Published in final edited form as:
Nat Rev Mol Cell Biol. 2013 April ; 14(4): 237–248. doi:10.1038/nrm3542.

Diversity in the origins of proteostasis networks- a driver for protein function in evolution

Evan T. Powers[1] and **William E. Balch**[2,3,4]

[1]Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037.

[2]Department of Cell and Molecular Biology, The Scripps Research Institute, La Jolla, California 92037.

[3]Department of Chemical Physiology, The Scripps Research Institute, La Jolla, California 92037.

[4]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037.

Abstract

Although a protein's primary sequence largely determines its function, proteins can adopt different folding states in response to changes in the environment, some of which may be deleterious to the organism. All organisms, including *Bacteria*, *Archaea* and *Eukarya*, have evolved a protein homeostasis network, or proteostasis network, that consists of chaperones and folding factors, degradation components, signalling pathways and specialized compartmentalized modules that manage protein folding in response to environmental stimuli and variation. Surveying the origins of proteostasis networks reveals that they have co-evolved with the proteome to regulate the physiological state of the cell, reflecting the unique stresses that different cells or organisms experience, and that they have a key role in driving evolution by closely managing the link between the phenotype and the genotype.

Protein homeostasis, or proteostasis, is a key mechanism by which cells rapidly respond to their environment to maintain the proteins in the cell in a state that allows optimum biological activity[1]. Although the chemical, biochemical and biophysical properties of a protein's primary sequence largely direct its function, translated proteins are not static; their conformational folding state can dynamically change in response to the local environment (for example, in response to changes in temperature or metabolites). These changes can alter protein function, and can help to mediate a response to both short-term and long-term challenges that ensures cell survival [2-4]. Such changes may also bring long-term benefits by fixing new traits in the genome, leading to enhanced fitness [5, 6]

Protein folding *in vivo* is controlled by chaperones, folding factors (including enzymes involved in oxidative (disulfide) folding and isomerization of peptide bonds), degradation components and regulatory signaling pathways that respond to the intracellular and extracellular environment to control folding and hence, function [2, 7-11]. In the case of Eukarya, proteostasis biology also includes diverse endomembrane compartments, such as the membrane trafficking compartments including lysosomes as well as mitochondria and chloroplasts, that create specialized folding management environments which considerably expand the repertoire of protein function achievable by the polypeptide chain sequence [12]. These factors work together to control proteostasis, and are collectively known as the proteostasis network [2]. Within this network, some factors promote protein folding, some

*Correspondence: epowers@scripps.edu; webalch@scripps.edu.

prevent or correct misfolding[7, 8, 13, 14], some prevent and/or redirect aggregation [15-17] and others direct proteins to degradation pathways [18-26]. If changes in proteome function that promote survival provide long-term benefits, they can be stored in the genome to ensure optimal fitness for subsequent generations. As such, the proteostasis network has a key role in modulating the link between the phenotype and the environment, either intracellular or extracellular [12], and in continuously evolving variations in the protein sequence that affect the biophysical properties of the fold to faciliate diversity in biology[6, 14, 27]. Disruption of proteostasis underlies many human diseases, highlighting the importance of the proteostasis network in healthspan and ageing [9, 28].

In this Review we describe how the origins of components of the proteostasis network in the three kingdoms of life - Bacteria, Archaea and Eukarya [29, 30] – optimize protein function in response to changes in the environment, with a particular focus on the heat shock protein family of chaperones. We discuss evidence indicating that the proteostasis network and the proteome have co-evolved to promote organismal survival in different niches, and how the proteostasis network has worked as a driver of evolution, facilitating adaptation and natural selection [4, 31].

Proteostasis biology as a system

At the base of proteostasis biology is the ribosome, an ancient and conserved machine that Woese and colleagues [29, 30] recognized based on sequencing of 16S ribosomal RNA could be used to trace the origins of life, anticipating the impact of proteostasis biology on evolvability (Figure 1a, Box 1) [2, 32, 33]. In addition to the ribosome-based translational machinery and its many associated regulatory and chaperone factors, the proteostasis network includes components that direct folding through ATP-dependent mechanisms (such as the central molecular chaperones/co-chaperones belonging to the HSP40, HSP70 and HSP90 families, which assist both translational and post-translational folding) and the HSP60 family of ‘chaperonins’ (including the bacterial protein GroEL, the archaeal thermosome and the eukaryotic TriC/CCT) of folding chambers that function as cages to transiently retain proteins in productive folding environments [34]. Others include the small heat shock proteins (sHSPs) and proteins that modulate oxidative folding [8, 35]. Homologues of many conserved proteostasis network components are abundant in nearly all species, whereas others are specialized and provide folding assistance to specific proteins or proteins in specific environments. In addition to the cytosolic proteostasis network components found in all three kingdoms, proteostasis components found exculsively in Eukarya populate the many different endomembrane compartments comprising the exocytic and endocytic trafficking pathways (endoplasmic reticulum (ER), Golgi, endosomes) as well as mitochondria, chloroplasts and nuclear compartments, greatly expanding folding capacity and management of protein function by proteostasis [12]. In addition to these folding components, the proteostasis network includes a vast armament of components that degrade misfolded or aggregated proteins such as the ATP-driven AAA+ proteases in Bacteria and Archaea that function as disaggregases [17], and the cytosolic ubiquitin-proteasome system (UPS) [18] and membrane-sequestered autophagy-lysosome pathways in Eukarya [12, 20] (Figure 1a; Box 1).

Given the complexity of proteostasis network activity in managing protein folding (Figure 1a), it can be more simply viewed as a ‘cloud’ surrounding each protein that is responsive to the local environment to manage protein synthesis, folding, misfolding and/or degradation, and thus optimize protein function in the cell to promote organismal survival (Figure 1b). Thus, a central feature of proteostasis biology is that it adds a new layer to our conventional view of the function of the protein fold reflected in the primary (1°), secondary (2°), ternary (3°) and quarternary (4°) structural states (Figure 1b). Proteostasis actively manages these

structural states in close cooperation with the metabolic state of the cell, internal cellular pathways directing, for instance, development and differentiation and specialized function, and in response to numerous signaling pathways sensing the external environment (Figure 1a) [1, 14]. We refer to this set of interactions between the proteostasis network and a given protein as the quinary (5°) physiologic state [1, 12], a term that emphasizes that the 1° polypeptide chain sequence which is encoded in the genome and gives rise to the 2°–4° structural states, is heavily managed by the local proteostasis network in the cytosol found in all species and by numerous specialized pathways found in endomembrane compartments in Eukarya (Figure 1b). The latter include the extensive exocytic and endocytic membrane trafficking pathways [12] as well as mitochondria and chloroplasts. (Figure 1b). Quinary states (Figure 1b) provide a dynamic mechanism to link the structural features of a protein as it matures to management of its physiologic function by the cell and the environment (Figure 1a; Box 1). The quinary physiologic state of the fold dictated by the multi-layered proteostasis network (Figure 1a,b) provides a foundation to considerably expand the function to the primary polypeptide sequence encoded by the genome to promote organismal biology and hence survival, adaptation and evolvability.

The proteostasis network is unique for each species and for each cell type in the case of multicellular Eukarya [8]. For example, stem cell proteostasis biology is very different from that of a mature, differentiated neuron, which is, in turn, different from that of a primary liver hepatocyte or a primary polarized epithelial cell in the lung [8, 36]. This also applies to the components of the proteostasis network defining the quinary state in the different subcellular compartments of eukaryotic cells [12, 37], highlighting their likely unique roles in protein fold management to achieve advanced cellular, tissue and organismal function in Eukarya. In addition, a poorly studied, but likely robust proteostasis system that does not use ATP protects the protein fold outside the cell and encompasses, for example, the periplasmic space of Bacteria and Archaea or the extracellular matrix and fluids such as plasma and lymphatic systems in multicellular Eukarya [38]. As might be expected, there are many mechanisms to change the composition of the proteostasis cloud, and hence survival and ultimately fitness of the cell, through signaling pathways that manage proteostasis network composition and through the ability of proteostasis components to manage signaling (Figure 1a; Box 1). These can involve both cell autonomous and cell non-autonomous signaling pathways that optimize folding for function in response to changing conditions [28, 39]. The centrality of proteostasis is clear in the well-characterized role of proteostasis in the management (and mismanagement) of human disease and aging (Figure 1c, Box 2) [8, 9, 39, 40].

The origins of proteostasis components

It is now evident that the role of the proteostasis network as a dynamic manager of the quinary physiologic state of the fold arose early during the development of life on Earth to successfully adapt to the many challenges to protein folding in response to diverse chemical and physical environments, and to increasingly complex cell biological and developmental programmes. On the one hand, folding is strongly opposed by conformational entropy and the dehydration of the polar parts of the protein backbone and side chains. On the other hand, it is strongly favored by the hydrophobic effect [41], intramolecular hydrogen bonding[42], and, to a lesser extent, other intramolecular interactions [43].

Proteins are generally most stable in the environment in which they evolved to function [1], a crucial observation for understanding the effect of proteostasis on survival and fitness. Departures from this environment, through changes in temperature, co-solutes, pH and ionic strength, and in the balance between the forces of folding and unfolding, generally result in destabilization of the folded state. As the folded state becomes less dominant, the population

of the unfolded state increases, as do the populations of partially folded or misfolded states. These non-folded states are vulnerable to degradation (Figure 1b) and often are prone to aggregation[11, 17, 44-46]. The effect of stresses on the folding status of the proteome cannot be overemphasized. A change in temperature from 37°C to 41°C, which corresponds to an increase in thermal energy of only 1.3%, can decrease the stability of the native state substantially [47]. Imbalance in the status of the functional fold is not healthy for an organism: too much unfolding and/or degradation of a protein challenges the quinary state managed by the proteostasis network and could lead to the loss of that protein's function via degradation or to both toxicity and loss of function via aggregation. Indeed, aggregates are now recognized to use specialized quinary states to manage degradation [15, 16] (Figure 1b,c). Thus, it is the relative fragility of the folded state of a protein in biology that is actively managed by proteostasis to enhance its attributes for function or to remove it to protect the cell and/or allow the cell to acquire new function (Figure 1b,c). Indeed, the proteostasis network and its ability to manage the quinary state of the fold is so enormously beneficial that all known organisms have some kind of proteostasis programme - no organism that lacks one is known to have survived to the present day [48].

Although many proteostasis network components are expressed constitutively, they are particularly important when an organism encounters a protein folding stress from the environment. Many proteostasis components are strongly upregulated by such stresses. This upregulation is often only transient - long enough to solve the physical folding problem. By contrast, chronic upregulation and imbalance in proteostasis biology in the cell is thought to contribute to multiple pervasive human diseases, as it places the vast majority of the proteome at risk in an altered physiologic state (Box 2).

Co-evolution of proteostasis and the fold

Because the evolution of proteostasis biology is simply too large a subject to be covered in a single review, we focus on one class of prominent components of the proteostasis network managing the quinary state, the canonical members of the group of folding chaperones known as heat shock proteins (HSPs)[13] and their constitutively expressed homologues[13, 49]. Other proteostasis biology subsystems, including degradative [18] and endomembrane trafficking components [12] (Figure 1a) as well as the unique proteostasis networks found in mitochondria and chloroplasts [50, 51], all have likely undergone a similar course of evolutionary adaptation and specialization. Figure 2 illustrates the evolution the HSR components HSP40, HSP70, HSP90, HSP60 (chaperonins) and sHSPs) and HSP100, a disaggregase in the context of increasing number of genes expressed as a function of increasing genome size (see supplementary information s1 (table) and supplementary information s2 (figure) for details). It is apparent that the quinary state of the fold defined by proteostasis evolved with the increasing complexity of protein structure and function (Figure 2a; supplementary information s1 (table) and supplementary information s2 (figure)). A possible and purely speculative scenario for proteostasis evolution is illustrated in Box 3. However they arose, the lesson is that the canonical HSPs spread widely enough among the organisms present early during the development of life that they were all probably present in the last universal common ancestor (LUCA) [52].

If the canonical HSPs were in fact present in the LUCA, then every organism present today could have them all. However, although all of the canonical HSPs are indeed widely distributed (Figure 2), none is universal. For example, HSP90 is largely absent from Archaea, an entire kingdom of life [29, 30, 53] (Figure 2). This observation suggests that proteostasis biology is tailored to proteomes, and therefore that the composition of an organism's proteostasis network and its ability to manage the quinary state of the fold

(Figure 1b) can illuminate the challenges that it faces in its struggle for survival and its contribution to population fitness in the context of natural selection [4].

Responses to environmental stresses that threaten the integrity of an organism's proteome can take place on two time scales. Individual organisms can adapt temporarily to stress on a short time scale by rapidly (within minutes) changing the compositions of their local proteostasis network. These proteostasis stress responses include the ubiquitous heat shock response (HSR) [49] and the more specialized unfolded protein response (UPR) and mitochondrial unfolded protein response (mitoUPR) [54] in eukaryotic subcellular compartments (Box 1). Specialization in each cell or, in the case of Eukarya, subcellular compartments (Figure 1b), adds new and dynamic layers to the versatility of the proteostasis network, vastly increasing the biological utility of the elementary primary sequence encoded by the genome [12]. Although transient adaptation to acute stress obviously conveys an immediate survival advantage, the longer-term responses to chronic stress present a different problem. This involves the evolution of permanent changes in the proteostasis network, evolution of the proteome itself, or both through epigenetic and genetic mechanisms.

Some lessons derived from general trends in proteostasis biology by following its evolution and specific examples of proteostasis biology at work in distinct niches are discussed below.

Lessons learnt from the cytosolic HSPs

The number of representatives of each of the canonical HSPs (except the HSP100s) increases roughly linearly as the number of genes in the genome increases (Figure 2; supplementary information s1 (table) and supplementary information s2 (figure)). On average, an organism has one homologue of HSP70 (and five to six HSP40s) for every 2,000 genes; one HSP90 for every 6,000 genes; and one sHSP and one type of HSP60 subunit for every 2,000 genes.

In some respects, the correlation between genome size and proteostasis network size is puzzling. As canonical HSPs generally have weak substrate specificity [55-58], expressing a single type of HSP70, for example, at a high concentration should be as effective to maintain proteostasis as expressing several types of HSP70. This notion is supported in part by the observation that one of the three HSP70 homologues in *E. coli*, DnaK, is expressed at much higher levels than either of the other two, HscA and HscC[59]. DnaK is in fact the 'housekeeping' HSP70, which functions to maintain the general health of the proteome, and neither HscA nor HscC can complement its loss[59]. By contrast, HscA has a specific role in the proteostasis of iron-sulfur cluster proteins [59]. In other words, it has acquired specificity in the course of evolution to refine the quinary state of a particular class of proteins. Less is known about the function of HscC. Like DnaK, it has weak substrate specificity; it is induced by ultra-violet stress and protects against cadmium toxicity; but it is not induced by heat shock, suggesting that HscC serves a specialized purpose in the SOS response [59], perhaps in response to common environmental threats that required special assistance in the context of natural selection.

Among human HSP70s, the dominant homologue in the cytosol is HSPA8 (Hsc70) [60]. Interestingly, chaperone specialization in the human cytosolic HSP70s has reached the point where the housekeeping and stress response functions have diverged. HSPA8 is weakly induced by stress (although its localization is affected by stress), whereas HSPA1A and HSPA1B are strongly induced [60]. A similar situation is found in the human cytosolic HSP90s: HSP90AB (HSP90β) is constitutively expressed and weakly stress-induced, whereas HSP90AA (HSP90α) is strongly stress-induced and can be highly specialized and found in high levels constitutively for certain types of tissue such as neurons, perhaps reflecting specialized function in synaptic vesicle cycling [61-64].

The division of labour among the HSP70s and HSP90s is an example of a general theme in proteostasis observed across evolutionarily distant species (Figure 2; supplementary information s1 (table) and supplementary information s2 (figure)). When multiple homologues of a particular chaperone are present in an organism, a few serve as general housekeepers whereas the others have more specialized functions, reflecting the evolution of the proteostasis network to support increasingly complicated proteomes.

Lessons learnt from the proteostasis networks in subcellular compartments

The division of proteostasis labour was markedly elaborated in Eukarya, which have multiple subcellular compartments with specialized functions to facilitate folding biology [12] and that include both membrane trafficking organelles as well as energy-producing compartments including mitochondria and chloroplasts, which generate unique quinary states (Figure 1b). For example, the ER-specialized HSP70 homologue BiP (binding immunoglobulin protein) is critically responsive to UPR signaling [65]. Moreover, in mammals the HSP90 homologue associated with the ER, Grp94, has evolved for a restricted repertoire of folding pathways within the ER, including the one activated by insulin-growth factor receptor 2 (IGFR2) promoting improved healthspan and longevity [66]. Additional examples are observed in the chloroplasts of *Arabidopsis thaliana*. Here, three of the six HSP60 subunits (Cpn60α1, Cpn60β2 and Cpn60β3 (LEN1)) are highly expressed [67] and hetero-oligomerize to form the housekeeping chaperonin [68]. The other three (Cpn60α2, Cpn60β1, and Cpn60β4) are expressed at lower levels, and Cpn60β4 is specifically required for the folding of the NdhH subunit of the chloroplast NADH dehydrogenase-like complex [69]. The lesson learnt is having distinct proteostasis networks in distinct subcellular compartments [12] enables further specialization of the quinary state managing the protein fold for function.

Lessons learnt from increasing proteome complexity

The increase in the size and complexity of proteostasis network as a function of proteome size and number of genes expresed (Figure 2) suggests that proteome complexity is itself a proteostasis burden. Unlike short term stresses that organisms face for survival (such as sudden changes in temperature, pH and salinity) that reflect a global but temporary challenge to the folding environment, the stress from having to produce many different kinds of proteins at different times and at different rates during growth and development requires building a repertoire of specialist proteostasis network components that function to complement the generalists. This is well recognized in the immune system, in which highly evolved ER proteostasis networks are responsible for the differentiation of plasma cells used for high volume immunoglobulin expression [70] or β-cells in the pancreas involved in insulin production [71]. As another example, in organisms that do not precisely regulate their temperature, their heat shock response has evolved to handle external temperature changes. By contrast, temperature homeostatic species such as multicellular mammals can use heat shock responses to protect themselves while using temperature as a means to attack invading pathogens, as in fever-linked immune responses to pathogens and/or physical injury. Such responses are often quite specialized for the environmental threat and, therefore, the proteostasis network necessarily co-evolves uniquely with a particular organism in response to its local challenges.

Moreover, with increasing organismal complexity such as found in higher Eukarya, the proteostasis biology of the different cell types indicated above needs to be linked to integrate with overall organismal function. New evidence suggests that they are likely integrated through cell non-autonomous neuronal circuitries that can uniquely tailor the proteostasis response to the type of threat to the protein fold [72]. Such tailored responses could include responses to, for example, oxidative stress (cigarette smoke) [73] or physical stress (muscle

damage) [18, 74, 75], among others. The lesson learnt is that there is an evolving relationship between the quinary state and proteome of the cell in response to complex cell and organismal environments that is exploited by proteostasis to optimize biology.

Exceptions to the rule

It is useful to consider exceptions to the general trends in proteostasis network expansion and quinary state management (Figure 1b) as discussed above, as they highlight the flaws in protein folding that proteostasis exists to correct.

The first example are mycoplasmas, a family of pathogenic Bacteria characterized by extremely small genomes (Figure 2; supplementary information s1 (table) and supplementary information s2 (figure)). Several mycoplasmas lack GroEL, the bacterial HSP60[76]. No organisms outside this family have been found to lack HSP60s, which are usually crucial components of proteostasis networks [76]. Given their extent of genome reduction, mycoplasmas could have made up for the loss of GroEL by selectively losing the proteins that are obligate GroEL clients. However, this is not the case [77]. It has instead been shown that several proteins that are GroEL-dependent in *E. coli* have homologues in the mycoplasma *Ureaplasma urealyticum*, but the *U. urealyticum* versions of these proteins have evolved to fold without GroEL [78]. Thus, mycoplasmas have adapted to the loss of GroEL through evolution of their proteome to match the constraints of the local proteostasis network.

As a second example, all Bacteria and Eukarya have an HSP70 system, but this chaperone system was lost by Archaea [79]. It was later reacquired by some Archaea, possibly by horizontal gene transfer from Bacteria, but thermophilic and hyperthermophilic Archaea continue to thrive in its absence [79]. How could the Archaea have lost such an important proteostasis network component and survived? Perhaps other proteostasis network components in these Archaea, particularly their HSP60s, could have shouldered the proteostasis burden left behind by HSP70. In fact, overproduction of GroEL and GroES can compensate for the loss of DnaK in *E. coli* [80]. A second possibility is that the proteomes of the thermophilic and hyperthermophilic Archaea could have evolved so that they were not as prone to the pitfalls that are corrected by the HSP70 system in other organisms. This most likely reflects their highly specialized niche, in which high temperatures put unusual thermodynamic constraints on the protein fold. Thus, such Archaea may have reached an evolutionary endpoint in terms of quinary management of the fold by proteostasis through ‘extreme’ niche adaptation [81].

A third example is organisms that lack a heat shock response (Figure 1a; Box 1). The heat shock response is as widely distributed in nature as the canonical HSPs [82]. It is retained by organisms with highly compacted genomes such as mycoplasmas and by hyperthermophilic Bacteria[83] and Archaea [84]. However, *Hydra oligactis*, a freshwater invertebrate, and a few of its close relatives lack a heat shock response [85]. Since this discovery, the absence of a heat shock response has been noted in several other organisms, most of which are Antarctic marine animals [86]. The loss of the heat shock response in Antarctic marine animals is easily explained in terms of their environment: the extraordinarily stable temperature in the Southern Ocean makes a heat shock response superfluous (although perhaps not inconsequential to artic ecology as a consequence of ongoing climate change). This explanation does not, however, apply to *Hydra oligactis*, as temperature variations at the surfaces of lakes where this organism lives are much greater [87]. It does, however, raise the possibility that this species lacks a heat shock response either because it evolved a temperature-resistant proteome or because responding to the temperature fluctuations that it experiences has a maladaptive effect on survival and fitness in its particular niche, as could be the case in prolonged chronic stress diseases in humans.

Above we have provided some examples of lessons that can be learnt from surveying proteostasis networks. Optimization of the proteostasis network of an organism, whether it be Bacteria, Archaea or Eukarya, reflects not only an evolutionary decision driven by natural selection in the context of niche adaptation, but anchors proteostasis and its capacity to define the quinary state of the fold as a primary driving force for subsequent organismal evolution, as discussed below.

Evolution of the proteostasis network

Adaptation of the function of proteostasis components to different environmental conditions is often readily apparent as it is, for example, in a comparison of the DnaK (HSP70) homologues from *E. coli* (which grows between 20 and 45°C [81]) and *Thermus thermophilus* (which grows between 40 and 85 °C [81]). The ATPase activity of DnaK from *E. coli* is strongly temperature dependent, increasing 70-fold between 20°C and 53°C and then dropping precipitously [88]. By contrast, the ATPase activity of DnaK from *T. thermophilus* depends weakly on temperature, increasing by only 3-fold between 25°C and 75°C [89]. These DnaK operating ranges are consistent with the organisms' growth temperature ranges: a narrow range of medium temperatures for *E. coli*, and a broad range of high temperatures for *T. thermophilus*. Chaperone adaptation can be apparent at the other end of the thermal range as well. By comparing HSP70 orthologues of three species of cod, two Antarctic and one temperate, it was found that the HSP70s from the Antarctic species were more active at temperatures close to the freezing point of water[90].

Adaptation extends to regulatory elements of the proteostasis network as well (Figure 1b, Box 1). The effects of environmental adaptation on the regulation of HSPs are evident from the differences in the function of the heat shock response among organisms from environments with different temperatures, consistent with the fact, as indicated above, that some organisms lack a heat shock response. For example, organisms living in environments with highly variable temperatures have evolved to activate their heat shock responses within the range of body temperatures that they regularly experience [91]. By contrast, for organisms found in environments with stable temperatures, the thermal limits of proteostasis have been set by the habitats in question and they are unable to respond to changes in temperature. In principle, this limits their ability to adapt to more extreme changes, should they occur. These evolutionary lessons emphasize that the proteostasis network sets the quinary state of the fold on a knife edge within a particular proteostasis cellular program [92], a condition that may significantly contribute to the biological control of survival and fitness in a given niche.

The molecular mechanisms responsible for the changes in the regulation of the canonical HSPs are known in a few cases. In *Hydra oligactis*, the transcription of HSP70 can be induced by high temperature, but rapid degradation of HSP70 mRNA at elevated temperature negates this increase in transcription[93]. In other cases, hints about changes in the regulation of the canonical HSPs can be found in the arrangement of heat shock elements in their promoter regions [94, 95] or from the (re)arrangement of HSP genes within the genome that suggest that they are differentially responsive to stimuli [96].

The changes described above in the regulation of canonical HSPs may require long evolutionary time frames to accumulate, but faster mechanisms for change must exist as well. For example, differences in thermotolerance arise quickly enough in natural populations to be frequently observed between different populations of the same species [97]. A particularly striking example comes from two populations of *Drosophila melanogaster* from "Evolution Canyon" in the Lower Nahal Oren, Mt. Carmel, Israel. Strains of *D. melanogaster* from the south-facing slope (which has higher levels of direct sunlight) are more thermotolerant than those from the north-facing slope, even after being removed from

their natural habitats and kept under laboratory conditions for many generations [98]. The lower thermotolerance of the flies from the North-facing slope is at least partly due to their lower expression levels of HSP40 and HSP70 compared with the flies from the South-facing slope [99, 100]. In the case of HSP70, the decrease in expression was probably due to the insertion of a *P* transposable element into the promoter region of *HSP70Ba*, the inducible form of HSP70 in *D. melanogaster*, in the flies form the South-facing slope [99]. Insertions of transposable elements into the promoters of heat shock genes are endemic in natural populations of *D. melanogaster*, providing a ready mechanism for them to quickly adapt to environments of different temperatures [101]. As the diverse mechanisms that evolve to solve the folding problem in extant species are frequently challenged by misfolding stress and disease, it will be useful to understand how the quinary state of the proteostasis network can be managed to optimize human healthspan in response to disease (Box 2). It could serve as a major platform for therapeutics that adjust the quinary state protecting the fold in acute and/or chronic disease where evolution would simply be too slow to promote survival.

Proteostasis and its impact on evolution

Above we focused on how proteostasis evolves and adapts to the needs of the proteomes for a given species and niche to promote survival and fitness. To be complete, we must now also ask the converse question: how is the process of natural selection influenced by proteome-specific proteostasis biology managing the fold? Given that the proteostasis network is different in each species and operates differentially even within the cell types in multicellular tissues and organisms, it is likely that the combination of protein folding dynamics and the proteostasis network uniquely determine a protein's evolvability – its ability to acquire genetic diversity to facilitate natural selection [4]. Hence, the quinary state of the fold provides a substantial addition to the repertoire of genetic forces that are currently thought to have a role in driving evolution.

It has previously been established [6, 14, 102], in yeast, fly and plant models, that HSP90 facilitates evolvability and potentiates genetic variation by having a crucial role in buffering (protecting) protein folds that harbor destabilizing mutations. Reducing the buffering capacity of HSP90, as occurs during stress, can lead to the emergence of new traits by allowing the misfolded protein to be functional in the physiological pathways operating in the altered cell environment. Thus, HSP90 can allow mutations in the genome to produce immediate cryptic phenotypes that may solve a temporary problem which may be lost when HSP90 function returns to normal- that is, when the stress response is abated as occurs in acute responses. In this way, HSP90 temporarily preserves the survival and/or fitness of the species in fluctuating circumstances. By contrast, repeated insults that sustain an activated chronic stress response by the proteostasis network can result in phenotypes that become assimilated in subsequent generations through the acquisition of additional mutations in the pathway or altered proteostasis capacity that can support new protein or pathway function. In this case, through Hsp90 buffering capacity, repeated insults allow the existence of different folded states that are then captured by genetic and epigenetic mechanisms to provide a new standard for both survival and fitness in subsequent generations [103].

There are many additional examples of the ability of proteostasis network components to augment evolvability. Its role in adaptive evolution is supported by the well-recognized role of HSF1 [104] and the proteostasis component HSP90 in promoting the survival of cancer cells [105-107], or the ability of a virus to acquire resistance to the host during infection through its dependence on HSP90 for maturation [108, 109]. Analysis of overexpression of proteostasis components suggests that the proteostasis network confers a more robust acceptance of a mutation [5, 110]. HSP90 can manage transposon-based mechanisms, as functional alterations of HSP90 affect the germ-line-specific silencing mechanisms

involving small RNAs that ultimately lead to transposon activation and the induction of morphological mutants [111,112]. HSP90 can also act as an modifier (inhibitor or promoter) of evolvability through its effects on chromatin structure and/or chromatin remodeling and by managing histone-based nucleosome structure-function relationships and telomere stability [113, 114]. Indeed, through its effects on nucleosome structure, HSP90 canaffect chromatin binding of histone acetylases (HATs) and histone deacetylases (HDACs), two key regulators of transcription. Interestingly, HATs and HDACs themselves regulate the expression of HSP40[115], HSP70[116], HSP90[117] and the activity of heat-shock transcription factor HSF1, the latter of which can attenuate the HSR response [118]. Thus, the growing relationship between the activity of HATs and HDACs, proteostasis biology, and their combined impact on survival and fitness through methylation pathways suggests that acetylation-deacetylation pathways have a central role in managing both short-term and long-term the quinary state of the fold and hence evolability (Figure 1a).

Perspective

The lessons learnt from surveying the origins of proteostasis suggests that proteostasis biology sets standards for a given species to manage protein function in complex environments. The proteostasis network promotes environment-specific evolvability because it allows functionally suboptimal folds to exist and to become dominant in response to a particular change in that environment. Although the genome cannot rapidly respond to either the environment or the change encoded by a protein variant, the proteostasis network can carry out both tasks by managing the quinary state of the fold; it can detect and rapidly adjust the folding environment [2] to allow protein function to meet the immediate folding stress challenges that can ultimately can be translated into population fitness. Proteostasis biology must use chemical, biochemical and biophysical rules that have not changed since the LUCA to optimize the protein primary polypeptide sequence for biological function [29, 30]- ancient rules that provide a foundation for the continued evolution of extant lifeforms.

Many questions remain unanswered. How does the proteostasis network operate in each cell type or species residing in particular niche at a particular moment in time (and space) to integrate the folding biology through a common quinary physiologic state? How does proteostasis achieve integration between cells within a multicellular species (for example, different tissue systems) to achieve a balance that can sustain survival and generate fitness of the species? Moreover, how do the ATP-independent proteostasis networks that operate outside the cell contribute to evolution[38]- perhaps the smoking gun that would provide us with new biomarkers for monitoring disease states, particularly those relevant to managing human disease and aging [2, 9, 14, 40, 73, 119]. By understanding proteostasis biology, its origins and its diversity, and the co-evolution of the proteostasis network and the proteome, we will begin to gain new insights on how proteostasis optimizes the temporary (survival) with the long-term inheritable (fitness) of the population through natural selection [4, 31].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to gratefully acknowledge our colleagues Dr. J. W. Kelly (The Scripps Research Institute, Departments of Molecular and Experimental Medicine and Chemistry), Dr. R. Morimoto (Northwestern University, Department of Biochemistry, Molecular Biology and Cell Biology) and Dr. A. Dillin (Salk Institute for Biological Studies) for helpful discussions. ETP was supported by grants from the NIH (AG03109); WEB was supported by grants from the NIH (GM42236, GM33301, AG03109, HL079442, HL095524) and the Cystic Fibrosis Foundation.

This article is dedicated the life and work of Carl R. Woese for his seminal insights into our understanding of diversity in the origins of the species, a body of work that upended our traditional views of the evolution of life and that anticipated the role of proteostasis as a molecular force in evolvability.

Glossary terms

Cage	The space within the oligomeric, ring-like structures of HSP60-type chaperones in which proteins can fold, isolated the other components of the cytosol. Often referred to as an Anfinsen cage.
Cryptic phenotype	An inactive, masked or hidden activity/state that becomes functional in response to folding stress.
Evolvability	The capacity of a biological system for adaptive evolution in response to the environment- that is, the ability of population to acquire adaptive genetic diversity to facilitate natural selection.
Healthspan	The period during which an organism, particularly a human, is healthy and free of diseases that compromise longevity.
Horizontal gene transfer	The transfer of genes between organisms in a manner other than traditional reproduction. This can occur through transformation (uptake and expression of foreign genetic material), transduction (movement of genetic material by viruses), bacterial conjugation (cell-to-cell contact resulting in delivery of genetic material) and gene transfer agents that include virus-like elements encoded by the host.
LUCA	(last universal common ancestor). The organism from which all organisms now living on Earth descend, estimated to have lived 3.5 to 3.8 billion years ago.
Membrane trafficking compartments	Subcellular organelles enriched in homologues of cytosolic proteostasis network components that fold and export transmembrane and extracellular cargo proteins to and from the cell surface. Movement between these compartments is directed by an extensive array of coat proteins (which generate transport vesicles), tethers (which direct compartment transfer) and fusion complexes (which facilitate transfer of cargo between compartments).
Nucleosome	The basic unit of DNA packaging in the chromatin structure of Eukarya, consisting of a segment of DNA wound in around histone protein cores. Their compactness and hence access for transcription are subject the epigenetic activity of acetylation and deacetylation by HATs and HDACs, respectively.
***P* transposable element**	A type of transposon (a genetic element that can jump to different sites in a genome) present specifically in *Drosophila melanogaster*.
SOS response	A stress response triggered by DNA damage that induces the expression of proteins involved in DNA repair.

References Cited

1. Anfinsen CB. Principles that govern the folding of protein chains. Science. 1973; 181:223–30. [PubMed: 4124164]
2. Balch WE, Morimoto RI, Dillin A, Kelly JW. Adapting proteostasis for disease intervention. Science. 2008; 319:916–9. [PubMed: 18276881]

3. Gidalevitz T, Prahlad V, Morimoto RI. The stress of protein misfolding: from single cells to multicellular organisms. Cold Spring Harb Perspect Biol. 2011; 3
4. Darwin, C. The Origin of the Species. 1867.
5. Tokuriki N, Tawfik DS. Stability effects of mutations and protein evolvability. Curr Opin Struct Biol. 2009; 19:596–604. [PubMed: 19765975]
6. Lindquist S. Protein folding sculpting evolutionary change. Cold Spring Harb Symp Quant Biol. 2009; 74:103–8. [PubMed: 20375316]
7. Broadley SA, Hartl FU. The role of molecular chaperones in human misfolding diseases. FEBS Lett. 2009; 583:2647–53. [PubMed: 19393652]
8. Powers ET, Morimoto RI, Dillin A, Kelly JW, Balch WE. Biological and chemical approaches to diseases of proteostasis deficiency. Annu Rev Biochem. 2009; 78:959–91. [PubMed: 19298183]
9. Douglas PM, Dillin A. Protein homeostasis and aging in neurodegeneration. J Cell Biol. 2010; 190:719–29. [PubMed: 20819932]
10. Voisine C, Pedersen JS, Morimoto RI. Chaperone networks: tipping the balance in protein folding diseases. Neurobiol Dis. 2010; 40:12–20. [PubMed: 20472062]
11. Lindquist SL, Kelly JW. Chemical and biological approaches for adapting proteostasis to ameliorate protein misfolding and aggregation diseases: progress and prognosis. Cold Spring Harb Perspect Biol. 2011; 3
12. Hutt DM, Balch WE. Expanding proteostasis by membrane trafficking networks. Cold Spring Harbor Perspect. Biology. 2013 Epub. **Jan25th**.
13. Hartl FU, Bracher A, Hayer-Hartl M. Molecular chaperones in protein folding and proteostasis. Nature. 2011; 475:324–32. [PubMed: 21776078]
14. Taipale M, Jarosz DF, Lindquist S. HSP90 at the hub of protein homeostasis: emerging mechanistic insights. Nat Rev Mol Cell Biol. 2010; 11:515–28. [PubMed: 20531426]
15. Chen B, Retzlaff M, Roos T, Frydman J. Cellular strategies of protein quality control. Cold Spring Harb Perspect Biol. 2011; 3:a004374. [PubMed: 21746797]
16. Kaganovich D, Kopito R, Frydman J. Misfolded proteins partition between two distinct quality control compartments. Nature. 2008; 454:1088–95. [PubMed: 18756251]
17. Tyedmers J, Mogk A, Bukau B. Cellular strategies for controlling protein aggregation. Nat Rev Mol Cell Biol. 2010; 11:777–88. [PubMed: 20944667]
18. Finley D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu Rev Biochem. 2009; 78:477–513. [PubMed: 19489727]
19. Watts C. The endosome-lysosome pathway and information generation in the immune system. Biochim Biophys Acta. 2012; 1824:14–21. [PubMed: 21782984]
20. Yang Z, Klionsky DJ. Mammalian autophagy: core molecular machinery and signaling regulation. Curr Opin Cell Biol. 2010; 22:124–31. [PubMed: 20034776]
21. Rubinsztein DC, Shpilka T, Elazar Z. Mechanisms of autophagosome biogenesis. Curr Biol. 2012; 22:R29–34. [PubMed: 22240478]
22. Muller S, Dennemarker J, Reinheckel T. Specific functions of lysosomal proteases in endocytic and autophagic pathways. Biochim Biophys Acta. 2012; 1824:34–43. [PubMed: 21767668]
23. Lee J, Giordano S, Zhang J. Autophagy, mitochondria and oxidative stress: cross-talk and redox signalling. Biochem J. 2012; 441:523–40. [PubMed: 22187934]
24. Lamark T, Johansen T. Aggrephagy: selective disposal of protein aggregates by macroautophagy. Int J Cell Biol. 2012; 2012:736–905.
25. Kaushik S, Cuervo AM. Chaperone-mediated autophagy: a unique way to enter the lysosome world. Trends Cell Biol. 2012; 22:407–17. [PubMed: 22748206]
26. Husnjak K, Dikic I. Ubiquitin-Binding Proteins: Decoders of Ubiquitin-Mediated Cellular Functions. Annu Rev Biochem. 2012
27. Tokuriki N, Tawfik DS. Protein dynamism and evolvability. Science. 2009; 324:203–7. [PubMed: 19359577]
28. Kikis EA, Gidalevitz T, Morimoto RI. Protein homeostasis in models of aging and age-related conformational disease. Adv Exp Med Biol. 2010; 694:138–59. [PubMed: 20886762]
29. Fox GE, et al. The phylogeny of prokaryotes. Science. 1980; 209:457–63. [PubMed: 6771870]

30. Balch WE, Magrum LJ, Fox GE, Wolfe RS, Woese CR. An ancient divergence among the bacteria. J Mol Evol. 1977; 9:305–11. [PubMed: 408502]
31. Darwin, C. The Voyage of the Beagle. 1856.
32. Roberts E, Sethi A, Montoya J, Woese CR, Luthey-Schulten Z. Molecular signatures of ribosomal evolution. Proc Natl Acad Sci U S A. 2008; 105:13953–8. [PubMed: 18768810]
33. Vetsigian K, Woese C, Goldenfeld N. Collective evolution and the genetic code. Proc Natl Acad Sci U S A. 2006; 103:10696–701. [PubMed: 16818880]
34. Horwich AL, Fenton WA, Chapman E, Farr GW. Two families of chaperonin: physiology and mechanism. Annu Rev Cell Dev Biol. 2007; 23:115–45. [PubMed: 17489689]
35. Margittai E, Sitia R. Oxidative protein folding in the secretory pathway and redox signaling across compartments and cells. Traffic. 2011; 12:1–8. [PubMed: 20716108]
36. Vilchez D, et al. Increased proteasome activity in human embryonic stem cells is regulated by PSMD11. Nature. 2012; 489:304–8. [PubMed: 22972301]
37. Hutt DM, Powers ET, Balch WE. The proteostasis boundary in misfolding diseases of membrane traffic. FEBS Lett. 2009; 583:2639–46. [PubMed: 19708088]
38. Powers ET, Balch WE. Protein folding: Protection from the outside. Nature. 2011; 471:42–3. [PubMed: 21368816]
39. Morimoto RI. The heat shock response: systems biology of proteotoxic stress in aging and disease. Cold Spring Harb Symp Quant Biol. 2011; 76:91–9. [PubMed: 22371371]
40. Ong DS, Kelly JW. Chemical and/or biological therapeutic strategies to ameliorate protein misfolding diseases. Curr Opin Cell Biol. 2011; 23:231–8. [PubMed: 21146391]
41. Pace CN, et al. Contribution of hydrophobic interactions to protein stability. J Mol Biol. 2011; 408:514–28. [PubMed: 21377472]
42. Bolen DW, Rose GD. Structure and energetics of the hydrogen-bonded backbone in protein folding. Annu Rev Biochem. 2008; 77:339–62. [PubMed: 18518824]
43. Bartlett GJ, Choudhary A, Raines RT, Woolfson DN. n-->pi* interactions in proteins. Nat Chem Biol. 2010; 6:615–20. [PubMed: 20622857]
44. Goldberg AL. Protein degradation and protection against misfolded or damaged proteins. Nature. 2003; 426:895–9. [PubMed: 14685250]
45. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. Annu Rev Biochem. 2006; 75:333–66. [PubMed: 16756495]
46. Kelly JW. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. Curr Opin Struct Biol. 1998; 8:101–6. [PubMed: 9519302]
47. Ghosh K, Dill K. Cellular proteomes have broad distributions of protein stability. Biophys J. 2010; 99:3996–4002. [PubMed: 21156142]
48. Kultz D. Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. J Exp Biol. 2003; 206:3119–24. [PubMed: 12909693]
49. Akerfelt M, Morimoto RI, Sistonen L. Heat shock factors: integrators of cell stress, development and lifespan. Nat Rev Mol Cell Biol. 2010; 11:545–55. [PubMed: 20628411]
50. Pellegrino MW, Nargund AM, Haynes CM. Signaling the mitochondrial unfolded protein response. Biochim Biophys Acta. 2012
51. Nordhues A, Miller SM, Muhlhaus T, Schroda M. New insights into the roles of molecular chaperones in Chlamydomonas and Volvox. Int Rev Cell Mol Biol. 2010; 285:75–113. [PubMed: 21035098]
52. Kultz D. Molecular and evolutionary basis of the cellular stress response. Annu Rev Physiol. 2005; 67:225–57. [PubMed: 15709958]
53. Chen B, Zhong D, Monteiro A. Comparative genomics and evolution of the HSP90 family of genes across all kingdoms of organisms. BMC Genomics. 2006; 7:156. [PubMed: 16780600]
54. Haynes CM, Ron D. The mitochondrial UPR - protecting organelle protein homeostasis. J Cell Sci. 2010; 123:3849–55. [PubMed: 21048161]
55. Rudiger S, Germeroth L, Schneider-Mergener J, Bukau B. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. EMBO J. 1997; 16:1501–7. [PubMed: 9130695]

56. Wang Q, Zhuravleva A, Gierasch LM. Exploring weak, transient protein--protein interactions in crowded in vivo environments by in-cell nuclear magnetic resonance spectroscopy. Biochemistry. 2011; 50:9225–36. [PubMed: 21942871]
57. Wang Z, Feng H, Landry SJ, Maxwell J, Gierasch LM. Basis of substrate binding by the chaperonin GroEL. Biochemistry. 1999; 38:12537–46. [PubMed: 10504222]
58. Scheibel T, Weikl T, Buchner J. Two chaperone sites in Hsp90 differing in substrate specificity and ATP dependence. Proc Natl Acad Sci U S A. 1998; 95:1495–9. [PubMed: 9465043]
59. Genevaux P, Georgopoulos C, Kelley WL. The Hsp70 chaperone machines of Escherichia coli: a paradigm for the repartition of chaperone functions. Mol Microbiol. 2007; 66:840–57. [PubMed: 17919282]
60. Brocchieri L, Conway de Macario E, Macario AJ. hsp70 genes in the human genome: Conservation and differentiation patterns predict a wide array of overlapping and specialized functions. BMC Evol Biol. 2008; 8:19. [PubMed: 18215318]
61. Chen B, Piel WH, Gui L, Bruford E, Monteiro A. The HSP90 family of genes in the human genome: insights into their divergence and evolution. Genomics. 2005; 86:627–37. [PubMed: 16269234]
62. Sreedhar AS, Kalmar E, Csermely P, Shen YF. Hsp90 isoforms: functions, expression and clinical importance. FEBS Lett. 2004; 562:11–5. [PubMed: 15069952]
63. Chen CY, Balch WE. The Hsp90 chaperone complex regulates GDI-dependent Rab recycling. Mol Biol Cell. 2006; 17:3494–507. [PubMed: 16687576]
64. Sakisaka T, Meerlo T, Matteson J, Plutner H, Balch WE. Rab-alphaGDI activity is regulated by a Hsp90 chaperone complex. EMBO J. 2002; 21:6125–35. [PubMed: 12426384]
65. Walter P, Ron D. The unfolded protein response: from stress pathway to homeostatic regulation. Science. 2011; 334:1081–6. [PubMed: 22116877]
66. Ostrovsky O, Eletto D, Makarewich C, Barton ER, Argon Y. Glucose regulated protein 94 is required for muscle differentiation through its control of the autocrine production of insulin-like growth factors. Biochim Biophys Acta. 2010; 1803:333–41. [PubMed: 19914304]
67. Weiss C, Bonshtien A, Farchi-Pisanty O, Vitlin A, Azem A. Cpn20: siamese twins of the chaperonin world. Plant Mol Biol. 2009; 69:227–38. [PubMed: 19031045]
68. Peltier JB, et al. The oligomeric stromal proteome of Arabidopsis thaliana chloroplasts. Mol Cell Proteomics. 2006; 5:114–33. [PubMed: 16207701]
69. Peng L, et al. A chaperonin subunit with unique structures is essential for folding of a specific substrate. PLoS Biol. 2011; 9:e1001040. [PubMed: 21483722]
70. Brewer JW, Hendershot LM. Building an antibody factory: a job for the unfolded protein response. Nat Immunol. 2005; 6:23–9. [PubMed: 15611778]
71. Back SH, Kang SW, Han J, Chung HT. Endoplasmic reticulum stress in the beta- cell pathogenesis of type 2 diabetes. Exp Diabetes Res. 2012; 2012:618396. [PubMed: 21915177]
72. Prahlad V, Cornelius T, Morimoto RI. Regulation of the cellular heat shock response in Caenorhabditis elegans by thermosensory neurons. Science. 2008; 320:811–4. [PubMed: 18467592]
73. Bouchecareilh M, Balch WE. Proteostasis, an emerging therapeutic paradigm for managing inflammatory airway stress disease. Curr Mol Med. 2012; 12:815–26. [PubMed: 22697348]
74. Demasi M, Laurindo FR. Physiological and Pathological Role of the Ubiquitin- Proteasome System in the Vascular Smooth Muscle Cell. Cardiovasc Res. 2012
75. Altun M, et al. Muscle wasting in aged, sarcopenic rats is associated with enhanced activity of the ubiquitin proteasome pathway. J Biol Chem. 2010; 285:39597–608. [PubMed: 20940294]
76. Lund PA. Multiple chaperonins in bacteria--why so many? FEMS Microbiol Rev. 2009; 33:785–800. [PubMed: 19416363]
77. Williams TA, Fares MA. The effect of chaperonin buffering on protein evolution. Genome Biol Evol. 2010; 2:609–19. [PubMed: 20660109]
78. Fujiwara K, Ishihama Y, Nakahigashi K, Soga T, Taguchi H. A systematic survey of in vivo obligate chaperonin-dependent substrates. EMBO J. 2010; 29:1552–64. [PubMed: 20360681]

79. Macario AJ, Brocchieri L, Shenoy AR, Conway de Macario E. Evolution of a protein-folding machine: genomic and evolutionary analyses reveal three lineages of the archaeal hsp70(dnaK) gene. J Mol Evol. 2006; 63:74–86. [PubMed: 16788741]
80. Vorderwulbecke S, et al. Low temperature or GroEL/ES overproduction permits growth of Escherichia coli cells lacking trigger factor and DnaK. FEBS Lett. 2004; 559:181–7. [PubMed: 14960329]
81. Sawle L, Ghosh K. How do thermophilic proteins and proteomes withstand high temperature? Biophys J. 2011; 101:217–27. [PubMed: 21723832]
82. Richter K, Haslbeck M, Buchner J. The heat shock response: life on the verge of death. Mol Cell. 2010; 40:253–66. [PubMed: 20965420]
83. Pysz MA, et al. Transcriptional analysis of dynamic heat-shock response by the hyperthermophilic bacterium Thermotoga maritima. Extremophiles. 2004; 8:209–17. [PubMed: 14991425]
84. Boonyaratanakornkit BB, Miao LY, Clark DS. Transcriptional responses of the deep-sea hyperthermophile Methanocaldococcus jannaschii under shifting extremes of temperature and pressure. Extremophiles. 2007; 11:495–503. [PubMed: 17332989]
85. Bosch TC, Krylow SM, Bode HR, Steele RE. Thermotolerance and synthesis of heat shock proteins: these responses are present in Hydra attenuata but absent in Hydra oligactis. Proc Natl Acad Sci U S A. 1988; 85:7927–31. [PubMed: 3186697]
86. Clark MS, Peck LS. HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review. Mar Genomics. 2009; 2:11–8. [PubMed: 21798167]
87. Bussieres N, Granger RJ. Estimation of water temperature of large lakes in cold climate regions during the period of strong coupling between water and air temperature fluctuations. Journal of Atmospheric and Oceanic Technology. 2007; 24:285–296.
88. McCarty JS, Walker GC. DnaK as a thermometer: threonine-199 is site of autophosphorylation and is critical for ATPase activity. Proc Natl Acad Sci U S A. 1991; 88:9513–7. [PubMed: 1835085]
89. Klostermeier D, Seidel R, Reinstein J. Functional properties of the molecular chaperone DnaK from Thermus thermophilus. J Mol Biol. 1998; 279:841–53. [PubMed: 9642065]
90. Place SP, Hofmann GE. Comparison of Hsc70 orthologs from polar and temperate notothenioid fishes: differences in prevention of aggregation and refolding of denatured proteins. Am J Physiol Regul Integr Comp Physiol. 2005; 288:R1195–202. [PubMed: 15637165]
91. Tomanek L. Variation in the heat shock response and its implication for predicting the effect of global climate change on species' biogeographical distribution ranges and metabolic costs. J Exp Biol. 2010; 213:971–9. [PubMed: 20190122]
92. Gidalevitz T, Krupinski T, Garcia S, Morimoto RI. Destabilizing protein polymorphisms in the genetic background direct phenotypic expression of mutant SOD1 toxicity. PLoS Genet. 2009; 5:e1000399. [PubMed: 19266020]
93. Brennecke T, Gellner K, Bosch TC. The lack of a stress response in Hydra oligactis is due to reduced hsp70 mRNA stability. Eur J Biochem. 1998; 255:703–9. [PubMed: 9738911]
94. Tian S, Haney RA, Feder ME. Phylogeny disambiguates the evolution of heat-shock cis-regulatory elements in Drosophila. PLoS One. 2010; 5:e10669. [PubMed: 20498853]
95. Garbuz DG, et al. Functional organization of hsp70 cluster in camel (Camelus dromedarius) and other mammals. PLoS One. 2011; 6:e27205. [PubMed: 22096537]
96. Garbuz DG, et al. Organization and evolution of hsp70 clusters strikingly differ in two species of Stratiomyidae (Diptera) inhabiting thermally contrasting environments. BMC Evol Biol. 2011; 11:74. [PubMed: 21426536]
97. Keller I, Seehausen O. Thermal adaptation and ecological speciation. Mol Ecol. 2012; 21:782–99. [PubMed: 22182048]
98. Rashkovetsky E, et al. Adaptive differentiation of thermotolerance in Drosophila along a microclimatic gradient. Heredity (Edinb). 2006; 96:353–9. [PubMed: 16552433]
99. Michalak P, et al. Genetic evidence for adaptation-driven incipient speciation of Drosophila melanogaster along a microclimatic contrast in "Evolution Canyon," Israel. Proc Natl Acad Sci U S A. 2001; 98:13195–200. [PubMed: 11687637]

100. Carmel J, Rashkovetsky E, Nevo E, Korol A. Differential expression of small heat shock protein genes Hsp23 and Hsp40, and heat shock gene Hsr-omega in fruit flies (Drosophila melanogaster) along a microclimatic gradient. J Hered. 2011; 102:593–603. [PubMed: 21505045]
101. Walser JC, Chen B, Feder ME. Heat-shock promoters: targets for evolution by P transposable elements in Drosophila. PLoS Genet. 2006; 2:e165. [PubMed: 17029562]
102. Jarosz DF, Lindquist S. Hsp90 and environmental stress transform the adaptive value of natural genetic variation. Science. 2010; 330:1820–4. [PubMed: 21205668]
103. Jarosz DF, Taipale M, Lindquist S. Protein homeostasis and the phenotypic manifestation of genetic diversity: principles and mechanisms. Annu Rev Genet. 2010; 44:189–216. [PubMed: 21047258]
104. Dai C, Whitesell L, Rogers AB, Lindquist S. Heat shock factor 1 is a powerful multifaceted modifier of carcinogenesis. Cell. 2007; 130:1005–18. [PubMed: 17889646]
105. Trepel J, Mollapour M, Giaccone G, Neckers L. Targeting the dynamic HSP90 complex in cancer. Nat Rev Cancer. 2010; 10:537–49. [PubMed: 20651736]
106. Santagata S, et al. Using the heat-shock response to discover anticancer compounds that target protein homeostasis. ACS Chem Biol. 2012; 7:340–9. [PubMed: 22050377]
107. Mendillo ML, et al. HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. Cell. 2012; 150:549–62. [PubMed: 22863008]
108. Geller R, Taguwa S, Frydman J. Broad action of Hsp90 as a host chaperone required for viral replication. Biochim Biophys Acta. 2012; 1823:698–706. [PubMed: 22154817]
109. Tokuriki N, Oldfield CJ, Uversky VN, Berezovsky IN, Tawfik DS. Do viral proteins possess unique biophysical features? Trends Biochem Sci. 2009; 34:53–9. [PubMed: 19062293]
110. Tokuriki N, Tawfik DS. Chaperonin overexpression promotes genetic variation and enzyme evolution. Nature. 2009; 459:668–73. [PubMed: 19494908]
111. Specchia V, et al. Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. Nature. 2010; 463:662–5. [PubMed: 20062045]
112. Abdelhakim AH, Oakes EC, Sauer RT, Baker TA. Unique contacts direct high-priority recognition of the tetrameric Mu transposase-DNA complex by the AAA+ unfoldase ClpX. Mol Cell. 2008; 30:39–50. [PubMed: 18406325]
113. Tariq M, Nussbaumer U, Chen Y, Beisel C, Paro R. Trithorax requires Hsp90 for maintenance of active chromatin at sites of gene expression. Proc Natl Acad Sci U S A. 2009; 106:1157–62. [PubMed: 19144915]
114. DeZwaan DC, Freeman BC. HSP90 manages the ends. Trends Biochem Sci. 2010; 35:384–91. [PubMed: 20236825]
115. Hageman J, et al. A DNAJB chaperone subfamily with HDAC-dependent activities suppresses toxic protein aggregation. Mol Cell. 2010; 37:355–69. [PubMed: 20159555]
116. Marinova Z, et al. Valproic acid induces functional heat-shock protein 70 via Class I histone deacetylase inhibition in cortical neurons: a potential role of Sp1 acetylation. J Neurochem. 2009; 111:976–87. [PubMed: 19765194]
117. Boyault C, Sadoul K, Pabion M, Khochbin S. HDAC6, at the crossroads between cytoskeleton and cell signaling by acetylation and ubiquitination. Oncogene. 2007; 26:5468–76. [PubMed: 17694087]
118. Westerheide SD, Anckar J, Stevens SM Jr. Sistonen L, Morimoto RI. Stress-inducible regulation of heat shock factor 1 by the deacetylase SIRT1. Science. 2009; 323:1063–6. [PubMed: 19229036]
119. Morimoto RI, Cuervo AM. Protein homeostasis and aging: taking care of proteins from the cradle to the grave. J Gerontol A Biol Sci Med Sci. 2009; 64:167–70. [PubMed: 19228787]
120. Boyault C, et al. HDAC6 controls major cell response pathways to cytotoxic accumulation of protein aggregates. Genes Dev. 2007; 21:2172–81. [PubMed: 17785525]
121. Araki K, Inaba K. Structure, mechanism, and evolution of Ero1 family enzymes. Antioxid Redox Signal. 2012; 16:790–9. [PubMed: 22145624]
122. Braakman I, Bulleid NJ. Protein folding and modification in the mammalian endoplasmic reticulum. Annu Rev Biochem. 2011; 80:71–99. [PubMed: 21495850]

123. Giudice A, Arra C, Turco MC. Review of molecular mechanisms involved in the activation of the Nrf2-ARE signaling pathway by chemopreventive agents. Methods Mol Biol. 2010; 647:37–74. [PubMed: 20694660]

124. Rabinowitz JD, White E. Autophagy and metabolism. Science. 2010; 330:1344–8. [PubMed: 21127245]

125. Carrano AC, Liu Z, Dillin A, Hunter T. A conserved ubiquitination pathway determines longevity in response to diet restriction. Nature. 2009; 460:396–9. [PubMed: 19553937]

126. Fadini GP, Ceolotto G, Pagnin E, de Kreutzenberg S, Avogaro A. At the crossroads of longevity and metabolism: the metabolic syndrome and lifespan determinant pathways. Aging Cell. 2011; 10:10–7. [PubMed: 21040402]

127. Durieux J, Wolff S, Dillin A. The cell-non-autonomous nature of electron transport chain-mediated longevity. Cell. 2011; 144:79–91. [PubMed: 21215371]

128. Piper RC, Lehner PJ. Endosomal transport via ubiquitination. Trends Cell Biol. 2011; 21:647–55. [PubMed: 21955996]

129. Weissman AM, Shabek N, Ciechanover A. The predator becomes the prey: regulating the ubiquitin system by ubiquitylation and degradation. Nat Rev Mol Cell Biol. 2011; 12:605–20. [PubMed: 21860393]

130. Balch WE, Roth DM, Hutt DM. Emergent properties of proteostasis in managing cystic fibrosis. Cold Spring Harb Perspect Biol. 2011; 3

131. Westermark P, Andersson A, Westermark GT. Islet amyloid polypeptide, islet amyloid, and diabetes mellitus. Physiol Rev. 2011; 91:795–826. [PubMed: 21742788]

132. Back SH, Kaufman RJ. Endoplasmic reticulum stress and type 2 diabetes. Annu Rev Biochem. 2012; 81:767–93. [PubMed: 22443930]

133. Whitesell L, Lindquist S. Inhibiting the transcription factor HSF1 as an anticancer strategy. Expert Opin Ther Targets. 2009; 13:469–78. [PubMed: 19335068]

134. Vinciguerra M, Musaro A, Rosenthal N. Regulation of muscle atrophy in aging and disease. Adv Exp Med Biol. 2010; 694:211–33. [PubMed: 20886766]

135. Musial K, Zwolinska D. Heat shock proteins in chronic kidney disease. Pediatr Nephrol. 2011; 26:1031–7. [PubMed: 21193931]

136. Wu J, Kaufman RJ. From acute ER stress to physiological roles of the Unfolded Protein Response. Cell Death Differ. 2006; 13:374–84. [PubMed: 16397578]

137. Wang M, Caetano-Anolles G. The evolutionary mechanics of domain organization in proteomes and the rise of modularity in the protein world. Structure. 2009; 17:66–78. [PubMed: 19141283]

138. Marsh JA, Teichmann SA. How do proteins gain new domains? Genome Biol. 2010; 11:126. [PubMed: 20630117]

139. Dekker C, Willison KR, Taylor WR. On the evolutionary origin of the chaperonins. Proteins. 2011; 79:1172–92. [PubMed: 21322032]

140. Sharma SK, De los Rios P, Christen P, Lustig A, Goloubinoff P. The kinetic parameters and energy cost of the Hsp70 chaperone as a polypeptide unfoldase. Nat. Chem. Biol. 2010; 6:914–20. [PubMed: 20953191]

141. Doyle SM, Wickner S. Hsp104 and ClpB: protein disaggregating machines. Trends Biochem Sci. 2009; 34:40–8. [PubMed: 19008106]

Box 1

Pathways managing proteostasis biology

Acetylation-deacetylation system

The acetylation-deactylation system comprises the extensive family of histone acetylases (HATs) and histone deacetylastes (HDACs), which control deacetylation of surface Lys resides. HDACs control activation of HSF1 through the acetylation status of Hsp90 by HDAC6[120] and control attenuation of the HSR by class III sirtuins through acetylation of HSF1[118]. The HAT-HDAC system may also affect the activity of a broad range of chaperones and co-chaperones.

Anti-oxidant response pathways

Anti-oxidant response signaling pathways are sensitive to both cytosolic and compartment-specific oxidative stress. NRF2 and protein disulfide isomerases (PDIs) as well as other accessory factors have an important role in mitigating the impact of oxidative stress on the folding environment[121-123].

Autophagy-lysosome pathways

Cytosolic aggregates and subcellular organelles are captured and degraded by autophagy-lysosome pathways. Autophagy can be induced by metabolome-linked mTOR signaling cascades, among others, and causes large substrates to be engulfed (such as protein aggregates, ER fragments and mitochondria) for delivery to the lysosomes, an acidic, protease-rich membrane trafficking compartment [20, 21, 124]. Chaperone-mediated autophagy (CMA) can deliver misfolded cytosolic proteins directly to the lysosome [25].

Heat shock response

The levels of cytosolic heat shock proteins (HSPs) in the cell are managed by the evolutionarily conserved heat shock response (HSR) pathway through the activity of the transcription factor HSF1[49]. This pathway involves rapid phosphorylation and trimerization of cytosolic monomeric HSF1 and HSF1 transit to the nucleus where it binds heat shock promoters, upregulating >1,000 proteins depending the nature of the stress response and the level of cytosolic misfolding. The heat shock response is generally rapidly attenuated (hours), but can remain elevated, an effect modulated by the acetylation-deacetylation pathways (see below), and the insulin growth factor 1 receptor (IGF1-R)-HSF1-linked FoxO transcription pathways[3] and caloric restriction pathways[125, 126].

Ubiquitin-proteasome system

Degradation in the cytosol of eukaryotes is managed by the ubiquitin proteasome system (UPS). This consists of well-characterized mono- and polyubiquitylation pathways facilitated by protein target specific ubiquitin ligases. These deliver misfolded proteins to the cytosolic proteasome complex, a conserved multi-subunit machine that deubiquitylates, unfolds and lyses the targeted protein in a concerted manner [18, 128, 129].

Unfolded protein response

The unfolded protein response (UPR) is a membrane trafficking compartment-associated signaling pathway that involves stress activated transcription factors generated through the Ire1, Perk and ATF6 pathways[65]. Folding factors and membrane trafficking components are upregulated and translation is downregulated to restore ER-associated and downstream misfolding stress. A specialized type of UPR occurs in mitochondria,

resulting generation of small peptides that signal upregulation of proteostasis components to alleviate the misfolding burden [54, 127].

Box 2

Proteostasis and human healthspan

Disruption of proteostasis is a central issue in many of the most challenging human diseases today, spanning from neurodegenerative diseases [28] and ageing [9], to inherited misfolding diseases (cystic fibrosis [130]), sporadic diseases (such as chronic obstructive pulmonary disease (COPD) [73] and type 2 diabetes[131, 132]), many forms of cancer [8, 107, 133], as well as the progression of systemic diseases of ageing including muscle pathologies [8, 134]_ENREF_8. Why do we have sporadic (environment-triggered) and/or inherited misfolding diseases given the potential depth and flexibility of a highly evolved proteostasis network to manage human physiology? Unfortunately, it is clear that there are trade-offs that proteostasis biology uses to optimize protein health to ensure immediate survival, thus compromising overal lifespan in response to both acute and, in particular, chronic challenges.

It is now apparent that the proteostasis network adapts the folding environment to a particular niche and lifestyle. Thus, a limiting step in maintaining healthspan is the need to optimize function within a rather narrow range of biological, chemical and biophysical constraints according to the rules of natural selection [4, 31]. In humans, this is exemplified by the many different cell types that put very specialized and dynamic demands on their proteostasis programmes. For example: cells in the renal medulla experience high concentrations of solutes that other cell types never encounter [135]; plasma cells have to secrete enormous quantities of antibodies [70]; β-cells manage insulin production in response to rapid changes in metabolic load[132, 136]; and stem cells and neurons need strict proteostasis management to maintain the function of the cell for an organism's entire lifespan- but for different purposes [36].

For each of these folding challenges there is a distinct multi-layered proteostasis network (Figure 1a). This suggests that the evolved human cell/species specific proteostasis biologies dispense with buffering capacity (general 'robustness') to carry out their specialized functions. Hence, they are susceptible to different types of physical and biological pathologies, such as occurs, for example, in response to high fat or high glucose diets (type 2 diabetes), environmental pollutants such as smoking (COPD) and ageing. In all cases, environmental change has outpaced the evolutionary capacity of the proteostasis network to make the appropriate adjustment in quinary state (the folding environment) in a time-frame that preserves survival, comprising fitness.

Box 3

A possible scenario for proteostasis evolution

Canonical HSPs bind to stretches of hydrophobic amino acids [13]. Because such stretches are more often exposed in non-native protein states- such as the unfolded, partially folded and misfolded states - chaperones can protect these states from degradation and aggregation. This preferential binding could have arisen during evolution through mutations that would have created hydrophobic patches on 'proto-chaperones' (see the figure), giving the organism that had them a substantial competitive advantage. Proto-chaperones would have been limited in function because they were bound to non-folded proteins without changing their conformation; that is, they would have been 'holdases'. To recover proteins from misfolded and aggregated states ('recoverases') or to prod proteins toward the folded state ('foldases'), the proto-chaperones would have required an energy input. The ability to use cellular energy stores could have arisen by a domain fusion or insertion event [137, 138] between a proto-chaperone and an ATPase. A scenario like this has been proposed for the origin of the HSP60s [139]. The next step in the evolution of energy-using proto-chaperones could have been to recruit co-factors, creating chaperone systems, a proteostasis network that more precisely manages client folding and function. The best example of this is HSP70, which early on acquired HSP40 and a nucleotide exchange factor as co-chaperones. Managing ATP delivery, utilization and removal provides a remarkable level of fine-tuning to folding management and small aggregates to be converted to the unfolded state and given another chance to fold [140]. The HSP70 system was further expanded in many organisms to include HSP100, which can disperse large aggregates that are inaccessible to the HSP70 system by itself [141]. The final step in the evolution of proteostasis could have been their integration with each other, for example, through the TPR-domain containing proteins such as HOP (heat shock organizing protein) that links HSP70 clients to HSP90.

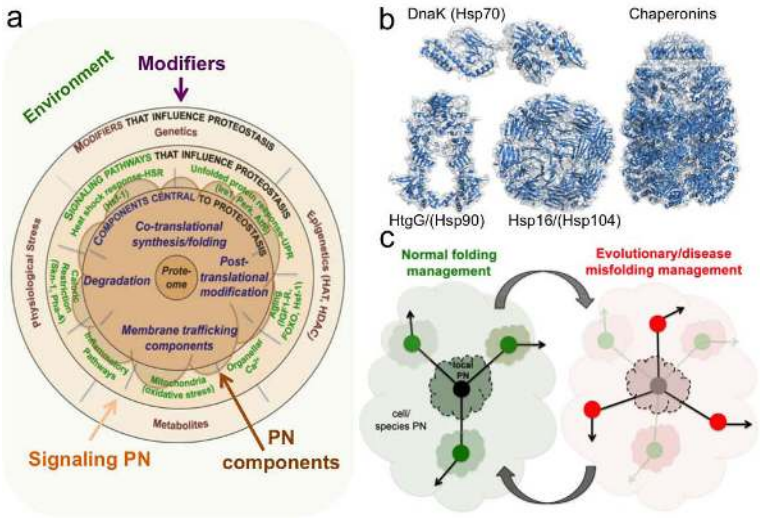

Figure 1. Proteostasis operates as a cloud

a | Illustrated is the hierarchy of proteostasis biology components and the molecular signaling and trafficking pathways that direct the response of a cell or a species to the environment [8]. The grey cloud icon highlights the local proteostasis network in a given cell type or species managing the fold. **b |** Illustrated is the impact of the quinary (5°) physiologic state (grey cloud) managed by proteostasis network on the generation the different secondary (2°, domain folded), ternary (intra-domain folded), and quaternary (multi-protein complex) structural states encoded by the primary polypeptide (1°) sequence encoded by the genome. The many different quinary state folding management environments generated by endomembrane compartments, including membrane trafficking compartments, mitochondria and chloroplasts among others, are illustrated by a blue oval cloud. **c |** Proteostasis serves as a folding buffer that surrounds every protein to manage the biological protein fold in evolution, and in health and disease. In the left hand panel, a protein (illustrated as black and green nodes in a protein interaction network) interacts with a local set of proteostasis network components (green cloud) to manage its normal function in a cell harboring a normal proteostasis network. In response to an inherited single nucleotide polymorphism or mutation and/or environmental stress (upper curved arrow), a protein misfolds (grey node) resulting in a disruption of binding of its normal partners (light green nodes) and new linkages to other proteins (pink nodes). Such misfolding cascades challenge proteostasis biology to fix the problem by changing its composition (pink cloud). Once fixed (lower arrow) (if fixable) the system returns to normal.

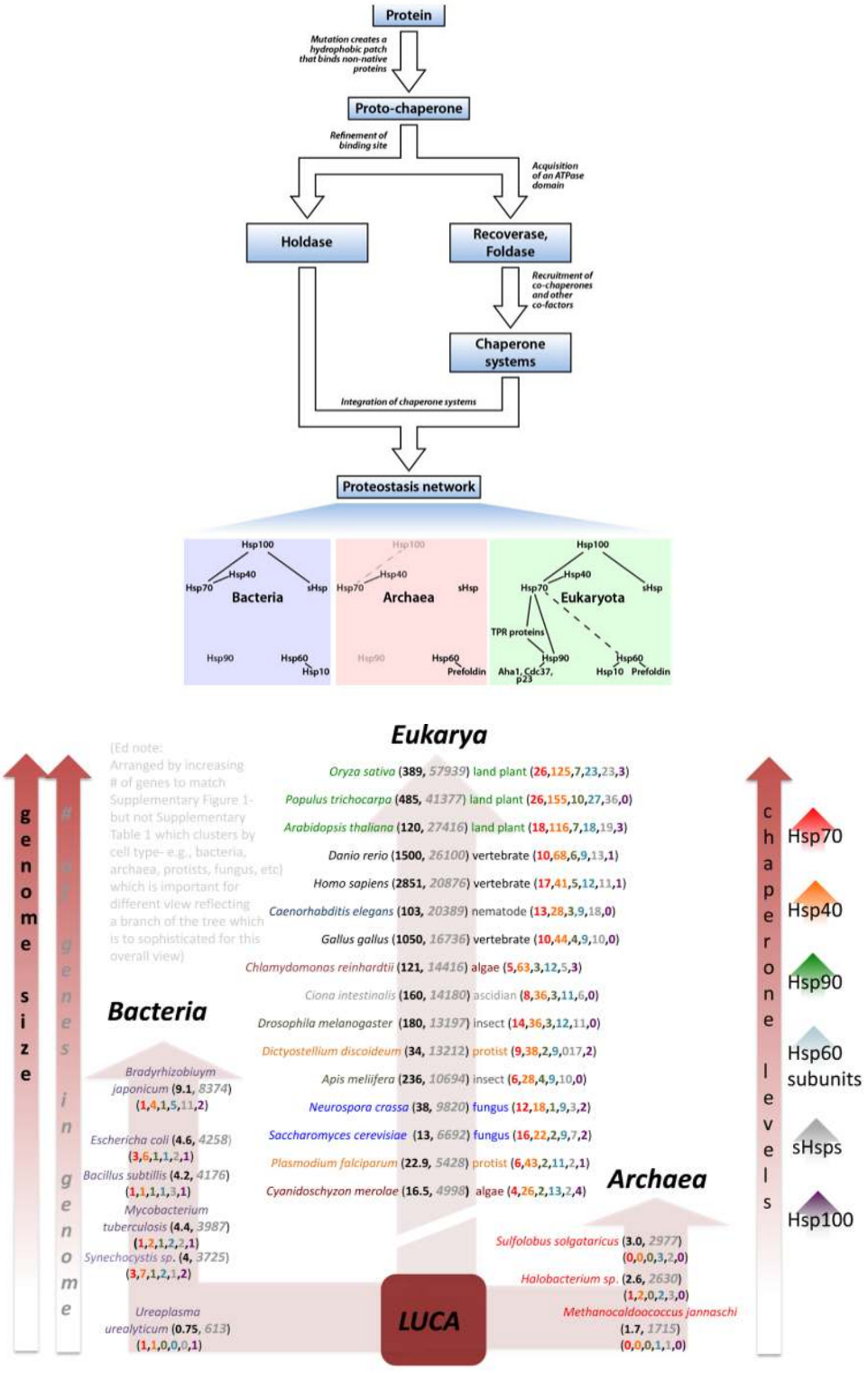

Figure 2. Diversity in origins of proteostasis

Shown is the evolution of HSPs among the extant three kingdoms of life- Bacteria, Archaea and Eukarya. Vertical arrows indicate increasing genome size and in number of genes expressed. The number of HSP40, HSP70 and HSP90 chaperones, of HSP60 chaperonin

family subunits and of sHSPs and HSP100s found in each species are indicated. See Supplemental Table S1 and Figure S1 for further details and references.

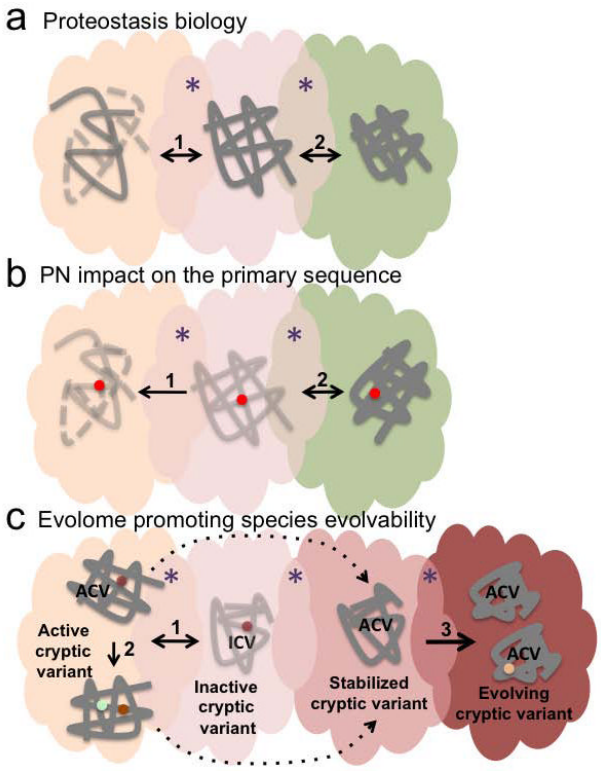

Figure 3. Role of proteostasis biology in evolvability

a | Proteostasis is normally optimized to support protein function (pink). However, it can also be limiting (fail to support folding) in response to metabolic, physiologic and environmental folding stress, and/or an inherited SNPs or mutations (orange). Alternatively, it can be bolstered (green) to support misfolding events in response to stress signaling pathways. **b |** A destabilizing mutation (red sphere) could not only fail to fold and lose function, but also challenge what was normally an 'acceptable' local folding environment to create one which has lost some aspect of its function in response to proteostasis stress[92] (path 1 to orange)[10, 39]. A protein carrying such a destabilizing mutation could be degraded because it is not protected by proteostasis. Alternatively, a rapid change accommodated by proteostasis biology in response to a need for survival can occur through either acute and/or chronic changes in proteostasis network composition via signaling pathways (path 2 to green), thereby providing some level of protection to the cell and species. **c |** An inactive cryptic variant protected by proteostasis biology in a species becomes an active cryptic variant in response to a stress challenge (path 1 to orange). Such a variant could be incorporated into the genome in subsequent generations if the change in function provided an immediately more favorable survival outcome. Alternatively, the selective pressure for use of the active cryptic variant could be sustained through either genetic (green sphere)) and/or epigenetic mechanisms that could modify the proteostasis network[5, 110], creating a new fold tolerance state (dotted paths to light maroon). Further evolution reflecting the new functional status of the cell would consolidate the proteostasis network function (path 3 to dark maroon). Thus, the proteostasis network strongly impacts the function of folding variants hidden in a population, and their short and long-term management by the proteostasis contributes to species evolability. Destabilized proteins are depicted as more transparent.