complement experimental work to examine the cellular pathways involved in maintaining protein homeostasis in simple and tractable but physiologically relevant bacterial models. This would open up new avenues towards a global, unambiguous and accurate understanding of protein folding, misfolding and aggregation in the cell. The basic knowledge gained at this level might well assist the development of new generic strategies to tackle devastating human conformational disorders.

**Acknowledgments** Work in our laboratory has been supported by grants BIO2007-68046 (Ministerio de Ciencia e Innovación, Spain) and 2009-SGR-760 (Generalitat de Catalunya).

## References

- Jahn TR, Radford SE (2008) Folding versus aggregation: polypeptide conformations on competing pathways. *Arch Biochem Biophys* 469:100–117
- Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75:333–366
- Puchtler H, Sweat F (1966) A review of early concepts of amyloid in context with contemporary chemical literature from 1839 to 1859. *J Histochem Cytochem* 14:123–134
- Sipe JD, Cohen AS (2000) Review: history of the amyloid fibril. *J Struct Biol* 130:88–98
- Westermarck P, Benson MD, Buxbaum JN, Cohen AS, Frangione B, Ikeda S, Masters CL, Merlini G, Saraiva MJ, Sipe JD (2007) A primer of amyloid nomenclature. *Amyloid* 14:179–183
- Makin OS, Serpell LC (2005) X-ray diffraction studies of amyloid structure. *Methods Mol Biol* 299:67–80
- Westermarck P, Benson MD, Buxbaum JN, Cohen AS, Frangione B, Ikeda S, Masters CL, Merlini G, Saraiva MJ, Sipe JD (2005) Amyloid: toward terminology clarification. Report from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid* 12:1–4
- Buxbaum JN (2003) Diseases of protein conformation: what do *in vitro* experiments tell us about *in vivo* diseases? *Trends Biochem Sci* 28:585–592
- Ventura S, Villaverde A (2006) Protein quality in bacterial inclusion bodies. *Trends Biotechnol* 24:179–185
- Ramakrishnan V (2008) What we have learned from ribosome structures. *Biochem Soc Trans* 36:567–574
- Pfanner N (1999) Who chaperones nascent chains in bacteria? *Curr Biol* 9:R720–R724
- Morell M, Bravo R, Espargaro A, Sisquella X, Aviles FX, Fernandez-Busquets X, Ventura S (2008) Inclusion bodies: specificity in their aggregation process and amyloid-like structure. *Biochim Biophys Acta* 1783:1815–1825
- Wang L, Maji SK, Sawaya MR, Eisenberg D, Riek R (2008) Bacterial inclusion bodies contain amyloid-like structure. *PLoS Biol* 6:e195
- Tartaglia GG, Vendruscolo M (2009) Correlation between mRNA expression levels and protein aggregation propensities in subcellular localizations. *Mol Biosyst* 5:1873–1876
- Cech TR (2000) Structural biology. The ribosome is a ribozyme. *Science* 289:878–879
- Yusupov MM, Yusupova GZ, Baucom A, Lieberman K, Earnest TN, Cate JH, Noller HF (2001) Crystal structure of the ribosome at 5.5 Å resolution. *Science* 292:883–896
- Voisset C, Thuret JY, Tribouillard-Tanvier D, Saupe SJ, Blondel M (2008) Tools for the study of ribosome-borne protein folding activity. *Biotechnol J* 3:1033–1040
- Simonetti A, Marzi S, Jenner L, Myasnikov A, Romby P, Yusupova G, Klaholz BP, Yusupov M (2009) A structural view of translation initiation in bacteria. *Cell Mol Life Sci* 66:423–436
- Simonetti A, Marzi S, Myasnikov AG, Fabbretti A, Yusupov M, Gualerzi CO, Klaholz BP (2008) Structure of the 30S translation initiation complex. *Nature* 455:416–420
- Nissen P, Hansen J, Ban N, Moore PB, Steitz TA (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science* 289:920–930
- Yusupova GZ, Yusupov MM, Cate JH, Noller HF (2001) The path of messenger RNA through the ribosome. *Cell* 106:233–241
- Ban N, Nissen P, Hansen J, Capel M, Moore PB, Steitz TA (1999) Placement of protein and RNA structures into a 5 Å-resolution map of the 50S ribosomal subunit. *Nature* 400:841–847
- Ban N, Nissen P, Hansen J, Moore PB, Steitz TA (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289:905–920
- Schubert U, Anton LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR (2000) Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* 404:770–774



25. Lindner AB, Madden R, Demarez A, Stewart EJ, Taddei F (2008) Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc Natl Acad Sci USA* 105:3076–3081
26. Rokney A, Shagan M, Kessel M, Smith Y, Rosenshine I, Oppenheim AB (2009) *E. coli* transports aggregated proteins to the poles by a specific and energy-dependent process. *J Mol Biol* 392:589–601
27. Stewart EJ, Madden R, Paul G, Taddei F (2005) Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol* 3:e45
28. Wickner S, Maurizi MR, Gottesman S (1999) Posttranslational quality control: folding, refolding and degrading proteins. *Science* 286:1888–1893
29. Tartaglia GG, Pechmann S, Dobson CM, Vendruscolo M (2009) A relationship between mRNA expression levels and protein solubility in *E. coli*. *J Mol Biol* 388:381–389
30. Das D, Das A, Samanta D, Ghosh J, Dasgupta S, Bhattacharya A, Basu A, Sanyal S, Das Gupta C (2008) Role of the ribosome in protein folding. *Biotechnol J* 3:999–1009
31. Komar AA (2009) A pause for thought along the co-translational folding pathway. *Trends Biochem Sci* 34:16–24
32. Kramer G, Boehringer D, Ban N, Bukau B (2009) The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nat Struct Mol Biol* 16:589–597
33. Crombie T, Swaffield JC, Brown AJ (1992) Protein folding within the cell is influenced by controlled rates of polypeptide elongation. *J Mol Biol* 228:7–12
34. Tsai CJ, Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM, Nussinov R (2008) Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. *J Mol Biol* 383:281–291
35. Irwin B, Heck JD, Hatfield GW (1995) Codon pair utilization biases influence translational elongation step times. *J Biol Chem* 270:22801–22806
36. Marin M (2008) Folding at the rhythm of the rare codon beat. *Biotechnol J* 3:1047–1057
37. Zhang G, Hubalewska M, Ignatova Z (2009) Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat Struct Mol Biol* 16:274–280
38. Krasheninnikov IA, Komar AA, Adzhubei IA (1988) Role of the rare codon clusters in defining the boundaries of polypeptide chain regions with identical secondary structures in the process of co-translational folding of proteins. *Dokl Akad Nauk SSSR* 303:995–999
39. Purvis IJ, Bettany AJ, Santiago TC, Coggins JR, Duncan K, Eason R, Brown AJ (1987) The efficiency of folding of some proteins is increased by controlled rates of translation in vivo. A hypothesis. *J Mol Biol* 193:413–417
40. Zhang G, Ignatova Z (2009) Generic algorithm to predict the speed of translational elongation: implications for protein biogenesis. *PLoS One* 4:e5036
41. Barak Z, Lindsley D, Gallant J (1996) On the mechanism of leftward frameshifting at several hungry codons. *J Mol Biol* 256:676–684
42. Roche ED, Sauer RT (1999) SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *EMBO J* 18:4579–4589
43. Withey JH, Friedman DI (2003) A salvage pathway for protein structures: tmRNA and trans-translation. *Annu Rev Microbiol* 57:101–123
44. Angov E, Hillier CJ, Kincaid RL, Lyon JA (2008) Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS One* 3:e2189
45. Argent RH, Parrott AM, Day PJ, Roberts LM, Stockley PG, Lord JM, Radford SE (2000) Ribosome-mediated folding of partially unfolded ricin A-chain. *J Biol Chem* 275:9263–9269
46. Chattopadhyay S, Pal S, Pal D, Sarkar D, Chandra S, Das Gupta C (1999) Protein folding in *Escherichia coli*: role of 23S ribosomal RNA. *Biochim Biophys Acta* 1429:293–298
47. Das B, Chattopadhyay S, Das Gupta C (1992) Reactivation of denatured fungal glucose 6-phosphate dehydrogenase and *E. coli* alkaline phosphatase with *E. coli* ribosome. *Biochem Biophys Res Commun* 183:774–780
48. Bera AK, Das B, Chattopadhyay S, Dasgupta C (1994) Refolding of denatured restriction endonucleases with ribosomal preparations from *Methanosarcina barkeri*. *Biochem Mol Biol Int* 32:315–323
49. Harms J, Schluenzen F, Zarivach R, Bashan A, Gat S, Agmon I, Bartels H, Franceschi F, Yonath A (2001) High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* 107:679–688
50. Lu J, Deutsch C (2005) Secondary structure formation of a transmembrane segment in Kv channels. *Biochemistry* 44:8230–8243
51. Woolhead CA, McCormick PJ, Johnson AE (2004) Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* 116:725–736
52. Jenni S, Ban N (2003) The chemistry of protein synthesis and voyage through the ribosomal tunnel. *Curr Opin Struct Biol* 13:212–219
53. Chattopadhyay S, Das B, Dasgupta C (1996) Reactivation of denatured proteins by 23S ribosomal RNA: role of domain V. *Proc Natl Acad Sci USA* 93:8284–8287
54. Samanta D, Mukhopadhyay D, Chowdhury S, Ghosh J, Pal S, Basu A, Bhattacharya A, Das A, Das D, DasGupta C (2008) Protein folding by domain V of *Escherichia coli* 23S rRNA: specificity of RNA–protein interactions. *J Bacteriol* 190:3344–3352
55. Lesley SA, Graziano J, Cho CY, Knuth MW, Klock HE (2002) Gene expression response to misfolded protein as a screen for soluble recombinant protein. *Protein Eng* 15:153–160
56. Smith HE (2007) The transcriptional response of *Escherichia coli* to recombinant protein insolubility. *J Struct Funct Genomics* 8:27–35
57. Nonaka G, Blankschien M, Herman C, Gross CA, Rhodius VA (2006) Regulon and promoter analysis of the *E. coli* heat-shock factor, sigma32, reveals a multifaceted cellular response to heat stress. *Genes Dev* 20:1776–1789
58. Gamer J, Bujard H, Bukau B (1992) Physical interaction between heat shock proteins DnaK, DnaJ, and GrpE and the bacterial heat shock transcription factor sigma 32. *Cell* 69:833–842
59. Tomoyasu T, Ogura T, Tatsuta T, Bukau B (1998) Levels of DnaK and DnaJ provide tight control of heat shock gene expression and protein repair in *Escherichia coli*. *Mol Microbiol* 30:567–581
60. Genevaux P, Keppel F, Schwager F, Langendijk-Genevaux PS, Hartl FU, Georgopoulos C (2004) In vivo analysis of the overlapping functions of DnaK and trigger factor. *EMBO Rep* 5:195–200
61. Hartl FU, Hayer-Hartl M (2002) Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 295:1852–1858
62. Lakshminpathy SK, Tomic S, Kaiser CM, Chang HC, Genevaux P, Georgopoulos C, Barral JM, Johnson AE, Hartl FU, Etschells SA (2007) Identification of nascent chain interaction sites on trigger factor. *J Biol Chem* 282:12186–12193

63. Maier R, Eckert B, Scholz C, Lilie H, Schmid FX (2003) Interaction of trigger factor with the ribosome. *J Mol Biol* 326:585–592
64. Maier R, Scholz C, Schmid FX (2001) Dynamic association of trigger factor with protein substrates. *J Mol Biol* 314:1181–1190
65. Deuerling E, Schulze-Specking A, Tomoyasu T, Mogk A, Bukau B (1999) Trigger factor and DnaK cooperate in folding of newly synthesized proteins. *Nature* 400:693–696
66. Teter SA, Houry WA, Ang D, Tradler T, Rockabrand D, Fischer G, Blum P, Georgopoulos C, Hartl FU (1999) Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell* 97:755–765
67. Thomas JG, Baneyx F (2000) ClpB and HtpG facilitate de novo protein folding in stressed *Escherichia coli* cells. *Mol Microbiol* 36:1360–1370
68. Schroder H, Langer T, Hartl FU, Bukau B (1993) DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J* 12:4137–4144
69. Szabo A, Langer T, Schroder H, Flanagan J, Bukau B, Hartl FU (1994) The ATP hydrolysis-dependent reaction cycle of the *Escherichia coli* Hsp70 system DnaK, DnaJ, and GrpE. *Proc Natl Acad Sci USA* 91:10345–10349
70. Bertelsen EB, Chang L, Gestwicki JE, Zuiderweg ER (2009) Solution conformation of wild-type *E. coli* Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proc Natl Acad Sci USA* 106:8471–8476
71. Bhattacharya A, Kurochkin AV, Yip GN, Zhang Y, Bertelsen EB, Zuiderweg ER (2009) Allostery in Hsp70 chaperones is transduced by subdomain rotations. *J Mol Biol* 388:475–490
72. Bukau B, Horwich AL (1998) The Hsp70 and Hsp60 chaperone machines. *Cell* 92:351–366
73. Hartl FU (1996) Molecular chaperones in cellular protein folding. *Nature* 381:571–579
74. Ueguchi C, Kakeda M, Yamada H, Mizuno T (1994) An analogue of the DnaJ molecular chaperone in *Escherichia coli*. *Proc Natl Acad Sci USA* 91:1054–1058
75. Genevaux P, Wawrzynow A, Zylicz M, Georgopoulos C, Kelley WL (2001) DjlA is a third DnaK co-chaperone of *Escherichia coli*, and DjlA-mediated induction of colanic acid capsule requires DjlA–DnaK interaction. *J Biol Chem* 276:7906–7912
76. Braig K, Adams PD, Brunger AT (1995) Conformational variability in the refined structure of the chaperonin GroEL at 2.8 Å resolution. *Nat Struct Biol* 2:1083–1094
77. Braig K, Otwinowski Z, Hegde R, Boisvert DC, Joachimiak A, Horwich AL, Sigler PB (1994) The crystal structure of the bacterial chaperonin GroEL at 2.8 Å. *Nature* 371:578–586
78. Bochkareva ES, Lissin NM, Flynn GC, Rothman JE, Girshovich AS (1992) Positive cooperativity in the functioning of molecular chaperone GroEL. *J Biol Chem* 267:6796–6800
79. Gray TE, Fersht AR (1991) Cooperativity in ATP hydrolysis by GroEL is increased by GroES. *FEBS Lett* 292:254–258
80. Yifrach O, Horovitz A (1995) Nested cooperativity in the ATPase activity of the oligomeric chaperonin GroEL. *Biochemistry* 34:5303–5308
81. Chaudhuri TK, Verma VK, Maheshwari A (2009) GroEL assisted folding of large polypeptide substrates in *Escherichia coli* : Present scenario and assignments for the future. *Prog Biophys Mol Biol* 99:42–50
82. Kanno R, Koike-Takeshita A, Yokoyama K, Taguchi H, Mitsuoka K (2009) Cryo-EM structure of the native GroEL–GroES complex from thermus thermophilus encapsulating substrate inside the cavity. *Structure* 17:287–293
83. Nojima T, Murayama S, Yoshida M, Motojima F (2008) Determination of the number of active GroES subunits in the fused heptamer GroES required for interactions with GroEL. *J Biol Chem* 283:18385–18392
84. Inbar E, Horovitz A (1997) GroES promotes the T to R transition of the GroEL ring distal to GroES in the GroEL–GroES complex. *Biochemistry* 36:12276–12281
85. Farr GW, Fenton WA, Chaudhuri TK, Clare DK, Saibil HR, Horwich AL (2003) Folding with and without encapsulation by *cis*- and *trans*-only GroEL–GroES complexes. *EMBO J* 22:3220–3230
86. Chaudhuri TK, Farr GW, Fenton WA, Rospert S, Horwich AL (2001) GroEL/GroES-mediated folding of a protein too large to be encapsulated. *Cell* 107:235–246
87. Anfinsen CB (1973) Principles that govern the folding of protein chains. *Science* 181:223–230
88. Ellis RJ (1996) Revisiting the Anfinsen cage. *Fold Des* 1:R9–R15
89. Xu Z, Horwich AL, Sigler PB (1997) The crystal structure of the asymmetric GroEL–GroES-(ADP)<sub>7</sub> chaperonin complex. *Nature* 388:741–750
90. Lin Z, Madan D, Rye HS (2008) GroEL stimulates protein folding through forced unfolding. *Nat Struct Mol Biol* 15:303–311
91. Taguchi H (2005) Chaperonin GroEL meets the substrate protein as a “load” of the rings. *J Biochem* 137:543–549
92. Wang JD, Weissman JS (1999) Thinking outside the box: new insights into the mechanism of GroEL-mediated protein folding. *Nat Struct Biol* 6:597–600
93. Chennubhotla C, Yang Z, Bahar I (2008) Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL. *Mol BioSyst* 4:287–292
94. Hyeon C, Lorimer GH, Thirumalai D (2006) Dynamics of allosteric transitions in GroEL. *Proc Natl Acad Sci USA* 103:18939–18944
95. Tehver R, Chen J, Thirumalai D (2009) Allostery wiring diagrams in the transitions that drive the GroEL reaction cycle. *J Mol Biol* 387:390–406
96. Tehver R, Thirumalai D (2008) Kinetic model for the coupling between allosteric transitions in GroEL and substrate protein folding and aggregation. *J Mol Biol* 377:1279–1295
97. Paul S, Singh C, Mishra S, Chaudhuri TK (2007) The 69 kDa *Escherichia coli* maltodextrin glucosidase does not get encapsulated underneath GroES and folds through *trans* mechanism during GroEL/GroES-assisted folding. *FASEB J* 21:2874–2885
98. Jones H, Preuss M, Wright M, Miller AD (2006) The mechanism of GroEL/GroES folding/refolding of protein substrates revisited. *Org Biomol Chem* 4:1223–1235
99. Ying BW, Taguchi H, Ueda T (2006) Co-translational binding of GroEL to nascent polypeptides is followed by post-translational encapsulation by GroES to mediate protein folding. *J Biol Chem* 281:21813–21819
100. West MW, Wang W, Patterson J, Mancias JD, Beasley JR, Hecht MH (1999) De novo amyloid proteins from designed combinatorial libraries. *Proc Natl Acad Sci USA* 96:11211–11216
101. Arie JP, Miot M, Sassoon N, Betton JM (2006) Formation of active inclusion bodies in the periplasm of *Escherichia coli*. *Mol Microbiol* 62:427–437
102. Carrio MM, Corchero JL, Villaverde A (1998) Dynamics of in vivo protein aggregation: building inclusion bodies in recombinant bacteria. *FEMS Microbiol Lett* 169:9–15
103. Gonzalez-Montalban N, Garcia-Fruitos E, Ventura S, Aris A, Villaverde A (2006) The chaperone DnaK controls the fractioning of functional protein between soluble and insoluble cell fractions in inclusion body-forming cells. *Microb Cell Fact* 5:26
104. Bowden GA, Paredes AM, Georgiou G (1991) Structure and morphology of protein inclusion bodies in *Escherichia coli*. *Biotechnology (N Y)* 9:725–730

105. Carrio MM, Cubarsi R, Villaverde A (2000) Fine architecture of bacterial inclusion bodies. *FEBS Lett* 471:7–11
106. Carrio MM, Villaverde A (2002) Construction and deconstruction of bacterial inclusion bodies. *J Biotechnol* 96:3–12
107. Schrodel A, de Marco A (2005) Characterization of the aggregates formed during recombinant protein expression in bacteria. *BMC Biochem* 6:10
108. Rinas U, Hoffmann F, Betiku E, Estape D, Marten S (2007) Inclusion body anatomy and functioning of chaperone-mediated in vivo inclusion body disassembly during high-level recombinant protein production in *Escherichia coli*. *J Biotechnol* 127:244–257
109. Carrio MM, Villaverde A (2005) Localization of chaperones DnaK and GroEL in bacterial inclusion bodies. *J Bacteriol* 187:3599–3601
110. Wilson MR, Yerbury JJ, Poon S (2008) Potential roles of abundant extracellular chaperones in the control of amyloid formation and toxicity. *Mol BioSyst* 4:42–52
111. Lybarger SR, Maddock JR (2001) Polarity in action: asymmetric protein localization in bacteria. *J Bacteriol* 183:3261–3267
112. Shapiro L, McAdams HH, Losick R (2002) Generating and exploiting polarity in bacteria. *Science* 298:1942–1946
113. Dougan DA, Mogk A, Bukau B (2002) Protein folding and degradation in bacteria: to degrade or not to degrade? That is the question. *Cell Mol Life Sci* 59:1607–1616
114. Schlieker C, Tews I, Bukau B, Mogk A (2004) Solubilization of aggregated proteins by ClpB/DnaK relies on the continuous extraction of unfolded polypeptides. *FEBS Lett* 578:351–356
115. Ben-Zvi A, De Los Rios P, Dietler G, Goloubinoff P (2004) Active solubilization and refolding of stable protein aggregates by cooperative unfolding action of individual hsp70 chaperones. *J Biol Chem* 279:37298–37303
116. De Los Rios P, Ben-Zvi A, Slutsky O, Azem A, Goloubinoff P (2006) Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proc Natl Acad Sci USA* 103:6166–6171
117. Gonzalez-Montalban N, Carrio MM, Cuatrecasas S, Aris A, Villaverde A (2005) Bacterial inclusion bodies are cytotoxic in vivo in absence of functional chaperones DnaK or GroEL. *J Biotechnol* 118:406–412
118. Carrio MM, Villaverde A (2003) Role of molecular chaperones in inclusion body formation. *FEBS Lett* 537:215–221
119. Werbeck ND, Kellner JN, Barends TR, Reinstein J (2009) Nucleotide binding and allosteric modulation of the second AAA+ domain of ClpB probed by transient kinetic studies. *Biochemistry* 48:7240–7250
120. Werbeck ND, Schlee S, Reinstein J (2008) Coupling and dynamics of subunits in the hexameric AAA+ chaperone ClpB. *J Mol Biol* 378:178–190
121. Acebron SP, Martin I, del Castillo U, Moro F, Muga A (2009) DnaK-mediated association of ClpB to protein aggregates. A chaperone network at the aggregate surface. *FEBS Lett* 583:2991–2996
122. Lewandowska A, Matuszewska M, Liberek K (2007) Conformational properties of aggregated polypeptides determine ClpB-dependence in the disaggregation process. *J Mol Biol* 371:800–811
123. Thomas JG, Baneyx F (1996) Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat-shock proteins. *J Biol Chem* 271:11141–11147
124. Watanabe YH, Nakazaki Y, Suno R, Yoshida M (2009) Stability of the two wings of the coiled-coil domain of ClpB chaperone is critical for its disaggregation activity. *Biochem J* 421:71–77
125. Diemand AV, Lupas AN (2006) Modeling AAA+ ring complexes from monomeric structures. *J Struct Biol* 156:230–243
126. Doyle SM, Wickner S (2009) Hsp104 and ClpB: protein disaggregating machines. *Trends Biochem Sci* 34(1):40–48
127. Mogk A, Bukau B (2004) Molecular chaperones: structure of a protein disaggregase. *Curr Biol* 14:R78–R80
128. Mogk A, Deuerling E, Vorderwulbecke S, Vierling E, Bukau B (2003) Small heat shock proteins, ClpB and the DnaK system form a functional triade in reversing protein aggregation. *Mol Microbiol* 50:585–595
129. Mogk A, Schlieker C, Friedrich KL, Schonfeld HJ, Vierling E, Bukau B (2003) Refolding of substrates bound to small Hsps relies on a disaggregation reaction mediated most efficiently by ClpB/DnaK. *J Biol Chem* 278:31033–31042
130. Mogk A, Schlieker C, Strub C, Rist W, Weibezahn J, Bukau B (2003) Roles of individual domains and conserved motifs of the AAA+ chaperone ClpB in oligomerization, ATP hydrolysis, and chaperone activity. *J Biol Chem* 278:17615–17624
131. Weibezahn J, Bukau B, Mogk A (2004) Unscrambling an egg: protein disaggregation by AAA+ proteins. *Microb Cell Fact* 3:1
132. Carrio MM, Corchero JL, Villaverde A (1999) Proteolytic digestion of bacterial inclusion body proteins during dynamic transition between soluble and insoluble forms. *Biochim Biophys Acta* 1434:170–176
133. Carrio MM, Villaverde A (2001) Protein aggregation as bacterial inclusion bodies is reversible. *FEBS Lett* 489:29–33
134. Cubarsi R, Carrio MM, Villaverde A (2005) A mathematical approach to molecular organization and proteolytic disintegration of bacterial inclusion bodies. *Math Med Biol* 22:209–226
135. de Groot NS, Espargaro A, Morell M, Ventura S (2008) Studies on bacterial inclusion bodies. *Future Microbiol* 3:423–435
136. Villaverde A, Carrio MM (2003) Protein aggregation in recombinant bacteria: biological role of inclusion bodies. *Biotechnol Lett* 25:1385–1395
137. Chen Y, Song J, Sui SF, Wang DN (2003) DnaK and DnaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA overexpressed in *Escherichia coli*. *Protein Expr Purif* 32:221–231
138. Zhang Z, Li ZH, Wang F, Fang M, Yin CC, Zhou ZY, Lin Q, Huang HL (2002) Overexpression of DsbC and DsbG markedly improves soluble and functional expression of single-chain Fv antibodies in *Escherichia coli*. *Protein Expr Purif* 26:218–228
139. Garcia-Fruitos E, Martinez-Alonso M, Gonzalez-Montalban N, Valli M, Mattanovich D, Villaverde A (2007) Divergent genetic control of protein solubility and conformational quality in *Escherichia coli*. *J Mol Biol* 374:195–205
140. Corchero JL, Cubarsi R, Enfors S, Villaverde A (1997) Limited in vivo proteolysis of aggregated proteins. *Biochem Biophys Res Commun* 237:325–330
141. Vera A, Aris A, Carrio M, Gonzalez-Montalban N, Villaverde A (2005) Lon and ClpP proteases participate in the physiological disintegration of bacterial inclusion bodies. *J Biotechnol* 119:163–171
142. Martinez-Alonso M, Garcia-Fruitos E, Villaverde A (2008) Yield, solubility and conformational quality of soluble proteins are not simultaneously favored in recombinant *Escherichia coli*. *Biotechnol Bioeng* 101:1353–1358
143. Martinez-Alonso M, Gonzalez-Montalban N, Garcia-Fruitos E, Villaverde A (2009) Learning about protein solubility from bacterial inclusion bodies. *Microb Cell Fact* 8:4
144. Garcia-Fruitos E, Gonzalez-Montalban N, Morell M, Vera A, Ferraz RM, Aris A, Ventura S, Villaverde A (2005) Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins. *Microb Cell Fact* 4:27
145. Schlieker C, Weibezahn J, Patzelt H, Tessarz P, Strub C, Zeth K, Erbse A, Schneider-Mergener J, Chin JW, Schultz PG, Bukau B, Mogk A (2004) Substrate recognition by the AAA+ chaperone ClpB. *Nat Struct Mol Biol* 11:607–615

146. Garcia-Fruitos E, Aris A, Villaverde A (2007) Localization of functional polypeptides in bacterial inclusion bodies. *Appl Environ Microbiol* 73:289–294
147. Speed MA, Wang DI, King J (1996) Specific aggregation of partially folded polypeptide chains: the molecular basis of inclusion body composition. *Nat Biotechnol* 14:1283–1287
148. Carrio M, Gonzalez-Montalban N, Vera A, Villaverde A, Ventura S (2005) Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol* 347:1025–1037
149. Hart RA, Rinas U, Bailey JE (1990) Protein composition of *Vitreoscilla* hemoglobin inclusion bodies produced in *Escherichia coli*. *J Biol Chem* 265:12728–12733
150. Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 424:805–808
151. de Groot NS, Aviles FX, Vendrell J, Ventura S (2006) Mutagenesis of the central hydrophobic cluster in Aβ42 Alzheimer's peptide. Side-chain properties correlate with aggregation propensities. *FEBS J* 273:658–668
152. Rousseau F, Serrano L, Schymkowitz JW (2006) How evolutionary pressure against protein aggregation shaped chaperone specificity. *J Mol Biol* 355:1037–1047
153. Ventura S, Lacroix E, Serrano L (2002) Insights into the origin of the tendency of the PI3-SH3 domain to form amyloid fibrils. *J Mol Biol* 322:1147–1158
154. Conchillo-Sole O, de Groot NS, Aviles FX, Vendrell J, Daura X, Ventura S (2007) AGGRESCAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics* 8:65
155. Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22:1302–1306
156. Sanchez de Groot N, Pallares I, Aviles FX, Vendrell J, Ventura S (2005) Prediction of “hot spots” of aggregation in disease-linked polypeptides. *BMC Struct Biol* 5:18
157. Thompson MJ, Sievers SA, Karanicolas J, Ivanova MI, Baker D, Eisenberg D (2006) The 3D profile method for identifying fibril-forming segments of proteins. *Proc Natl Acad Sci USA* 103:4074–4078
158. Monsellier E, Ramazzotti M, Taddei N, Chiti F (2008) Aggregation propensity of the human proteome. *PLoS Comput Biol* 4:e1000199
159. Castillo V, Ventura S (2009) Amyloidogenic regions and interaction surfaces overlap in globular proteins related to conformational diseases. *PLoS Comput Biol* 5:e1000476
160. Vendruscolo M, Paci E, Karplus M, Dobson CM (2003) Structures and relative free energies of partially folded states of proteins. *Proc Natl Acad Sci USA* 100:14817–14821
161. Chiti F, Dobson CM (2009) Amyloid formation by globular proteins under native conditions. *Nat Chem Biol* 5:15–22
162. Espargaro A, Castillo V, de Groot NS, Ventura S (2008) The in vivo and in vitro aggregation properties of globular proteins correlate with their conformational stability: the SH3 case. *J Mol Biol* 378:1116–1131
163. Shorter J, Lindquist S (2008) Hsp104, Hsp70 and Hsp40 interplay regulates formation, growth and elimination of Sup35 prions. *EMBO J* 27:2712–2724
164. Wang L (2009) Towards revealing the structure of bacterial inclusion bodies. *Prion* 3:139–145
165. Wasmer C, Benkemoun L, Sabate R, Steinmetz MO, Coulary-Salin B, Wang L, Riek R, Saupé SJ, Meier BH (2009) Solid-state NMR spectroscopy reveals that *E. coli* inclusion bodies of HET-s(218–289) are amyloids. *Angew Chem Int Ed Engl* 48:4858–4860
166. Sabate R, Espargaro A, Saupé SJ, Ventura S (2009) Characterization of the amyloid bacterial inclusion bodies of the HET-s fungal prion. *Microb Cell Fact* 8:56
167. Hammer ND, Wang X, McGuffie BA, Chapman MR (2008) Amyloids: friend or foe? *J Alzheimers Dis* 13:407–419
168. Barnhart MM, Chapman MR (2006) Curli biogenesis and function. *Annu Rev Microbiol* 60:131–147
169. Fowler DM, Koulov AV, Balch WE, Kelly JW (2007) Functional amyloid—from bacteria to humans. *Trends Biochem Sci* 32:217–224
170. Gebbink MF, Claessen D, Bouma B, Dijkhuizen L, Wosten HA (2005) Amyloids—a functional coat for microorganisms. *Nat Rev Microbiol* 3:333–341
171. Otzen D, Nielsen PH (2008) We find them here, we find them there: functional bacterial amyloid. *Cell Mol Life Sci* 65:910–927
172. Austin JW, Sanders G, Kay WW, Collinson SK (1998) Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. *FEMS Microbiol Lett* 162:295–301
173. Vidal O, Longin R, Prigent-Combaret C, Dorel C, Hooreman M, Lejeune P (1998) Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. *J Bacteriol* 180:2442–2449
174. Olsen A, Arnqvist A, Hammar M, Sukupolvi S, Normark S (1993) The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Mol Microbiol* 7:523–536
175. Loferer H, Hammar M, Normark S (1997) Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG. *Mol Microbiol* 26:11–23
176. Nennering AA, Robinson LS, Hultgren SJ (2009) Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. *Proc Natl Acad Sci USA* 106:900–905
177. Kanamaru S, Kurazono H, Terai A, Monden K, Kumon H, Mizunoe Y, Ogawa O, Yamamoto S (2006) Increased biofilm formation in *Escherichia coli* isolated from acute prostatitis. *Int J Antimicrob Agents* 28(Suppl 1):S21–S25
178. Claessen D, Rink R, de Jong W, Siebring J, de Vreugd P, Boldersma FG, Dijkhuizen L, Wosten HA (2003) A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils. *Genes Dev* 17:1714–1726
179. Elliot MA, Karoonuthaisiri N, Huang J, Bibb MJ, Cohen SN, Kao CM, Buttner MJ (2003) The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. *Genes Dev* 17:1727–1740
180. Sakellaris H, Hannink NK, Rajakumar K, Bulach D, Hunt M, Sasakawa C, Adler B (2000) Curli loci of *Shigella* spp. *Infect Immun* 68:3780–3783
181. Claessen D, Stokroos I, Deelstra HJ, Penninga NA, Bormann C, Salas JA, Dijkhuizen L, Wosten HA (2004) The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlets and chaplins. *Mol Microbiol* 53:433–443
182. Oh J, Kim JG, Jeon E, Yoo CH, Moon JS, Rhee S, Hwang I (2007) Amyloidogenesis of type III-dependent harpins from plant pathogenic bacteria. *J Biol Chem* 282:13601–13609
183. Destoumieux-Garzon D, Thomas X, Santamaria M, Goulard C, Barthelemy M, Boscher B, Bessin Y, Molle G, Pons AM, Letellier L, Peduzzi J, Rebuffat S (2003) Microcin E492 antibacterial activity: evidence for a TonB-dependent inner membrane permeabilization on *Escherichia coli*. *Mol Microbiol* 49:1031–1041

184. Bieler S, Estrada L, Lagos R, Baeza M, Castilla J, Soto C (2005) Amyloid formation modulates the biological activity of a bacterial protein. *J Biol Chem* 280:26880–26885
185. Dang TX, Hotze EM, Rouiller I, Tweten RK, Wilson-Kubalek EM (2005) Prepore to pore transition of a cholesterol-dependent cytolysin visualized by electron microscopy. *J Struct Biol* 150:100–108
186. Lashuel HA, Hartley D, Petre BM, Walz T, Lansbury PT Jr (2002) Neurodegenerative disease: amyloid pores from pathogenic mutations. *Nature* 418:291
187. Plomp M, Leighton TJ, Wheeler KE, Hill HD, Malkin AJ (2007) In vitro high-resolution structural dynamics of single germinating bacterial spores. *Proc Natl Acad Sci USA* 104:9644–9649
188. Alteri CJ, Xicohtencatl-Cortes J, Hess S, Caballero-Olin G, Giron JA, Friedman RL (2007) *Mycobacterium tuberculosis* produces pili during human infection. *Proc Natl Acad Sci USA* 104:5145–5150
189. Lundmark K, Westermark GT, Olsen A, Westermark P (2005) Protein fibrils in nature can enhance amyloid protein A amyloidosis in mice: cross-seeding as a disease mechanism. *Proc Natl Acad Sci USA* 102:6098–6102
190. Broxmeyer L (2002) Parkinson's: another look. *Med Hypotheses* 59:373–377
191. Diaz-Corrales FJ, Colasante C, Contreras Q, Puig M, Serrano JA, Hernandez L, Beaman BL (2004) *Nocardia otitidiscaviarum* (GAM-5) induces Parkinsonian-like alterations in mouse. *Braz J Med Biol Res* 37:539–548
192. Miklossy J, Kis A, Radenovic A, Miller L, Forro L, Martins R, Reiss K, Darbinian N, Darekar P, Mihaly L, Khalili K (2006) Beta-amyloid deposition and Alzheimer's type changes induced by *Borrelia spirochetes*. *Neurobiol Aging* 27:228–236
193. Fowler DM, Koulov AV, Alory-Jost C, Marks MS, Balch WE, Kelly JW (2006) Functional amyloid formation within mammalian tissue. *PLoS Biol* 4:e6
194. Maji SK, Perrin MH, Sawaya MR, Jessberger S, Vadodaria K, Rissman RA, Singru PS, Nilsson KP, Simon R, Schubert D, Eisenberg D, Rivier J, Sawchenko P, Vale W, Riek R (2009) Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 325:328–332
195. de Groot NS, Ventura S (2010) Protein aggregation profile of the bacterial cytosol. *PLoS One* 5(2):e9383