BIOCHEMISTRY
including biophysical chemistry & molecular biology

Current Topic

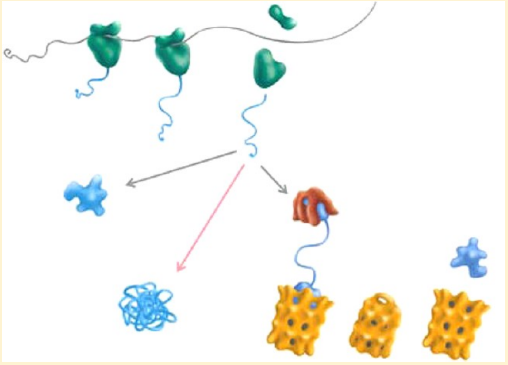
pubs.acs.org/biochemistry

Cumulative Impact of Chaperone-Mediated Folding on Genome Evolution

David Bogumil and Tal Dagan*

Institute for Genomic Microbiology, Heinrich-Heine University of Düsseldorf, Düsseldorf, Germany

ABSTRACT: Molecular chaperones support protein folding and unfolding along with assembly and translocation of protein complexes. Chaperones have been recognized as important mediators between an organismal genotype and phenotype as well as important maintainers of cellular fitness under environmental conditions that induce high mutational loads. Here we review recent studies revealing that the folding assistance supplied by chaperones is evident in genomic sequences implicating chaperone-mediated folding as an influential factor during protein evolution. Interaction of protein with chaperones ensures a proper folding and function, yet an adaptation to obligatory dependence on such assistance may be irreversible, representing an evolutionary trap. A correlation between the requirement for a chaperone and protein expression level indicates that the evolution of substrate−chaperone interaction is bounded by the required substrate abundance within the cell. Accumulating evidence suggets that the utility of chaperones is governed by a delicate balance between their help in mitigating the risks of protein misfolding and aggregate formation on one hand and the slower rate of protein maturation and the energetic cost of chaperone synthesis on the other.



Molecular chaperones were first described as proteins that assist in the assembly of other proteins into their functional conformation.[1,2] Besides the assembly of protein complexes and de novo folding of nascent polypeptides, chaperones play a role in protein translocation across membranes,[3] stabilization of protein−protein interactions,[4,5] and ribosome biogenesis,[6] but regardless of their exact function, different chaperones provide assistance in the same assignment: proteins have to maintain their designated function in the right place at the right time.

CHAPERONE-MEDIATED PROTEIN FOLDING IN THREE DOMAINS OF LIFE

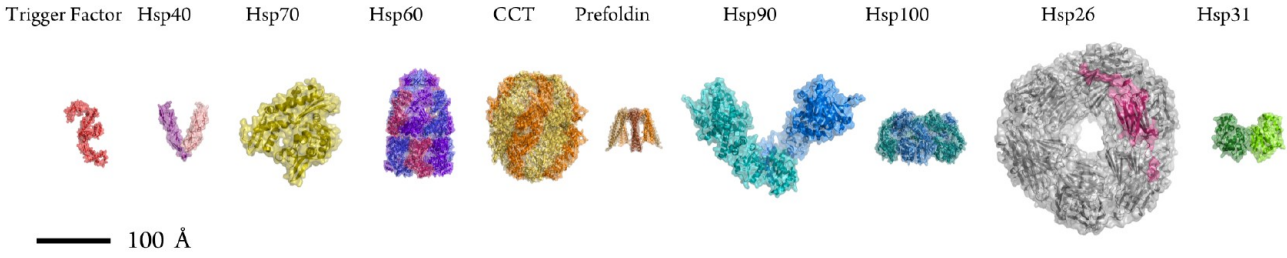
Species in the three domains of life (eubacteria, archaebacteria, and eukaryotes) utilize slightly different chaperones that assemble into diverse protein folding pathways. The major chaperone families in eubacteria are Trigger Factor (TF), DnaJ (Hsp40), DnaK (Hsp70), GrpE (nucleotide exchange factor), and GroEL/GroES (Hsp60/Hsp10). Trigger Factor is the first chaperone that binds to the nascent polypeptide chain emerging from the ribosome, and its function is to shield hydrophobic (especially aromatic) stretches of the translated protein to keep it soluble[7] (Figure 1). Members of the DnaK and DnaJ chaperone families assist in protein folding by forming a complex with their substrate proteins. The substrate binding specificity of the ATPase-like DnaK chaperone is determined by the DnaJ cochaperone.[48,49] Experimental data show that DnaJ in *Escherichia coli* binds to hydrophobic protein surfaces and initiates the functional cycle of the DnaK system by targeting the DnaK to hydrophobic patches within the substrate.[50] DnaK then stabilizes the intermediate conformational state of the substrate using ATP. The nucleotide exchange factor GrpE is involved in binding and release of ATP and ADP. Chaperonin systems comprise barrel-like structures that assist in protein folding by providing an isolated environment for the protein to fold.[51] GroE is a eubacterial chaperonin complex composed of two proteins: GroEL, a barrel-like structure consisting of two heptameric rings (Figure 1), and GroES, also a heptameric ring, which functions as a lid for the GroEL barrel.

Archaeal species utilize chaperones of the Hsp70,[52] Hsp40,[53] GrpE,[54] and TriC/CCT[55] families. Interestingly, almost all thermophilic species lack DnaK and its cochaperone, DnaJ.[56] It is unclear if archaea rely on proteins other than DnaK and DnaJ to remove the cellular debris caused by heat shock or if they rely on proteases instead.[56] The existence of a nascent chain-associated complex in archaebacteria has been experimentally confirmed, and it was found to be associated with a ribosome-like eukaryotic NAC homologue.[57] The archaeal chaperonin system is termed thermosomes. It forms an octameric double-ring structure with an apical loop instead of a capping cofactor like GroES.[58] *Methanosarcina mazei* is an exception among archaea as it also encodes homologues of eubacterial GroEL and HtpG (Hsp90), which were acquired via a horizontal gene transfer.[59] The substrate sets of the two chaperonin systems in *M. mazei* largely overlap; however, the GroE substrates are biased toward proteins with complex α/β domains, while the substrates of the thermosome include a wider range of different domain folds.[60] On the other hand, several eubacterial species, including clostridial and cyanobacterial representatives, encode a CCT-like chaperonine.[61] This chaperonine forms a structure

Received: October 5, 2012
Revised: November 19, 2012
Published: November 20, 2012

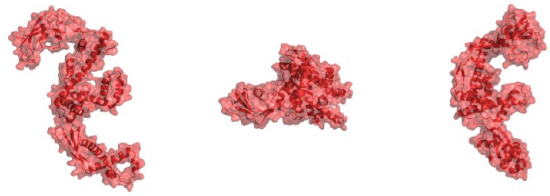
   dx.doi.org/10.1021/bi3013643 | *Biochemistry* 2012, 51, 9941−9953

a)



b)

Trigger Factor

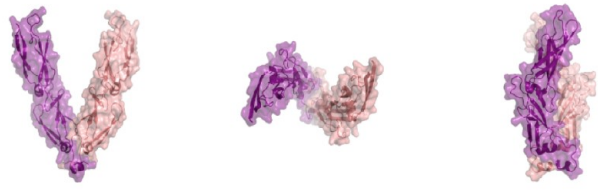


☒ Eubacteria
☐ Archaea
☐ Eukaryotes

Monomer
97 kDa
PDB: 1W26

The eubacterial trigger factor (TF) is associated to nascent polypeptides emerging from the ribosome (8). It projects the extended domains over the exit of the ribosomal tunnel, creating a protected folding space where nascent polypeptides may be shielded from proteases and aggregation (9).

Hsp40 (synonym: DnaJ)

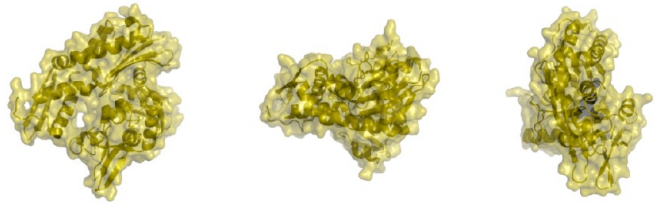


☒ Eubacteria
☒ Archaea
☒ Eukaryotes

Dimer
2 x 19 kDa
PDB: 1C3G

Hsp40 is a U-shaped Dimer (10). It functions as a Co-chaperone that stimulates the ATPase activity of the HSP70 chaperone (11). It is involved in protein translocation and the proteolysis of misfolded proteins (12-14).

Hsp70 (synonym: DnaK)



☒ Eubacteria
☒ Archaea
☒ Eukaryotes

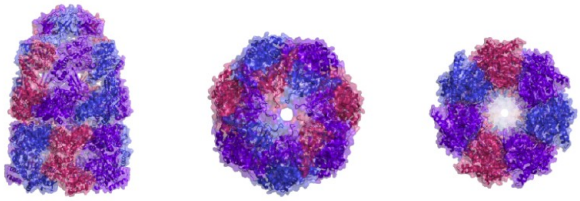
Monomer
78 kDa
PDB: 3QFP

Hsp70 is a cytoplasmic ATPase (15). It is a ribosome-associated molecular chaperone that is involved in folding of newly made polypeptide chains (16). It functions with a J-protein (Hsp40) partner in the ribosome-associated complex (RAC), which is involved in translocation of proteins into mitochondria as well (17).

Figure 1. continued

dx.doi.org/10.1021/bi3013643 | *Biochemistry* 2012, 51, 9941−9953

Hsp60 (synonyms: GroEL, GroE, Cpn60, mtHsp60)

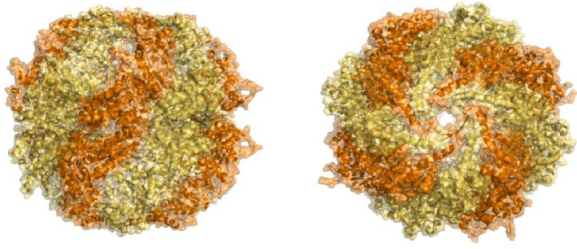


☒ Eubacteria
☒ Archaea
☒ Eukaryotes

Tetradecamer
14 x 57 kDa
PDB: 1AON

GroEL is a eubacterial chaperonin that prevents aggregation after heat shock and is required for ATP-dependent folding of precursor polypeptides and complex assembly (18). It has a barrel like structure consisting of 2 heptameric rings composed of 14 identical subunits (19). In Eukaryotes, Hsp60 is localized in the mitochondria and is involved in both protein import (20) and mtDNA transmission (21). GroEL/Hsp60 functions with a capping cofactor termed GroES/Hsp10, which is a ring consisting of seven identical subunits (22) and encapsulates the substrate inside the GroEL molecule (23).

CCT (Synonyms: TriC, Tcp-1, Thermosome)

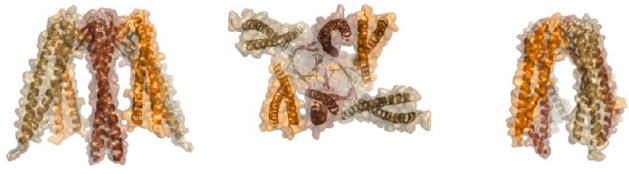


☐ Eubacteria
☒ Archaea
☒ Eukaryotes

Hexadecamer
16 x 57..64 kDa
PDB: 3P9D

CCT is the archaeal chaperonin type and is present also in the eukaryotic cytosol. It consists of 2 octameric rings, with 8 or 9 subunits encoded by different genes (24). Unlike GroEL, CCT functions without a cofactor, using a loop in the apical domain to encapsulate substrate proteins (25). CCT is involved in the folding of actin (26) and tubulin (27).

Prefoldin (Synonym: PFD)

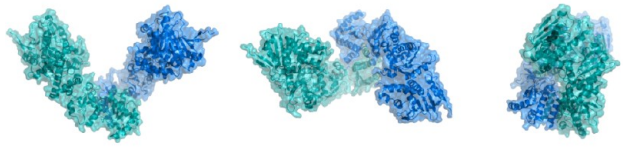


☐ Eubacteria
☒ Archaea
☒ Eukaryotes

Heterohexamer
6 x 14..23 kDa
PDB: 1FXK

Prefoldin (PFD) is a heterohexameric chaperone. The archaeal homolog is composed of two subunits, while the eukaryotic PFD is composed of six subunits (28). Prefoldin is present in eukaryotes and archaea, where it binds specifically to cytosolic chaperonin (CCT) and transfers target proteins to it (28).

Hsp90 (synonym: HtpG)



☒ Eubacteria
☐ Archaea
☒ Eukaryotes

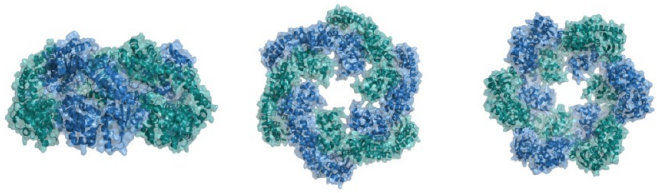
Dimer
2 x 71 kDa (E. coli)
2 x 82 kDa (Yeast)
PDB: 2IOQ

Hsp90 is a dimer that binds its substrates like a molecular clamp (29). It is present in eukaryotes and eubacteria where it is termed htpG (30). Hsp90 is essential in yeast (31). Besides protein folding, it is required for pheromone signaling (32) and preprotein delivery to Tom70p and subsequent translocation into mitochondria (33). Hsp90 also promotes telomerase DNA binding and nucleotide addition (34).

Figure 1. continued

dx.doi.org/10.1021/bi3013643 | *Biochemistry* 2012, 51, 9941−9953

Hsp100 (synonym: Clp)

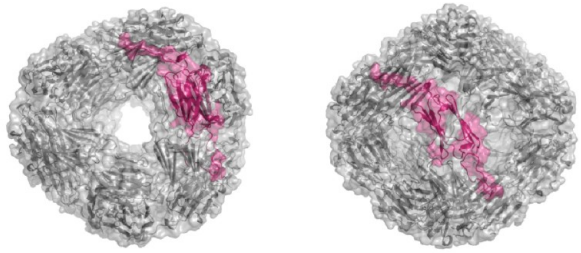


☒ Eubacteria
☐ Archaea
☒ Eukaryotes

Hexamer
6 x 96 kDa
PDB: 3PXG

Hsp100 is a heat shock protein that refolds and reactivates previously denatured aggregated proteins (35). It cooperates with Ydj1p (Hsp40) and Ssa1p (Hsp70) (36). Homologs of Hsp100 are present in eubacteria, eukaryotes and Mitochondria (37-39). Structurally it is a two-tiered hexamer (40).

Hsp26 (synonym: Hsp16.5)

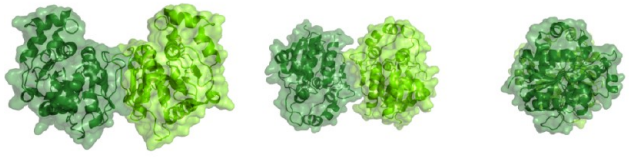


☐ Eubacteria
☒ Archaea
☒ Eukaryotes

24mer
24 x 16.5 kDa (Archaea)
24 x 23.7 kDa (Yeast)
PDB: 1SHS

Hsp26 forms hollow oligomers that suppress unfolded protein aggregation. The oligomer activation requires heat induced conformational change. Hsp26 oligomers dissociate into dimers under heat stress. Each dimer binds a substrate monomer. After the substrate binding the dimers are newly assembled forming an Hsp-substrate complex. Hsp26 also has mRNA binding activity. The archaeal homolog is hsp16.5 (41).

Hsp31



☒ Eubacteria
☒ Archaea
☒ Eukaryotes

Dimer
2 x 26.5 kDa
PDB: 1R7W

Hsp31 is a dimeric chaperone and cysteine protease in eukaryotes and bacteria (42-44). It is a member of the DJ-1/ThiJ/PfpI superfamily (45), which includes the human DJ-1 protein that is involved in Parkinson's disease (46). The archaeal homolog is PfpI (47).

Figure 1. Structural properties of chaperones. (a) Relative complex sizes of chaperones. (b) Chaperone structures are shown in an upright orientation (left), with a 90° rotation along the *X*-axis (middle), and with a 90° rotation along the *Y*-axis (right). Complexes having a radial symmetry are rotated by 90° along the *X*-axis to show the top (middle) and bottom (right) of the molecule. Chaperone plots were generated using the PyMOL Molecular Graphics System, version 1.5.0.4 (Schrödinger, LLC).

that is similar to that of the archaeal CCT, and it is thought to be acquired by an ancient horizontal gene transfer during Firmicute evolution.[62] A survey for chaperones in archaeal genome sequences led to the interesting finding that Hsp90 and Hsp100 are absent from nearly all species.[56]

The eukaryotic chaperone repertoire reflects the hybrid origin of the eukaryotic cell. Eukaryotes evolved from endosymbiosis of a eubacterium engulfed within an archaebecterial host.[63−65] The eubacterium gave rise to the mitochondrion organelle, and the holobiont became what we know today as eukaryotic cells.[66] Subsequently, most of the endosymbiont genomic material was either lost or transferred to the host nucleus. Plastids of photosynthetic eukaryotes originated via a similar evolutionary event in which a cyanobacterial endosymbiont was engulfed within a eukaryotic host (see ref 67 for a review). Eukaryotic proteomes are thus mosaics of archaebacterial and eubacterial homologues representing the contribution of the host and organellar ancestors to eukaryotic evolution.[68] Chaperones comprising the eukaryotic protein folding pathway are no exception to that rule, and they include homologues of both eubacterial and archaeal chaperones.[58,69] The ribosome-associated complex (RAC) is the first chaperone complex that interacts with most newly synthesized proteins in *Saccharomyces cerevisiae*.[17] It consists of the Hsp40 chaperone Zuo1 and the Hsp70 partner Ssz1.[70−72] Further folding of completely translated peptides can be assisted by other Hsp40−Hsp70 complexes, as well as the Hsp90 system and TriC/CCT class chaperones and their prefoldin cochaperones.[69] Prefoldin operates mainly on cytoskeleton-associated substrate proteins and assists in their targeting to TriC/CCT for folding.[28] The TriC/CCT is a chaperonin system consisting of two rings, but in a manner different from that of GroE, each ring is formed by eight subunits.[73] This hexadecameric barrel structure is the same as in the archaeal thermosomes, comprising the group of type II chaperonins as opposed to the group I chaperonins GroEL in eubacteria or the mitochondrial Hsp60 (see ref 74 for a review). The eukaryotic TriC/CCT consist of eight different subunits,[73] whereas the archaeal thermosome is composed of only two types, the

dx.doi.org/10.1021/bi3013643 | *Biochemistry* 2012, 51, 9941−9953

α- and β-type subunits.[75] Another difference between TriC/CCT and GroE is that CCT is found to interact with nascent peptide chains more frequently[76] and does not utilize a capping cofactor for the ATP-dependent, GroE-like enclosure and folding process.[75] Nevertheless, both chaperonin types are thought to share similar substrate recognition: GroE- and TriC-type chaperonins have a substrate overlap of 80% when presented with denatured protein extract of human fibroblasts.[77] The eukaryotic organelles, mitochondria and chloroplasts, utilize Hsp60 and Hsp10 chaperones. These are homologues of the eubacterial GroEL and GroES chaperones.[78] The eukaryotic Hsp60 has a "double-doughnut" structure similar to that of GroEL, and its level of expression in the mitochondrion is increased under heat stress conditions.[20] The Hsp60 chaperone interacts with Hsp10 which serves as a capping protein similar to eubacterial GroES.[79]

CHAPERONE-MEDIATED PROTEIN UNFOLDING

Molecular chaperones are functional also in unfolding and refolding of previously misfolded proteins.[36] Protein synthesis is energetically the most expensive process within living cells. In bacteria, it has been estimated that ∼60% of the ATP molecules required for the formation of a whole cell are consumed by protein translation.[80] A recent study that aimed to quantify gene expression control in mammalian cells suggested that protein synthesis consumes ∼90% of the energy that is needed to maintain the cellular protein levels and determined that protein translation is the limiting factor in protein production.[81] Proteins that fail to fold into their native (functional) state represent an energetic burden of wasted "translation energy". The ability to unfold previously misfolded proteins and reinsert them into the folding pathway is an important process considering the energetic balance of protein production. Protein unfolding and refolding compensate for the fitness cost impaired by the toxicity of protein aggregates in the cell, but maybe even more importantly, the refolding of misfolded proteins means that the energy invested in the synthesis of a misfolded protein will not be wasted. For example, an in vitro measurement of the energetic investment in unfolding a luciferase protein by a DnaK−DnaJ−GrpE complex into its intermediate state revealed that only five ATP molecules are required in this process.[82] Achieving the same outcome by hydrolysis and resynthesis of the hydrogen bonds in luciferase (550 residues, Swiss-Prot entry P08659.1[83]) is estimated to require ∼3000 ATP molecules; hence, the rescue of this protein by the chaperones is 3 orders of magnitude energetically cheaper than its recycling.[82]

CHAPERONE-MEDIATED PROTEIN TRANSLOCATION

In addition to providing cotranslational folding mechanisms, molecular chaperones are also involved in the translocation of protein across membranes, by assiting in stabilizing transported proteins. For example, the Hsc70 (Hsp70) chaperone in mammals keeps unfolded mitochondrial proteins soluble on their way to the mitochondrial import receptor Tom70.[13] Cytosolic chaperones of the Hsp70 and Hsp90 families can guide preproteins to the Tom70 import receptors in the outer membrane of mitochondria and induce the import process by binding to Tom70 themselves.[33] Mitochondrial Hsp70 forms a motor complex with Tim44 and Mge1 on the inner membrane to facilitate the movement and unfolding of preprotein domains. Together with Mdj1 (Hsp40), Hsp60 and Hsp10 chaperones are also involved in the refolding of already imported proteins in the mitochondria.[84]

A similar protein import mechanism was observed in chloroplasts of plant cells involving Hsp70,[85] Tic40,[86] and chloroplast Hsp60.[87] Chaperones play an important role also in the trafficking of protein between neighboring cells. For example, in *Arabidopsis thaliana*, the chaperonin TriC-Cct8 was found to be involved in the translocation of KNOTTED1 (KN1) protein through the plasmodesmata channels.[88] KN1 is an essential transcription factor for the establishment and maintenance of stem cells.[89]

PROTEIN MISFOLDING AND FITNESS

The folding of translated polypeptides into a functional protein is thought to be determined by intrinsic features of the primary sequence as well as environmental factors within the cell.[90] In most cases, the native structure of a protein is the one that is also the most stable thermodynamically.[91] Studies of the folding dynamics of small proteins (60−100 residues), which convert from their unfolded to their native (functional) state without the complication of highly populated intermediate states, suggest that a few residue interactions within the sequence form a stable folding nucleus around which the rest of the polypeptide rapidly condenses.[92] Misfolding of proteins or protein structure instability is disadvantageous to the cell not only because the protein function is lacking but also because of the formation of protein aggregates. Misfolded proteins tend to cluster in the cell and form long unbranched, and often twisted, fibers that are a few nanometers in diameter. A prominent example is the amyloid fibril.[93] The structural characteristics of proteins involved in the formation of amyloid fibers vary from intact globular proteins to large unstructured polypetides, but they all share the same common organization with a core structure of β-sheets whose strands run perpendicular to the fibril axis.[94] The formation of misfolded protein aggregates is known to hinder cell viability. For example, both Alzheimer's disease and Parkinson's disease are founded in the deposition of protein aggregates in neuronal tissues.[95,96]

A recent study[97] quantified the impact of misfolded proteins on organism fitness by expressing different variants of structurally destabilized yellow fluorescent protein (YFP) in yeast cells and measuring their growth rates. The results revealed that an induction of a small amount of YFP aggregates leads to a significant reduction in growth rate. Because the YFP is a gratuitous protein, i.e., its function is not essential in yeast, this result indicates that the presence of protein aggregates alone, regardless of the protein function, imposes a selective cost on the organism.

If protein misfolding imposes a selective cost, could it be that the effects of this selective pressure are imprinted in genomes? It has been long known that the protein expression level, codon usage, and evolutionary rate are correlated (see the Glossary for an explanation of terms). For almost every sequenced genome tested so far, the proportion of optimal codons within a protein-coding gene is negatively correlated with amino acid substitution rate.[98−104] An analysis of protein abundance in model organisms revealed a significant positive correlation between protein expression level and codon adaptation both in *E. coli*[105] and in yeast[100] (Figure 2). The consistency of this expression−codon adaptation−conservation (ECC for short) covarion structure led to the suggestion that a single factor underlies these correlations.[106] It has been suggested that protein network properties[107] or protein essentiality[108] plays a key role as a determinant of the ECC covarion. However, none of these factors provides a plausible explanation for the correlation between protein expression level and the selection against synonymous nucleotide substitutions. On the basis of the comparison of yeast paralogs having similar protein sequence but different expression levels, Drummond et al.[102]

dx.doi.org/10.1021/bi3013643 | *Biochemistry* 2012, 51, 9941−9953

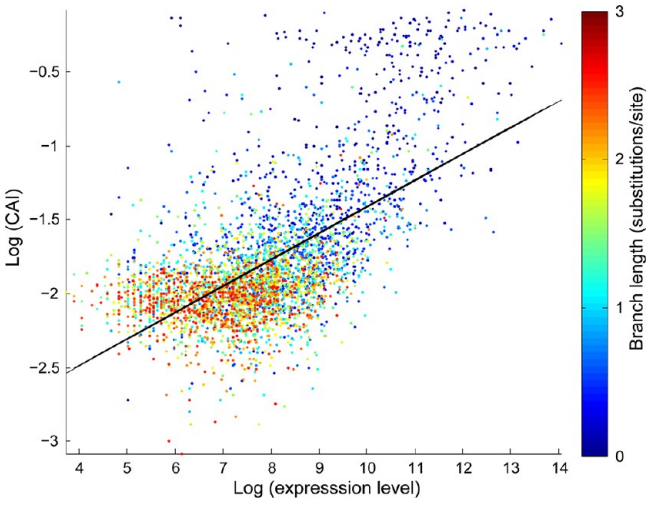


Figure 2. Three-way correlation among expression level, codon adaptation index (CAI), and evolutionary rate calculated for yeast (data from ref 138).

concluded that protein expression level is the major determinant of the ECC covarion and suggested that the selection at the DNA level acts against ribosome infidelity during translation to minimize protein misfolding. According to their model, the selection against synonymous nucleotide substitutions maintains protein translation accuracy while the selection against nonsynonymous nucleotide substitutions maintains translation robustness.[102] Thus, in a model in which the fitness is determined by protein translation efficiency, the ECC covarion is determined by selection against mistranlation-induced protein misfolding.[104]

A strong impact of protein mistranslation on protein folding robustness is extremely important in in vitro systems (e.g., ref 109). In living cells, however, misfolding of mistranslated proteins can be compensated by chaperones. Molecular chaperones lower the energetic barrier for a stable conformation, thus allowing polypeptides that contain destabilizing residues to fold into a functional protein. The chaperone expression level is increased in the presence of unfolded polypeptides, regardless of the type of intracellular and/or environmental stress condition.[110,111] This mechanism of action was exemplified in the recent work by Geiler-Samerotte et al.,[97] who observed an elevation of chaperone expression level, including members of the Hsp70, Hsp40, and Hsp90 families, in the presence of misfolded YFP aggregates. The transcription of many of these chaperones in yeast is activated by heat shock transcription factor 1 (HSF1) whose availability is conditioned by the presence of misfolded proteins in a negative feedback loop.[112] The role of molecular chaperones in the stress response (especially to heat stress) has long been studied.[113] By providing proper folding of translated proteins, chaperones mitigate the decrease in fitness caused by stress-induced protein misfolding.

CHAPERONE-MEDIATED PROTEIN FOLDING AND FITNESS

Experiments in which chaperone activity in whole organisms was repressed highlighted the extent to which living cells depend upon chaperone-mediated protein folding under normal conditions. A decrease in Hsp90 activity in *Drosophila* by crossing over with a weak Hsp90 allele (*Hsp83*) or by feeding the flies with an inhibitor of Hsp90 revealed phenotypic deformities that were much more abundant than expected by chance.[115] Applying an inhibitor of Hsp90 activity to *A. thaliana* seedlings revealed phenotypes similar to those observed under heat stress conditions.[116] The resulting phenotypic deformities in these experiments are attributed to the misfolding of Hsp90 clients; many of them are involved in signal transduction.[115,116] The Hsp90 chaperone inhibition in these experiments revealed phenotypic variation that was encoded in the genome but masked by chaperone activity. This leads to the conclusion that chaperones have a significant impact on the organism's fitness as buffers of phenotypic variation.[116,117] In other words, some genetic variation in protein-encoding genes has a negligible effect on the phenotype (i.e., it is neutral) as long as the protein conformation, and consequently its function, is kept constant by the crucial assistance of chaperones.[116,117]

Populations facing high mutational loads are prone to suffer from reduced fitness caused by destabilizing mutations in protein folding genes leading to protein misfolding. The implication of chaperones as mediators of phenotypic stability suggests that they might be useful for survival under such conditions.[117−119] Indeed, experimental studies of *E. coli* and *Salmonella typhimurium* populations that have been exposed to random mutagenesis

Glossary

Gene expression: Process by which the information from a protein-coding or an RNA-specifying gene is used in the synthesis of a gene product. Gene expression is usually measured on a genome-wide scale by using DNA microarrays, which measure the relative amounts of RNA transcripts from thousands of genes at once.

Protein expression: Translation of an mRNA into a protein. Protein expression levels are measured with technologies such as antibody arrays that target specific known proteins or liquid chromatography associated with tandem mass spectrometry. The units of measure are usually the number of protein molecules per cell.

Preferred codon: Most frequently used codon for a particular amino acid.

Codon usage bias: Degree with which codon usage in a protein-encoding gene deviates from equal frequencies of occurrence of synonymous codons.

Codon adaptation index (CAI): Measure quantifying the deviation of actual codon usage from optimal codon usage. An optimal codon is the one, among several that encode the same amino acid, having the highest concentration of corresponding tRNA in the cell. CAI is calculated assuming that because of selection the most abundant codon for each amino acid is the optimal one.[98]

Protein conservation: Degree of similarity between homologous proteins.

Amino acid replacement rate: Number of amino acid replacements per site per unit time. In comparative studies, the unit time is the divergence time between the two sequences.

Protein connectivity: Number of links that a protein node has to other nodes in the protein−protein interaction network.

Synonymous substitution: Substitutions of a nucleotide in the reading frame of a protein-encoding gene that results in a change from one codon to a synonymous one. A synonymous substitution does not alter the amino acid encoded by the codon, unless the substitution affects a splicing site or an RNA editing site.

Experimental evolution: Propagating organisms under controlled conditions with the objective of studying phenotypic and genotypic changes over time (for a review, see ref 114).

revealed that an overexpression of the GroE−chaperonin complex restored[119] or improved[120] their fitness. Fares et al.[119] demonstrated the buffering effect of GroE chaperonin by using a mutator *E. coli* strain[121] that accumulates mutations 3.3-fold faster than the wild type. After 3240 generations of random mutation accumulation, the fitness (measured by growth rate) of the mutated strains decreased by 50% compared to that of the ancestral line. Cloning a constitutive GroE operon into the mutated strain resulted on average in 86-fold higher levels of the chaperonin and led to a restored fitness that was only 20% lower than that of the ancestral strain. The improved fitness was, however, conditioned by supplementing the growth media with ample amino acids that were probably required for the translation of GroE in massive quantities.[119] This result demonstrates that chaperone overexpression is useful for overcoming high mutational loads, yet a trade-off between the beneficial impact of the chaperones and the resources required for their production exists.

A later study by Maisnier-Patin et al.[120] showed that a modest increase in the GroEL expression level is sufficient to improve the fitness. Mutagenesis in *S. typhimurium* was induced by expressing an error-prone DNA polymerase at different levels, and the accumulation of random mutations led to decreased fitness. Samples under high mutational loads showed increased levels of expression of the DnaK and GroEL chaperones at a level 2−3-fold higher than that of the ancestral strain. The chaperones were probably upregulated because of the presence of misfolded proteins resulting from the accumulation of destabilizing mutations. Furthermore, an artificial induction of GroEL expression by a factor of ∼1.5 improved the fitness substantially.[120] These results supply further evidence that chaperones contribute to antagonistic epistasis where the cumulative effect of mutations in the genome is mitigated.[120]

Natural populations evolving under high mutational loads supply further evidence of the buffering effect of chaperones. Microbial endosymbionts are characterized by small populations and effectively no recombination, leading to an increased rate of fixation of deleterious mutations in their genes.[122] Measurements of GroEL concentrations in the bacterium *Buchnera aphidicola*, an intracellular endosymbiont of aphids, showed that it is expressed at a level 7.5-fold higher than that of *E. coli* under normal conditions.[123] This naturally induced overexpression of the chaperonin probably evolved as a compensatory mechanism to maintain protein stability under high mutational loads.[118,122]

An analysis of the chaperone repertoire in eukaryotic endosymbionts supplies further evidence of the importance of chaperones during reductive evolution. Microsporidia are unicellular eukaryotes, a sister group of fungi, that evolved into obligate intracellular parasites infecting most eukaryotic phyla.[124] Members of the group are characterized by highly reduced genomes encoding very few genes. The number of open reading frames (ORFs) in currently sequenced microsporidal genomes ranges between 1997 in *Encephalitozoon cuniculi*[125] and 2633 in *Trachipleistophora hominis*.[126] A comparison of the microsporida chaperone repertoire to that of yeast reveals an extreme reduction in the populations of the Hsp40 and Hsp70 protein families, while all eight genes encoding the TriC/CCT subunits have been retained. This may suggest that the CCT/TriC chaperones have an essential role in maintaining eukaryotic protein stability under high mutational loads that are typical in reductive genome evolution.[127]

CHAPERONE-MEDIATED FOLDING AND PROTEIN EVOLUTION

The observations that chaperone expression under high mutational loads can restore or improve the organism's fitness led to the suggestion that interaction of protein with chaperones enlarges the spectrum of neutral mutations and consequently increases protein evolvability.[117,128] Using an experimental evolution approach, Tokuriki and Tawfik[129] examined the impact of GroE-mediated folding on protein evolution. Various enzymes whose folding (i.e., function) depends upon the GroE chaperonin were exposed to random mutagenesis using an error-prone polymerase chain reaction, and the resulting variants were selected for a further mutagenesis round according to their enzymatic activity. The experiment was performed both under normal conditions and in the presence of overexpressed GroE. The results revealed that overexpression of GroE facilitated the accumulation of significantly more mutations in comparison to the normal mutational drift and led to the conclusion that interaction of protein with the chaperones indeed promotes enzyme evolution.

The finding that GroE increases protein evolvability has been evaluated in an experimental setting. If chaperone-mediated protein evolution also occurs in nature, we might be able to find evidence of it in sequenced genomes. To test this hypothesis, one has to compare the evolutionary dynamics of proteins whose folding is assisted by chaperones with those of proteins that fold in a manner independent of the chaperones. Proteins that interact with GroE in *E. coli* can be divided into three classes based on their dependency upon the GroE for folding.[130] GroE-independent proteins (class I) fold spontaneously under standard conditions (37 °C) and attain on average 55% of their activity in a manner independent of chaperones, GroE or otherwise. GroE partially dependent proteins (class II) require GroEL assistance, in addition to other chaperones, at 37 °C but do not require GroES at 25 °C, where spontaneous folding is observed. GroE obligatory proteins (class III) fail to fold spontaneously at 37 °C and have an obligate requirement for GroE to attain activity.[130,131] A comparison of *E. coli* proteins to their orthologs in 446 proteobacterial genomes revealed that obligatory substrates of GroE (class III) evolve 35% faster than GroE-independent substrates (class I).[132,133] The significant difference in amino acid substitution rate among the three GroE dependency classes could not be explained by other correlates of protein evolutionary rates such as expression level, protein essentiality, or the number of interactions with other proteins (protein centrality).[132] These results suggest that during evolution, GroE-mediated folding increases the evolutionary rate of substrate proteins by buffering the deleterious effects of misfolding-related mutations.[132,133]

A comparison of codon usage across the three GroE dependency classes revealed that casual GroE substrates (class I) exhibit a higher level of codon and tRNA adaptation than obligate GroE substrates (class III).[132,134,135] Constraining the comparison of codon usage to buried sites only, which are considered to be structurally sensitive, revealed that the enrichment in optimal codon usage within casual GroE substrates is even more pronounced.[134] The optimal codon enrichment within the coding sequences of casual GroE substrates indicates that they are less prone to mitranslation-induced misfolding,[134] which fits well with their weakened dependency upon the GroE for folding. Because codon usage and protein expression level are positively correlated, the difference in codon adaptation among the GroE dependency

classes means that casual GroE substrates are predicted to be more strongly expressed than obligatory substrates.[134] A comparison of protein expression levels measured in *E. coli* strain K12 MG1655[136] among the GroE dependency classes revealed that this is indeed the case.[132]

However, proteins that depend on chaperones for folding also have different physiochemical properties according to the chaperones with which they interact and the degree of their dependency. In an analysis of the impact of chaperone buffering capacity on genome evolution in *E. coli*, strictly dependent substrates of GroE were found to be enriched with positively charged amino acids and with cysteine and proline, and their genes were found to have higher GC content. In addition, the number of protein−protein interactions decreased with the dependency upon GroE.[132] Similarly, protein solubility experiments revealed enriched levels of glycine and alanine in proteins that belong to the most strictly GroE dependent substrate class in *E. coli*.[131] These proteins are also characterized by inherent aggregation propensities that were significantly higher than those of proteins less dependent on GroE for folding.

GroE dependence also correlates with patterns of protein interactions. Casual GroE interactors (class I) have more protein interactions[132] and are more central in the *E. coli* metabolic network[137] than obligatory substrates. Hence, proteins that depend upon GroE for folding are found in the periphery of the protein−protein interaction network and the metabolic network.[132,137] These observations led to the suggestion that interaction of protein with GroE facilitates the expansion of the metabolic network by allowing substrate proteins to explore their conformational space and evolve novel functions.[132,137]

Studying the correlation between chaperone-mediated folding and protein evolution in eukaryotes is complicated by the wealth of chaperones encoded in eukaryotic genomes and the many different folding pathways in which they interact with substrate proteins. A recent large-scale survey of chaperone interactors in *S. cerevisiae* using the TAP-tag approach revealed that ∼60% of the yeast proteome interacts with one or more chaperones.[79] The number of chaperones interacting with a single protein can reach a total of 25 as in the example of Hca4, a putative nucleolar DEAD box RNA helicase. Many chaperones overlap in their subsets of interacting proteins. For example, 63% of the proteins that interact with Ssb1 (Hsp70) interact also with its paralog, Ssa1.[79] On the other hand, some chaperones, especially those of the Hsp70 family, can interact with a multitude of substrate proteins, with Sse1 (Hsp70) having the largest number of interacting proteins (2705 of the 5880 proteins encoded in yeast).[79] The global chaperone−protein interaction pattern revealed a positive correlation between the number of interacting chaperones per substrate−protein pair and the number of hydrophobic stretches in the protein sequence, suggesting the frequency of hydrophobic regions as the phenotypic signal of structurally vulnerable proteins.[79]

Within the cytosolic chaperone repertoire, the TriC/CCT chaperonin complex was found to have a significant substrate overlap with the eubacterial GroE complex.[60,77] This raises the question of whether TriC/CCT influences protein evolution in a manner similar to that of GroE. Warnecke and Hurst[134] searched for detectable evidence of the evolutionary impact of TriC/CCT in substrate protein sequences. They found proteins that interact with TriC/CCT to be longer than proteins that do not interact with that chaperone,[134] yet no correlation between protein expression level and CCT interaction could be observed, despite the fact that longer genes encode less abundant proteins.[134] However, the large substrate overlap and complex interaction patterns in the eukaryotic chaperone interactions network are likely to mask the effect of any single chaperone.

A recent examination of the yeast chaperone−substrate interaction patterns using tools from the field of network science revealed a remarkable order in the complex chaperone interaction network.[138] An application of modularity function[139] that seeks to divide the network into the most connected components (termed also communities) revealed 10 communities of proteins and their dedicated chaperones. Five Hsp70 chaperones were not grouped into any community; those interact with more than 1000 proteins each, and 3275 proteins in total,[79] indicating a low substrate specificity in their interaction. Substrate proteins in the 10 communities were found to be significantly different in their physiochemical properties such as protein length, the proportion of negative and polar amino acids, aromaticity, and the proportion of α-helical and coiled-coil secondary structures.[138] Proteins with more chaperone interactions in yeast are longer, heavier, and enriched with aspartate, glutamate, and lysine amino acids.[79] Proteins with fewer chaperone interactions were found to exhibit higher aromaticity and hydrophobicity and were enriched with cysteine and phenylalanine.[79] However, the number of hydrophobic stretches between one and five residues in length was increasing with the number of chaperone interactions. Substrates of the chaperonin TriC/CCT are enriched with β-sheets.[77] Proteins with high β-sheet content were found to fold slowly and to be vulnerable to misfolding and aggregation.[77] In the network analysis of chaperone−protein interactions in yeast, the substrate proteins in the modules were found to be significantly different not only in the above-mentioned biochemical properties but also in the usage of many single amino acids. The usage of aspartate, glutamate, glycine, isoleucine, leucine, phenylalanine, proline, and valine was significantly different among the 10 modules after a false discovery rate test for multiple comparisons.[138]

Using a network approach to analyze the yeast chaperone−substrate interaction network revealed that proteins that interact with different sets of chaperones are significantly different in their levels of expression, codon adaptations, and levels of sequence conservation. Ranking the chaperone−substrate communities by these three properties shows that they are intercorrelated like the ECC covarion observed in whole genomes.[138] Communities of proteins that are strongly expressed are also the communities that evolve with the slowest substitution rates and are encoded by a high proportion of preferred codons. Conversely, communities of proteins that have the lowest expression level also evolve with the highest substitution rate and show a decreased level of codon adaptation. Much of the variability in protein substitution rates among the communities is explained by protein expression level, signifying protein abundance within the cell as a major determinant in the ECC covarion.

Chaperones from the Hsp70 family are mostly unspecific in their interaction, but many other chaperones, such as the members of the Hsp40 family[140] and the Hsp90 system,[141] are. The exact mechanism of recognition of the substrate by the chaperones is not yet fully understood.[140,141] This is a difficult question to tackle because proteins whose functional folding depends upon the chaperones are probably recognized by the characteristics of their intermediate, relatively unstable, structure that is difficult to document using the existing techniques for the determination of protein structure. Nevertheless, the biased amino acid usage and overrepresentation of particular secondary structure elements in substrates of several chaperone families

(e.g., GroE and CCT/TriC) suggest that the information underlying substrate recognition is encoded within the protein sequence. Consequently, proteins that interact with similar chaperones are expected to have common features within their primary and secondary structures. Comparative genomics of proteins classified by their interaction with chaperones revealed that those are significantly different not only in their physiochemical properties[77,79,131] but also in their evolutionary properties.[132,138] These studies implicate interaction of protein with chaperones as a major force that shapes the genomic landscape during evolution.

EVOLUTION OF INTERACTION OF PROTEIN WITH CHAPERONES

The impact of chaperone-mediated folding on genomic architecture should be placed in an evolutionary context. How can we make sense of interaction of protein with chaperones in light of evolution? We suggest that the origin of molecular chaperones and the evolution of their interaction with substrate proteins can be explained by the constructive neutral evolution model,[142] which supplies a possible explanation for the origin of complex biological systems while accounting for the lack of an advantage from their intermediate stages.[143]

Spontaneous folding of protein into a stable structure most probably preceded the origin of chaperones. Thus, chaperones evolved in the presence of spontaneously folding proteins to prevent the aggregation of misfolded polypeptides[144] and functioned at their origin more as "holders" than as "folders". At this stage, the novel function supplied by the chaperones was either beneficial or neutral, imposing only the production costs of the chaperones themselves. The folding assistance provided by chaperones doubtless became beneficial under stress conditions leading to protein structural destabilization (e.g., heat shock). Prokaryotes have been shown to evolve with increased mutation rates under stress conditions;[145,146] the buffering supplied by the chaperones could be an essential molecular mechanism in this case. Thus, environmental instability must have played an important role in the emergence of chaperones and their fixation during evolution. Chaperones and their interacting proteins coevolved, and some proteins became obligatorily dependent on that interaction. The adaptation of protein to the folding assistance of distinct chaperones represents an evolutionary trap that is not easily escaped by random mutational process and drift.[147] Hence, chaperone-mediated folding allowed for an increased structural complexity at the cost of an obligatory requirement for the chaperones.

The translation of proteins that require the assistance of molecular chaperones for folding has to be coordinated with chaperone interaction. Recent studies revealed an important role of codon usage and codon usage distribution along the gene in controlling protein translation speed dynamics.[148,149] Casual GroE substrates in *E. coli*, which can also fold into their functional structure spontaneously, are encoded by a higher proportion of preferred codons in comparison to obligatory substrates and are also more abundant in the cell,[132,134] which fits well with the ECC covarion. We suggest that this bias stems from the requirement for synchronization between protein translation and cotranslational folding.[150] Nascent polypeptides that are able to fold spontaneously into their functional conformation are free from that constraint and can be translated at a higher speed. Moreover, is it possible that to gain a stable conformation, the whole nascent polypeptide should be available before folding. However, with increasing translation speed, the fitness cost of misfolding also increases drastically. Consequently, accuracy becomes more important so that proteins that are translated at high speed should also be more conserved.[104]

The evolution of interaction of protein with chaperones should be inspected also from the systemic point of view. A recent survey of GroEL interactors in *E. coli* revealed that 794 proteins (~18% of the *E. coli* proteome) interact with the chaperonin.[151] Of the 5880 proteins in yeast, 595 where found to interact with the CCT/TriC chaperone.[79] We propose that the required protein abundance in the cell largely determines the kind and mode of interaction of that protein with molecular chaperones for folding. The first reason is the energetic investment in chaperone-mediated folding. Chaperone-mediated folding by itself does not require much ATP in comparison to the translation process. For example, the GroE chaperonin consumes seven ATP molecules in each round of substrate turnover,[152] while translation of a single amino acid costs four ATP molecules.[80] The average protein sequence length in *E. coli* is 316; hence, one round within the chaperonin will add only 0.5% to the ATP consumption of the protein production. However, if GroE is required for the folding of many proteins, then GroE by itself should be strongly expressed. Moreover, if it is required for the production of strongly expressed proteins, then it should be produced in even larger quantities. The GroE production costs amount to translation of seven GroES subunits (7×97 amino acids) and 14 GroEL subunits (14×548 amino acids). Apparently, the constitutive production of GroE creates an overload of the translation system and an arrest of cell growth.[119] Furthermore, each round of folding by GroE takes ~10 s,[152] which may considerably slow protein production. This indicates that chaperone attention should be limited according to the available energetic resources and temporal dynamics of protein synthesis within the cell. Large-scale analysis of chaperone interaction data supports that notion. A comparison of levels of expression between GroE dependence groups showed that casual substrates are more strongly expressed than obligatory interactors.[132−134] Similarly, yeast proteins that interact only with one of the promiscuous Hsp70 chaperones are more strongly expressed than proteins that interact with additional chaperones.[138]

Studies of hemoglobin polymerization in vitro showed that the polymer formation rate depends on the concentration of soluble monomers.[153] Existing polymers serve as a basis for the formation of heterogeneous polymers. Thus, hemoglobin polymerization is an autocatalytic process whose rate is log-linear proportional to monomer concentration.[153] This idea was recently adopted for the formation of amyloid fibrils.[154] Taken together, these studies indicate that the formation of protein aggregates within the cell largely depends on the abundance of misfolded proteins. This could act as an additional negative selection pressure that keeps highly expressed proteins from developing a dependency upon the chaperones for folding because failure in the folding stage will lead to a massive amount of misfolded proteins in a very short time.

In summary, chaperones are crucial in allowing many nascent polypeptides to attain their functional conformation and in providing an energetically efficient mechanism for the recycling of misfolded proteins. Genomic data reveal that chaperones have an important role in shaping genomic landscapes, stemming from the part they play in the intricate correlation among expression level, translation rate, codon usage, and sequence conservation. In a broader evolutionary context, molecular chaperones mitigate the deleterious effects of protein misfolding, thus allowing a wider range of genetic variability, the raw material for positive selection, adaptation, and innovation.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tal.dagan@hhu.de. Telephone: +49 211 811 0490.

Funding

We acknowledge support of the European Research Council (Grant 281357).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Giddy Landan for valuable discussions. We thank Sander Smits, Tom Williams, and Amnon Kohen for their helpful comments on the manuscript and Dan Graur for his contribution to the glossary of terms in molecular evolution.

REFERENCES

(1) Laskey, R., Honda, B., Mills, A., and Finch, J. (1978) Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature 275*, 416−420.

(2) Ellis, J. (1987) Proteins as molecular chaperones. *Nature 328*, 378−379.

(3) Becker, J., Walter, W., Yan, W., and Craig, E. (1996) Functional interaction of cytosolic hsp70 and a DnaJ-related protein, Ydj1p, in protein translocation in vivo. *Mol. Cell. Biol. 16*, 4378−4386.

(4) Hartl, F. U. (1996) Molecular chaperones in cellular protein folding. *Nature 381*, 571−580.

(5) Ellis, R. J. (2006) Molecular chaperones: Assisting assembly in addition to folding. *Trends Biochem. Sci. 31*, 395−401.

(6) Albanese, V., Reissmann, S., and Frydman, J. (2010) A ribosome-anchored chaperone network that facilitates eukaryotic ribosome biogenesis. *J. Cell Biol. 189*, 69−81.

(7) Patzelt, H., Rüdiger, S., Brehmer, D., Kramer, G., Vorderwülbecke, S., Schaffitzel, E., Waitz, A., Hesterkamp, T., Dong, L., Schneider-Mergener, J., Bukau, B., and Deuerling, E. (2001) Binding specificity of *Escherichia coli* trigger factor. *Proc. Natl. Acad. Sci. U.S.A. 98*, 14244−14249.

(8) Hesterkamp, T., Hauser, S., Lütcke, H., and Bukau, B. (1996) *Escherichia coli* trigger factor is a prolyl isomerase that associates with nascent polypeptide chains. *Proc. Natl. Acad. Sci. U.S.A. 93*, 4437−4441.

(9) Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004) Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature 431*, 590−596.

(10) Sha, B., Lee, S., and Cyr, D. M. (2000) The crystal structure of the peptide-binding fragment from the yeast Hsp40 protein Sis1. *Structure 8*, 799−807.

(11) Cyr, D., Langer, T., and Douglas, M. (1994) Trends in Biochemical Sciences: DnaJ-like proteins: Molecular chaperones and specific regulators of Hsp70. *Trends Biochem. Sci. 19*, 176−181.

(12) Caplan, A. J., Cyr, D., and Douglas, M. (1992) Ydj1p Facilitates Polypeptide Translocation Across Different Intracellular Membranes by a Conserved Mechanism. *Cell 71*, 1143−1155.

(13) Deshaies, R. J., Koch, B. D., Werner-Washburne, M., Craig, E. A., and Schekman, R. (1988) A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. *Nature 332*, 800−805.

(14) Walsh, P., Bursać, D., Law, Y. C., Cyr, D., and Lithgow, T. (2004) The J-protein family: Modulating protein assembly, disassembly and translocation. *EMBO Rep. 5*, 567−571.

(15) Zylicz, M., LeBowitz, J. H., McMacken, R., and Georgopoulos, C. (1983) The dnaK protein of *Escherichia coli* possesses an ATPase and autophosphorylating activity and is essential in an in vitro DNA replication system. *Proc. Natl. Acad. Sci. U.S.A. 80*, 6431−6435.

(16) Beckmann, R. P., Mizzen, L., and Welch, W. J. (1990) Interaction of Hsp 70 with newly synthesized proteins: Implications for protein folding and assembly. *Science 248*, 850−854.

(17) Gautschi, M., Lilie, H., Funfschilling, U., Mun, A., Ross, S., Lithgow, T., Rucknagel, P., and Rospert, S. (2001) RAC, a stable ribosome-associated complex in yeast formed by the DnaK-DnaJ homologs Ssz1p and zuotin. *Proc. Natl. Acad. Sci. U.S.A. 98*, 3762−3767.

(18) Thirumalai, D., and Lorimer, G. H. (2001) Chaperonin-mediated protein folding. *Annu. Rev. Biophys. Biomol. Struct. 30*, 245−269.

(19) Hendrix, R. W. (1979) Purification and Properties of Groe, a Host Protein Involved in Bacteriophage Assembly. *J. Mol. Biol. 129*, 375−392.

(20) Cheng, M. Y., Hartl, F. U., Martin, J., Pollock, R. A., Kalousek, F., Neupert, W., Hallberg, E. M., Hallberg, R. L., and Horwich, A. (1989) Mitochondrial heat-shock protein Hsp60 is essential for assembly of proteins imported into yeast mitochondria. *Nature 337*, 620−625.

(21) Kaufman, B. A. (2003) A function for the mitochondrial chaperonin Hsp60 in the structure and transmission of mitochondrial DNA nucleoids in *Saccharomyces cerevisiae*. *J. Cell Biol. 163*, 457−461.

(22) Hunt, J. F., Weaver, A. J., Landry, S. J., Gierasch, L., and Deisenhofer, J. (1996) The crystal structure of the GroES co-chaperonin at 2.8 Å resolution. *Nature 379*, 37−45.

(23) Langer, T., Pfeifer, G., Martin, J., Baumeister, W., and Hartl, F. U. (1992) Chaperonin-Mediated Protein Folding: Groes Binds to One End of the Groel Cylinder, Which Accommodates the Protein Substrate Within Its Central Cavity. *EMBO J. 11*, 4757−4765.

(24) Kubota, H., Hynes, G., Carne, A., Ashworth, A., and Willison, K. (1994) Identification of six Tcp-1-related genes encoding divergent subunits of the TCP-1-containing chaperonin. *Curr. Biol. 4*, 89−99.

(25) Klumpp, M., Baumeister, W., and Essen, L. O. (1997) Structure of the substrate binding domain of the thermosome, an archaeal group II chaperonin. *Cell 91*, 263−270.

(26) Gao, Y., Thomas, J. O., Chow, R. L., Lee, G. H., and Cowan, N. J. (1992) A cytoplasmic chaperonin that catalyzes β-actin folding. *Cell 69*, 1043−1050.

(27) Frydman, J., Nimmesgern, E., Erdjument-Bromage, H., Wall, J. S., Tempst, P., and Hartl, F. U. (1992) Function in Protein Folding of Tric, a Cytosolic Ring Complex Containing Tcp-1 and Structurally Related Subunits. *EMBO J. 11*, 4767−4778.

(28) Vainberg, I., Lewis, S., Rommelaere, H., Ampe, C., Vandekerckhove, J., Klein, H., and Cowan, N. (1998) Prefoldin, a chaperone that delivers unfolded proteins to cytosolic chaperonin. *Cell 93*, 863−873.

(29) Prodromou, C., Roe, S. M., Piper, P. W., and Pearl, L. H. (1997) A molecular clamp in the crystal structure of the N-terminal domain of the yeast Hsp90 chaperone. *Nat. Struct. Mol. Biol. 4*, 477−482.

(30) Bardwell, J. C., and Craig, E. A. (1987) Eukaryotic Mr 83,000 heat shock protein has a homologue in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A. 84*, 5177−5181.

(31) Borkovich, K. A., Farrelly, F. W., Finkelstein, D. B., Taulien, J., and Lindquist, S. (1989) Hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. *Mol. Cell. Biol. 9*, 3919−3930.

(32) Louvion, J. F., Abbas-Terki, T., and Picard, D. (1998) Hsp90 is required for pheromone signaling in yeast. *Mol. Biol. Cell 9*, 3071−3083.

(33) Young, J., Hoogenraad, N., and Hartl, F. (2003) Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. *Cell 112*, 41−50.

(34) Toogun, O. A., DeZwaan, D. C., and Freeman, B. C. (2007) The Hsp90 Molecular Chaperone Modulates Multiple Telomerase Activities. *Mol. Cell. Biol. 28*, 457−467.

(35) Parsell, D. A., Kowal, A. S., Singer, M. A., and Lindquist, S. (1994) Protein disaggregation mediated by heat-shock protein Hsp104. *Nature 372*, 475−478.

(36) Glover, J., and Lindquist, S. (1998) Hsp104, Hsp70, and Hsp40: A novel chaperone system that rescues previously aggregated proteins. *Cell 94*, 73−82.

(37) Mogk, A., Tomoyasu, T., Goloubinoff, P., Rüdiger, S., Röder, D., Langen, H., and Bukau, B. (1999) Identification of thermolabile *Escherichia coli* proteins: Prevention and reversion of aggregation by DnaK and ClpB. *EMBO J. 18*, 6934−6949.
(38) Queitsch, C., Hong, S. W., Vierling, E., and Lindquist, S. (2000) Heat shock protein 101 plays a crucial role in thermotolerance in *Arabidopsis*. *Plant Cell 12*, 479−492.
(39) Schmitt, M., Neupert, W., and Langer, T. (1996) The molecular chaperone Hsp78 confers compartment-specific thermotolerance to mitochondria. *J. Cell Biol. 134*, 1375−1386.
(40) Lee, S., Sowa, M. E., Watanabe, Y. H., Sigler, P. B., Chiu, W., Yoshida, M., and Tsai, F. (2003) The structure of clpB: A molecular chaperone that rescues proteins from an aggregated state. *Cell 115*, 229−240.
(41) Haslbeck, M., Walke, S., Stromer, T., Ehrnsperger, M., White, H., Chen, S., Saibil, H., and Buchner, J. (1999) Hsp26: A temperature-regulated chaperone. *EMBO J. 18*, 6744−6751.
(42) Quigley, P. M., Korotkov, K., Baneyx, F., and Hol, W. G. J. (2003) The 1.6-Å crystal structure of the class of chaperones represented by *Escherichia coli* Hsp31 reveals a putative catalytic triad. *Proc. Natl. Acad. Sci. U.S.A. 100*, 3137−3147.
(43) Lee, S. J. (2003) Crystal Structures of Human DJ-1 and *Escherichia coli* Hsp31, Which Share an Evolutionarily Conserved Domain. *J. Biol. Chem. 278*, 44552−44559.
(44) Zhao, Y., Liu, D., Kaluarachchi, W. D., Bellamy, H. D., White, M. A., and Fox, R. O. (2003) The crystal structure of *Escherichia coli* heat shock protein YedU reveals three potential catalytic active sites. *Protein Sci. 12*, 2303−2311.
(45) Wilson, M. A., Amour, C. V. S., Collins, J. L., Ringe, D., and Petsko, G. A. (2004) The 1.8-Å resolution crystal structure of YDR533Cp from *Saccharomyces cerevisiae*: A member of the DJ-1/ThiJ/PfpI superfamily. *Proc. Natl. Acad. Sci. U.S.A. 101*, 1531−1536.
(46) Bonifati, V. (2002) Mutations in the DJ-1 Gene Associated with Autosomal Recessive Early-Onset Parkinsonism. *Science 299*, 256−259.
(47) Halio, S. B., Blumentals, I. I., Short, S. A., Merrill, B. M., and Kelly, R. M. (1996) Sequence, expression in *Escherichia coli*, and analysis of the gene encoding a novel intracellular protease (PfpI) from the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol. 178*, 2605−2612.
(48) Zylicz, M., Ang, D., Liberek, K., and Georgopoulos, C. (1989) Initiation of lambda DNA replication with purified host- and bacteriophage-encoded proteins: The role of the dnaK, dnaJ and grpE heat shock proteins. *EMBO J. 8*, 1601−1608.
(49) Hoffmann, H. J., Lyman, S. K., Lu, C., Petit, M. A., and Echols, H. (1992) Activity of the Hsp70 chaperone complex—DnaK, DnaJ, and GrpE—in initiating phage lambda DNA replication by sequestering and releasing lambda P protein. *Proc. Natl. Acad. Sci. U.S.A. 89*, 12108−12111.
(50) Rüdiger, S., Schneider-Mergener, J., and Bukau, B. (2001) Its substrate specificity characterizes the DnaJ co-chaperone as a scanning factor for the DnaK chaperone. *EMBO J. 20*, 1042−1050.
(51) Chen, S., Roseman, A. M., Hunter, A. S., Wood, S. P., Burston, S. G., Ranson, N. A., Clarke, A. R., and Saibil, H. R. (1994) Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy. *Nature 371*, 261−264.
(52) Macario, A. J., Dugan, C. B., and Conway de Macario, E. (1991) A dnaK homolog in the archaebacterium *Methanosarcina mazei* S6. *Gene 108*, 133−137.
(53) Macario, A. J., Dugan, C. B., Clarens, M., and Conway de Macario, E. (1993) dnaJ in Archaea. *Nucleic Acids Res. 21*, 2773.
(54) Conway de Macario, E., Dugan, C. B., and Macario, A. J. (1994) Identification of a grpE heat-shock gene homolog in the archaeon *Methanosarcina mazei*. *J. Mol. Biol. 240*, 95−101.
(55) Rommelaere, H., Van Troys, M., Gao, Y., Melki, R., Cowan, N. J., Vandekerckhove, J., and Ampe, C. (1993) Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven related subunits. *Proc. Natl. Acad. Sci. U.S.A. 90*, 11975−11979.
(56) Large, A. T., Goldberg, M. D., and Lund, P. A. (2009) Chaperones and protein folding in the archaea. *Biochem. Soc. Trans. 37*, 46−51.
(57) Spreter, T. (2005) The Crystal Structure of Archaeal Nascent Polypeptide-associated Complex (NAC) Reveals a Unique Fold and the Presence of a Ubiquitin-associated Domain. *J. Biol. Chem. 280*, 15849−15854.
(58) Large, A. T., and Lund, P. A. (2009) Archaeal chaperonins. *Front. Biosci. 14*, 1304−1324.
(59) Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R., Martinez-Arias, R., Henne, A., Wiezer, A., Baumer, S., Jacobi, C., Bruggemann, H., Lienard, T., Christmann, A., Bomeke, M., Steckel, S., Bhattacharyya, A., Lykidis, A., Overbeek, R., Klenk, H., Gunsalus, R., Fritz, H., and Gottschalk, G. (2002) The genome of *Methanosarcina mazei*: Evidence for lateral gene transfer between bacteria and archaea. *J. Mol. Microbiol. Biotechnol. 4*, 453−461.
(60) Hirtreiter, A. M., Calloni, G., Forner, F., Scheibe, B., Puype, M., Vandekerckhove, J., Mann, M., Hartl, F. U., and Hayer-Hartl, M. (2009) Differential substrate specificity of group I and group II chaperonins in the archaeon *Methanosarcina mazei*. *Mol. Microbiol. 74*, 1152−1168.
(61) Williams, T. A., Codoñer, F. M., Toft, C., and Fares, M. A. (2010) Two chaperonin systems in bacterial genomes with distinct ecological roles. *Trends Genet. 26*, 47−51.
(62) Techtmann, S. M., and Robb, F. T. (2010) Archaeal-like chaperonins in bacteria. *Proc. Natl. Acad. Sci. U.S.A. 107*, 20269−20274.
(63) Embley, T. M., and Martin, W. (2006) Eukaryotic evolution, changes and challenges. *Nature 440*, 623−630.
(64) Cox, C. J., Foster, P. G., Hirt, R. P., Harris, S. R., and Embley, T. M. (2008) The archaebacterial origin of eukaryotes. *Proc. Natl. Acad. Sci. U.S.A. 105*, 20356−20361.
(65) Cotton, J. A., and McInerney, J. O. (2010) Eukaryotic genes of archaebacterial origin are more important than the more numerous eubacterial genes, irrespective of function. *Proc. Natl. Acad. Sci. U.S.A. 107*, 17252−17255.
(66) Esser, C., Ahmadinejad, N., Wiegand, C., Rotte, C., Sebastiani, F., Gelius-Dietrich, G., Henze, K., Kretschmann, E., Richly, E., Leister, D., Bryant, D., Steel, M. A., Lockhart, P. J., Penny, D., and Martin, W. (2004) A Genome Phylogeny for Mitochondria Among Proteobacteria and a Predominantly Eubacterial Ancestry of Yeast Nuclear Genes. *Mol. Biol. Evol. 21*, 1643−1660.
(67) Gould, S. B., Waller, R. F., and McFadden, G. I. (2008) Plastid Evolution. *Annu. Rev. Plant Biol. 59*, 491−517.
(68) Thiergart, T., Landan, G., Schenk, M., Dagan, T., and Martin, W. F. (2012) An Evolutionary Network of Genes Present in the Eukaryote Common Ancestor Polls Genomes on Eukaryotic and Mitochondrial Origin. *Genome Biol. Evol. 4*, 466−485.
(69) Young, J., Agashe, V., Siegers, K., and Hartl, F. (2004) Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol. 5*, 781−791.
(70) Michimoto, T., Aoki, T., Toh-e, A., and Kikuchi, Y. (2000) Yeast Pdr13p and Zuo1p molecular chaperones are new functional Hsp70 and Hsp40 partners. *Gene 257*, 131−137.
(71) Yan, W., Schilke, B., Pfund, C., Walter, W., Kim, S., and Craig, E. (1998) Zuotin, a ribosome-associated DnaJ molecular chaperone. *EMBO J. 17*, 4809−4817.
(72) Gautschi, M., Mun, A., Ross, S., and Rospert, S. (2002) A functional chaperone triad on the yeast ribosome. *Proc. Natl. Acad. Sci. U.S.A. 99*, 4209−4214.
(73) Valpuesta, J., Martin-Benito, J., Gomez-Puertas, P., Carrascosa, J., and Willison, K. (2002) Structure and function of a protein folding machine: The eukaryotic cytosolic chaperonin CCT. *FEBS Lett. 529*, 11−16.
(74) Lund, P., Large, A., and Kapatai, G. (2003) The chaperonins: Perspectives from the Archaea. *Biochem. Soc. Trans. 31*, 681−685.
(75) Ditzel, L., Lowe, J., Stock, D., Stetter, K., Huber, H., Huber, R., and Steinbacher, S. (1998) Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. *Cell 93*, 125−138.

(76) McCallum, C., Do, H., Johnson, A., and Frydman, J. (2000) The interaction of the chaperonin tailless complex polypeptide 1 (TCP1) ring complex (TRiC) with ribosome-bound nascent chains examined using photo-cross-linking. *J. Cell Biol. 149*, 591−601.
(77) Yam, A. Y., Xia, Y., Lin, H.-T. J., Burlingame, A., Gerstein, M., and Frydman, J. (2008) Defining the TRiC/CCT interactome links chaperonin function to stabilization of newly made proteins with complex topologies. *Nat. Struct. Mol. Biol. 15*, 1255−1262.
(78) Hemmingsen, S. M., Woolford, C., van der Vies, S. M., Tilly, K., Dennis, D. T., Georgopoulos, C. P., Hendrix, R. W., and Ellis, R. J. (1988) Homologous plant and bacterial proteins chaperone oligomeric protein assembly. *Nature 333*, 330−334.
(79) Gong, Y., Kakihara, Y., Krogan, N., Greenblatt, J., Emili, A., Zhang, Z., and Houry, W. A. (2009) An atlas of chaperone-protein interactions in *Saccharomyces cerevisiae*: Implications to protein folding pathways in the cell. *Mol. Syst. Biol. 5*, 1−14.
(80) Stouthamer, A. H. (1973) A theoretical study on the amount of ATP required for synthesis of microbial cell material. *Antonie van Leeuwenhoek 39*, 545−565.
(81) Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011) Global quantification of mammalian gene expression control. *Nature 473*, 337−342.
(82) Sharma, S. K., De Los Rios, P., Christen, P., Lustig, A., and Goloubinoff, P. (2010) The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. *Nat. Chem. Biol. 6*, 914−920.
(83) De Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R., and Subramani, S. (1987) Firefly luciferase gene: Structure and expression in mammalian cells. *Mol. Cell. Biol. 7*, 725−737.
(84) Voos, W., and Röttgers, K. (2002) Molecular chaperones as essential mediators of mitochondrial biogenesis. *Biochim. Biophys. Acta 1592*, 51−62.
(85) Zhang, X., and Glaser, E. (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends Plant Sci. 7*, 14−21.
(86) Chou, M., Fitzpatrick, L., Tu, S., Budziszewski, G., Potter-Lewis, S., Akita, M., Levin, J., Keegstra, K., and Li, H. (2003) Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. *EMBO J. 22*, 2970−2980.
(87) Lubben, T. H., Donaldson, G. K., Viitanen, P. V., and Gatenby, A. A. (1989) Several proteins imported into chloroplasts form stable complexes with the GroEL-related chloroplast molecular chaperone. *Plant Cell 1*, 1223−1230.
(88) Xu, X. M., Wang, J., Xuan, Z., Goldshmidt, A., Borrill, P. G. M., Hariharan, N., Kim, J. Y., and Jackson, D. (2011) Chaperonins Facilitate KNOTTED1 Cell-to-Cell Trafficking and Stem Cell Function. *Science 333*, 1141−1144.
(89) Lucas, W. J., Bouché-Pillon, S., Jackson, D. P., Nguyen, L., Baker, L., Ding, B., and Hake, S. (1995) Selective trafficking of KNOTTED1 homeodomain protein and its mRNA through plasmodesmata. *Science 270*, 1980−1983.
(90) Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884−890.
(91) Dobson, C., Sali, A., and Karplus, M. (1998) Protein folding: A perspective from theory and experiment. *Angew. Chem., Int. Ed. 37*, 868−893.
(92) Fersht, A. (2000) Transition-state structure as a unifying basis in protein-folding mechanisms: Contact order, chain topology, stability, and the extended nucleus mechanism. *Proc. Natl. Acad. Sci. U.S.A. 97*, 1525−1529.
(93) Colon, W., and Kelly, J. W. (1992) Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry 31*, 8654−8660.
(94) Sunde, M., Serpell, L. C., Bartlam, M., Fraser, P. E., Pepys, M. B., and Blake, C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol. 273*, 729−739.
(95) Hardy, J., and Allsop, D. (1991) Amyloid deposition as the central event in the aetiology of Alzheimer's disease. *Trends Pharmacol. Sci. 12*, 383−388.
(96) Braak, H., and Braak, E. (1990) Cognitive impairment in Parkinson's disease: Amyloid plaques, neurofibrillary tangles, and neuropil threads in the cerebral cortex. *J. Neural Transm.: Parkinson's Dis. Dementia Sect. 2*, 45−57.
(97) Geiler-Samerotte, K. A., Dion, M. F., Budnik, B. A., Wang, S. M., Hartl, D. L., and Drummond, D. A. (2011) Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. *Proc. Natl. Acad. Sci. U.S.A. 108*, 680−685.
(98) Sharp, P. M., and Li, W. H. (1987) The Codon Adaptation Index: A measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res. 15*, 1281−1295.
(99) Grantham, R., Gautier, C., Gouy, M., Jacobzone, M., and Mercier, R. (1981) Codon catalog usage is a genome strategy modulated for gene expressivity. *Nucleic Acids Res. 9*, 43−74.
(100) Pál, C., Papp, B., and Hurst, L. D. (2001) Highly expressed genes in yeast evolve slowly. *Genetics 158*, 927−931.
(101) Krylov, D. M. (2003) Gene Loss, Protein Sequence Divergence, Gene Dispensability, Expression Level, and Interactivity Are Correlated in Eukaryotic Evolution. *Genome Res. 13*, 2229−2235.
(102) Drummond, D. A., Bloom, J. D., Adami, C., Wilke, C. O., and Arnold, F. H. (2005) Why highly expressed proteins evolve slowly. *Proc. Natl. Acad. Sci. U.S.A. 102*, 14338−14343.
(103) Pál, C., Papp, B., and Lercher, M. J. (2006) An integrated view of protein evolution. *Nat. Rev. Genet. 7*, 337−348.
(104) Drummond, D. A., and Wilke, C. O. (2008) Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution. *Cell 134*, 341−352.
(105) Rocha, E. P. C. (2003) An Analysis of Determinants of Amino Acids Substitution Rates in Bacterial Proteins. *Mol. Biol. Evol. 21*, 108−116.
(106) Drummond, D. A., Raval, A., and Wilke, C. O. (2006) A single determinant dominates the rate of yeast protein evolution. *Mol. Biol. Evol. 23*, 327−337.
(107) Fraser, H. B., Hirsh, A. E., Steinmetz, L. M., Scharfe, C., and Feldman, M. W. (2002) Evolutionary rate in the protein interaction network. *Science 296*, 750−752.
(108) Hirsh, A., and Fraser, H. (2001) Protein dispensability and rate of evolution. *Nature 411*, 1046−1049.
(109) Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. (2006) Protein stability promotes evolvability. *Proc. Natl. Acad. Sci. U.S.A. 103*, 5869−5874.
(110) Ananthan, J., Goldberg, A. L., and Voellmy, R. (1986) Abnormal proteins serve as eukaryotic stress signals and trigger the activation of heat shock genes. *Science 232*, 522−524.
(111) Baler, R., Welch, W. J., and Voellmy, R. (1992) Heat shock gene regulation by nascent polypeptides and denatured proteins: hsp70 as a potential autoregulatory factor. *J. Cell Biol. 117*, 1151−1159.
(112) Zou, J., Guo, Y., Guettouche, T., Smith, D., and Voellmy, R. (1998) Repression of heat shock transcription factor HSF1 activation by HSP90 (HSP90 complex) that forms a stress-sensitive complex with HSF1. *Cell 94*, 471−480.
(113) Lindquist, S. (1986) The heat-shock response. *Annu. Rev. Biochem. 55*, 1151−1191.
(114) Hindré, T., Knibbe, C., Beslon, G., and Schneider, D. (2012) New insights into bacterial adaptation through in vivo and in silico experimental evolution. *Nat. Rev. Microbiol. 10*, 352−365.
(115) Rutherford, S. L., and Lindquist, S. (1998) Hsp90 as a capacitor for morphological evolution. *Nature 396*, 336−342.
(116) Queitsch, C., Sangster, T., and Lindquist, S. (2002) Hsp90 as a capacitor of phenotypic variation. *Nature 417*, 618−624.
(117) Rutherford, S. L. (2003) Between genotype and phenotype: Protein chaperones and evolvability. *Nat. Rev. Genet. 4*, 263−274.
(118) Moran, N. A. (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. *Proc. Natl. Acad. Sci. U.S.A. 93*, 2873−2878.
(119) Fares, M. A., Ruiz-González, M. X., Moya, A., Elena, S. F., and Barrio, E. (2002) GroEL buffers against deleterious mutations. *Nature 417*, 398.
(120) Maisnier-Patin, S., Roth, J. R., Fredriksson, Å., Nyström, T., Berg, O. G., and Andersson, D. I. (2005) Genomic buffering mitigates

the effects of deleterious mutations in bacteria. *Nat. Genet. 37*, 1376–1379.
(121) de Visser, J. A. G. M., Zeyl, C. W., Gerrish, P. J., Blanchard, J. L., and Lenski, R. E. (1999) Diminishing returns from mutation supply rate in asexual populations. *Science 283*, 404–406.
(122) McCutcheon, J. P., and Moran, N. A. (2011) Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol. 10*, 13–26.
(123) Baumann, P., Baumann, L., and Clark, M. (1996) Levels of *Buchnera aphidicola* chaperonin GroEL during growth of the aphid *Schizaphis graminum*. *Curr. Microbiol. 32*, 279–285.
(124) Corradi, N., and Slamovits, C. H. (2011) The intriguing nature of microsporidian genomes. *Briefings Funct. Genomics 10*, 115–124.
(125) Katinka, M. D., Duprat, S., Cornillot, E., Méténier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretaillade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., and Vivarès, C. P. (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature 414*, 450–453.
(126) Heinz, E., and Lithgow, T. (2012) Back to basics: A revealing secondary reduction of the mitochondrial protein import pathway in diverse intracellular parasites. *Biochim. Biophys. Acta*, DOI: DOI: 10.1016/j.bbamcr.2012.02.006.
(127) Heinz, E., Williams, T. A., Nakjang, S., Noel, C. J., Swan, D. C., Goldberg, A. V., Harris, S. R., Weinmaier, T., Markert, S., Becher, D., Bernhardt, J., Dagan, T., Hacker, C., Schweder, T., Rattei, T., Hirt, R. P., and Embley, T. M. (2012) The Genome of the Obligate Intracellular Parasite *Trachipleistophora hominis*: New Insights into Microsporidian Genome Dynamics and Reductive Evolution. *PLoS Pathog. 8*, e1002979.
(128) Tokuriki, N., and Tawfik, D. S. (2009) Stability effects of mutations and protein evolvability. *Curr. Opin. Struct. Biol. 19*, 596–604.
(129) Tokuriki, N., and Tawfik, D. S. (2009) Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature 459*, 668–673.
(130) Kerner, M. J., Naylor, D. J., Ishihama, Y., Maier, T., Chang, H.-C., Stines, A. P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., and Hartl, F. U. (2005) Proteome-wide Analysis of Chaperonin-Dependent Protein Folding in *Escherichia coli*. *Cell 122*, 209–220.
(131) Fujiwara, K., Ishihama, Y., Nakahigashi, K., Soga, T., and Taguchi, H. (2010) A systematic survey of in vivo obligate chaperonin-dependent substrates. *EMBO J. 29*, 1552–1564.
(132) Bogumil, D., and Dagan, T. (2010) Chaperonin-Dependent Accelerated Substitution Rates in Prokaryotes. *Genome Biol. Evol. 2*, 602–608.
(133) Williams, T. A., and Fares, M. A. (2010) The Effect of Chaperonin Buffering on Protein Evolution. *Genome Biol. Evol. 2*, 609–619.
(134) Warnecke, T., and Hurst, L. D. (2010) GroEL dependency affects codon usage-support for a critical role of misfolding in gene evolution. *Mol. Syst. Biol. 6*, 1–11.
(135) Noivirt-Brik, O., Unger, R., and Horovitz, A. (2007) Low folding propensity and high translation efficiency distinguish in vivo substrates of GroEL from other *Escherichia coli* proteins. *Bioinformatics 23*, 3276–3279.
(136) Lu, P., Vogel, C., Wang, R., Yao, X., and Marcotte, E. M. (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol. 25*, 117–124.
(137) Takemoto, K., Niwa, T., and Taguchi, H. (2011) Difference in the distribution pattern of substrate enzymes in the metabolic network of *Escherichia coli*, according to chaperonin requirement. *BMC Syst. Biol. 5*, 98.
(138) Bogumil, D., Landan, G., Ilhan, J., and Dagan, T. (2012) Chaperones Divide Yeast Proteins into Classes of Expression Level and Evolutionary Rate. *Genome Biol. Evol. 4*, 618–625.
(139) Newman, M. E. J. (2006) Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E 74*, 0605087.
(140) Kampinga, H. H., and Craig, E. A. (2010) The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nat. Rev. Mol. Cell Biol. 11*, 579–592.
(141) Taipale, M., Jarosz, D. F., and Lindquist, S. (2010) HSP90 at the hub of protein homeostasis: Emerging mechanistic insights. *Nat. Rev. Mol. Cell Biol. 11*, 515–528.
(142) Stoltzfus, A. (1999) On the possibility of constructive neutral evolution. *J. Mol. Evol. 49*, 169–181.
(143) Gray, M. W., Lukes, J., Archibald, J. M., Keeling, P. J., and Doolittle, W. F. (2010) Cell biology. Irremediable complexity? *Science 330*, 920–921.
(144) Lansbury, P. T. (1999) Evolution of amyloid: What normal protein folding may tell us about fibrillogenesis and disease. *Proc. Natl. Acad. Sci. U.S.A. 96*, 3342–3344.
(145) Bjedov, I. (2003) Stress-Induced Mutagenesis in Bacteria. *Science 300*, 1404–1409.
(146) Rosenberg, S. M., and Hastings, P. J. (2003) Modulating mutation rates in the wild. *Science 300*, 1382–1383.
(147) Geller, R., Vignuzzi, M., Andino, R., and Frydman, J. (2007) Evolutionary constraints on chaperone-mediated folding provide an antiviral approach refractory to development of drug resistance. *Genes Dev. 21*, 195–205.
(148) Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010) An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. *Cell 141*, 344–354.
(149) Cannarozzi, G., Schraudolph, N. N., Faty, M., von Rohr, P., Friberg, M. T., Roth, A. C., Gonnet, P., Gonnet, G., and Barral, Y. (2010) A Role for Codon Order in Translation Dynamics. *Cell 141*, 355–367.
(150) O'Brien, E. P., Vendruscolo, M., and Dobson, C. M. (2012) Prediction of variable translation rate effects on cotranslational protein folding. *Nat. Commun. 3*, 868.
(151) Niwa, T., Kanamori, T., Ueda, T., and Taguchi, H. (2012) Global analysis of chaperone effects using a reconstituted cell-free translation system. *Proc. Natl. Acad. Sci. U.S.A. 109*, 8937–8942.
(152) Horwich, A. L., Apetri, A. C., and Fenton, W. A. (2009) The GroEL/GroES cis cavity as a passive anti-aggregation device. *FEBS Lett. 583*, 2654–2662.
(153) Ferrone, F. A., Hofrichter, J., and Eaton, W. A. (1985) Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. *J. Mol. Biol. 183*, 611–631.
(154) Lorenzen, N., Cohen, S. I. A., Nielsen, S. B., Herling, T. W., Christiansen, G., Dobson, C. M., Knowles, T. P. J., and Otzen, D. (2012) Role of Elongation and Secondary Pathways in S6 Amyloid Fibril Growth. *Biophys. J. 102*, 2167–2175.