Chapter 1
Analysis of Chaperone Network Throughput

Craig Lawless and Simon J. Hubbard

Abstract The regulation of protein folding is an important aspect of systems biology that is often overlooked in the modern age of post-genomics. Although the transcriptome and proteome can now be relatively easily quantified, the protein complement in a cell must also be properly folded and delivered to the cognate site of action in order to carry out its function. To understand how a eukaryotic cell can accomplish this task requires an understanding of how the cell's chaperone complement acts to mediate the folding of its substrates. In this chapter, we examine and combine the data available from recent landmark studies to measure the chaperone interactome (the "chaperome") in the model eukaryote Baker's yeast with recent attempts to quantify the levels of yeast proteins. This computational analysis leads to an independent in silico assessment of the workload placed upon the chaperones in a cell, and shows there is a direct relationship between chaperone abundance and properties of their targets. By further considering protein turnover data, we are able to consider the folding flux passing through individual chaperones and chaperone groups, enabling a revaluation of the workload placed upon them, which we estimate exceeds 60 % of the cell's protein complement. We also cluster chaperones into coherent groups based on a filtered set of targets. These clusters reproduce some well-known features of the chaperone classes, as well as showing biases in subcellular location of the chaperone targets by factoring in the flux. These integrated analyses show how systems approaches can shed light on proteostasis defined by throughput in the chaperone network.

1 Introduction

Quantitative proteomics is one of the most rapidly advancing fields in the post genomic era. Arguably, it has lagged behind the field of transcriptomics, via microarrays or more recently next generation sequencing (ribonucleic acid, RNA-seq),

S. J. Hubbard (✉) · C. Lawless
Faculty of Life Sciences, University of Manchester, Manchester, UK
e-mail: simon.hubbard@manchester.ac.uk

C. Lawless
e-mail: craig.lawless@manchester.ac.uk

W. A. Houry (ed.), *The Molecular Chaperones Interaction Networks in Protein Folding and Degradation,* Interactomics and Systems Biology, DOI 10.1007/978-1-4939-1130-1_1,

as it has struggled to describe the complement of proteins expressed in a cell in a "genome-wide" fashion. Also, and more importantly, mass spectrometry based proteomics is not inherently quantitative, whereas transcriptomics is. This is a crucial issue for the biologist wishing to understand a process or pathway at the systems level, since this usually requires knowledge of the levels of some of the components, either to parameterize models or to assess the merits of predictions which point to quantitative changes. Moreover, proteins are usually the primary molecules responsible for the delivery of biological function, and the existence of a transcript is not necessarily a guarantee of the presence of a folded, functional protein. Hence, there are many good reasons to want to study the proteome in a comprehensive and quantitative fashion, despite the challenges in doing so.

Most modern proteomics relies on mass spectrometry as the underpinning analytical technology. Advances in instrumentation, chromatography, and allied informatics support the identification and measurement of abundance of an increasingly larger fraction of the protein complement derived from a tissue or cellular context. Quantitation is usually achieved as a relative measure by a variety of techniques that use labelled or label-free approaches to estimate changes in the analytical signal between two samples. These quantitative measurements can then aide our understanding of the mechanisms by which gene products are regulated and organized in order to elicit their cellular effects. Such data is important for generating systems-level models of an organism or a functional pathway, which encompass the interactions of the genome, transcriptome, proteome, and metabolome to describe cellular regulation and reactions to environmental stress. It is this interdisciplinary "systems biology" approach that has allowed the expansion of the relatively simple "central dogma" termed by Crick where "DNA makes RNA makes protein," to complex systems level models that incorporate multiple isoforms and interactions.

Recent studies of the model organism *Saccharomyces cerevisiae*, Baker's yeast, have exemplified this paradigm, providing data for the entire transcriptome [1], interactome [2, 3], translational control rates [4], protein localisation [5], and protein turnover rates [6, 7]. In addition, yeast has also been chosen for several proteome-wide quantitative studies [8–10]. Integration of such data has been used to build a metabolic model in yeast, using glycolysis as proof of principle [11, 12]. Advances have also been made for mammalian systems where transcription and translation data has been integrated with both RNA and protein turnover to develop whole genome models [13], which support global estimates of the relative importance of translational control in regulating gene expression.

Although excellent progress has been made, systems biology still faces many challenges when trying to integrate the vast information that is required to build suitable models to adequately define cellular regulation. Indeed, although we can now measure protein levels relatively routinely, the final and proper function of a protein (or protein complex) requires that it has the correct structure and is localized to the correct part of the cell. The subset of proteins assigned this responsibility are the chaperones, which facilitate correct protein folding, help recognize mis-folding, and prevent protein aggregation within the cell. In eukaryotes, chaperones act in a translation-coupled mechanism and recognize nascent polypeptide chains, thereby ensuring correct de novo folding as translation progresses [14]. Once translated and correctly folded, these proteins are then passed through the general chaperone network and are transported to their correct subcellular localization [15]. Given their critical role in the general maintenance of proteostatsis, many chaperones are heavily involved in stress responses, such as the aptly named heat shock proteins (HSPs), due to their thermo-reactive response [16].

In yeast, there are 63 proteins commonly characterized as chaperones, which have been the subject of a variety of studies [14, 17, 18]. More recently, the focus has moved to a more global view of chaperone function. Houry and colleagues have undertaken a comprehensive affinity purification analysis coupled with mass spectrometry in an attempt to define protein-protein interactions (PPIs) between yeast chaperones and their protein targets [19]. This has expanded on previous network-based analyses that have revealed two distinct networks; de novo protein folding and stress response [14]. Further work on the chaperone interaction network from Gong et al. [19] focused on clustering chaperones based on their target interactions, uncovering 10 chaperone modules that share common features among chaperone targets such as evolutionary rates and expression levels [20].

However, as already suggested, one dimension that must also be considered is quantitation; individual proteins are present in different concentrations in the cell, and therefore the workload placed on individual chaperones will vary depending on the abundance of the substrate/targets. Hence, one should also factor in the target workload (or flux) of each chaperone (or chaperone complex) in terms of protein abundance, or more formally, the total synthesis rate of each protein which can be estimated from protein turnover rates at steady state. To this end, we extend upon previous work that has attempted to characterize the yeast proteome using quantitative approaches, by characterizing the level of the chaperones using the QconCAT approach [21] and integrating other publicly available quantitative datasets, including protein turnover [6]. The QconCAT approach uses stable isotopic heavy-labelled peptide surrogates which are used as an internal standard for specific yeast proteins to provide absolute quantitation via mass spectrometry (see next section for more details). We show that there is a correlation between the abundance of chaperones and the target "folding" workload based on the number of targets and the target abundance. This is expanded further to estimate folding flux through each chaperone (and chaperone complex) using protein turnover rates. We describe these findings in the context of biological annotation which include subcellular localization and protein essentiality.

2 Chaperone Quantification

As part of a larger project to quantify the proteome of *S. cerevisiae* using QconCAT technology [21, 22], three QconCATs were designed to quantify the 63 known chaperones in yeast. Each QconCAT was constructed as a concatamer of heavy-labelled surrogate peptides where by each chaperone protein is represented by two of these target Q-peptides. These recombinant proteins were expressed in *E. coli* grown in media containing stable isotope-labelled arginine and lysine to produce a heavy-labelled QconCAT. The "heavy" QconCAT was used as a reference and spiked-in in known amounts to the yeast protein sample to enable peptide quantification using a selective reaction monitoring (SRM) targeted approach. Each peptide was targeted by three transitions for both the light (target) peptide and the heavy (reference) peptide. Quantification values were calculated from the ratio between the heavy and light extracted ion chromatograms (XICs), averaged across four biological replicates to obtain peptide copies per cell (cpc). Peptide ratios were acquired using the mProphet pipeline [23], which provides false discovery rate (FDR) estimates for the sets of transitions used for each peptide and therefore an estimate of reliability. Final protein abundances were taken as either the average of the peptide quantification values if they were in agreement, the minimum of the two peptides if the higher abundant peptide contained tryptophan, or the maximum peptide value for the remaining cases. The protein quantification values were then assigned into one of three classes; type A proteins where acceptable data above noise for both the target and reference peptides is obtained, type B where only the reference peptide gave acceptable data, and type C where neither gave acceptable data. In the case of type B proteins, we are still able to provide an upper limit of quantification as we know the concentration of the heavy-labelled peptide surrogate and the minimum spike-in level where it is detected in the XIC.

Using our targeted quantification approach, we obtained abundance values for 51 of the 63 chaperones in yeast [24], over a dynamic range covering three orders of magnitude, from 250 to 440,000 cpc as shown in Fig. 1.1. This list contains some chaperones that have eluded previous epitope-tagging approaches, including those that are part of the TRiC/CCT complex, and some proteins that have escaped both label-mediated and label-free MS-based studies. We ascribe our improved ability to quantify these particular proteins to the targeted nature of the QconCAT-SRM approach, which not only eliminates ambiguity by selecting unique peptides and associated productions but is credited with the greatest sensitivity in targeting low abundance proteins [25]. While many proteins are open to quantitation via label-free means to a low level, we believe that targeted approaches are still the gold standard for accurate absolute quantitation of proteins, and in our hands are able to measure down to around 100 cpc with coefficient of variation values (CVs) across replicates routinely less than 20 %.

3 Chaperone Target Quantitation

At present, a complete set of absolute quantitation values for yeast proteins derived via SRM- targeted mass spectrometry is not available. However there are a wide range of proteome-wide quantification data sets in yeast derived from other approaches, catalogued in the protein abundance database PaxDb [26]. We have used this useful data resource for our studies. PaxDb supplies all protein abundance

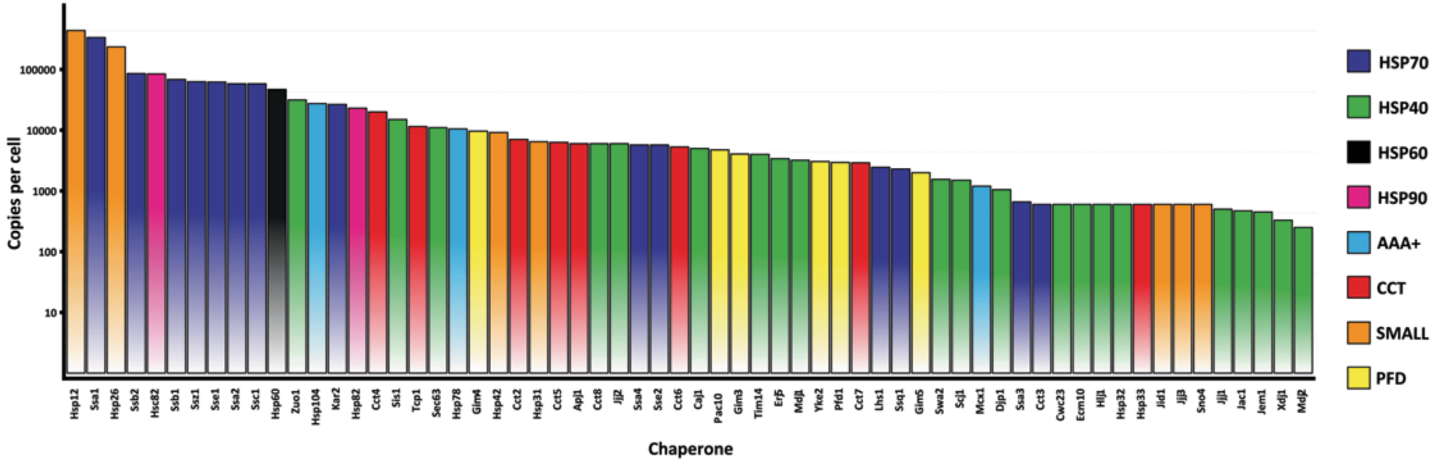

Fig. 1.1 Protein abundance estimates of yeast chaperones in yeast measured using the QconCAT-SRM method. Abundance values range from ~250 copies per cell (*cpc*) to >400,000 cpc. Bars are colored according to the chaperone class they belong to, shown in the accompanying key

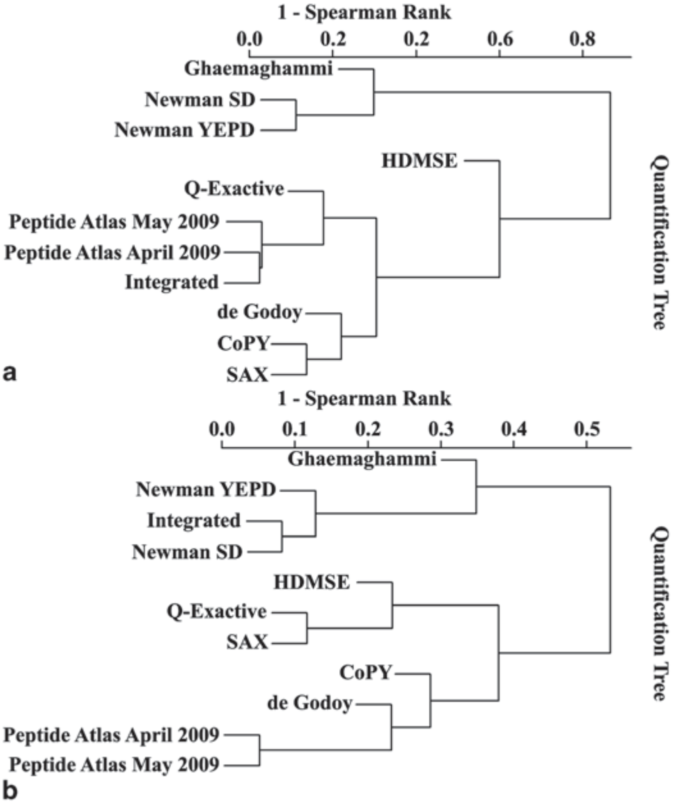

Fig. 1.2 **a** Dendrograms highlighting the differences between protein quantitation datasets for chaperones only and **b** for the 1000 proteins acquired via QconCAT-SRM on an internal project to date. External datasets were taken from the PaxDb database [26], where all protein values were normalized to parts per million (*ppm*) values

values in consistent units. Since protein quantitative values from different studies vary in nature, it is necessary to normalize relative or absolute values to parts per million (ppm), permitting comparisons between each of the methods. In order for our chaperone QconCAT abundance values to be compared to these other approaches they also need to be converted to ppm from cpc. To do this, we made the assumption of 60 million protein molecules per yeast cell and used this value to convert to ppm. While different analytical methods demonstrate positive correlations one to another (c.f. Brownridge et al. [22]), a comparison of all the quantitative methods with each other including our QconCAT approach and label-free acquisitions for the same yeast samples (HDMS^E, Q-Exactive, and SAX), shows considerable discrepancy between the techniques; both in the case of just the chaperones (Fig. 1.2a) and the ~ 1000 proteins quantified by our project using QconCATs so far (Fig. 1.2b).

The HDMS^E and Q-Exactive label-free data were obtained using a Top3 approach, which is described in more detail in [27]. Briefly, the HDMS^E was acquired on a Waters Synapt^{TM} G2, processed using Protein Lynx Global Server v2.5 using the Hi^3 approach [28], and the Q-Exactive was acquired on a Thermo Scientific Q-Exactive^{TM} instrument, processed using Progenesis (Nonlinear Dynamics) and Progenesis PostProcessor [29]. The SAX method was obtained by fractionating the sample using Off-Gel fractionation followed by label-free data acquisition for each fraction on a Thermo Scientific Q-Exactive and processed using Progenesis (Nonlinear Dynamics) and Progenesis PostProcessor [29]. We are indebted to our colleagues Dean Hammond, Philip Brownridge, and Rob Beynon at the University of Liverpool, UK for these data sets.

The most striking outcome of the clustering is the separation between mass spectrometry-based approaches and epitope-tagging-based approaches (Ghaemaghammi and Newman datasets), which co-cluster in unique clades. This is perhaps not surprising given the common analytical technique employed in the mass spectrometry grouping, which operates in a fundamentally different way to quantitative tagging methods that exploit antibody technology. Both have their advantages and disadvantages, but it is worth noting that the mass spectrometry community have recently thrown down the gauntlet to the antibody-based methods [30]. Within the mass spectrometry-based clade, we note there is an apparent separation of the label-free methods from the label-mediated approaches. These features are not perfectly conserved when considering the chaperone-only subset (Fig. 1.2a) compared to ~1000 proteins (Fig. 1.2b), but overall, it is noticeable that the major difference between all of the result sets appears to be methodological and not biological. This is exemplified by the GFP-tagging (green fluorescent protein) approach by Newman and colleagues [10] where despite being grown in both yeast extract peptone dextrose (YEPD) and sucrose deficient media (YMD)u, the two datasets cluster very closely together, both for the chaperone and 1000 protein subsets. This is in contrast to our label-free data acquisition on biologically identical yeast samples, but undertaken by three different label-free methods using different mass spectrometers (HDMSE, Q-Exactive, and SAX on Figs. 1.2a and 1.2b); these are not so highly correlated, particularly for the chaperone subset. These dendograms suggest that there is less inherent variance in biological samples than that currently obtained by comparison of two alternate quantitative methods. The absolute data derived from the QconCAT approach produces different abundance values to both the label-free approaches, adding further evidence to this, though interestingly, it clusters closely to another label-mediated data set using stable isotope labeling by amino acids in cell culture (SILAC) [8].

Our experience when comparing these datasets suggests that methodological variance contributes more strongly than biological variance to the protein abundance, and we urge caution in attempting to merge or analyze quantitative data derived from multiple techniques, or even on different instruments. Although this appears to be a somewhat disappointing outcome, it should be noted that the studies considered here have not used the same yeast strain cultured in identical conditions, and despite this, there is still a good correlation between even the most disparate of results. Nevertheless, these findings provide further evidence that there are still large variations between quantitative methodologies despite recent advancements in the field.

To examine this further, we extracted all the unique protein quantitation yeast datasets from PaxDb. These included both tagging based datasets from Ghaemmaghami

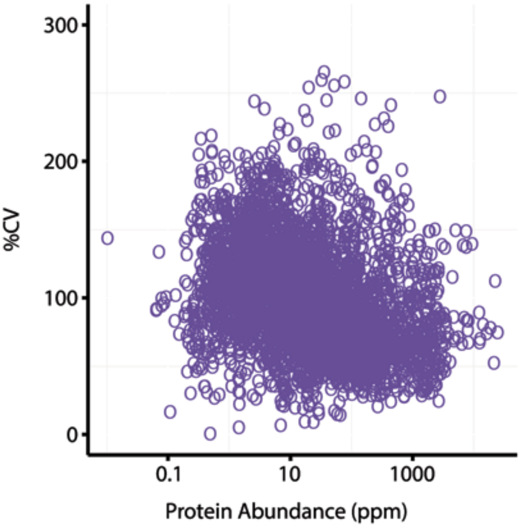

Fig. 1.3 Scatter plot of the coefficient of variances (% CV) of all yeast protein abundance measurements, in parts per million, from QconCAT-SRM, HDMS^E, Q-Exactive, SAX, and all PaxDb datasets plotted against a representative protein abundance. The average (mean) % CV across all datasets is 97 %. As can be seen, there is a modest, but significant, negative correlation ($R^2=0.11$, $P<2.2e^{-16}$) between protein abundance and % CV across independent determinations of protein abundance. This suggests there is a weak negative relationship between protein abundance and the ability of independent methods to quantify its abundance as there appears to be greater variance for low abundance proteins, as would be expected

et al. [9] and Newman et al. [10], the MS-based datasets from de Godoy et al. [8], Lu et al. [31], global proteome machine (GPM) [32], Peptide Atlas [33], and the PaxDb yeast-integrated dataset [26]. All datasets were collated by common protein in order to estimate the variation by method. We plot this data in Fig. 1.3 which shows the estimated CV percentage (CV%) plotted against protein abundance in ppm taken from the integrated PaxDB dataset, which is the most comprehensive, as it is a weighted aggregate dataset of the other quantitative datasets [26]. This striking plot shows that generally the CV % exceeds 100 % of the protein ppm value; equivalent to a 2x fold change in the protein abundance estimate. This might seem like a large value, but this represents a slightly artificial calculation using protein quantitative data spanning almost 10 years of research, albeit on a well-characterized model eukaryote. It suggests that using such a range of methods one can expect to estimate protein abundance fold changes to within two-fold accuracy. This contrasts with the typical 10–25 % CV cited by targeted proteomics studies using SRM approaches [22, 24, 34, 35] which appear more robust. Regardless, we suggest that quantitative proteomics still has work to do to define a definitive absolute quantitative dataset for the complete yeast proteome.

A closer look at specific chaperones and chaperone complexes across the methods highlights some of their shortcomings. Taking the chaperonin-containing TCP1 (CCT) complex as an example, we note that this set of chaperones is covered poorly by the epitope-tagging methods. In fact, CCT4 is the only protein of the eight members of this complex quantified in any of these tagging datasets. Conversely, the mass spectrometry based methods fared much better and consistently acquired abundances for seven or all eight of the CCT chaperones. This outcome is not surprising when taken in context with the structure and function of the CCT chaperones. The CCT proteins form an 8 protein heteromeric ring structure; all in a one-to-one stoichiometry [36]. It seems reasonable to assume that this structure could be disturbed or disrupted by epitope-tagging methods, when an additional protein is tagged on to any one of the CCT chaperones. Indeed, given that the CCT chaperone complex mediates the folding of a large majority of essential proteins with the cell [37], one would expect that any perturbations of this folding mechanism could lead to incorrect folding of these essential proteins and as a result be lethal to the cell. Interestingly, the CV calculated across all the quantitative methods is lowest for the CCT chaperones (44% CV, compared to 84–177% for other classes) when compared to the chaperone groups defined in [19], which is unsurprising given the one-to-one structural stoichiometry.

4 A Correlation Between Chaperone and Substrate Abundance

Chaperones operate by interacting with substrate proteins to ensure correct folding during protein synthesis, to stabilize protein structure during environmental stress, and to mediate their transport to their correct locations within the cell [38]. To identify chaperone substrate interactions, Gong and colleagues [19] undertook a proteome-wide affinity purification study in *S. cerevisiae* and produced a chaperone-interaction map containing 4340 candidate chaperone substrates. However, it is widely known that tandem affinity experiments have a tendency to contain false-positives due to nonspecific binding or common contaminants. Indeed, a variety of sophisticated informatics approaches have been generated to help deal with this [39–41], as well as a dedicated database for biologists to filter their data [42]. Here, to circumvent this, we used a simple consistency based approach and constructed a "high-quality" interaction dataset using three public PPI repositories (STRING [43], BioGRID [44], and MIPS [45]). All chaperone interactor pairs were obtained from each repository (in yeast), retaining only those pairings where the reciprocal (substrate-chaperone) interaction was also observed. For this study, we also excluded chaperone-chaperone interactions, focusing solely on chaperone-target interactions. Additionally, for STRING PPI pairs, a confidence score cutoff of 0.7 was used, retaining only those with superior scores. The three PPI datasets were then combined into a "high-quality" dataset where each interaction had to be present in at least two of the three datasets. This combined dataset contained 60 of the 63 known chaperones, 1711 chaperone substrates, 3649 interactions, and contained all but one of the reciprocal interactions identified by Gong et al. [19]. This qualitative interaction dataset, however, lacks protein abundances required to understand

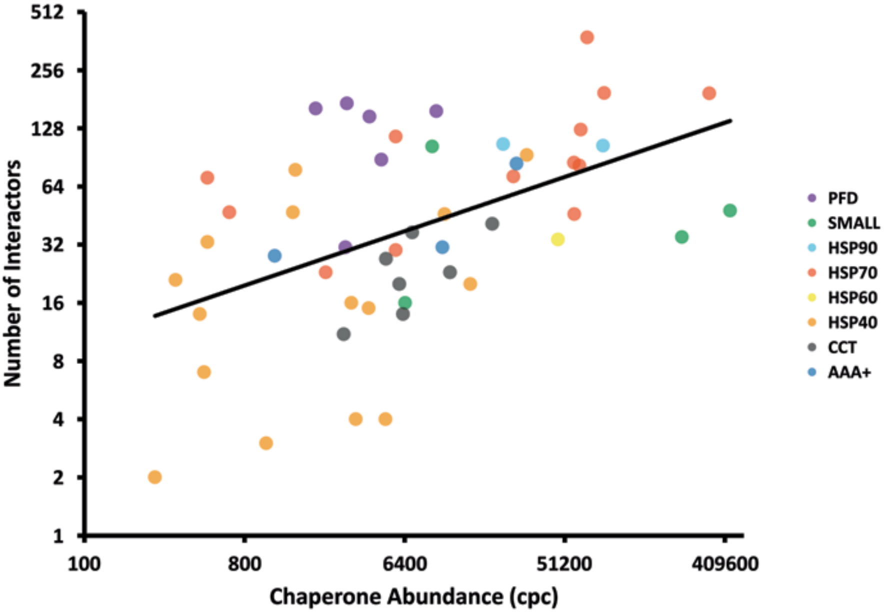

Fig. 1.4 Scatter plot of the number of chaperone interactors against chaperone abundance as measured by QconCAT-SRM mass spectrometry in our (CoPY) lab. There is a good linear correlation between the number of interactors and chaperone abundance ($R^2 = 0.25$, $P < 0.0002$), suggesting that chaperones with more clients are themselves more abundant in steady state conditions. The chaperones have been colored according to their chaperone class (as shown in the key) to visualize the spread of chaperone classes across the chaperone-interaction landscape. (Adapted from Brownridge [24])

the quantitative workload of each chaperone (and chaperone complex) and only provides an account of the different substrates that each chaperone interacts with.

In order to better characterize the workload placed on individual chaperones and chaperone classes, the abundances of their substrate protein targets in the cell needs to be taken into account. In the first instance, we examined chaperone workload with a simple measure, by considering the relationship between the abundance of the chaperone itself compared to the number of different substrate interactions it has, under the assumed logic that the more abundant chaperones would have a higher number of interactors. No prior correlation has been observed by Gong and colleagues [19]. As we reported previously [24], using our filtered interaction dataset we find a significant correlation using the QconCAT chaperone abundance data, as shown in Fig. 1.4, resulting in a Spearman Rank (R_{sp}) correlation of 0.49 ($P < 0.00002$). This correlation is also apparent across other quantitative datasets available from PaxDB, with R_{sp} ranging 0.08–0.54 (data not shown), and a significant and positive correlation is present regardless of whether we filter the interactome set or the quantitation data set selected from PaxDB. This supports the general hypothesis that those chaperones responsible for mediating the most folding in the cell are themselves generally highly abundant.

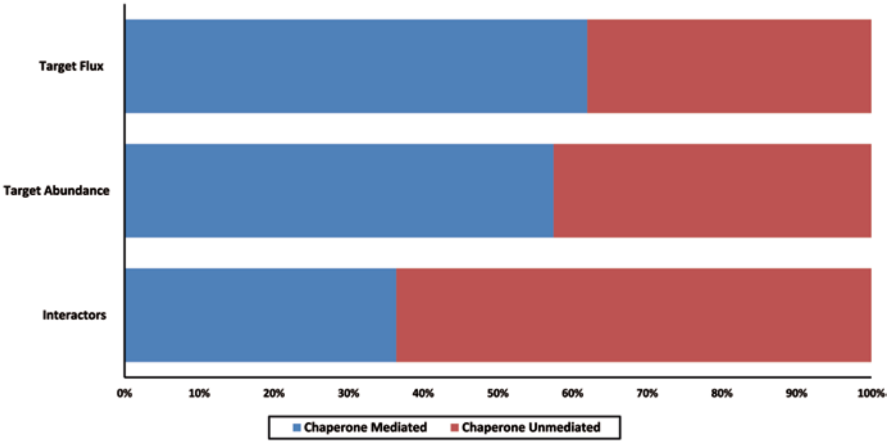

Fig. 1.5 Stacked bar charts showing the proportion of proteome folding mediated by chaperones, as determined by three different measures; the total number of interactors, the aggregated target abundance, and the aggregated target flux (molecules per minute) over the whole proteome. The latter is more representative of the true workload placed on chaperones in the cell, and is the largest relative fraction

Figure 1.4 also shows how breaking the chaperones into classes, as reported by Gong et al. [19], reveals general trends within the classes. For example, HSP40 chaperones are observed here to be low-abundance proteins with fewer interactions than many others. This is interesting, as they are generally considered as co-chaperones of the more abundant HSP70 chaperones, which bind to Hsp70 partners via the ribosome-associated complex (RAC), mediating the folding of the majority of nascent peptides, and thereby regulating the various roles of their Hsp70 partners [15, 38, 46]. Indeed, we observe that the HSP70 class (and similarly SMALL class) chaperones are more abundant and have more interactions, which ties in with their known promiscuity and substantive role [38].

Although looking at the number of different substrate interactions of each chaperone yields valuable information regarding their diversity and specificities, it does not provide a full picture of the true workload placed on any given chaperone. In order to better characterize chaperone workload, we need to have a measure of the total amount or "volume" of protein, in terms of total cpc, that is being mediated by a chaperone. The volume of protein (V_c) of a chaperone (c) can easily be estimated from the total abundance (CPC_n) of all n substrates for c, as shown below.

$$V_c = \sum_{1}^{n} cpc_n$$

Applying this methodology to consider the total cellular workload fundamentally changes the proportion of proteins mediated by chaperones when taking substrate volume and not just the number of substrates into account (Fig. 1.5). When considering just the number of interactors (taken from the unfiltered list from Gong and colleagues), this shows that only 36 % of all yeast proteins undergo folding regulation mediated by chaperones. However, when rescaling this calculation by incorporating protein abundance, using measurements taken from de Godoy et al. [8], this fraction dramatically increases to 57 % of the total protein volume in the cell. This reinforces the important role that chaperones play in proteostasis, since they are clearly responsible for the majority of protein folding by absolute molecular count.

Although this provides a better representation of chaperone workload, it ignores one important additional criterion; the workload of a chaperone cannot be expressed by a protein volume alone, as this is solely a static measure that ignores protein dynamics. Proteins are naturally synthesized and degraded at different rates. In order to calculate the folding workload, we also need to take into account the protein turnover rates of a chaperone's substrates. Fortunately, the majority of these have been measured using epitope-tagging techniques [6], and we were able to use their protein degradation rates (k_{deg}) to estimate protein synthesis rates (k_{syn}) for the substrates of each chaperone.

$$k_{syn} = cpc_n \times k_{deg}$$

The details are described more fully elsewhere [24], but briefly, this calculation presumes the cell is at steady state where protein synthesis and degradation rates are equal $\left(\frac{dCPC_n}{dt} = 0\right)$. From this, we can estimate the rate of synthesis of individual proteins and presume that this flux in molecules per unit time is the responsibility of individual chaperones. This allows us to estimate the total mediation flux on a per chaperone basis, which in turn can then be used to estimate the total workload, or folding flux (F_c), of a chaperone (c) in terms of molecules per minute per cell.

$$F_c = \sum_{1}^{n} k_{syn}$$

To circumvent the issue of substrates being mediated by multiple chaperones, the flux values were divided equally pro rata among each chaperone. Any missing values were replaced by the geometric mean across the entire turnover dataset; this was to ensure the mean k_{deg} of the dataset remained unchanged. To calculate the individual flux values F_c, we used data from the de Godoy estimation of yeast protein abundance taken from PaxDb [26], as opposed to our own QconCAT-SRM quantitation, since this currently does not fully cover the yeast proteome.

We recognize that cellular growth rates have not been formally taken into account in our model, primarily because the protein abundances and turnover rates used here have been obtained using different experimental yeast studies where strains and growth rates are not directly matched. However, to consider growth rate formally would simply add an error constant to the calculated k_{deg} values that is equivalent to the dilution rate of cells grown under controlled system (e.g., chemostat culture). Therefore, we believe the current calculations are still representative of the relative split of flux across the chaperone complement in yeast and despite the

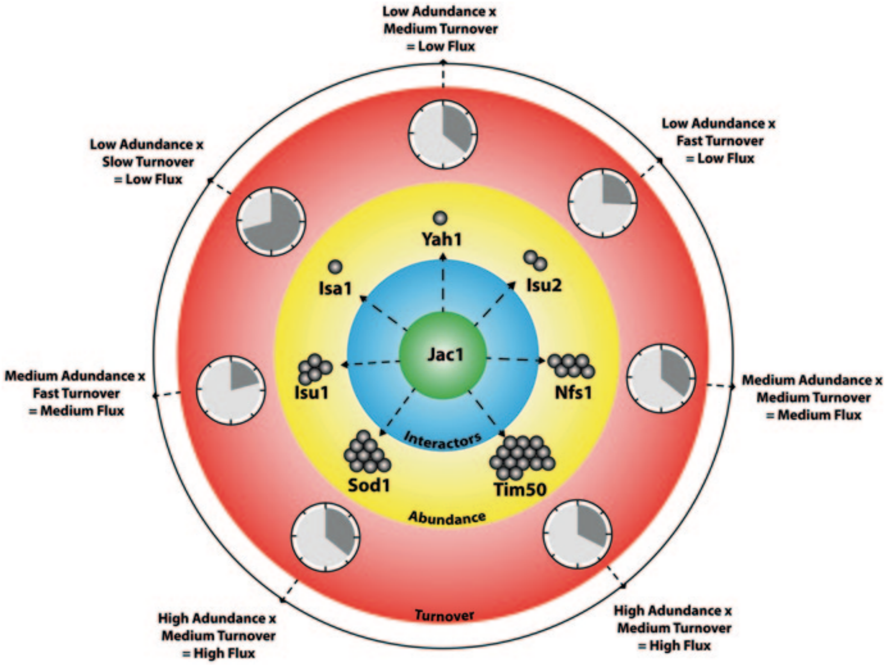

Fig. 1.6 Different levels of chaperone flux, using *Jac1* as an example. The concentric circles show an increasingly representative view of the true workload placed on a chaperone, which takes into account the number of chaperone interactors, the abundance of each target, and the turnover of the targets. Only by taking all three into account can the workload of the chaperone be properly estimated

assumptions made, these flux values are the most accurate proteome-wide estimates currently available.

Before considering the results of these calculations, it is worth putting this into a simple theoretical context describing how proteostasis can be maintained in the cell. This is schematized in Fig. 1.6, which shows different scenarios for regulating protein levels in the cell for a single example chaperone, *Jac1*. This chaperone has seven interactors in our filtered dataset, shown in the yellow concentric circle. The grey spheres broadly correspond to the protein abundance of each substrate with *Tim50* being the highest. As can be seen in the red concentric circle, this protein also apparently has one of the shorter half-lives (shown by the stopwatch graphic), and hence, there is a particularly high-folding flux required. *Tim50* is essential for protein translocation across the mitochondrial inner membrane and its loss can lead to programmed cell death [47, 48]. In contrast, *Isa1*, involved in iron-sulphur assembly displays a low flux, being a low-abundance protein with a relatively longer half-life, and unlike *Tim50* is not essential. The three “levels” of control (chaperone interactions, substrate abundance, and substrate turnover) all relate to the overall folding workload on the cell, and we argue that it is the integrated view on the outside of the circles in Fig. 1.6 which best represents the potential impact on chaperone function and workload.

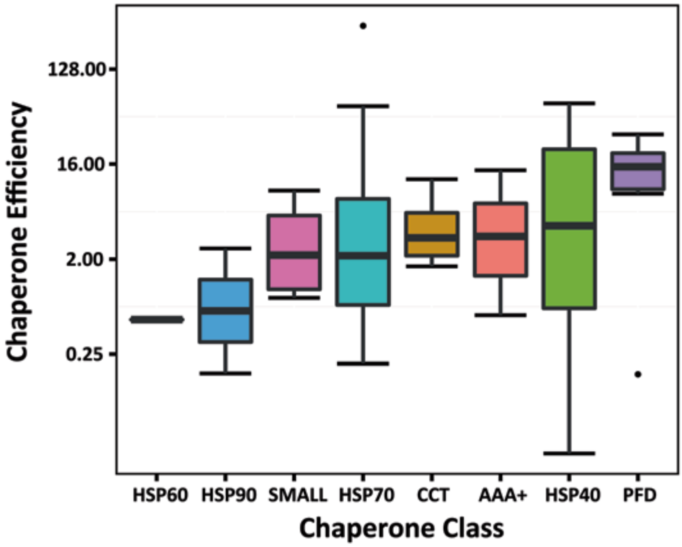

Fig. 1.7 Boxplot showing the range of mediation efficiencies of chaperones in each chaperone class. The efficiency of all chaperones was calculated and the grouped by chaperone class to show the distribution differences by class. The pre-foldin (*PFD*) class is shown to be generally the most efficient with a compact distribution. (Adapted from Brownridge [24])

When applying this theory to estimate substrate flux values (F_c) for each chaperone, we observe that the correlation with chaperone protein abundance generally increases; with R_{sp} ranging from 0.26 to 0.69 across the quantitative datasets [24]. As indicated by the strong Spearman correlation values, the more abundant proteins typically have a larger mediation workload in the cell, i.e., they fold/translocate more substrates per minute within the cell, as one might expect. Introducing the concept of chaperone "efficiency" for chaperone-substrate pairs and calculating a median efficiency for each chaperone class allows us to examine this data further. We estimate efficiency by dividing the substrate flux of a chaperone (F_c) by the abundance of the chaperone itself, reasoning that this effectively describes the number of substrate molecules per minute mediated by a single chaperone molecule. The ranges of efficiencies within each chaperone class are shown as a boxplot in Fig. 1.7. We see a broad range of values, but this suggests that most individual chaperone protein molecules are able to mediate the folding of more than one substrate molecule per minute. The prefoldin class appears to be particularly effective, acting as a major player in transporting proteins to the CCT complex, whose primary substrates are tubulin and actin. Here, the efficiency exceeds 10 molecules per chaperone per minute. The Hsp60 chaperones, which facilitate the transport and import of mitochondrial proteins, appear the least efficient by these criteria, perhaps because they are limited by mitochondrial import processes that do not want to be overloaded. Despite this apparently slow rate, Hsp60 is an essential gene [49].

5 Chaperone Clustering and Annotation

As calculated above, the substrate flux of the chaperones provides a good estimate of the workload diversity placed on them. To extend this analysis further, it is interesting to see how this workload is dispersed within the cell, in terms of function and localization of chaperone targets. It is already well recognized that many chaperones operate on substrates in, or destined for, different subcellular locations and we wished to examine this trend using the flux data. To do this, we used the high-quality interaction dataset to compare substrate specificities across the chaperones, by first grouping the chaperones in to common clusters based on their substrate profiles. The chaperones were assigned a vector specifying each substrate (of the total substrates within the interaction dataset) as either a target (1), or not a target (0), for that chaperone. The chaperone vectors were then submitted to a standard hierarchical clustering algorithm based on the binary distance between the vectors. We then manually assigned the chaperones to nine distinct clusters using a single threshold value, as shown in Fig. 1.8. This produces a set of clusters which closely resemble the chaperone family classifications as described in Gong et al. [19] although it is worth noting that these were arrived at completely independently, based only on chaperone target presence/absence. The clustering procedure carried out here is considerably simpler than that undertaken by Bogumil and colleagues [20], though we filter the interactome data prior to clustering, which derived a similar set of chaperone modules containing both chaperones and their substrate targets for common properties. Their modules were found to be enriched for common gene expression levels and evolutionary rates. Our simple procedure also generally recapitulates these well-known chaperone families and known relationships. For example, the CCT complex members (an octamer of subunits) cluster together and form a single clade, as do the prefoldins. We also observed clustering of coupled chaperone sets, such as the Hsp70 and Hsp40 groups which function as cognate partners. For example, the red clade in the bottom right of the wheel contains the Ssb1 and Ssz1 (two Hsp70s) and Zuo1 (a cognate Hsp40) that are tightly coupled to the ribosome. These clades form coherent groups of chaperones/co-chaperones that can be considered to be acting on common substrate groups.

Another noteworthy feature is the “cluster” of singletons. These are formed as a group that do not share particularly tight clustering with any other chaperones, mostly possessing relatively few targets. They are predominantly Hsp40s, which is not unexpected given their regulatory role in mediating Hsp70 function. Hence, one would expect them to have relatively few direct substrates and our filtering steps when applied to the entire Gong data set would be expected to remove many of the weaker, indirect interactions. Similarly, our procedures currently do not include chaperone-chaperone interactions. These findings are also supported by Fig. 1.4, which highlights the reduced numbers of chaperone clients for most Hsp40s.

We next considered the subcellular location of the chaperones targets, which were assigned to subcellular compartments using the Yeast Gene Ontology Slim from the *Saccharomyces cerevisiae* database (SGD) [50]. This allowed us to

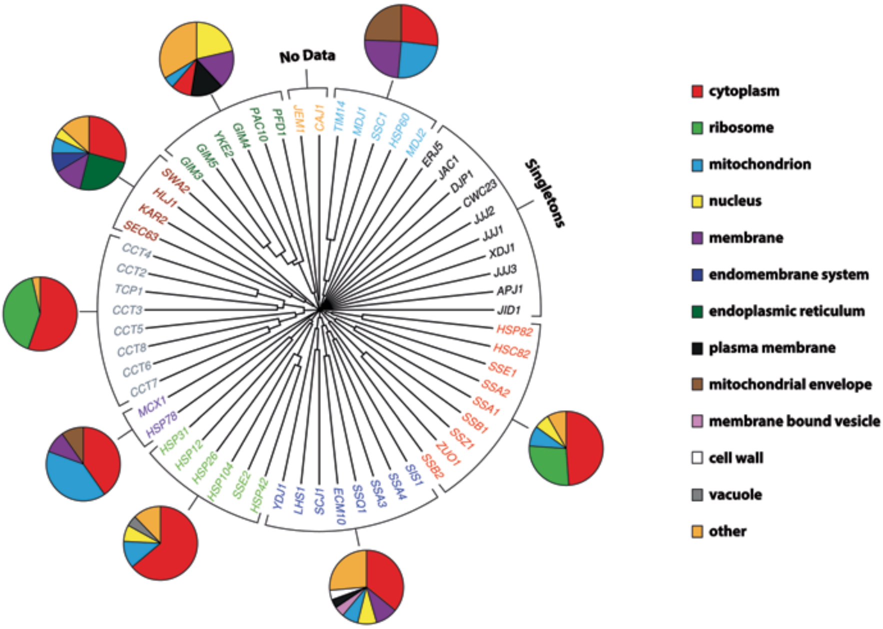

Fig. 1.8 Dendrogram wheel showing chaperones clustered by common target interactors. The chaperones were clustered using hierarchical clustering of the binary distance between their targets. Without any a priori information the clades formed are similar to the chaperone classes defined by Gong et al. [19], such as the eight CCT chaperones highlighted in *grey*. The heat shock protein (*HSP*)40 and HSP70 proteins generally cluster together (*dark blue, light blue,* and *maroon*) supporting the understanding that HSP40 are co-chaperones to HSP70 chaperones. In addition to the dendrogram wheel, each cluster is associated with a pie chart representing the subcellular localization of the cluster targets, weighted by the protein flux (as opposed to counts or protein abundance). These uncover diverse subcellular roles of chaperone mediation by cluster, such as the large ribosome compartment of the *grey* (CCT) and *red* (largely HSP70) clusters suggesting that these chaperones are heavily involved in mediation of proteins required for translation

estimate the spread of chaperone workload across subcellular locations based on the k_{syn} values for folding flux estimated previously. For any substrates that were assigned to more than one cellular localization, the k_{syn} values were split pro rata between the parent chaperones. Although previous analyses had not observed any strong trends in terms of subcellular bias for given chaperones or classes [19], we reasoned this might be different when considering the total folding flux rather than simply counting the number of different targets.

Figure 1.8 shows the target workload/flux for eight of the nine clusters and illustrates not only the differing roles between the clusters but also how the workload is distributed within a cluster. Indeed, there are several notable features of Fig. 1.8. As expected, some of the clusters show strong biases towards substrates destined for that location, such as the top blue cluster containing *Hsp60*, as well as other well-known mitochondrially active chaperones. We also see other strong biases, including the CCT complex which folds cytosolic actins and tubulins, and the *Sec63* cluster that are coupled to endoplasmic reticulum (ER)-associated folding and import.

6 Summary

The work showcased in this chapter highlights how systems biology approaches can shed light on the mechanisms by which the cell maintains proteostasis, and manages the complement of chaperones to ensure correct protein folding on a cellular basis. This is only possible through a lot of hard work and diligence on the part of many groups, to determine the chaperone interactome network via a variety of techniques, coupled to the determination of extensive abundance and turnover data sets. The latter is particularly important, since it represents a tractable way to measure protein-folding rates under steady state conditions and allow us to infer rates of protein synthesis. In other words, we consider this to be the folding flux or workload placed on chaperones within the network, which provides a more holistic understanding of the demand placed on these components than simply counting the number of different protein clients each chaperone has.

The clustered groups of chaperones also show important features that tally with expectation in the literature. We have probably only scratched the surface of the computational analyses that can be performed. To extend this further requires the construction of integrated models of proteostasis that can predict how the cell deals with these processes under perturbation, such as a stress (e.g., heat shock or oxidative stress). This is a current focus of research in our laboratory, coupled to quantitation of the chaperome players via mass spectrometry. We believe that the recent advances made in characterizing the chaperome network structure, coupled to the increasing capacity and decreasing cost of quantitative proteomics, will make this tractable and lead to a much better understanding of how cells manage their protein complement in changing environments.

Acknowledgements We would like to thank Robert Beynon, Philip Brownridge, Dean Hammond, and Stephen Holman from the University of Liverpool for providing the quantitative datasets used in the analysis. We would also like to acknowledge the Biotechnology and Biological Sciences Research Council (BBSRC) for the support of the CoPY project, via grants BB/G009112/1, and BB/G009058/1.

References

1. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. Science 320(5881):1344–1349. doi:10.1126/science.1158441
2. Gavin AC, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dumpelfeld B, Edelmann A, Heurtier MA, Hoffman V, Hoefert C, Klein K, Hudak M, Michon AM, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B, Bork P, Russell RB, Superti-Furga G (2006) Proteome survey reveals modularity of the yeast cell machinery. Nature 440(7084):631–636. doi:10.1038/nature04532
3. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro

A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P, Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, Collins SR, Chandran S, Haw R, Rilstone JJ, Gandi K, Thompson NJ, Musso G, St Onge P, Ghanny S, Lam MH, Butland G, Altaf-Ul AM, Kanaya S, Shilatifard A, O'Shea E, Weissman JS, Ingles CJ, Hughes TR, Parkinson J, Gerstein M, Wodak SJ, Emili A, Greenblatt JF (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. Nature 440(7084):637–643. doi:10.1038/nature04670

4. Vogel C, Silva GM, Marcotte EM (2011) Protein expression regulation under oxidative stress. Mol Cell Proteomics 10(12):M111.009217. doi:10.1074/mcp.M111.009217
5. Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK (2003) Global analysis of protein localization in budding yeast. Nature 425(6959):686–691. doi:10.1038/nature02026
6. Belle A, Tanay A, Bitincka L, Shamir R, O'Shea EK (2006) Quantification of protein half-lives in the budding yeast proteome. Proc Natl Acad Sci U S A 103(35):13004–13009. doi:10.1073/pnas.0605420103
7. Helbig AO, Daran-Lapujade P, van Maris AJA, de Hulster EAF, de Ridder D, Pronk JT, Heck AJR, Slijper M (2011) The diversity of protein turnover and abundance under nitrogen-limited steady-state conditions in *Saccharomyces cerevisiae*. Mol Biosyst 7(12):3316–3326. doi:10.1039/C1MB05250K
8. de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, Frohlich F, Walther TC, Mann M (2008) Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature 455(7217):1251–1254. doi:10.1038/nature07341
9. Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003) Global analysis of protein expression in yeast. Nature 425(6959):737–741. doi:10.1038/nature02046
10. Newman JR, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS (2006) Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. Nature 441(7095):840–846. doi:10.1038/nature04785
11. Smallbone K, Messiha HL, Carroll KM, Winder CL, Malys N, Dunn WB, Murabito E, Swainston N, Dada JO, Khan F, Pir P, Simeonidis E, Spasić I, Wishart J, Weichart D, Hayes NW, Jameson D, Broomhead DS, Oliver SG, Gaskell SJ, McCarthy JEG, Paton NW, Westerhoff HV, Kell DB, Mendes P (2013) A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. FEBS Lett 587(17):2832–2841. doi:10.1016/j.febslet.2013.06.043
12. Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Bluthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novere N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasic I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kirdar B, Penttila M, Klipp E, Palsson BO, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB (2008) A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. Nat Biotechnol 26(10):1155–1160. doi:10.1038/nbt1492
13. Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M (2011) Global quantification of mammalian gene expression control. Nature 473(7347):337–342. doi:10.1038/nature10098
14. Albanèse V, Yam AY-W, Baughman J, Parnot C, Frydman J (2006) Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. Cell 124(1):75–88. doi:10.1016/j.cell.2005.11.039
15. Vabulas RM, Raychaudhuri S, Hayer-Hartl M, Hartl FU (2010) Protein folding in the cytoplasm and the heat shock response. Cold Spring Harb Perspect Biol 2(12):a004390. doi:10.1101/cshperspect.a004390
16. Parsell DA, Lindquist S (1993) The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. Annu Rev Genet 27(1):437–496. doi:10.1146/annurev.ge.27.120193.002253

17. Trotter EW, Kao CM-F, Berenfeld L, Botstein D, Petsko GA, Gray JV (2002) Misfolded proteins are competent to mediate a subset of the responses to heat shock in *Saccharomyces cerevisiae*. J Biol Chem 277(47):44817–44825. doi:10.1074/jbc.M204686200
18. Siegers K, Bolter B, Schwarz JP, Bottcher UMK, Guha S, Hartl FU (2003) TRiC/CCT cooperates with different upstream chaperones in the folding of distinct protein classes. EMBO J 22(19):5230–5240
19. Gong Y, Kakihara Y, Krogan N, Greenblatt J, Emili A, Zhang Z, Houry WA (2009) An atlas of chaperone-protein interactions in *Saccharomyces cerevisiae*: implications to protein folding pathways in the cell. Mol Syst Biol 5:275. doi:10.1038/msb.2009.26
20. Bogumil D, Landan G, Ilhan J, Dagan T (2012) Chaperones divide yeast proteins into classes of expression level and evolutionary rate. Genome Biol Evol 4(5):618–625. doi:10.1093/gbe/evs025
21. Pratt JM, Simpson DM, Doherty MK, Rivers J, Gaskell SJ, Beynon RJ (2006) Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. Nat Protoc 1(2):1029–1043. doi:10.1038/nprot.2006.129
22. Brownridge P, Holman SW, Gaskell SJ, Grant CM, Harman VM, Hubbard SJ, Lanthaler K, Lawless C, O'Cualain R, Sims P, Watkins R, Beynon RJ (2011) Global absolute quantification of a proteome: challenges in the deployment of a QconCAT strategy. Proteomics 11(15):2957–2970. doi:10.1002/pmic.201100039
23. Reiter L, Rinner O, Picotti P, Huttenhain R, Beck M, Brusniak MY, Hengartner MO, Aebersold R (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. Nat Methods 8(5):430–435. doi:10.1038/nmeth.1584
24. Brownridge P, Lawless C, Payapilly AB, Lanthaler K, Holman SW, Harman VM, Grant CM, Beynon RJ, Hubbard SJ (2013) Quantitative analysis of chaperone network throughput in budding yeast. Proteomics 13(8):1276–1291. doi:10.1002/pmic.201200412
25. Picotti P, Rinner O, Stallmach R, Dautel F, Farrah T, Domon B, Wenschuh H, Aebersold R (2010) High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. Nat Methods 7(1):43–46. doi:10.1038/nmeth.1408
26. Wang M, Weiss M, Simonovic M, Haertinger G, Schrimpf SP, Hengartner MO, von Mering C (2012) PaxDb, a database of protein abundance averages across all three domains of life. Mol Cell Proteomics. doi:10.1074/mcp.O111.014704
27. Silva JC, Denny R, Dorschel CA, Gorenstein M, Kass IJ, Li G-Z, McKenna T, Nold MJ, Richardson K, Young P, Geromanos S (2005) Quantitative proteomic analysis by accurate mass retention time pairs. Anal Chem 77(7):2187–2200. doi:10.1021/ac048455k
28. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ (2006) Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. Mol Cell Proteomics 5(1):144–156. doi:10.1074/mcp.M500230-MCP200
29. Qi D, Brownridge P, Xia D, Mackay K, Gonzalez-Galarza FF, Kenyani J, Harman V, Beynon RJ, Jones AR (2012) A software toolkit and interface for performing stable isotope labeling and Top3 quantification using progenesis LC-MS. Omics 16:489–495. doi:10.1089/omi.2012.0042
30. Aebersold R, Burlingame AL, Bradshaw RA (2013) Western blots versus selected reaction monitoring assays: time to turn the tables? Mol Cell Proteomics 12(9):2381–2382. doi:10.1074/mcp.E113.031658
31. Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotech 25(1):117–124. doi:10.1038/nbt1270
32. Craig R, Cortens JP, Beavis RC (2004) Open source system for analyzing, validating, and storing protein identification data. J Proteome Res 3(6):1234–1242. doi:10.1021/pr049882h
33. Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J, Chen S, Eddes J, Loevenich SN, Aebersold R (2006) The PeptideAtlas project. Nucleic Acids Res 34(suppl 1):D655–D658. doi:10.1093/nar/gkj040

34. Anderson L, Hunter CL (2006) Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. Mol Cell Proteomics 5(4):573–588. doi:10.1074/mcp.M500331-MCP200

35. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham A-JL, Keshishian H, Hall SC, Allen S, Blackman RK, Borchers CH, Buck C, Cardasis HL, Cusack MP, Dodder NG, Gibson BW, Held JM, Hiltke T, Jackson A, Johansen EB, Kinsinger CR, Li J, Mesri M, Neubert TA, Niles RK, Pulsipher TC, Ransohoff D, Rodriguez H, Rudnick PA, Smith D, Tabb DL, Tegeler TJ, Variyath AM, Vega-Montoto LJ, Wahlander A, Waldemarson S, Wang M, Whiteaker JR, Zhao L, Anderson NL, Fisher SJ, Liebler DC, Paulovich AG, Regnier FE, Tempst P, Carr SA (2009) Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. Nat Biotech 27(7):633–641. doi:10.1038/nbt.1546

36. Dekker C, Roe SM, McCormack EA, Beuron F, Pearl LH, Willison KR (2011) The crystal structure of yeast CCT reveals intrinsic asymmetry of eukaryotic cytosolic chaperonins. EMBO J 30(15):3078–3090. doi:10.1038/emboj.2011.208

37. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. Nature 418(6896):387–391. doi:10.1038/nature00935

38. Kim YE, Hipp MS, Bracher A, Hayer-Hartl M, Ulrich Hartl F (2013) Molecular chaperone functions in protein folding and proteostasis. Annu Rev Biochem 82(1):323–355. doi:10.1146/annurev-biochem-060208-092442

39. Choi H, Larsen B, Lin Z-Y, Breitkreutz A, Mellacheruvu D, Fermin D, Qin ZS, Tyers M, Gingras A-C, Nesvizhskii AI (2011) SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat Methods 8(1):70–73. doi:10.1038/nmeth.1541

40. Sardiu ME, Cai Y, Jin J, Swanson SK, Conaway RC, Conaway JW, Florens L, Washburn MP (2008) Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. Proc Natl Acad Sci U S A 105(5):1454–1459. doi:10.1073/pnas.0706983105

41. Sowa ME, Bennett EJ, Gygi SP, Harper JW (2009) Defining the human deubiquitinating enzyme interaction landscape. Cell 138(2):389–403. doi:10.1016/j.cell.2009.04.042

42. Mellacheruvu D, Wright Z, Couzens AL, Lambert J-P, St-Denis NA, Li T, Miteva YV, Hauri S, Sardiu ME, Low TY, Halim VA, Bagshaw RD, Hubner NC, al-Hakim A, Bouchard A, Faubert D, Fermin D, Dunham WH, Goudreault M, Lin Z-Y, Badillo BG, Pawson T, Durocher D, Coulombe B, Aebersold R, Superti-Furga G, Colinge J, Heck AJR, Choi H, Gstaiger M, Mohammed S, Cristea IM, Bennett KL, Washburn MP, Raught B, Ewing RM, Gingras A-C, Nesvizhskii AI (2013) The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. Nat Methods 10(8):730–736. doi:10.1038/nmeth.2557

43. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, Jensen LJ, Mering CV (2011) The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res 39(suppl 1):D561–D568. doi:10.1093/nar/gkq973

44. Stark C, Breitkreutz B-J, Chatr-aryamontri A, Boucher L, Oughtred R, Livstone MS, Nixon J, Van Auken K, Wang X, Shi X, Reguly T, Rust JM, Winter A, Dolinski K, Tyers M (2010) The BioGRID interaction database: 2011 update. Nucleic Acids Res 39(suppl 1):D698–D704

45. Güldener U, Münsterkötter M, Oesterheld M, Pagel P, Ruepp A, Mewes H-W, Stümpflen V (2006) MPact: the MIPS protein interaction resource on yeast. Nucleic Acids Res 34(suppl 1):D436–D441

46. Kampinga HH, Craig EA (2010) The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat Rev Mol Cell Biol 11(8):579–592. doi:10.1038/nrm2941
47. Sugiyama S, Moritoh S, Furukawa Y, Mizuno T, Lim YM, Tsuda L, Nishida Y (2007) Involvement of the mitochondrial protein translocator component tim50 in growth, cell proliferation and the modulation of respiration in *Drosophila*. Genetics 176(2):927–936. doi:10.1534/genetics.107.072074
48. Guo Y, Cheong N, Zhang Z, De Rose R, Deng Y, Farber SA, Fernandes-Alnemri T, Alnemri ES (2004) Tim50, a component of the mitochondrial translocator, regulates mitochondrial integrity and cell death. J Biol Chem 279(23):24813–24825. doi:10.1074/jbc.M402049200
49. Cheng MY, Hartl FU, Martin J, Pollock RA, Kalousek F, Neupert W, Hallberg EM, Hallberg RL, Horwich AL (1989) Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria. Nature 337(6208):620–625. doi:10.1038/337620a0
50. Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, Fisk DG, Hirschman JE, Hitz BC, Karra K, Krieger CJ, Miyasato SR, Nash RS, Park J, Skrzypek MS, Simison M, Weng S, Wong ED (2012) *Saccharomyces* genome database: the genomics resource of budding yeast. Nucleic Acids Res 40(Database issue):D700–D705. doi:10.1093/nar/gkr1029