Kinetic analysis of protein stability reveals age-dependent degradation

Kinetic analysis of protein stability reveals age-dependent degradation

Erik McShane[1], Celine Sin[2], Henrik Zauber[1], Jonathan N. Wells[3], Neysan Donnelly[4], Xi Wang[1], Jingyi Hou[1], Wei Chen[1,7], Zuzana Storchova[4,5], Joseph A. Marsh[3], Angelo Valleriani[2] and Matthias Selbach[1,6,*]

1 Max Delbrück Center for Molecular Medicine, Robert-Rössle-Str.13, 13092 Berlin, Germany

2 Max Planck Institute of Colloids and Interfaces, Department of Theory and Bio-Systems, 14424 Potsdam, Germany

3 MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4, UK

4 Max-Planck-Institut für Biochemie, Am Klopferspitz 18, 82152 Martinsried, Germany

5 TU Kaiserslautern, Paul-Ehrlich Str. 24, 67663 Kaiserslautern, Germany

6 Charité - Universitätsmedizin Berlin, Germany

7 present address: Department of Biology, South University of Science and Technology of China, No. 1088, Xueyuan Rd., Xili, Nanshan District, Shenzhen, Guangdong 518055 China

* Lead Contact

Correspondence should be addressed to matthias.selbach@mdc-berlin.de

Abstract

Do young and old protein molecules have the same probability to be degraded? We addressed this question using metabolic pulse-chase labeling and quantitative mass spectrometry to obtain degradation profiles for thousands of proteins. We find that over 10% of proteins are degraded non-exponentially. Specifically, proteins are less stable in the first few hours of their life and stabilize with age. Degradation profiles are conserved and similar in two cell types. Many non-exponentially degraded (NED) proteins are subunits of complexes that are produced in super-stoichiometric amounts relative to their exponentially degraded (ED) counterparts. Within complexes, NED proteins have larger interaction interfaces and assemble earlier than ED subunits. Amplifying genes encoding NED proteins increases their initial degradation. Consistently, decay profiles can predict protein level attenuation in aneuploid cells. Together, our data show that non-exponential degradation is common, conserved and has important consequences for complex formation and regulation of protein abundance.

Introduction

Pioneering experiments by Rudolph Schönheimer established that proteins are in a dynamic state of synthesis and degradation (Schoenheimer, 1942). The subsequent discovery of lysosomes and the ubiquitin-proteasome system (UPS) provided detailed insights into the molecular mechanisms of cellular protein homeostasis. It is now well-established that proteins are extensively turned over, that this process is specific, and that the stability of individual proteins can vary under different physiological conditions (Ciechanover, 2005).

Despite these mechanistic insights, the kinetics of cellular protein degradation are still not well understood. Early analyses indicated that intracellular protein degradation follows first order kinetics (Goldberg and Dice, 1974; Schimke and Doyle, 1970). Accordingly, protein degradation is thought to be an exponential decay process in which young and old proteins have the same degradation probability per unit time (i.e., degradation rate) (Fig. 1 A). However, there is substantial evidence that protein degradation does not always follow first-order kinetics. Pulse-chase experiments by Wheatley et al. indicated that a substantial fraction of proteins are degraded within the first 2 h after synthesis (Wheatley et al., 1980). In addition, the newly synthesized immature forms of proteins, like the cystic fibrosis transmembrane conductance regulator (CFTR) and basigin (CD147), were found to be rapidly degraded while the “older” mature forms were stable (Tyler et al., 2012; Ward and Kopito, 1994). It was also observed that proteins can be ubiquitinated co-translationally (Duttler et al., 2013) and that most ubiquitinated proteins in a cell are relatively young (Kim et al., 2011). Finally, it has been estimated that around 30% of newly synthesized proteins are quickly degraded (Schubert et al.,

2000), although this number was questioned in later studies (Vabulas and Hartl, 2005). Collectively, these studies suggest that decay probabilities of proteins can vary as a function of their molecular age. However, to the best of our knowledge, the degradation kinetics of individual proteins has not yet been investigated globally.

To systematically assess cellular protein degradation kinetics we sought to perform pulse-chase experiments on a proteome-wide scale. The general idea is to metabolically label a population of proteins with a short pulse and then to quantify how much of this population is left after different lengths of chase. We and others have previously used stable isotope labeling by amino acids in cell culture (SILAC) to study protein synthesis and turnover (Andersen et al., 2005; Doherty et al., 2009; Hinkson and Elias, 2011; Jovanovic et al., 2015; Kristensen et al., 2013; Larance et al., 2013; Schwanhäusser et al., 2011; Selbach et al., 2008). A disadvantage of these approaches is that they require rather long labeling times. Metabolic labeling with bioorthogonal amino acids has emerged as an attractive alternative (Dieterich et al., 2006). Cells can incorporate the artificial amino acid azidohomoalanine (AHA) into newly synthesized proteins instead of methionine (Kiick et al., 2002). AHA contains an azide group enabling capture of proteins via click chemistry (Dieterich et al., 2006). Combining AHA with SILAC enables relatively short pulse times (Eichelbaum and Krijgsveld, 2014; Eichelbaum et al., 2012).

Here, we combined AHA and SILAC to obtain a global survey of protein degradation kinetics. We find that a sizable fraction of proteins are degraded non-exponentially. Many non-exponentially degraded (NED) proteins are members of heteromeric protein complexes that are over-produced relative to other members of the same complex. Thus, in contrast to recent

findings in bacteria (Li et al., 2014), disproportional protein synthesis appears to be common and evolutionarily conserved in metazoans. Our data allowed us to predict how protein levels change in response to gene copy number alterations in aneuploid cells. Global quantification of protein degradation kinetics reveals an unexpected layer of posttranslational regulation with important functional implications.

Results

Combining metabolic pulse labeling and click-chemistry for global pulse-chase experiments

To perform proteome-wide pulse-chase experiments we combined AHA and SILAC labeling (Fig. 1 B). First, cells are fully labeled heavy, medium-heavy or light using SILAC. Second, all three cell populations are pulse labeled with AHA for 1 h. This relatively long pulse labeling time was chosen to allow sufficient label incorporation. Heavy cells are harvested immediately after the pulse while medium-heavy and light cells are chased in AHA-free medium for different lengths of time. All three cell populations are then combined and lysed. AHA-containing proteins are purified from the mixed lysate. After digestion into peptides, SILAC-based quantification reveals how much of the pulse labeled fraction remains at different time points.

We first confirmed that click chemistry-based capturing of heavy AHA labeled proteins is highly specific and reproducible (Figure S1 A-C). Next, we checked if AHA might itself affect protein degradation kinetics and found no evidence for this (Fig. S1 D-G). Also, peptides derived from both the N-terminal and C-terminal halves of AHA-labeled proteins had similar intensities, suggesting that AHA does not induce premature termination (Fig. S1 H). Collectively, these data indicate that AHA pulse-chase (AHA p-c) enables specific enrichment of newly synthesized proteins with no apparent impact on protein stability, consistent with previous reports (Cohen et al., 2013; Dieterich et al., 2006; Howden et al., 2013; tom Dieck et al., 2015).

For global quantification of protein degradation kinetics we performed three parallel triple-SILAC experiments with different chase times (1, 2, 4, 8, 16 and 32 h) in mouse fibroblasts (NIH 3T3). Heavy cells were always harvested immediately after the pulse and served as a common reference point. Exemplary mass spectra for a stable protein (filamin A, Flna) and an unstable protein (cathepsin L, Ctsl1) show expected slow and fast degradation, respectively (Fig. 1 C). The levels of basigin (Bsg) quickly decreased after the chase but then stabilized after approximately 4 h. This is consistent with the observation that most of the newly synthesized immature basigin is degraded while the mature form of the protein is stable (Tyler et al., 2012). We combined the data from the three triple-SILAC experiments to obtain kinetic profiles with seven time points. To compensate for differences in cell numbers we normalized the data using a selected set of very stable proteins (Fig. S2). We also subtracted background signals which could otherwise give rise to erroneous degradation profiles (see Method and Resources for details). The entire large scale experiment was carried out three times, thus yielding data from three independent biological replicates. In total, we obtained profiles for 5,247 proteins (Fig. 1 D). After applying several quality filters (Fig. S3 A) we kept 3,605 profiles for further analysis (Fig. S3 E). About half of these profiles were derived from at least two biological replicates, allowing us to assess reproducibility. Overall, reproducibility was very good with coefficients of variation (CVs, computed in log space) of less than 10% at time point 32 h for approximately 90% of the proteins (Fig. S3 F).

Stochastic modeling reveals extensive non-exponential degradation

To model our experimental protein degradation profiles we adapted a Markov chain-based approach previously used to study mRNA decay (Fig. 2 A) (Deneke et al., 2013). We considered two different models. In the first model proteins only exist in a single state "A" that is characterized by a constant decay probability. This "1-state model" therefore describes exponential degradation (ED). The second model has an additional state: Newly synthesized proteins first populate state A from where they can either be degraded or transit to state B, which is characterized by a different decay probability. This "2-state model" thus describes non-exponential degradation (NED). To distinguish between both scenarios we compared the relative quality of both models for each protein degradation profile using the Akaike information criterion (AIC) (Akaike, 1974). This approach considers the trade-off between the goodness of fit and model complexity (that is, the number of parameters). Hence, the 2-state-model is only preferred when the improved fit outweighs the increased complexity. The AIC thus provides a conservative estimate of the fraction of NED proteins. Degradation profiles of Ctsl1 and Flna were better explained by the 1-state model (Fig. 2 B). In contrast, the degradation profile of basigin was better explained by the 2-state model.

Overall, we found that the profiles of 509 proteins are better explained by the 2-state model (Fig. 2 C, AIC probability > 0.8). This corresponds to approximately 14% of the 3,605 proteins that passed our quality criteria (Fig. S3). We conclude that a sizable number of proteins show a tendency towards NED.

In principle, the 2-state model can describe two different scenarios: When the degradation rate in state A is higher than in state B ($k_A > k_B$) the model describes proteins that become more stable as they age. Alternatively, when the degradation rate in state B is higher than in state A ($k_B > k_A$) the model describes destabilization of older proteins. Intriguingly, we only observed age-dependent stabilization (Fig. 2 D). This is surprising, as we originally expected proteins to become more unstable over time due to the cumulative effects of age-related damage. However, in our cell line model and within the time period monitored (32 h) we did not find any evidence for this.

While our AIC probability describes the relative quality of both models, it does not provide information about the extent of NED for individual proteins. We therefore also measured the distance of intermediate data points from the linear fit in the semi log plot (Fig. 2 E). This delta score (Δ-score) is a simple measure for the extent of non-exponentiality of individual degradation profiles. We then defined NED and ED proteins based on their AIC probabilities and Δ-scores (Methods and Resources). With these filters, ~10% of proteins were classified as NED and ~50% as ED. The remaining proteins were classified as undefined ("UN", Figure 2 F). For all subsequent analyses we relied on this classification (Supplemental Table S1). In summary, our global quantification of cellular protein degradation kinetics reveals that many proteins become more stable once they have survived the first few hours of their existence.

NED can be validated by independent methods

Although AHA labeling does not appear to globally affect protein degradation (Fig. S1), it is still possible that labeling with an artificial amino acid introduces systematic biases. We therefore wanted to validate our data with independent methods. First, we compared protein degradation rates in the present study with previously published dynamic SILAC data from the same cell line (Schwanhäusser et al., 2011) and found overall good agreement (Figure S4 A-D). Second, we used classical radioactive pulse-chase and immunoprecipitation to confirm degradation kinetics of an ED and an NED protein (Figure S4 E). Finally, to systematically validate our classification, we designed a novel SILAC-based strategy (Fig. 3 A): We pulse labeled light cells with heavy SILAC medium for 4 hours. At this time point, the proteome consists of two populations of protein molecules: Light proteins, which are older than 4 hours, and heavy proteins with an age between 0 and 4 hours. We harvested half of the cells to quantify the H/L ratio at this time point. Since NED proteins become more stable with age, the younger heavy protein molecules should degrade faster than the older light ones. We therefore cultivated the other half of cells on medium-heavy growth medium for 8 additional hours. This resulted in the expected decrease in H/L ratios for NED but not for ED proteins (Fig. 3 B-D). Label swap experiments and different chase times confirmed this finding (Fig. S4 F-H). We computed a NED validation score based on the average rank of proteins across these SILAC pulse-chase experiments (Fig. 3 E, F). This score indicates how well non-exponential degradation could be validated for individual NED proteins and is included in Supplemental Table S1. Since these validation experiments have longer pulse times and thus lower sensitivity

a low validation score of a NED protein should not be interpreted as evidence for erroneous classification.

The ubiquitin proteasome system mediates NED

Having shown that a sizable number of proteins are non-exponentially degraded, we next asked how degradation occurs. The two major cellular protein degradation systems are the proteasome and the lysosome. We therefore used drugs, MG132 or wortmannin in combination with bafilomycin A1, to inhibit proteasomal or lysosomal degradation, respectively. To assess non-exponential degradation we performed AHA p-c experiments with 4 and 8 h chase times in the presence of the inhibitor or carrier controls (DMSO). The impact of both pathway inhibitions on NED was quantified by measuring their impact on the Δ-score (Fig. 4 A). Proteasome inhibition reduced Δ-scores of most NED proteins (Fig. 4 B and D, and Supplemental Table 2). In contrast, inhibition of the lysosome did not have an observable impact (Fig. 4 C and E, and Supplemental Table 2). We conclude that the ubiquitin proteasome system is involved in the initial degradation of most NED proteins.

Many NED proteins are members of multiprotein complexes

We next characterized features that distinguish ED and NED proteins and found that NED proteins are on average more abundant and have a higher degree of secondary structure (Fig. S5 A). Moreover, many NED proteins are members of annotated heteromeric complexes (Fig. S5 A and B). Mapping the data to protein structures showed that 70% of NED proteins are

members of heteromeric protein complexes, which is a significant enrichment relative to ED and UN proteins (Fig. 5 A, S5 C and Supplemental Table 3).

Many complexes contain both NED and ED proteins. Therefore, we investigated if the properties of NED and ED proteins in complexes differ. We found that NED proteins have significantly larger interaction interfaces within complexes (Fig. 5 B and Supplemental Table 3). It has been shown that complexes form via ordered and evolutionarily conserved assembly pathways (Marsh et al., 2013). We found that NED subunits assemble significantly earlier than ED subunits (Fig. 5 C and Supplemental Table 3). Finally, we found that the NED subunits show stronger correlated coexpression with other subunits of the same complex, while the ED subunits show less coherent expression (Fig 5D, S5D, Supplemental Table 3). Together, these data indicate that NED proteins are typically core components of multiprotein complexes while ED proteins tend to be more peripheral.

A long-standing hypothesis is that proteins are stabilized by complex formation (Goldberg, 2003). While several individual examples support this idea (Blikstad et al., 1983; Lam YW, 2007; Toyama et al., 2013) it has, to the best of our knowledge, not yet been investigated systematically. We reasoned that complex formation could explain non-exponential degradation (Fig. 5 E): If NED proteins were produced in excess relative to ED proteins, the degradation of the non-assembled subunits would lead to non-exponential decay. We therefore analyzed the abundance of newly synthesized proteins directly after the pulse. To this end, we normalized the data in a complex-centric manner (Fig. 5 F). We found that NED proteins are indeed over-synthesized relative to ED proteins in the same complex. Moreover, while the

initial degradation rates (i.e. the degradation rates of state A) of NED proteins within a complex varied considerably, their second (state B) degradation rates were more similar and close to the degradation rates of the ED subunits (Fig. S6). Thus, many NED proteins in complexes appear to be produced in super-stoichiometric amounts relative to their ED counterparts.

To independently validate these findings, we performed the same complex-centric analysis on ribosome profiling data (Subtelny et al., 2014) and observed the same trend (Fig. 5 G). To directly assess the link between complex assembly and NED we focused on the ribosome – a complex rich in NED proteins. Since the ribosome consists of proteins and rRNAs, ribosome assembly can be inhibited by blocking rRNA transcription. We found that actinomycin D treatment increased the initial degradation of ribosomal proteins (Fig. 5 H), consistent with our hypothesis and with previous data (Lam YW, 2007).

NED is evolutionarily conserved

All data presented so far are based on the analysis of mouse fibroblasts (NIH 3T3). We therefore asked if our findings are due to specific features of this model system. For example, even though NIH 3T3 cells are derived from primary cells, a recent cytogenetic study revealed a complex rearranged karyotype (Leibiger et al., 2013). It is therefore possible that the super-stoichiometric synthesis of NED proteins is due to genomic amplification of the corresponding genes. In this case, our data would have little relevance for other model systems.

To test this possibility, we analyzed the diploid human retinal pigmented epithelial cell line RPE-1. Low coverage whole genome sequencing of this cell line confirmed that, with the exception

of partial trisomies for chromosome 10 and 12, it is mostly diploid (Fig. S7). We performed three independent large AHA p-c experiments to obtain degradation profiles for 4,079 proteins. 47% and 9.4% of the proteins that passed the quality filters (n=3,133) were classified as ED and NED, respectively (Fig. 6 A left bar and Supplemental Table 4). These fractions are similar to the mouse fibroblast data. We then grouped the human proteins according to the degradation profiles of their mouse orthologs (Fig. 6 A). Human orthologs of mouse NED proteins were enriched in NED proteins. Conversely, human orthologs of mouse ED proteins were enriched in ED proteins. In addition, the Δ-scores of the human proteins and their mouse orthologs were significantly correlated (Fig. 6 B). This is surprising, especially since there are many reasons why degradation profiles may actually differ (cell cycle properties, overall proteome composition which may drastically differ between both cell types). We also found that human NED proteins that are part of multiprotein complexes tend to be produced in super-stoichiometric amounts (Fig. 6 C), consistent with the mouse data (Fig. 5 F). This observation still holds when human proteins are classified according to the degradation profile of their mouse orthologs, further supporting the notion of conservation.

To further assess the conservation of NED we used ribosome profiling data from different species (human, HEK293 (Liu et al., 2013), mouse (Shalgi et al., 2013), zebrafish (Chew et al., 2013) and *C. elegans* (Nedialkova and Leidel, 2015)). We classified proteins in these datasets according to the degradation profiles of their human orthologs (in RPE-1 cells). We then normalized the ribosome protected fragment (RPF) reads in a complex-centric manner. We found that orthologs of human NED proteins tend to be over-synthesized in mouse and zebrafish (Fig. 6 D). The trend also holds in *C. elegans* although it is not significant. Together,

these data show that non-exponential degradation is not mainly due to “erroneous” protein overproduction caused by genomic rearrangements in a specific cell line. Instead, our findings show that protein degradation kinetics are – at least partially – conserved between species and independent of the cell type.

NED predicts protein level attenuation in aneuploidy

Based on our findings we propose a simple model that explains the relationship between protein degradation kinetics and complex formation (Fig. 7 A). Accordingly, NED proteins are overproduced relative to the ED proteins in the same complex. Therefore, only a fraction of the overproduced proteins are stabilized by complex formation while the rest are degraded. Importantly, this overproduction occurs in disomic cells and is thus not generally due to aneuploidy. However, we reasoned that aneuploid cells would allow us to test the model: Genomic amplification of NED proteins should increase their over-production and thus the initial degradation (Fig. 7 A). Consequently, amplification of genes encoding NED proteins should not lead to correspondingly increased protein levels. Instead, NED proteins should be attenuated, i.e. their protein levels should remain relatively constant despite the genomic amplification (Dephoure et al., 2014; Geiger et al., 2010; Stingele et al., 2012).

To test this prediction we took advantage of RPE-1 cells that were engineered to carry one additional copy of specific chromosomes (Stingele et al., 2012). Low coverage genome sequencing (Fig. S7 A,B) and chromosome painting (Fig. S7 C and D) verified that these cells are trisomic for chromosome 5 and part of chromosome 11. We therefore also performed AHA-pc

experiments with these “RPE-1 trisomic” cells (Supplemental Table 5). We then compared protein production in RPE-1 and RPE-1 trisomic cells using abundance levels after the pulse as a proxy. As expected, proteins encoded by trisomic regions were up-regulated in trisomic cells (Fig. 7 B). Note that chromosome 10 and 12 were also partially trisomic in the parental cell line and therefore excluded from subsequent analyses.

We next compared the extent of non-exponential degradation in RPE-1 and RPE-1 trisomic cells using the Δ-score. Proteins encoded in disomic regions of the genome had similar Δ-scores in both cell lines (Fig. 7 C). However, consistent with our prediction, NED proteins in trisomic regions displayed significantly increased non-exponential degradation as measured by the change in their Δ-scores. Importantly, this behavior was specific for NED proteins and not observed for ED proteins. We conclude that increasing over-production of NED proteins tends to increase their initial degradation.

Aneuploidy has severe developmental effects, it is the leading cause of mental retardation and spontaneous abortions, as well as a hallmark of cancer (Santaguida and Amon, 2015). However, the functional consequences of aneuploidy are only beginning to emerge. Studies in yeast and mammalian cell lines have shown that mRNA levels generally scale with gene copy numbers. However, protein levels are sometimes attenuated towards the euploid state. It has also been noted that most attenuated proteins are members of multiprotein complexes (Dephoure et al., 2014; Geiger et al., 2010; Stingele et al., 2012). However, not all proteins in multiprotein complexes are attenuated and not all proteins that are attenuated are part of (known) multiprotein complexes.

Our model predicts that NED proteins should be attenuated. To test this idea, we directly quantified relative changes in steady-state protein levels in RPE-1 and RPE-1 trisomic cells in a separate SILAC experiment. We found that NED proteins encoded in trisomic regions were more attenuated than ED proteins (Fig. 7 D). This is consistent with our model, even though interpreting the data is complicated by the rather small number of NED proteins in trisomic regions. We also observed this effect in the subset of proteins that are part of annotated complexes (Fig. 7 E). Hence, protein degradation kinetics can explain why some proteins in complexes show more attenuation (NED proteins) than others (ED proteins). Moreover, even for the subset of proteins that are not part of an annotated complex, NED proteins showed more attenuation than ED proteins (Fig. 7 F). Collectively, these data show that NED can help to predict protein level attenuation in aneuploidy.

Discussion

Our kinetic analysis of protein stability reveals widespread age-dependent degradation. We find that kinetic profiles are similar in two different cell types and conserved between mouse and humans. Many non-exponentially degraded proteins are subunits of multiprotein complexes that are produced in excess. Accordingly, we propose a simple model in which only a fraction of newly synthesized NED proteins is stabilized by complex formation while the rest are degraded. This model can help to predict how changes in DNA copy-number affect protein levels.

While the AHA-pc method employed here has many advantages, it is also important to keep its limitations in mind. For example, due to the technical challenges of AHA-pc our data is not comprehensive. We do not know how our findings extrapolate to the uncovered part of the proteome. Also, the pulse time of 1 h is too long to capture events on the timescale of minutes. Therefore, our data cannot be used to estimate the extent of co- or peri-translational degradation (Duttler et al., 2013; Schubert et al., 2000; Wheatley et al., 1980). Moreover, since our longest chase time is 32 h, we cannot tell what happens to older proteins. It is possible that longer chase times would reveal age-dependent destabilization. Additionally, the seven time points we covered allowed us to distinguish between a 1-state and a 2-state model, whereas more complex multi-stage models would require measurements at more time points. Increasing the number of data points per profile would also decrease the number of proteins we could not classify as NED or ED. We classified these unclear cases as “undefined” – a purely technical definition that has probably no biological significance. Finally, despite validation by independent methods, we cannot fully rule out biases in our AHA-pc data.

We find that some proteins become very stable after they have survived the first few hours of their molecular life. We have experienced that this surprises many scientists since they intuitively believe that all proteins are constantly turned over. However, our finding is consistent with other observations. For example, SILAC labeling of post-mitotic cells typically remains incomplete despite long labeling times (Liao et al., 2008). Moreover, using metabolic labeling of rats, Toyama and co-workers identified several proteins with extraordinarily long lifespans in the rat brain (Toyama et al., 2013). The latter study also provided evidence that the high stability of one protein, Nup96, is due to its deposition into a stable complex.

The observation that mammalian cells overproduce specific subunits of multiprotein complexes is surprising and in marked contrast to *E. coli* where subunits were reported to be made in precise proportion to their required stoichiometry (Li et al., 2014). Why are mammalian cells producing excess amounts of specific proteins? We would like to discuss four potential answers to this question:

(i) It is possible that only a fraction of newly synthesized proteins adopt a functional state while the rest are terminally misfolded and thus degraded. However, since NED proteins tend to be relatively short and well-structured (Fig. S5 A) this explanation does not seem to generally hold.

(ii) The observation that ED proteins tend to be disordered (Fig. S5 A) suggests an alternative explanation: Disordered proteins are often harmful when overexpressed, which is probably due to their tendency to make promiscuous interactions (Vavouri et al., 2009). Overproduction of NED proteins may have evolved to ensure that

potentially harmful ED proteins are never alone: The super-stoichiometric synthesis of “benign” NED subunits would be a failsafe mechanism against deleterious effects of unbalanced production of the more harmful ED proteins. NED proteins would thus have a chaperone-like function towards their cognate ED proteins.

(iii) ED proteins might have evolved as limiting factors to facilitate the coordinated regulation of protein complex abundance: Up-regulating an ED protein would stabilize interacting NED proteins and thus increase the abundance of the entire complex. Conversely, reducing the expression of the limiting ED protein would decrease complex abundance. This explanation resonates well with the finding that mRNAs encoding ED proteins show less co-expression with mRNAs encoding other subunits and have longer 5’ and 3‘ UTRs (Fig. 5 D and data not shown). It also fits to data from yeast showing that most complexes consist of both constitutive and periodically expressed subunits (de Lichtenberg et al., 2005) and to the finding that complex stoichiometry can vary across tissues in mammals (Ori et al., 2016).

(iv) Overproduction of NED proteins may be important for ordered complex assembly (Marsh et al., 2013; Matalon et al., 2014): Since the formation of protein-protein interactions depends on protein concentration, the relative abundance of proteins may in part determine the assembly order. This explanation is consistent with our observation that NED proteins tend to assemble earlier (Fig. 5 C).

Which (if any) of these four possibilities explains protein overproduction and NED remains to be investigated. It is likely that different reasons are relevant for different proteins. It is also important to note that not all NED proteins are components of (known) multiprotein

complexes. NED of monomers could be due to many different molecular mechanisms. For example, NED of CFTR and basigin appears to be due to failed protein folding in the ER (Tyler et al., 2012; Ward and Kopito, 1994). More generally, biphasic degradation can be due to the existence of distinct pools of a protein. For example, these pools may reflect residency in different compartments (cytosol, nucleus, mitochondria, extracellular etc.), different posttranslational modification states (glycosylation, phosphorylation etc.) or assembly into different complexes. While our manuscript focusses on the latter possibility, this is by no means the only possible explanation. Finally, while we have interpreted the age-dependent rate as giving information on the ageing of every individual molecules, we would like to note that an alternative interpretation based on frailty theory (Aalen, 1994) would give similar fitting quality.

Our results have significant implications for aneuploidy. First, we confirm the previous observation that amplified genes encoding members of multiprotein complexes are often attenuated at the protein level (Dephoure et al., 2014; Geiger et al., 2010; Stingele et al., 2012). This finding was interpreted in the light of the longstanding idea that unassembled subunits of complexes are unstable (Goldberg, 2003). Accordingly, overproduction caused by gene amplification is attenuated for proteins in complexes. Our findings considerably extend this concept. The important difference is that we already observe unbalanced subunit production at baseline (Fig. 7 A). Consequently, our model can explain why only some subunits of complexes show attenuation (Fig. 7 E). More broadly, our data helps to predict how protein levels change in response to altered protein production. We expect that this will turn out to be useful for comprehending the complex cellular phenotypes of aneuploidy and somatic copy number

alterations in cancer. The data might also help to explain posttranslational buffering (Battle et al., 2015).

Contributions

E.M. performed most wet lab experiments and contributed to study design, data analysis and data interpretation. C.S. and A.V. developed and implemented the normalization, mathematical model, model calibration and model selection procedure. H.Z. performed most of the bioinformatic analyses such as validation score-, complex centered-, AUC-, aneuploidy-, ribosome profiling- and conservation analysis. J.N.W. and J.A.M. analyzed the complex assembly data (such as in Fig. 5 A-D). N.D. and Z.S. generated the trisomic RPE-1 cell line, performed chromosome paintings and helped in interpreting the aneuploidy data. X.W., J.H. and W.C. sequenced RPE-1 and RPE-1 (trisomic) cells and computed copy number estimates. M.S. conceived and supervised the study and wrote the manuscript with input from all authors.

Acknowledgement

We thank Katrin Eichelbaum for help with establishing AHA labeling, Piotr Grabowski for setting up SCX chromatography and Koshi Imami (all MDC) for establishing RP-HPLC on monolithic columns. We also thank Thomas Sommer and his lab members (MDC) as well as Thomas Langer and Simon Tröder (University of Cologne) for fruitful discussions. Christian Sommer, Eva Kärgel and Yvonne Kraus provided excellent technical assistance and Rebecca Eccles (all MDC) helped proofreading the manuscript. This work was in part funded by a grant of the Helmholtz

Association to M.S. and by a grant of the European Union (ITN “NICHE”) to C.S. and A.V.. J.M. is supported by a Medical Research Council Career Development Award (MR/M02122X/1).

METHODS AND RESOURCES

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for reagents may be directed to, and will be fulfilled by the corresponding author Matthias Selbach (matthias.selbach@mdc-berlin.de).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Lines

NIH 3T3 Mouse fibroblasts cells were acquired from ATCC. The human retinal pigmented epithelium cell line RPE-1 hTERT, referred to as “RPE-1”, was a kind gift from Stephen Taylor (University of Manchester, UK). The trisomic cell line, referred to as “RPE-1 trisomic”, was generated using microcell-mediated chromosome transfer (Stingele et al., 2012) but has spontaneously gained a part of chromosome 11 specific to this study.

Cells were, if nothing else is stated, cultured in SILAC DMEM (life technologies) complemented with glutamine (Glutamax, life technologies), 1 % Penicillin and Streptomycin (life technologies) and 10 % dialyzed fetal calf serum (Pan-Biotech). The SILAC DMEM was supplemented with standard L-arginine (Arg0, Sigma-Aldrich) and L-lysine (Lys0, Sigma-Aldrich) as in (Schwanhäusser et al., 2011) and referred to as “Light SILAC DMEM” below. Alternatively, Arg6 and Lys4 or Arg10 and Lys8 were added in place of their light counterparts these media are referred to as “Medium-heavy SILAC DMEM” and “Heavy SILAC DMEM”, respectively (Ong et al., 2002). Cells were cultured at 37°C and 5% CO_2.

For the radioactive pulse-chase in combination with AHA or methionine experiments cells were cultured in Light SILAC DMEM. Confluent cells in 6-well plate wells were washed in pre-warmed PBS before being starved of methionine and cysteine for 45 min in DMEM free of both amino acids (Sigma-Aldrich) supplemented with glutamine, 1 % Penicillin and Streptomycin and 10 % dialyzed fetal calf serum referred to as “methionine and cysteine starvation DMEM”. Cells were then pulsed for 1 h with 80 µCi final concentration of ^{35}S-Cysteine (Perkin Elmer) in combination with either 1 mM AHA or 1 mM methionine (Dieterich et al., 2006). Cells were washed twice in Light SILAC DMEM before either being directly lyzed or chased for 6 or 24 h in “cold” medium

with either 50 µM cycloheximide or 10 fold cysteine (Sigma-Aldrich) added to prevent re-incorporation of the radiolabeled amino acids.

For the radioactive pulse-chase coupled to immunoprecipitation experiments the mouse cells were grown in 15 cm plates in Light SILAC DMEM. Cells were washed twice in pre-warmed PBS before being starved for 1 h in methionine and cysteine starvation DMEM supplemented with 5% HEPES buffer (life technologies) added to compensate for the incubator lacking a CO_2 source. Cells were then pulsed for 1 h in the same medium supplemented with radioactive ^{35}S methionine and cysteine (Perkin Elmer) at a 125 µCi/mL final concentration. After the pulse cells were washed twice in pre-warmed PBS before being chased in Light SILAC DMEM containing 10 fold cysteine and methionine.

For the enrichment specificity experiments mouse cells were cultured in a 15 cm plate with Heavy SILAC DMEM. The cells were washed twice in pre-warmed PBS before being starved of methionine for 1h in Heavy SILAC DMEM depleted of methionine (custom made from Biosera) referred to as “methionine starvation SILAC DMEM” below. The methionine starvation was followed by a 2.5 h long incubation with 1mM AHA.

For the AHA pulse-chase experiments mouse and human cells were cultured in either Light SILAC DMEM, Medium-heavy SILAC DMEM or Heavy SILAC DMEM until fully labelled. Experiments were started when cells reached ~25 % cell density. For the first two biological replicates two 10 cm plates were used per time point and for the third biological replicate two 15 cm plates were used per time point to increase the starting material. After 1 h in the respective Light, Medium-heavy or Heavy methionine starvation SILAC DMEM cells were labeled with 1 mM AHA for 1 h. After the pulse Medium and Light cells were washed in PBS then Light or Medium-heavy SILAC DMEM, respectively, before being chased in the same medium.

The AHA pulse-chase of mouse fibroblasts in combination with inhibitor treatments experiments were performed in a manner similar to the AHA p-c experiments. Methionine starved cells were pulsed with 1 mM AHA for 1 h before being washed twice in pre-warmed PBS and then chased in Light or Medium-heavy SILAC. In these experiments different inhibitors or the vector control dimethyl sulfoxide (DMSO, Biomol) were added as follows. Proteasomes were blocked using 20 µM (S)-MG-132 (MG132, Cayman chemical) and inhibition of autophagy was secured by a combination of 250 nM Bafilomycin A1 (Invivogen) and 500 nM wortmannin (Calbiochem) both

treatments were added only during the chase. In contrast, 100 nM Actinomycin D (Sigma-Aldrich) was added both during the pulse and chase.

For SILAC pulse-chase the mouse fibroblasts were grown to 80 % confluency in 15 cm plates in Light SILAC DMEM. Cells were washed three times in PBS before being pulsed in Heavy SILAC DMEM for 4 h (or as annotated in fig. S4 F-H). Cells were then washed in PBS before being trypsinated for 2 min at 37 °C. Cells were resuspended in PBS before half of the cells were transferred to a 10 cm plate containing Medium-heavy SILAC DMEM and the other half spun down and pellet then frozen.

For measurements of relative protein levels at steady state in RPE-1 and RPE-1 trisomic cells fully Heavy SILAC labeled RPE-1 and Light labeled RPE-1 trisomic cells were grown to 70 % confluency in 10 cm plates as described above.

For the chromosome spreads RPE-1 and trisomic RPE -1 cells were maintained at 37 °C with 5 % CO_2 in DMEM GlutaMax (Gibco) containing 10 % fetal bovine serum (FBS), 100U penicillin and 100U streptomycin. The cells were grown to 70-80 % confluency before being treated with 50 ng/ml colchicine for 3-5 h.

METHOD DETAILS

<u>^{35}S cysteine pulse-chase in combination with AHA or methionine</u>

NIH 3T3 mouse fibroblast cells were grown in Light SILAC DMEM. Confluent cells in 6-well plate wells were washed in pre-warmed PBS before being starved of methionine and cysteine for 45 min in methionine starvation SILAC DMEM. Cells were then pulsed for 1h with 80 µCi final concentration of ^{35}S-Cysteine in combination with either 1mM AHA or 1mM methionine. Cells were then washed twice in Light SILAC DMEM before either being directly lyzed or chased for 6 or 24 h in “cold” medium with either 50 µM cycloheximide or 10 fold cysteine (Sigma-Aldrich) added to prevent re-incorporation of the radiolabeled amino acids. After chase, cells were scraped and lyzed in modified radio-immunoprecipitation buffer (50 mM Tris HCl (pH 7.4), 1 mM EDTA, 150 mM NaCl, 1 % Nonidet P-40, 0.25 % Na-deoxycholate and 0.1 % SDS) containing 2 fold protease inhibitor cocktail (Roche). All samples were frozen at -80 °C before being thawed on ice for 30 min in the presence of an endonuclease (Benzonase, Merck). Samples were spun down to clear cell debris. The resulting supernatant was diluted in LDS

sample buffer (Invitrogen) complemented with DL-Dithiothreitol (DTT, Sigma-Aldrich) before being boiled at 95 °C for 5 min. Proteins were resolved by SDS-PAGE using a 10 % polyacrylamide gel. Proteins were fixed in the gel by 5 % Acetic acid and 50 % methanol and then Commassie stained (colloid blue stain kit, Novex). The gel was vacuum dried for 2 h using a gel drying system (Bio Rad) at 75°C. Vacuum dried gels were scanned using a scanner (Cannon) for quantification of total loaded protein. The radioisotope signal was measured by exposing the gel to a magnetic photostimulable phosphor plate overnight. The plate was then scanned on a phosphorimager (Typhoon FLA 9500, GE Healthcare). Radioactive and Coomassie images were quantified using the ImageQuant software (GE Healthcare). Each lane was quantified separately and background signal was estimated by marking a lane with no proteins loaded. The measured background was subtracted from the signal. The radioactive signal was then further normalized to total protein input estimated by the Coomassie staining. Three biological replicates were performed each containing technical triplicates. Statistics and plotting was performed in Excel (Microsoft).

Enrichment specificity of AHA labeled proteins

Fully Heavy SILAC labelled mouse fibroblasts were cultured in a 15 cm plate until 50 % confluent. The cells were washed twice in pre-warmed PBS before being starved of methionine in Heavy methionine starvation SILAC DMEM for 1h. The starvation was followed by a 2.5 h long incubation with 1 mM AHA. The AHA-labeled heavy cells were then washed and scraped in ice cold PBS, spun down and mixed 1:1 with fully light labeled cells not labeled by AHA (see Fig S1A for experimental design). Two reversed experiment were also performed in where the Light cells were AHA-labeled and the Heavy cells not. The cells were lyzed using urea lysis buffer from the Click-iT protein enrichment-kit supplemented with two-fold protease inhibitors (Roche). The lysate was treated with benzonase for 10 min before being sonicated in a water bath. Samples were spun down at 20,000 rcf and supernatants were transferred to new tubes containing alkyne agarose beads. The click reaction was performed overnight following the “Click-iT protein enrichment kit”-protocol as described earlier (Hou et al., 2015). Proteins were reduced by heating to 70 °C in the presence of 10 mM DTT in SDS buffer and later alkylated by the addition of 40 mM iodoacetamide (Sigma-Aldrich) final concentration. Beads were sequentially washed in SDS buffer, 8 M Urea in 100 mM Tris (pH 8) and 80 % acetonitrile by centrifugation and decanting supernatant. Proteins were digested “on bead” in 5 % acetonitrile

in ABC buffer first 3 h by LysC and then over night with trypsin. The peptide solution was acidified by addition of trifluoroacetic acid (TFA, Sigma-Aldrich) before peptides were desalted and stored on StageTips (Rappsilber et al., 2003). In short, stageTips were prepared by inserting 3 discs of C18 material (3M) into 200 µL pipette tips. The C18 was activated by methanol. Organic solvents were washed away by Buffer A (5 % acetonitrile and 0.1 % formic acid) before peptides were loaded onto the stageTips. Salts were washed away by washing the retained peptides in Buffer A. Peptides were eluted using Buffer B (80 % Acetonitrile and 0.1 % formic acids) and organic solvent was evaporated using a speedvac (Eppendorf). Samples were diluted in Buffer A (5 % acetonitrile and 0.1 % formic acid) before being put on SCX tips (3 M) (Kulak et al., 2014). In brief, 3 discs of SCX material (3 M) were added to a 200 µl pipette tip and activated by sequentially washing in methanol, buffer B and 500 mM ammonium acetate in 15 % acetonitrile and 0.5 % formic acid. SCX tips were washed three times in 0 mM salt buffer (15 % acetonitrile and 0.5 % formic acid) before the samples were added. Flow through was collected and the remaining peptides were eluted by increasing the salt concentration in three fractions using 50, 150 and 500 mM ammonium acetate. Eluted peptides were once again desalted on stageTips.

In parallel, peptides from the input to the click chemistry (lyzed cells mixed 1:1) were prepared by Wessel-Flügge precipitation and “on pellet” digest as previously described (Sheean et al., 2014). In short proteins from the input were precipitated by sequentially adding, MS-grade methanol, chloroform and finally water before spinning down the samples at 10,000 rcf (Puchades et al., 1999). The upper water phase was removed, methanol was added, the sample was centrifuged again and supernatant discarded. The retrieved protein pellet was air dried before being resuspended in 6 M Urea, 2 M Thiourea in 10 mM Hepes (pH8). Proteins were denatured by adding 10 mM DTT in 50 mM ammoniumbicarbonate (ABC buffer) and then alkylated by the addition of 55 mM iodoacetamide in ABC buffer. Proteins were then digested by the addition of LysC for 3 h. The sample was diluted in ABC buffer until the concentration of urea was less than 2 M before trypsin was added over night. Resulting peptides were desalted on stageTips.

Input and on bead digested peptides were separated on a 2,000 mm monolithic column with a 100-µm inner diameter filled with C18 material that was kindly provided by Yasushi Ishihama (Kyoto University) (from now on referred to as “2 m monolithic column”) with a flow rate of 300 nl/min using a 4 h linear gradient with 250 nl/min flow rate of increasing Buffer B concentration on a High Performance Liquid Chromatography (HPLC) system

(ThermoScientific). Peptides were ionized using an electrospray ionization (ESI) source (ThermoScientific) and analyzed on a Q Exactive mass spectrometer (ThermoScientific). The mass spectrometer was run in data dependent mode selecting the top 10 most intense ions in the MS full scans (Orbitrap resolution: 70,000; target value: 3,000,000 ions; maximum injection time of 20 ms) for higher energy collision induced dissociation. The resulting MS/MS spectra from the Orbitrap had a resolution of 17,500 after a maximum ion collection time of 60 ms with a target of reaching 1,000,000 ions.

The resulting raw files were analyzed using MaxQuant software version 1.5.1.2 (Cox and Mann, 2008). Default settings were kept except that `match between runs` was turned on. Lys8 and Arg10 were set as labels and oxidation of methionines, n-terminal acetylation and deamidation of aspargine and glutamine residues were defined as variable modifications. Carbamidomethyl of c-termini was set as fixed modification. The *in silico* digests of the mouse Uniprot database (2014-10) and a database containing common contaminants were done with Trypsin/P. The false discovery rate was set to 1 % at both the peptide and protein level and was assessed by in parallel searching a database containing the reversed sequences from the Uniprot database. The resulting text files were filtered to exclude reverse database hits, potential contaminants and proteins only identified by site. Plotting and statistics were done using R and figures were modified in Illustrator (Adobe). The median H/L ratios from the input samples were used to estimate the mixing ratio of the input and the H/L ratios after the enrichment were adjusted correspondingly.

AHA pulse-chase of SILAC labelled NIH 3T3 mouse fibroblasts

Fully Light, Medium and Heavy SILAC labelled mouse fibroblast were grown as in the “enrichment efficacy experiment”. Experiments were performed when cells reached ~25 % cell density so that full confluency would not be reached during the 32 hours of chase time (See Figure 1B for experimental design). For the first two replicates two 10 cm plates were used per time point and for the third replicate two 15 cm plates were used to increase the starting material. During the whole experiment cells were grown in the presence of arginine and lysine in their respective labeling sate (Light, Medium-heavy or Heavy). After 1h in methionine starvation SILAC DMEM cells were labeled with 1 mM AHA for 1h. After the pulse Medium and Light cells were washed first in PBS then SILAC DMEM before being chased in the same medium. Heavy cells used for time point 0 h were instead washed in ice cold PBS before being scraped in the

same, spun down and cell pellets were frozen. After the chase the medium and light cells were also scraped and frozen.

The frozen pellets were thawed and lyzed as described for “enrichment efficacy experiment”. Also the click reaction and washing of beads, denaturation, alkylation and digestion was performed as above. In one, out of the three experiments, the peptides were pre-fractionated by isoelectric focusing into 12 fractions as described in (Eravci et al., 2014). In two experiments the peptides were separated using strong anion exchange (SAX). The SAX protocol was performed as in (Wisniewski et al., 2009). In short SAX material (3M) was put in 200 µL pipette tips and activated by methanol. The SAX tip was then washed by high pH buffer (20 mM Acetic acid, 20 mM phosphoric acid, 20 mM boric acid, pH was adjusted to 11 by titrating in 1 M sodium hydroxide) before peptides were loaded onto the tip. The peptides were then eluted stepwise by decreasing the pH of the buffer in discrete steps (pH of 11 (flow though), 8, 5 and 3 all prepares as above with the addition of 0.25 M NaCl to the pH 3 buffer). The eluted peptides were stored on stageTips.

IEF and SAX fractionated peptides were separated on a HPLC system as described above by either 4 or 2 h gradients with a 250 nl/min flow rate on a 15 cm column with an inner diameter of 75 µm packed in house with ReproSil-Pur C18-AQ material (Dr. Maisch, GmbH). Peptides were ionized using an ESI source and analyzed on a Q-exactive with the above described settings.

The acquired raw-files were analyzed using MaxQuant with the same settings as for the enrichment specificity experiment but with Arg10 and Lys8 set as heavy labels and Arg6 and Lys4 as medium-heavy labels.

For all downstream analysis we used non-normalized SILAC ratios (see below for normalization procedure) with a minimum of 2 SILAC counts. Reverse database hits, potential contaminants and proteins only identified by site were all excluded.

Data normalization

Normalization is a common challenge for experiments measuring abundances – differences in starting material, labeling efficiency, instrument sensitivity, etc. are all contributors to deviations in the scale of measurements. A common normalization strategy is to normalize the data to the median value in each experiment or replicate, assuming that the median values should not

change through the series of experiments. However, for pulse chase experiments, we expect the measured quantities (and thus also the median) to decay over time, thus this strategy cannot be used. In our data, we expect each time point to have an unknown and potentially different multiplicative factor which affects all measurements at that time point. Thus, we aim to estimate the average value of the multiplicative factor that affects the real data at each time point without assuming any protein degradation rates *a priori* while being robust to experimental errors.

Our normalization scheme is based on the assumption that there are stable proteins within the pool of proteins measured, whose amounts decay very little during the time course of the experiment (Fig. S4) (Schwanhäusser et al., 2011; Toyama et al., 2013). Without noise, the signals corresponding to these proteins would remain unchanged and equal to 100 % left throughout the experiment. With noise, these very stable proteins can still be identified; the Medium/Heavy and Light/Heavy SILAC ratios of these proteins should be among the highest throughout the experiment. Using this method, we identify the most stable proteins and then calculate the multiplicative factor necessary to normalize the data for each time point such that the geometric mean of the measurements of these very stable proteins will have a signal of 100 % (see Fig S2 for overview of normalization strategy).

To find the most stable proteins, we consider proteins with data at all time points in all replica – one reason for this is that it ensures that all the potential candidates are able to contribute to the normalization factor. Furthermore, being in this subset suggests that these proteins are reliably measurable. For each of these proteins, we assign a score, defined as

$$\text{score}_i = \sum_{t \in \{8,16,32\}} \text{PercentileRank}_i(t)$$

Where the index i denotes the protein and $\text{PercentileRank}_i(t)$ maps the rank of each protein's signal strength (from smallest to largest, at time t) to the interval (0, 1). Proteins with higher signals at each time point will have higher scores. Thus for a protein who has the highest signal at all time points would have a score equal to the number of time points, which we call maxScore (i.e. the range of scores is (0, maxScore]). Each protein has up to 3 scores, one from each replica. From the three scores, we calculate the deviation of the score from the maximum score:

$$dev_i = \sum_{j \in \#replicas} \left(\text{maxScore} - score_{i,j} \right)^2$$

Candidates for normalization are those proteins with the lowest deviations. This normalization scheme is based on 4 key assumptions:

1. All groups of cells (heavy/medium/light) produce and degrade proteins equally.
2. Proteins degrade at different rates, which can be differentiated in the time scale of our experiments)
3. Proteins degrading the slowest have the highest Medium and Light to Heavy ratios (and thus lowest deviations)
4. The slowest degrading proteins do not degrade at all in the time scale of our experiment. Note that protein dilution due to cell division does not impact our data since we harvest the entire cell population.

From the data, we find the population of proteins with the lowest score deviations (LSD, n = 200, <5% of total population) and deem these to be the stable proteins (i.e. the LSD-proteins). This set is chosen intentionally large in order to mitigate the effects of outlying data points.

In addition, based on the enrichment efficacy experiments (Fig S1 A-C) described above we applied a stringent cut-off excluding all data points smaller than 10 % protein remaining after normalization. Alternatively, we also tried to subtract protein specific background based on the fact that the background was highly reproducible (Fig. S1C). This did not have any major impact on the protein classification and we therefore stayed with the simpler 10% cut-off.

Parameter fitting

In this study, we consider two simple models: a 1-state model (exponential decay, ED) and a 2-state model (non-exponential decay, NED). In the 1-state model, proteins are in state A just after synthesis. From state A, they are degraded at the rate k_A. The system is memoryless, meaning that the life expectancy for any single protein molecule does not change as the molecules age. For true exponential decay, the data should resemble a straight line when plotted in a log-linear plot. While the one-state model is a good approximation for some decay patterns, other decay patterns have dynamics that are not well described by a one-state model (Deneke et al., 2013).

In the two-state model, proteins are in state A after synthesis. From state A, the molecule can immediately degrade at the rate of k_A, or it can transition to state B with the rate k_{AB}. Molecules that reach state B are degraded with rate k_B. From the analysis point of view, one important distinction between the 1-state and 2-state models is that we lose the property of memorylessness; for 2-state models, the history of a specific molecule (which determines whether the molecule is in state A or state B) changes the expected residual life of the molecule. In short, the residual life of the molecule depends on the age of the molecule. In pulse-chase experiments, the duration of the pulse affects the composition of molecule ages at the beginning of the chase – for a very short pulse, the molecules synthesized in the pulse are likely to have the same age. However, for a longer pulse, molecules synthesized at the beginning of the pulse are "older") while there are some molecules which are just newly synthesized. In short, the length of the pulse must be taken into account for the calculations to accurately uncover the dynamics of degradation.

The derivation of the mathematical description consists of two steps: First is to translate the single molecule dynamics model (e.g. the one-state or two-state model) into the degradation from steady state at the level of population averages. The translation of single molecule dynamics to population averages has been covered in (Deneke et al., 2013). The second step takes the pulse into account and returns the degradation curve of the population averages (e.g. the measurements from the experiments). Calculation of the response of the system resulting from a pulse has been covered in (Sin et al., 2016).

In our formalism, the function (t) defines the theoretical decay pattern, namely the fraction of molecules left after a decay of t time units. This function is expressed in terms of parameters defined through the underlying degradation model. In our degradation model, we have assumed that the proteins follow either a 1-state model, in which there is only one degradation parameter, or a 2-state model, where there are three parameters. The equations used for fitting are as follows:

$$\Lambda(t) = e^{-k_A t} \qquad \text{for the 1-state model}$$

$$\Lambda(t_p + t) = \frac{G(t) - G(t_p + t)}{G(0) - G(t_p)} \qquad \text{for the 2-state model}$$

where $G(t) = k_{AB}(k_{AB} + k_A)e^{-k_B t} + k_B(k_A - k_B)e^{-(k_{AB}+k_A)t}$, and t is the measurement time after the end of the pulse whereas t_p is the pulse length (Sin et al., 2016).

Parameter estimation is performed by MATLAB through nonlinear fitting by minimizing the square deviation from the logarithm of the experimental data and the logarithm of the theoretical function. The routine employed for the nonlinear fit is *fmincon*.

After parameter fitting we applied two quality criteria for selection of proteins for downstream analysis. First, only proteins which had measurements for more than four data points were kept. Second, profiles with RSS > 0.05 were not considered for downstream analysis.

Model selection by the Akaike Information Criterion (Burnham K.P., 2002)

The Akaike Information Criterion (AIC) indicates the quality of the model for a given set of data. Based on information theory, the AIC aims to find the model with minimal Kullback-Leibler distance between the proposed model and the "true" model (as assessed from the data). Models with more parameters have more degrees of freedom during the parameter estimation process, and can often deliver a more accurate fit to the data. However, a more accurate fit to the data does not necessarily imply a higher quality model – instead of a model describing the system's dynamics, these "overfitted models" describe quantities not related to degradation, such as measurement noise. To decide which model we should adapt for each protein, we calculate the AIC for each model. The model resulting in the lowest AIC is the preferred model. We use the AIC with correction for small sample sizes to evaluate each of the two models fitted to each protein degradation pattern:

$$AIC = 2k + n\ln\left(\frac{RSS}{n}\right) + \frac{2k(k+1)}{n-k-1}$$

where n is the number of data points, k is the number of parameters, and RSS is the residual sum of squares. The AIC penalizes models with more parameters, worse fits, and less data. That is, the AIC quantifies the tradeoff between fit accuracy and model complexity.

Furthermore, we can calculate the probability that a particular model i is the preferred one (relative to the other models we consider) by:

$$\prod_{AIC_i} = \frac{\exp\left(\frac{AIC_{\min} - AIC_i}{2}\right)}{\sum_j \exp\left(\frac{AIC_{\min} - AIC_j}{2}\right)}$$

Our modeling approach connects to a more general approach based on the hazard rate, which is the key functional in survival analysis (Aalen and Gjessing, 2001). In our models, the hazard rate is the age-dependent degradation rate (Deneke et al., 2013). Based on biological reasoning, we have interpreted the age-dependent rate as giving information on the ageing of every individual molecule. We would like to point out that an alternative interpretation based on frailty theory (Aalen, 1994) would give similar fitting quality.

Δ-score calculations

For proteins that decay exponentially, one can derive the relative protein abundance at any time point by drawing a straight line between time point 0h (100% protein left) and any other measurement. Making allowances for measurement noise and quantification errors, all other measurements should fall on this line. If a measurement does not fall on this line, the protein is non-exponentially degraded. We used this relationship to estimate the size and direction (increased or decreased stability with age of the molecule) of non-exponentiality of degradation for each protein. We used the median log "protein remaining [%]" at time point 8 h (tp8) after chase to calculate the expected relative protein abundance at time point 4h assuming exponential degradation. For this we solved the straight line equation (y = mx + c) for x = 4 h, where the intercept c is log(100%), and the slope m is calculated using the value at tp8:

$$y(4h) = -\frac{(\log(100\%) - \mathrm{tp8})}{8\mathrm{h}} * 4\mathrm{h} + \log(100\%)$$

Finally, we calculated the distance from the measured median log "protein remaining [%]" at time point 4 h (tp4), to the expected value, y(4h):

$$\Delta - score = \ y(4h) - tp4$$

This calculation was repeated for all proteins. The time points 4 and 8 h were selected because of the observation that most of the initial degradation of NED proteins had already happened by after 4 h chase. Thereby we expected to be able to catch age-dependent stabilization (or destabilization) by comparing these two time points. Also, few proteins (see supplemental Table 1) had a half-life shorter than 2.5 h and could thereby theoretically not be detected at the 8 h

time point. In addition, these short lived proteins were almost exclusively exponentially degraded according to the AIC call.

SILAC pulse-chase (confirmation experiment)

To exclude issues related to using non-natural amino acids and to the enrichment process (e.g. background binders) we performed a pulse-chase experiment using only stable isotope labelled amino acids. Mouse fibroblasts were grown to 80 % confluency in 15 cm plates in Light SILAC DMEM. Cells were washed three times in PBS before being pulsed in Heavy SILAC DMEM for 4 h (or as annotated in fig. S4). Cells were then washed in PBS before being trypsinated for 2 min at 37 °C. Cells were resuspended in PBS before half of the cells were transferred to a 10 cm plate containing Medium SILAC DMEM and the other half spun down and pellet then frozen. After the Medium chase (see fig. 4 and S4, for different chase length) cells were spun down and frozen. In addition, “label-swap” experiments were also performed in this fashion. However, in the label-swap experiments the cells were pulsed with Medium-heavy and chased in Heavy amino acids. Cell pellets were lyzed and proteins denatured in 0.2 % SDS, 0.1 M DTT and 50 mM ABC (pH 8) by boiling for 10 min at 95 °C. After cooling, Benzonase was added for 10 min before cell lysates were spun down and supernatants were transferred to fresh tubes. Proteins were alkylated by adding iodoacetamide to a 0.25 M final concentration, in the dark, for 20 min. Proteins were precipitated by Wessel-Flügge precipitation as described above. The retrieved protein pellet was resuspended in 6 M Urea, 2 M Thiourea in 10 mM Hepes (pH8). Proteins were digested with LysC before being diluted in ABC buffer and trypsinated overnight. The resulting peptide solution was desalted on StageTips before being eluted in buffer B as described above. The peptides were resolved on a 4 m long monolithic column (2 x 2 m column combined) using a 12 h gradient of increasing buffer B concentration and a flow rate of 500 nl/min. Peptides were ionized by ESI and analyzed on a Q-exactive orbitrap all with previous settings. Resulting raw-files were analyzed with MaxQuant with the same parameter settings as above. Plotting and statistics were performed using R and figures were modified in Illustrator.

^{35}S cysteine and methionine pulse chase coupled with immunoprecipitation

NIH 3T3 mouse fibroblast were grown in 15 cm plates as described above. Cells were washed twice in pre-warmed PBS before being starved in methionine starvation SILAC DMEM for 1 h.

Cells were then pulsed for 1 h in the same medium supplemented with radioactive ^{35}S methionine and cysteine (Perkin Elmer) at a 125 µCi/mL final concentration. After the pulse cells were either washed twice in pre-warmed PBS before being chased in medium containing 10 fold cysteine and methionine or scraped in ice cold PBS, spun down and pellets frozen (0 h time point). After being chased, either for 4 or 8 h, cells were collected as the 0 h time point. Cell pellets were lyzed for 15 min in RIPA buffer (50 mM Tris (pH8), 150 mM NaCl, 0.1 % SDS, 0.5 % sodium deoxycholate, mM EDTA, 1 % NP40) supplemented with benzonase and two fold protease inhibitors. Lysates were spun down at 15 0000 rcf for 5 min and supernatant was transferred to new tubes. Lysates were precleared for 30 min by incubation with protein-A sepharose beads (Biovision) at 4 °C while turning over head. Beads were spun down and supernatant was split into two fresh tubes (i.e. time point 0 h were split in two, time point 4 h was split in two and so forth). Polyclonal antibodies raised in rabbit against either VCP (LSBio, LS-C287469) or CCT3 (Proteintech, 10571-1-AP) were added overnight to one tube each. Thereby the VCP and CCT3 immunoprecipitations were performed on the same lysate this to limit differences in radioactive labeling. Two tube had no antibody added and was subsequently used as a bead control. In the morning, protein-A sepharose beads were added to each tube and incubated while turning head over heel at 4°C for 3 h. Afterwards the beads were washed 3 times in lysis buffer. Supernatant was fully decanted and leftover liquid boiled off. Immunoprecipitated proteins were then eluted by cooking the beads in 1-fold LDS loading buffer with DTT. Beads were spun down and the same volume of supernatant was loaded onto a 4-12 % gradient SDS-polyacrylamide gel (ThermoFisher) and separated using electrophoresis as described above. The proteins were then further transferred to a PVDF membrane (Merck Millipore) using a wet western blot contraption (Invitrogen) set to a constant current of 250 mA for 2 h. The radioactive signal from metabolically labelled proteins was detected by exposing a magnetic phosphor plate overnight and then measuring in a PhosphorImager as described above. As loading control the same membranes were also probed by the same antibodies as used for the IPs. Briefly the membranes were first blocked by incubating in 1 % milk powder in Tris-buffered saline (TBS) and then incubated with the protein specific antibody diluted 1:5000 in 1 % milk in TBS overnight while rotating at 4 °C. Membranes were washed in TBS and 1% Tween before being incubated at RT for 1 h with protein-A conjugated to horseradish peroxidase (Merck Millipore). Membranes were washed again before chemiluminescence substrate (PerkinElmer) was added and x-ray films (Fujifilm) were exposed to the membranes and developed using an Optimax 2010 machine (Protec). Two biological replicates were performed.

Inhibitor treatments + controls

Inhibitor treatment experiments were performed as the AHA p-c experiments but only with three time points (0, 4 and 8 h). In addition to pulsing the cells with 1 mM AHA different inhibitors or vector control dimethyl sulfoxide (DMSO, Biomol) were added. Proteasomes were blocked using 20 µM MG132 (Cayman chemical) and a robust inhibition of autophagy was secured by a combination of 250 nM Bafilomycin A1 (Invivogen) and 500 nM wortmannin (Calbiochem) both treatments were added only during the chase. 100 nM Actinomycin D (Sigma-Aldrich) was added both during the pulse and chase.

Inhibition of autophagy by Bafilomycin A1/Wortmannin was monitored by in parallel taking samples for western blotting as previously described (Sury et al., 2015). In short, scraped cells were spun down and directly lyzed in LDS sample buffer supplemented with DTT. Samples were run on 4-12 % Bis-Tris gradient gels (NuPAGE, Invitrogen) before being blotted onto PVDF membrane (Immobilion-P, Millipore) using a wet blotting contraption (Invitrogen). The Autophagy blocked cells were probed against LC3 and afterwards the membrane was stripped at 37 °C for 15 min in stripping buffer (2 % SDS (Roth), 2 % β-mercaptoethanol in 65 mM Tris Base (pH 6.7, Roth)) before being re-blotted using an anti-β-actin antibody.

Treated cells for mass spec analysis were scraped, lyzed, and had their AHA labeled proteins clicked to alkyne-agarose beads as described above. Proteins were reduced with DTT and alkylated before beads were washed all as in the main AHA p-c experiment. Proteins were digested “on bead” by LysC and then trypsinated overnight. Peptide solution were put on 4 mm/1 ml C18 columns (Empore, 3 M) and washed in buffer A. Peptides were eluted in buffer B and vacuum dried.

MG132 treated samples were separated using an online SCX/WAX approach. Samples were loaded on a column packed first with C18 material “trap” and then with a 2:1 mixture of WAX 3 µm beads (PolyLC Inc. PolyWAX LP) and 3 µm SCX beads (PolyLC Inc. PolyWAX LP) (Motoyama et al., 2007). The peptides were subsequently eluted with increasing salt concentration (ammonium acetate in 4, 8, 16, 32, 64 and 500 mM steps) onto the C18 trap part of the pre-column. Each fraction eluted from the SAX/SCX material where then separated as normal on a 15 cm C18 column with 2 h gradients of increasing buffer B concentration with a 250 nl/min flow rate. Bafilomycin A1/ Wortmannin treated samples were put on SCX tips and

washed in no salt buffer, as described above, to minimize polymer contamination. Samples were then eluted with 500 mM ammonium acetate before being desalted on stageTips (Kulak et al., 2014). Samples were eluted from stageTips by Buffer B, vacuum dried, re-suspended in Buffer A and then separated on a HPLC system using a 2 m column and an 8 h gradient as previously described (Hou et al., 2015). Actinomycin D samples were also put on SCX tips but further manually fractionated by eluting in steps by increasing salt concentration (8, 16, 32 and 500 mM ammonium acetate) before being put back on stageTips (Kulak et al., 2014). Peptides eluted from stageTips were vacuum dried and resuspended in buffer A before being separated on a 15 cm C18 column as described above. In all three cases above, eluted peptides were ionized using an ESI source and analyzed on a Q-exactive with the above described settings. ESI and mass spectrometer setting were, for all samples, as described above. The resulting raw files were analyzed by MaxQuant with the same settings as the standard AHA p-c experiment.

Timelines were reassembled from non-normalized protein ratios, resulting into three time points (0, 4 and 8 h) for each inhibitor treatment and the corresponding DMSO control. Proteins were filtered for being represented by at least two peptide identification events. Each time point was normalized to the geometric mean of the identified intersection of the LSD proteins that were identified in the mouse dataset used for the mathematical modeling and then values below 10% were removed. From this normalized dataset Δ-scores were calculated for each, treatment and control, as described above. The difference of the Δ-scores between treatment and DMSO control was compared for the three protein subsets identified as NED, ED and UN. Differences between distributions were tested using the Wilcoxon rank-sum test. The corresponding p-value is reported in the figure legend for each treatment.

Degradation profile prediction from different protein features

The following features were selected to test each for prediction power of protein degradation profiles. The “Part of a Complex” feature distinguished proteins that are part of a complex from proteins that are not part of a complex (Ori et al., 2016). Proteins were defined as being part of a complex, if they are listed in a published manually curated protein complex database (unfiltered version; Ori et al 2016). Protein Length refers to the protein sequence length and was taken from the UniProt fasta table (version 10.2011). “Protein abundances in steady state” refer to average protein copy numbers per cell (Schwannhäuser et al. 2011) mapped by Uniprot accessions and gene names if the Uniprot accession was not mapped. The feature “Low

Complexity Region" were obtained from the "mmusculus_gene_ensembl" dataset from the biomart database (status 14.10.2015). Listed lengths of Low Complexity regions were summed up per protein. Disordered, Helix and Beta Sheet fractions per protein were obtained by secondary structure prediction using the s2d method (Sormanni et al., 2015). All structural features ("Low complexity", "Disorder", "Helix" and "Beta Sheet") were normalized to protein length. For each feature a ROC-curve was generated and the area under the curve calculated using the pracma R-package. The robustness of the calculated AUCs was tested by running 200 bootstrap repetitions. The 90 % confidence intervals of the resulting AUCs are shown as error bars in the corresponding bar plot. Each feature was additionally randomized resulting in a real AUC and a random AUC population. Each feature prediction was tested for being absent or present by reversing the sorting vector. In each case the positive AUC was reported and labeled with "absence" or "presence" of the corresponding feature.

Protein structural dataset

Starting from the entire set of protein structures in the Protein Data Bank on 2016-02-24, we searched for all polypeptide chains with >70 % sequence identity to a human or mouse gene. For genes that map to multiple chains, we selected a single chain sorting by sequence identity, then number of unique subunits in the complex, and then the number of atoms present in the chain. Pairwise interfaces were calculated between all pairs of subunits using AREAIMOL (Winn et al., 2011). The normalized assembly order was calculated for all complexes, excluding those containing nucleic acid chains, by first predicting the (dis)assembly pathway as previously described using all the pairwise interfaces from each heteromeric complex (Marsh et al.) and implemented in the *assembly-prediction* package (Wells et al., 2016). For subunits with multiple copies within a single complex, the average assembly order of each subunit type was considered. The normalized assembly order was defined so that the first subunit to assemble has a value of 0, the last has a value of 1, and the average value for all unique subunits in a complex is equal to 0.5.

Non-structural dataset

To complement the analysis of protein complexes of known structure, we also performed coexpression analyses on the non-redundant "core" set of mammalian complexes from CORUM

(Ruepp et al., 2008) (downloaded 2015-10-20). As CORUM preferentially uses human complexes in its non-redundant set, homologous mouse versions of each complex were generated by replacing each subunit/gene with its mouse counterpart, provided sequence identity was at least 70 %. Sequence identities were calculated by collecting all mouse sequences for which NED/ED classifications were available and running BLAST on these against all genes in the CORUM core set. In cases where the identity of a subunit was ambiguous (as defined by CORUM), the first possible subunit for which homology data was available was selected.

Coexpression analyses

Coexpression data was downloaded from COXPRESdb (Okamura et al., 2015) (mouse dataset: Mmu.v13-01.G20959-S31479; human dataset: Hsa.v13-01.G20280-S73083). For each complex, the mean coexpression of each available subunit was calculated, using all other subunits in the complex. Cases where fewer than three unique subunits were present in the complex were discarded, due to calculations of average coexpression being superficially identical.

Estimation of relative protein abundance after pulse (iBAQ)

To estimate the protein abundance after the pulse (i.e. the relative amount of newly synthesized proteins) we used intensity based absolute quantification (iBAQ, (Schwanhäusser et al., 2011)). First, all the intensities reported directly after the pulse, i.e. the H-intensities, for each protein group were divided by the number of observable peptides to correct for observability biases. Second, all the corrected H-intensities were normalized by using the median H-intensities for the LSD-proteins (see normalization strategy above). This allowed the combination of experiments. Finally, we reported the median H-Intensity from all experiments as the relative abundance. The median was used to avoid counting highly abundant proteins which show up in all replicates multiple times.

For a complex centered analysis of the relative protein abundances after pulse, identified proteins from the mouse dataset were mapped to a filtered version a protein complex database provided by (Ori et al., 2016)(see previous section) using gene names. Protein abundances

were normalized in a complex centered manner: First all proteins that mapped to a complex were extracted. Second, abundances of all proteins of a complex were normalized to the average abundance of each complex. Subsequent filtering was applied to complex-centered values. For each protein only the average value derived from the complex(es) with the highest number of subunits is reported. The resulting filtered complex-centered abundances were compared between the protein subsets ED, NED and UN. Only proteins from complexes with at least one ED or one NED subunit but with at least two different categories (ED, NED or UN) were considered for the analysis.

Preparation of chromosome spreads and chromosome painting

Cells were grown to 70-80 % confluency before treatment with 50 ng/ml colchicine for 3-5 h. Subsequently, cells were collected by trypsinization and centrifuged at 250rcf for 10min. Pellets were then resuspended in 75 mM KCl and incubated for 10-15 min at 37 °C. After centrifugation at 150 rcf for 10 min, cell pellets were resuspended in 3:1 methanol/acetic acid for fixation. Finally, cell pellets were washed several times in 3:1 methanol/acetic acid, spread on a wet glass slide and air-dried at 42 °C for 5 min. Each sample was labeled with probes for two different chromosomes. Probes (Chrombios GmbH, Raubling, Germany) for chromosomes 5 and 11 were tagged with FITC and TAMRA, respectively. The chromosomes were labeled according to the manufacturer's instructions and counterstained with DAPI. Images were obtained by a fully automated Zeiss inverted microscope.

Genomic DNA sequencing and copy number estimation of RPE-1 and RPE-1 trisomic cells

DNA was isolated using the Blood and Cell Culture DNA kit (Qiagen) according to the manufacturer's recommendations. 1µg genomic DNA was sheared following the SureSelect^{XT} Target Enrichment System for Illumina Paired-End Multiplexed Sequencing Library Protocol (Agilent Technologies, Publication Number G7530-90000). Genomic DNA sequencing library was prepared with 100 ng sheared genomic DNA using TruSeq ChIP Library Prep Kit according to the manufacturer's guidance (Illumina). The libraries were sequenced in 1x 100 nt manner on HiSeq 2000 platform with a depth of ~30 million reads per library (Illumina). Sequencing reads were aligned to the human reference genome (hg19) using Bowtie (version 2.1.0) with default parameters, and only uniquely mapped reads were kept for downstream analysis. With a sliding

window of size 100Kb and a step size of 50-Kb, mapped reads in each window were then counted and used for copy number estimation. With the assumption that most genomic regions for the cells were diploid, we took C_i given by the following formula as the copy number estimates for genomic location at the *i*th window:

$$C_i = 2 \times \frac{R_i}{\underset{j \in I}{\text{median}}\ R_j}$$

where R_i is the read counts of the *i*th window. To avoid underestimating copy numbers for regions with multi-aligned reads, we adjusted for mappability based on mfapping of simulated reads with uniform coverage across the genome. The original read counts were divided by the read counts in the same window obtained from the simulation data, and the adjusted read counts were instead used for copy number estimation.

AHA pulse chase of SILAC labeled RPE-1 and RPE-1 trisomic cells

RPE-1 and RPE-1 trisomic cells were grown, methionine starved, AHA pulsed, chased and lyzed as described for the mouse fibroblast. Experiments were started when cells reach ~30 % confluency and two 15 cm plates were used per time point. Click chemistry, denaturation, alkylation, washing and digestion were performed as described for the mouse cells. Peptides were stageTipped on 4 mm/1 ml C18 columns (3M). Peptides were eluted using 500 µl buffer B and dried in a speed-vac until a few µl liquid was left. For two of the samples Buffer A was added to 10 µl final volume. 5 µl of sample was loaded onto a 15 cm column and 5 µl onto a 2 m monolithic column using a HPLC system. The 15 cm column and 2 m monolithic column samples were analyzed on a Q-Exactive orbitrap system, as described above, deploying 4 and 6h gradient of increasing Buffer B, respectively. For one sample the peptides were further SCX fractionated into 2 fractions 125 mM and 500 mM ammonium acetate as described above. These samples were analyzed using 4 h gradients of increasing Buffer B concentration over a 15 cm column. The resulting raw files were analyzed using MaxQuant with the previously described parameter settings with the exception that Andromeda search engine was matching the MS/MS to the human Uniprot database (2014-10) and unmodified counterpart peptides were kept for quantification. 3 biological replicates were performed per cell line.

Normalization, fitting of models, Δ-score and abundance after pulse calculations were performed as for the mouse fibroblasts. For all downstream analysis proteins derived from

genes located on autosomes were used except when from chromosome 10 (fully trisomic in both parental and trisomic cell line) and chromosome 12 (clonal expansion of trisomic cells among control cells).

Relative protein levels at steady state in RPE-1 and RPE-1 trisomic cells

Fully Heavy SILAC labeled RPE-1 and Light labeled RPE-1 trisomic cells were grown to 70 % confluency in 10 cm plates. Cells were scraped in ice cold PBS before being spun down at 1000rcf and PBS decanted. Cell pellets were lyzed in 1.3 % SDS, 0.1 M DTT in 50 mM ammonium bicarbonate solution. Samples were heated to 95 °C for 10 min. After cooling the samples, Benzonase was added for another 10 min. The samples from the two cell lines were then mixed 1:1 and spun down at 20,000 rcf to clear cell debris. Proteins in supernatant were alkylated by the addition of 0.25 M iodoacetamide, final concentration, and left in the dark at room temperature for 20 min. After alkylation proteins were directly precipitated, to get rid of SDS, by Wessel-Flügge precipitation as described above. The upper water phase was discarded and more methanol was added to the precipitated proteins and the samples were spun down again. Supernatant was discarded and pellet air dried. The pellet was solubilized by shaking the sample in 6 M Urea/2 M thiourea in 10 mM Hepes (ph8). Proteins were digested “on pellet” by Lys-C for 3 h at room temperature before the sample was diluted in ABC buffer and Trypsin was added overnight. Peptides were acidified by triflouroacetic acid before being stored on stageTips. Peptides were prepared for HPLC as described above and analyzed using a 6 h gradient on a 15 cm column packed with C18 material as described above. The Q-exactive was run with standard setting and the raw files were analyzed as described for the AHA enrichment specificity experiment. MaxQuant output was filtered as described above but this time normalized SILAC ratios were used for downstream analysis. A label swap experiment was also performed in where RPE-1 cells were grown in light SILAC medium and RPE-1 trisomic cells were grown in heavy SILAC medium. The analysis used the average SILAC ratio for the two experiments.

Bioinformatics of RPE-1 cells

Δ-scores and abundances after pulse were calculated for the datasets of the RPE-1 and the RPE-1 trisomic cell line as described in the previous sections. For the conservation analysis

proteins identified in RPE-1 were mapped to proteins from the mouse data using always the first entry in the gene name column from both protein-groups tables (proteinGroups.txt, as provided by MaxQuant). Mapped proteins from mouse that fell into the categories ED, UN and NED were compared to the fraction of mapped proteins identified in RPE-1: RPE-1 proteins that mapped to the mouse dataset (all, mapped to mouse orthologues), RPE-1 proteins that mapped to the ED subset of mouse (ED in mouse), and RPE-1 proteins that mapped to the NED subset (NED in mouse). Enrichment of RPE-1 ED or NED definitions in the fraction ED or NED in mouse against the corresponding fraction of all mapped proteins was tested applying a hypergeometric test (phyper function as implemented in R setting lower.tail to False). Δ-scores of mapped RPE-1 and Mouse Genes were compared and Pearson correlation of all available data points was calculated (using cor.test function as implemented in R). The corresponding p-value indicating the significance of the observed correlation coefficient is given in figure legend (Fig. 5). RPE-1 and RPE-1 trisomic datasets were merged using the leading protein ID in each protein-group. Proteins were linked to chromosome positions using the human reference genome (hg19). Proteins were first mapped to chromosome positions based on uniprot IDs. Remaining unmapped proteins were mapped using gene name entries as provided in the uniprot fasta file. Proteins were further grouped into the categories disomic, trisomic or ambiguous. Disomic proteins included all proteins that mapped to disomic chromosomes as identified by genome sequencing (see previous section). Trisomic regions included proteins mapping to chromosome 5 and chromosome 11 downstream from position 62 650 000. Chromosome 12 and 10 starting from nucleotide position 62 500 000 were considered to be ambiguous since they show partial aneuploidy already in the RPE-1 cell line. Only autosomes were considered for the subsequent analysis. Distributions of the Δ-scores and steady-state protein levels were compared between proteins from disomic and trisomic regions. P-values were derived from a one-sided Wilcoxon rank-sum test (wilcox.test as implemented in R). Steady state protein levels were further split into two subsets: proteins in a complex (listed in the manually curated protein complex database from (Ori et al., 2016)) and proteins not in a complex (not listed in the protein complex database).

Ribosomal Profiling Data analysis

The datasets for Human (SRA061778; HEK293 cells no treatment;(Liu et al., 2013)), Zebrafish (GSE46512; 28hours post-fertilization; (Chew et al., 2013)) and Worm (GSE67387; wildtype L4

stage; (Nedialkova and Leidel, 2015)) were downloaded as precompiled RPKM tables from RPFDB (Xie et al., 2016) If available, for the subsequent analysis CDS-RPKM values were used. For the worm dataset Gene-RPKM values were used. Datasets including ribosomal profiling data for mouse were obtained from (Subtelny et al., 2014) (Fig. 5) and Shalgi et al. 2013 (Shalgi et al., 2013) (Fig. 6). RPKM values with the same gene name were summed up and median was taken in log2 space across available replicates. Ribosomal datasets were mapped to degradation profile definitions (ED, NED or UN) from either the human or mouse dataset using official gene names. Mappings of orthologues genes between human and worm were obtained from “http://www.ensembl.org/biomart/martview”. Subsequently, the assembled data table was complex centered as described above using the filtered version of the protein complex database from (Ori et al., 2016).

QUANTIFICATION AND STATISTICAL ANALYSIS

The type of statistical test (e.g. Wilcoxon rank-sum or hypergeometric test) is annotated in the Figure legend and/or in the Methods and Resources segment specific to the analysis. In addition, statistical parameters such as the value of n, mean/median, SEM, SD and significance level are reported in the Figures and/or in the Figure Legends. An alpha of 0.05 was set for significance in all analyses. Higher p-values are sometimes displayed in the figures but are not referred to as significant. When * are used to signify the significance level the key is reported in the respective Figure legend. Statistical analyses were performed using MATLAB or R as described in Methods and Resources for each individual analysis.

DATA AND SOFTWARE AVAILABILITY

Data Resources

Raw data files for the DNA sequencing of the RPE-1 and RPE-1 trisomic cells have been deposited in the NCBI BioProject under accession number PRJNA339199.

Raw data files for the Proteomic analysis have been deposited in the PRIDE database under accession numbers TBA (human) and TBA (mouse).

References

Aalen, O. (1994). Effects of frailty in survival analysis. Statistical Methods in Medical Research *3*, 227-243.

Aalen, O.O., and Gjessing, H.K. (2001). Understanding the shape of the hazard rate: a process point of view (With comments and a rejoinder by the authors). 1-22.

Akaike, H. (1974). A new look at the statistical model identification. Automatic Control, IEEE Transactions on *19*, 716-723.

Andersen, J.S., Lam, Y.W., Leung, A.K., Ong, S.E., Lyon, C.E., Lamond, A.I., and Mann, M. (2005). Nucleolar proteome dynamics. Nature *433*, 77-83.

Battle, A., Khan, Z., Wang, S.H., Mitrano, A., Ford, M.J., Pritchard, J.K., and Gilad, Y. (2015). Impact of Regulatory Variation from RNA to Protein. Science (New York, NY) *347*, 664-667.

Blikstad, I., Nelson, W.J., Moon, R.T., and Lazarides, E. (1983). Synthesis and assembly of spectrin during avian erythropoiesis: stoichiometric assembly but unequal synthesis of alpha and beta spectrin. Cell *32*, 1081-1091.

Burnham K.P., A.D.R. (2002). model selection and multimodel inference: A practical information-Theoretic approach (Springer-Verlag New York Inc.).

Chew, G.L., Pauli, A., Rinn, J.L., Regev, A., Schier, A.F., and Valen, E. (2013). Ribosome profiling reveals resemblance between long non-coding RNAs and 5' leaders of coding RNAs. Development (Cambridge, England) *140*, 2828-2834.

Ciechanover, A. (2005). Aaron Ciechanover - Nobel Lecture: Intracellular Protein Degradation, Vol The Nobel Prizes 2004 (Stockholm: Science History Publications/Watson Publishing [Nobel Foundation]).

Cohen, L.D., Zuchman, R., Sorokina, O., MÃ¼ller, A., Dieterich, D.C., Armstrong, J.D., Ziv, T., and Ziv, N.E. (2013). Metabolic turnover of synaptic proteins: kinetics, interdependencies and implications for synaptic maintenance. PLoS One *8*, e63191.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

de Lichtenberg, U., Jensen, L.J., Brunak, S., and Bork, P. (2005). Dynamic complex formation during the yeast cell cycle. Science *307*, 724-727.

Deneke, C., Lipowsky, R., and Valleriani, A. (2013). Complex degradation processes lead to non-exponential decay patterns and age-dependent decay rates of messenger RNA. PLoS One *8*, e55442.

Dephoure, N., Hwang, S., O'Sullivan, C., Dodgson, S.E., Gygi, S.P., Amon, A., and Torres, E.M. (2014). Quantitative proteomic analysis reveals posttranslational responses to aneuploidy in yeast. eLife *3*, e03023.

Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc Natl Acad Sci U S A *103*, 9482-9487.

Doherty, M.K., Hammond, D.E., Clague, M.J., Gaskell, S.J., and Beynon, R.J. (2009). Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. J Proteome Res *8*, 104-112.

Duttler, S., Pechmann, S., and Frydman, J. (2013). Principles of cotranslational ubiquitination and quality control at the ribosome. Mol Cell *50*, 379-393.

Eichelbaum, K., and Krijgsveld, J. (2014). Rapid temporal dynamics of transcription, protein synthesis, and secretion during macrophage activation. Mol Cell Proteomics *13*, 792-810.

Eichelbaum, K., Winter, M., Berriel Diaz, M., Herzig, S., and Krijgsveld, J. (2012). Selective enrichment of newly synthesized proteins for quantitative secretome analysis. Nat Biotechnol *30*, 984-990.

Eravci, M., Sommer, C., and Selbach, M. (2014). IPG strip-based peptide fractionation for shotgun proteomics. Methods in molecular biology (Clifton, NJ) *1156*, 67-77.
Geiger, T., Cox, J., and Mann, M. (2010). Proteomic changes resulting from gene copy number variations in cancer cells. PLoS genetics *6*, e1001090.
Goldberg, A.L. (2003). Protein degradation and protection against misfolded or damaged proteins. Nature *426*, 895-899.
Goldberg, A.L., and Dice, J.F. (1974). Intracellular protein degradation in mammalian and bacterial cells. Annu Rev Biochem *43*, 835-869.
Hinkson, I.V., and Elias, J.E. (2011). The dynamic state of protein turnover: It's about time. Trends Cell Biol *21*, 293-303.
Hou, J., Wang, X., McShane, E., Zauber, H., Sun, W., Selbach, M., and Chen, W. (2015). Extensive allele-specific translational regulation in hybrid mice. Molecular systems biology *11*, 825.
Howden, A.J.M., Geoghegan, V., Katsch, K., Efstathiou, G., Bhushan, B., Boutureira, O., Thomas, B., Trudgian, D.C., Kessler, B.M., Dieterich, D.C.*, et al.* (2013). QuaNCAT: quantitating proteome dynamics in primary cells. Nat Methods *10*, 343-346.
Jovanovic, M., Rooney, M.S., Mertins, P., Przybylski, D., Chevrier, N., Satija, R., Rodriguez, E.H., Fields, A.P., Schwartz, S., Raychowdhury, R.*, et al.* (2015). Immunogenetics. Dynamic profiling of the protein life cycle in response to pathogens. Science *347*, 1259038.
Kiick, K.L., Saxon, E., Tirrell, D.A., and Bertozzi, C.R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proc Natl Acad Sci U S A *99*, 19-24.
Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J.*, et al.* (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol Cell *44*, 325-340.
Kristensen, A.R., Gsponer, J., and Foster, L.J. (2013). Protein synthesis rate is the predominant regulator of protein expression during differentiation. Molecular systems biology *9*, 689.
Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Methods *11*, 319-324.
Lam YW, L.A., Mann M, Andersen JS (2007). Analysis of Nucleolar Protein Dynamics Reveals the Nuclear Degradation of Ribosomal Proteins. Curr Biol *17*, 749–760.
Larance, M., Ahmad, Y., Kirkwood, K.J., Ly, T., and Lamond, A.I. (2013). Global subcellular characterization of protein degradation using quantitative proteomics. Mol Cell Proteomics *12*, 638-650.
Leibiger, C., Kosyakova, N., Mkrtchyan, H., Glei, M., Trifonov, V., and Liehr, T. (2013). First molecular cytogenetic high resolution characterization of the NIH 3T3 cell line by murine multicolor banding. The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society *61*, 306-312.
Li, G.W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell *157*, 624-635.
Liao, L., Park, S.K., Xu, T., Vanderklish, P., and Yates, J.R. (2008). Quantitative proteomic analysis of primary neurons reveals diverse changes in synaptic protein content in fmr1 knockout mice. Proceedings of the National Academy of Sciences *105*, 15281-15286.
Liu, B., Han, Y., and Qian, S.B. (2013). Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. Mol Cell *49*, 453-463.
Marsh, Joseph A., Hernández, H., Hall, Z., Ahnert, Sebastian E., Perica, T., Robinson, Carol V., and Teichmann, Sarah A. (2013). Protein Complexes Are under Evolutionary Selection to Assemble via Ordered Pathways. Cell *153*, 461-470.
Matalon, O., Horovitz, A., and Levy, E.D. (2014). Different subunits belonging to the same protein complex often exhibit discordant expression levels and evolutionary properties. Current opinion in structural biology *26*, 113-120.

Motoyama, A., Xu, T., Ruse, C.I., Wohlschlegel, J.A., and Yates, J.R., 3rd (2007). Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides. Analytical chemistry *79*, 3623-3634.
Nedialkova, D.D., and Leidel, S.A. (2015). Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. Cell *161*, 1606-1618.
Okamura, Y., Aoki, Y., Obayashi, T., Tadaka, S., Ito, S., Narise, T., and Kinoshita, K. (2015). COXPRESdb in 2015: coexpression database for animal species by DNA-microarray and RNAseq-based expression data with multiple quality assessment systems. Nucleic acids research *43*, D82-86.
Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.
Ori, A., Iskar, M., Buczak, K., Kastritis, P., Parca, L., Andrés-Pons, A., Singer, S., Bork, P., and Beck, M. (2016). Spatiotemporal variation of mammalian protein complex stoichiometries. Genome Biology *17*, 47.
Puchades, M., Westman, A., Blennow, K., and Davidsson, P. (1999). Removal of sodium dodecyl sulfate from protein samples prior to matrix-assisted laser desorption/ionization mass spectrometry. Rapid communications in mass spectrometry : RCM *13*, 344-349.
Rappsilber, J., Ishihama, Y., and Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. Analytical chemistry *75*, 663-670.
Ruepp, A., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Stransky, M., Waegele, B., Schmidt, T., Doudieu, O.N., Stumpflen, V.*, et al.* (2008). CORUM: the comprehensive resource of mammalian protein complexes. Nucleic acids research *36*, D646-650.
Santaguida, S., and Amon, A. (2015). Short- and long-term effects of chromosome mis-segregation and aneuploidy. Nature reviews Molecular cell biology *16*, 473-485.
Schimke, R.T., and Doyle, D. (1970). Control of enzyme levels in animal tissues. Annu Rev Biochem *39*, 929-976.
Schoenheimer, R. (1942). The Dynamic State of Body Constituents (Cambridge, MA: Harvard University Press).
Schubert, U., Antón, L.C., Gibbs, J., Norbury, C.C., Yewdell, J.W., and Bennink, J.R. (2000). Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. Nature *404*, 770-774.
Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.
Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58-63.
Shalgi, R., Hurt, J.A., Krykbaeva, I., Taipale, M., Lindquist, S., and Burge, C.B. (2013). Widespread regulation of translation by elongation pausing in heat shock. Mol Cell *49*, 439-452.
Sheean, M.E., McShane, E., Cheret, C., Walcher, J., Muller, T., Wulf-Goldenberg, A., Hoelper, S., Garratt, A.N., Kruger, M., Rajewsky, K.*, et al.* (2014). Activation of MAPK overrides the termination of myelin growth and replaces Nrg1/ErbB3 signals during Schwann cell development and myelination. Genes & development *28*, 290-303.
Sin, C., Chiarugi, D., and Valleriani, A. (2016). Degradation Parameters from Pulse-Chase Experiments. PLoS ONE *11*, e0155028.
Sormanni, P., Camilloni, C., Fariselli, P., and Vendruscolo, M. (2015). The s2D method: simultaneous sequence-based prediction of the statistical populations of ordered and disordered regions in proteins. J Mol Biol *427*, 982-996.

Stingele, S., Stoehr, G., Peplowska, K., Cox, J., Mann, M., and Storchova, Z. (2012). Global analysis of genome, transcriptome and proteome reveals the response to aneuploidy in human cells. Molecular systems biology *8*, 608.
Subtelny, A.O., Eichhorn, S.W., Chen, G.R., Sive, H., and Bartel, D.P. (2014). Poly(A)-tail profiling reveals an embryonic switch in translational control. Nature *508*, 66-71.
Sury, M.D., McShane, E., Hernandez-Miranda, L.R., Birchmeier, C., and Selbach, M. (2015). Quantitative proteomics reveals dynamic interaction of c-Jun N-terminal kinase (JNK) with RNA transport granule proteins splicing factor proline- and glutamine-rich (Sfpq) and non-POU domain-containing octamer-binding protein (Nono) during neuronal differentiation. Mol Cell Proteomics *14*, 50-65.
tom Dieck, S., Kochen, L., Hanus, C., Heumuller, M., Bartnik, I., Nassim-Assir, B., Merk, K., Mosler, T., Garg, S., Bunse, S*., et al.* (2015). Direct visualization of newly synthesized target proteins in situ. Nat Meth *12*, 411-414.
Toyama, B.H., Savas, J.N., Park, S.K., Harris, M.S., Ingolia, N.T., Yates, J.R., 3rd, and Hetzer, M.W. (2013). Identification of long-lived proteins reveals exceptional stability of essential cellular structures. Cell *154*, 971-982.
Tyler, R.E., Pearce, M.M.P., Shaler, T.A., Olzmann, J.A., Greenblatt, E.J., and Kopito, R.R. (2012). Unassembled CD147 is an endogenous endoplasmic reticulum-associated degradation substrate. Mol Biol Cell *23*, 4668-4678.
Vabulas, R.M., and Hartl, F.U. (2005). Protein synthesis upon acute nutrient restriction relies on proteasome function. Science *310*, 1960-1963.
Vavouri, T., Semple, J.I., Garcia-Verdugo, R., and Lehner, B. (2009). Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity. Cell *138*, 198-208.
Ward, C.L., and Kopito, R.R. (1994). Intracellular turnover of cystic fibrosis transmembrane conductance regulator. Inefficient processing and rapid degradation of wild-type and mutant proteins. J Biol Chem *269*, 25710-25718.
Wells, J.N., Bergendahl, L.T., and Marsh, J.A. (2016). Operon Gene Order Is Optimized for Ordered Protein Complex Assembly. Cell reports *14*, 679-685.
Wheatley, D.N., Giddings, M.R., and Inglis, M.S. (1980). Kinetics of degradation of "short-" and "long-lived" proteins in cultured mammalian cells. Cell Biol Int Rep *4*, 1081-1090.
Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A*., et al.* (2011). Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography *67*, 235-242.
Wisniewski, J.R., Zougman, A., and Mann, M. (2009). Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome. J Proteome Res *8*, 5674-5678.
Xie, S.Q., Nie, P., Wang, Y., Wang, H., Li, H., Yang, Z., Liu, Y., Ren, J., and Xie, Z. (2016). RPFdb: a database for genome wide information of translated mRNA generated from ribosome profiling. Nucleic acids research *44*, D254-258.

Figure Legends

Figure 1 I Global quantification of protein degradation kinetics by AHA pulse-chase (AHA p-c)

A) Exponential decay can be recognized as a straight line (in a semi log plot) indicating that the degradation rate is constant, i.e. young and old molecules have the same degradation probability per unit time. **B)** Experimental setup for global pulse-chase experiments. SILAC labeled cells are pulse-labeled with azidohomoalanine (AHA) and either directly harvested (time point 0 h) or chased in medium without AHA (cold chase). Samples are combined, AHA containing proteins are enriched, digested and analysed by LC-MS/MS. **C)** Measured MS1 spectra for three peptides representing the major types of decay profiles detected. Filamin alpha (Flna, ASGPGLNTTGVPASLPVEFTIDAK) shows slow exponential degradation, cathepsin L1 (Ctsl1, NLDHGVLLVGYGYEGTDSNK) shows fast exponential degradation, basigin (Bsg, VLQEDTLPDLHTK) shows non-exponential degradation. **D)** Decay profiles of individual proteins based on the median of three biological replicates (grey traces). Note that due to the experimental design not all proteins were detected at all time point. Increases in protein levels over time are theoretically impossible and probably reflect measurement noise. Highlighted profiles depict proteins shown in C) and are based on all three replicates (mean +/- sd). Outliers (> 130% protein left) were removed.

Figure 2 | Many proteins are degraded non-exponentially

A) Graphical representation of the two Markov models applied. The 1-state model and 2-state model reflect exponential degradation and non-exponential degradation, respectively. **B)** Fitting both models to the exemplary profiles from Fig. 1 D. For Flna and Ctsl1 both models have residual sum of squares (RSS) of similar size. The Akaike information criterion (AIC) therefore recommends the simpler 1-state model. The profile of Bsg is better explained by the 2-state model. **C)** Histogram of all probabilities for the 2-state model for all proteins that passed our quality criteria. **D)** All proteins with a 2-state probability > 0.8 had a larger initial (k_A) than subsequent degradation rate (k_B). **E)** The delta score (Δ-score) as a measure for the extent of non-exponential degradation. For each profile, a straight line is drawn between the 0 and 8 h time point (in semi log plot). The Δ-score corresponds to the distance of the measurement at 4 h from this line. Positive and negative Δ-scores indicate age-dependent stabilization and destabilization, respectively. **F)** Fractions of exponentially degraded (ED), non-exponentially degraded (NED) and undefined (UN) proteins defined by their AIC probabilities and Δ-scores.

Figure 3 | Validation of AHA p-c data

A) Experimental design for direct validation of non-exponential decay. Light (L) cells are pulsed with Heavy (H) SILAC medium for 4 h and split into two populations. The first population is harvested immediately while the second is chased for 8 h in Medium-heavy SILAC medium (M chase). If new proteins are less stable than old proteins their H/L ratio is expected to decrease during the chase. Two example spectra from an ED (**B**) and a NED protein (**C**) confirm this expectation. **D)** Proteins in the SILAC p-c experiment were classified according to their degradation profile. Only NED proteins show significantly (alpha = 0.05) reduced H/L ratios after the chase compared to all proteins. *** = p-value < 0.0001 from a one-sided Wilcoxon rank-sum test. **E**) Density distributions of ranked SILAC ratios for NED and ED proteins averaged across all four experiments (3 D and Fig S4 F-H). NED proteins were assigned a validation score (0 = low validation score; 5 = high validation score) that scales relative to the median of the ED protein distribution. **F**) Counts of NED proteins for the different validation score bins.

Figure 4 | NED is decreased by proteasome inhibition.

A) To quantify the impact of inhibitors on non-exponential degradation the Δ-score in treated and control cells is compared. **B-C)** Distributions of net Δ-scores for ED, UN and NED proteins displayed as boxplots. The proteasome inhibitor MG132 significantly reduced Δ-scores of NED proteins (one-sided Wilcoxon rank-sum test; *** = $p < 0.0001$) while the autophagy inhibitors wortmannin and bafilomycin A had no significant impact (alpha = 0.05). Numbers of proteins in each group are depicted. Classification of proteins is based on the original experiment (Figure 2 F). **D)** We estimated the effect of MG132 on protein on general degradation by plotting "% protein remaining" for all proteins with or without treatment. MG132 stabilized the majority of the measured proteins. **E)** To control for inhibition of lysosomal degradation samples acquired in parallel to the MS experiment were analyzed by western blot and probed against the autophagy marker LC3-II and for β-actin. "Autophagy" in the plot refers to the inhibitor combo.

Figure 5 | NED and protein complexes

A) NED proteins are significantly overrepresented in heteromeric protein complexes (Fisher's exact test, heteromeric vs monomeric or homomeric subunits). Numbers within bars represent raw subunit counts. The trend holds when ribosomal proteins are excluded (Fig. S6 C). **B)** NED proteins tend to form larger interfaces in complexes than ED proteins. Subunits were binned by the number of unique subunits per complex to control for the fact that NED proteins are overrepresented in larger complexes. P-values were calculated using Wilcoxon rank-sum tests, comparing NED to ED. Subunit counts are given along the bottom. **C)** NED subunits of large complexes (>5 unique subunits) tend to assemble earlier than ED subunits. Normalized assembly scores of 0-to-1 indicate the first-to-last steps of a given (dis)assembly pathway. P-values were calculated using Wilcoxon rank-sum tests. Raw subunit counts are given within each box. The difference in assembly order was not significant for smaller complexes. **D)** NED proteins show stronger coexpression (mRNA level). For each subunit, the average coexpression correlation coefficient is calculated with all other subunits within the same complex. P-value is calculated with the Wilcoxon rank-sum test comparing NED to ED. **E)** A simple model can explain non-exponential degradation of subunits of a heteromeric complex: NED proteins (turquoise) are synthesized in super-stoichiometric amounts relative to ED proteins (red). Only a fraction of the NED protein molecules is stabilized by complex formation while the excess is degraded. **F)** NED proteins tend to be produced in super-stoichiometric amounts. Protein abundances after the pulse (t = 0 h) were normalized in a complex centered manner. **G)** Complex-centered analysis of ribosome profiling data supports super-stoichiometric production of NED proteins. **H**) Inhibition of rRNA synthesis with actinomycin D selectively increased Δ-scores of ribosomal proteins. See Fig. 4 A for experimental design. P-values in F-H are based on one-sided Wilcoxon rank-sum tests.

Figure 6 | NED is evolutionarily conserved

A) Relative fractions of NED (turquoise), ED (red) and undefined (grey) proteins in the diploid human epithelial cell line RPE-1. We mapped human proteins to their mouse orthologs and grouped them according to their degradation profile in mouse fibroblasts. Human proteins with ED mouse orthologs are enriched in ED proteins. Similarly, human proteins with NED mouse orthologs are enriched in NED proteins. P-values are based on a hypergeometric test. **B)** Orthologous human and mouse proteins show significantly correlated Δ-scores. Pearson's correlation coefficient (R) is derived from all plotted Δ-score pairs. **C)** NED subunits of protein complexes are synthesized in super-stoichiometric amounts relative to other subunits in human (left). This is even the case when the mouse definitions (for ED, UN and NED) are used on the human dataset (right). **D)** Analysis of ribosome profiling data from several species confirms that the super-stoichiometric synthesis of NED proteins is evolutionarily conserved. Depicted p-values are based on one-sided Wilcoxon rank-sum tests.

Figure 7 | NED, aneuploidy and attenuation

A) A model depicting the expected impact of gene amplification on protein synthesis and degradation. In normal (that is, disomic) cells NED proteins are over-synthesized relative to ED proteins in the same complex. Degradation of the excess molecules gives rise to their NED profile. Genomic amplification of NED proteins further increases over-production and thus initial degradation. **B)** Log2 fold changes of protein abundances after pulse in RPE-1 cells and RPE-1 cells carrying extra copies of specific chromosomes (sorted by chromosome and genomic position). The data was divided into disomic (black), trisomic (orange) and ambiguous regions (grey) based on genome sequencing data (Fig. S8). Regions with significantly different protein abundance in comparison to the disomic cells are marked with asterisks (one-sided Wilcoxon rank-sum; ***: $p < 0.0001$). **C)** NED proteins show increased initial

degradation (Δ-scores) when the corresponding genes are in trisomic regions. Degradation of ED proteins is not affected. **D-F)** NED predicts protein level attenuation. We compared steady-state protein levels in RPE-1 and RPE-1 trisomic cells using standard SILAC. Boxplots show log2 fold changes for ED and NED proteins in trisomic regions compared to all proteins in disomic regions. This analysis is shown for all proteins **(D)**, proteins that are part of complexes **(E)** and proteins that are not part of complexes **(F)**. The number of analyzable protein pairs is displayed below each boxplot. P-values were computed using one-sided Wilcoxon rank-sum tests and are shown for significantly different distributions (alpha = 0.05).

Supplemental Figure Legends

Figure S1. Related to Figure 1

Establishing AHA pulse-chase

A) Enrichment specificity of AHA labeled proteins. SILAC labeled heavy mouse fibroblasts were labeled for 3 h with AHA. The cells were lyzed and mixed 1:1 with Light unlabeled proteins. The AHA labeled proteins were enriched and the samples were analyzed using LC-MS/MS. In theory, only heavy labeled proteins should be enriched. **B)** The majority of proteins show up only in the heavy form (infinite ratio (Inf), n=3034) with few exceptions that were exclusively identified in the light form (-Inf, n=47). Light proteins included proteins with very few peptides identified and extracellular proteins (e.g. serum albumin) probably representing exogenous contaminants. A number of abundant proteins (n=1556) showed up in both channels. **C)** Comparison of SILAC ratios between the experiment in B) and two label-swap experiments show that they are strongly negatively correlated. The median SILAC ratio was >11 (excluding Inf values) for the three biological replicates (inversed ratios were used for the label-swap experiments). Thus, enrichment is overall highly specific. **D)** AHA labeling has no apparent impact on protein degradation. NIH 3T3 mouse fibroblasts were starved for methionine and cysteine for 45 min before being pulsed with ^{35}S cysteine in combination with either methionine or azidohomoalanine for 1 h. Cells were then washed before being chased in either excess of cysteine (10 fold normal concentration) or cycloheximide (50 µM). Proteins were resolved on a SDS-PAGE. **E)** Radioactivity was measured by a phosphorImager and total protein was assessed by coomassie blue staining. Quantification of protein degradation using **F)** Cysteine chase (n=3) and **G)** Cycloheximide (n=3). Error bars: SE, n.s.: not significant based on two sided t-test (alpha 0.05). **H)** AHA does not induce premature ribosomal fall-off. All proteins in the mouse Uniprot database were split in half resulting in a C-terminal and N-terminal part of each protein. All peptides detected in one AHA p-c experiment were matched

back to the new database and sorted in the two categories N or C-terminal (peptides matching to the middle were excluded). Plotted is the distribution of heavy intensity (~abundance after the pulse) for the detected peptides. If AHA labeling would induce premature Ribosomal fall-off we expected to see a high N to C-terminal ratio in the number of detected peptides and a higher abundance of N-terminal peptides.

Figure S2. Related to Figure 1

Normalization strategy

The raw data required normalization due to input differences (caused by pipetting errors, cell density variation, and so forth). Plot **A)** displays the non-normalized raw data from one biological replicate. For normalization an assumption of non-changing data points over time or conditions are required (e.g. median normalization). In our case we look for very stable proteins. We defined the most stable proteins over all biological replicates as proteins with the consistently lowest (H/L or H/M) ratios at all time points (LSD proteins, see Material and Methods). Example LSD proteins for mouse are shown in the left grey arrow. Each time point was then independently normalized by setting the geometric means for the LSD proteins (red points in **B)**) to 100 % (i.e. no degradation). The resulting shift in the data can be seen in **C)** with new median values for the LSD proteins as red dots.

Figure S3. Related to Figure 1 and 2

AHA pulse-chase is reproducible over biological replicates.

The reproducibility of the mouse NIH 3T3 AHA p-c data was evaluated at each of the three different filtering steps **A)**: i) 1 % FDR filtering on PSM and protein level; SILAC Count > 1 and 10 % Cutoff (subfigures **B, C** and **D)**. ii) Proteins were filtered to have more than four data points and an RSS smaller 0.05 (subfigures **E, F** and **G)**. iii) Proteins were required to have a Δ-score based on the 4 and 8 h time points (subfigures **H, I** and **J**). The overlap of proteins across the three experiments is shown as weighted Venn diagrams (**B, E** and **H**). The coefficient of variation (CV) was calculated for each time point from the intersecting proteins across the experiments (**C, F** and **I**). The corresponding number of proteins with a CV is given in each plot. Within the outermost grey area in the plot 90 % of the decay profiles reside. The second area contains 70 % of decay profiles and so forth. Available data points per protein and time point are shown in D, G and J.

Figure S4. Related to Figure 3.

Validation experiments

A-D) Comparison of half-lives measured by AHA p-c to published half-lives measured by dynamic SILAC (Schwanhäusser et. al., 2011). A) We initially assumed exponential degradation and calculated protein half-lives from the exponential fit ("exponential half-life"). For ED proteins this resulted in good agreement with the dynamic SILAC data. This is reassuring, especially since dynamic SILAC does not involve artificial amino acids and the way in which half-lives are computed is different. **B)** In contrast, NED proteins had overall shorter half-lives in AHA p-c data when we assumed exponential degradation. This is expected since the exponential fit does not take into account that these proteins become stabilized at later time points. **C)** We therefore used our Markov chain-based model to compute "steady state half-lives" (Deneke et al., 2013). This half-life corresponds to the time it takes for half of the protein molecules present at *steady state* to be degraded (see Materials and Methods). **D)** Importantly, this led to better overall agreement with the dynamic SILAC data. Proteins with half-lives > 300 h were excluded in A-D because they cannot be accurately quantified. **E)** Radioactive pulse-chase experiments coupled to immunoprecipitation (IP). Mouse fibroblasts were pulsed for 1 h with ^{35}S cysteine and methionine and then either directly collected (0 h time point) or chased for 4 or 8 h in medium containing excess unlabeled methionine and cysteine. The cells were then lysed and VCP and Cct3 were immunoprecipitated from the same lysate. Eluted proteins were analyzed by western blotting and the membranes were probed for VCP or Cct3 as loading control ("Western blot" in figure). The amount of newly synthesized proteins surviving the chase were detected using a phosphorImager ("PhosphorImager"). The ED protein VCP was found to be very stable consistent with the half-life of >100 h reported by the AHA p-c experiment. In contrast, the NED protein CCT3 had decreased considerably after 4 h chase but then stayed stable until the 8 h time point. The figure shows one

representative experiment out of two. **F-H)** SILAC pulse-chase experiments confirm NED assignments. Confirmation experiments as shown in Figure 3. Differences from the experimental procedure in the main figure are highlighted in red before each corresponding plot. **F)** A label swap experiment in where the Heavy medium used for the pulse was swapped with the Medium-heavy used for the pulse. **G)** The chase time was limited to 4h. **H)** Also with 4 h chase but with the same label swap as in F). *** = p-value < 0.0001 from a one sided Wilcoxon rank-sum test. Stars are displayed on top for significantly different distributions (alpha = 0.05).

Figure S5. Related to Figure 5

Characteristics of NED and ED proteins

A) Prediction of ED and NED based on different features. Protein features were tested to predict either exponential (ED) or non-exponential degradation (NED) in mouse by calculating the area under the curve (AUC) of a receiver operating characteristic (ROC) curve. Error bars were derived from 200 bootstrap repetitions and indicate the 90 % confidence interval of the corresponding calculated AUCs. Features were additionally categorized as being present (blue) or absent (grey) in the tested protein fraction. For example, the bar for "protein length" is colored grey for NED proteins, indicating that NED proteins tend to be shorter than ED and undefined proteins. **B)** NED proteins are more likely to belong to multiprotein complexes. Fraction of proteins listed in a manually curated protein complex database (CORUM). Fractions are shown for all, exponentially degraded (ED), non-exponentially degraded (NED) and undefined (UN) proteins. The p-value is derived from a hypergeometric test for enrichment against all proteins. **C)** Due to the very large size of ribosomes, they can significantly skew small datasets. Here, we repeated the analysis shown in Fig. 5 A after excluding ribosomal proteins. The observed tendency of NED proteins to form heteromeric complexes is still apparent after excluding ribosomal proteins. **D)** Coexpression analyses repeated using CORUM complexes. In order to confirm that the observations displayed in Fig. 5 D are not an artifact of the structural data, we replicated the procedure using data from the "core" (i.e. non-redundant) set of CORUM complexes (Ruepp et al., 2008). To maximize available data, non-mouse protein complex subunits were mapped to mouse genes wherever sequence identity was at least 70% (see Methods).

Figure S6. Related to Figure 5

The second degradation rate of NED proteins is similar to the exponential degradation rate of ED proteins in the same complex

A) NED proteins have two degradation rates (see also Fig. 2 A and Supplemental Table 1): The initial one (k_A) is in our model related to the free subunit or subunits residing in partially assembled protein complexes. The second degradation rate (k_B) is associated with subunits in the fully assembled complex. ED proteins only have a single degradation rate which is equivalent to the degradation rate of the fully assembled complex. We plotted degradation rates A and B for NED proteins and the single ED degradation rate for selected complexes (**B-H**). In all cases, degradation rate k_B of NED proteins is lower and more similar than k_A to the ED degradation rates of proteins in the same complex.

Figure S7. Controls for Figure 6 and 7

Genomic characteristics of RPE-1 and RPE-1 trisomic cells

A-B) Low coverage whole genome sequencing of RPE-1 and RPE-1 trisomic cells. Gene copy number estimates based on sequencing of genomic DNA from RPE-1 parental and RPE-1 trisomic cells. Copy number estimates are based on mappability corrected read counts (see Materials and Methods) and ordered over the chromosomes. **A)** RPE-1 parental cell-line is trisomic for part of chromosome 10 and displays clonal expansion of a population of cells also trisomic for chromosome 12. **B)** The trisomic RPE-1 cells are fully trisomic for chromosome 5, 10 and 12 and a region of chromosome 11. For the downstream analysis only chromosome 5 and part of 11 where used while chromosomes 10 and 12 where ignored. **C-D)** Chromosome painting of parental and Trisomic RPE-1 cells. Chromosome paintings labeling chromosome 11 (magenta) and 5 (green). **C)** RPE-1 parental cell line is disomic for chromosome 5 and 11. **D)** RPE-1 trisomic cells are trisomic for chromosome 5 and part of 11. The trisomic part of chromosome 11 has fused to an unidentified chromosome.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit polyclonal anti-VCP	LifeSpan BioSciences	Cat#LS-C287469
Rabbit polyclonal anti-CCT3	Proteintech group	Cat# 10571-1-AP RRID:AB_2073658
Mouse monoclonal anti-beta-actin	Sigma-Aldrich	Cat# A5441 RRID:AB_476744
Rabbit polyclonal anti-LC3	Novus Biologicals	Cat# NB100-2220 RRID:AB_10003146
Chemicals, Peptides, and Recombinant Proteins		
Actinomycin D	Sigma-Aldrich	Cat#A9415 CAS: 50-76-0
(S)-MG-132 (MG132)	Cayman chemical	Cat#10012628 CAS: 133407-82-6
Bafilomycin A1	Invivogen	Cat#tlrl-baf1 CAS: 88899-55-2
Wortmannin	Merck Millipore	Cat#681676 CAS:19545-26-7
Cycloheximide	Sigma-Aldrich	Cat#C4859 CAS: 66-81-9
Demecolcine (Colchicine)	Sigma-Aldrich	Cat#D1925 CAS: 477-30-5
L-Azidohomoalanine (AHA)	Anaspec	Cat#AS-63699
L-Cysteine [35S]	PerkinElmer	Cat#NEG022T005MC
L-Lysine 4,4,5,5-D4 (Lys4)	Cambridge Isotope Laboratories	Cat#DLM-2640
L-Lysine $^{13}C_6$ $^{15}N_2$ (Lys8)	Cambridge Isotope Laboratories	Cat#CNLM-291-H
L-Arginine $^{13}C_6$ (Arg6)	Sigma-Aldrich	Cat#643440
L-Arginine $^{13}C_6{}^{15}N_4$ (Arg10)	Sigma-Aldrich	Cat#608033
Critical Commercial Assays		
Click-iT protein enrichment-kit	Invitrogen	Cat#C10416
EasyTag™ EXPRESS35S Protein Labeling Mix	PerkinElmer	Cat#NEG772014MC
TruSeq ChIP Library Prep Kit	Illumina	Cat#IP-202-1012

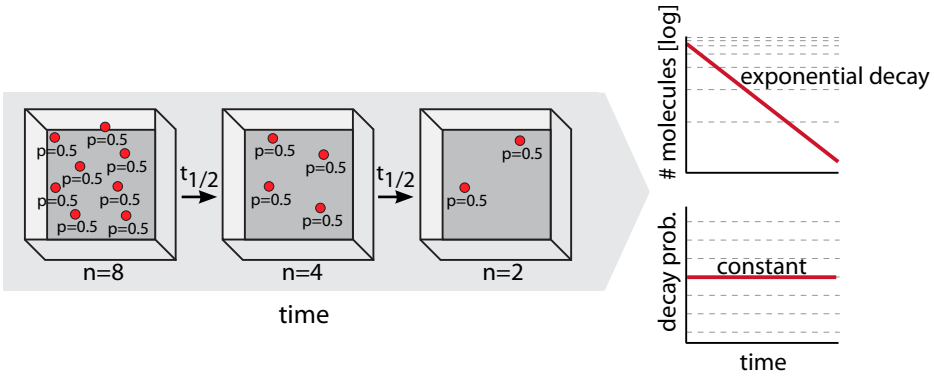
SureSelect^{XT} Target Enrichment System Kit for Illumina Multiplexed Sequencing	Agilent	Publication Number G7530-90000
Blood and Cell Culture DNA kit	Qiagen	Cat#13323
Deposited Data		
DNA sequencing data set	This paper	NCBI BioProject: ID PRJNA339199 http://www.ncbi.nlm.nih.gov/bioproject/339199
Proteomic data sets	This paper	TBA
Human UniprotKB/Swiss-Prot data base (Human UniProt 2014-10)		http://www.uniprot.org/proteomes/
Mouse UniprotKB/Swiss-Prot data base (Mouse UniProt 2014-10)		http://www.uniprot.org/proteomes/
Human reference genome NCBI (hg19)		http://genome.ucsc.edu/
Ribosome profiling data (Human)	(Liu et al., 2013)	http://sysbio.sysu.edu.cn/rpfdb/index.html
Ribosome profiling data (Zebrafish)	(Chew et al., 2013)	http://sysbio.sysu.edu.cn/rpfdb/index.html
Ribosome profiling data (Worm)	(Nedialkova and Leidel, 2015)	http://sysbio.sysu.edu.cn/rpfdb/index.html
Ribosome profiling data (Mouse)	(Shalgi et al., 2013; Subtelny et al., 2014)	
Co-expression data mouse dataset: Mmu.v13-01.G20959-S31479; human dataset: Hsa.v13-01.G20280-S73083	COXPRESdb (Okamura et al., 2015)	http://coxpresdb.jp/
Protein structure data sets	PDB	http://www.rcsb.org/
Protein complex annotation data set (CORUM downloaded 2015-10-20)	(Ruepp et al., 2008)	http://mips.helmholtz-muenchen.de/genre/proj/corum/index.html
Protein complex annotation data set	(Ori et al., 2016)	
Experimental Models: Cell Lines		

NIH 3T3 mouse fibroblast cells	ATCC	N/A
RPE-1 hTERT human retinal pigmented epithelial cells	Stephen Taylor (University of Manchester, UK)	N/A
RPE-1 5/3 11/3 12/3 human retinal pigmented epithelium cells	(Stingele et al., 2012) and this study	N/A
Software and Algorithms		
R version 3.2.2 (2015-08-14)		https://cloud.r-project.org/
R package: ggplot v2.0.0		http://ggplot2.org/
R package: dplyr v0.4.3		https://cran.r-project.org/web/packages/dplyr/index.html
R package: gridExtra v2.0.0		https://cran.r-project.org/web/packages/gridExtra/index.html
R package: dgof v1.2		https://cran.r-project.org/web/packages/dgof/index.html
R package: pracma v1.8.6		https://cran.r-project.org/web/packages/pracma/index.html
R package: Vennerable 3.0		https://r-forge.r-project.org/R/?group_id=474
R package: Uniprot.ws 2.10.4		https://bioconductor.riken.jp/packages/3.2/bioc/html/UniProt.ws.html
R package: grImport 0.9-0		https://cran.r-project.org/web/packages/grImport/index.html
R package: diagram 1.6.3		https://cran.r-project.org/web/packages/diagram/index.html
R package: biomaRt 2.26.1		https://bioconductor.org/packages/release/bioc/html/biomaRt.html

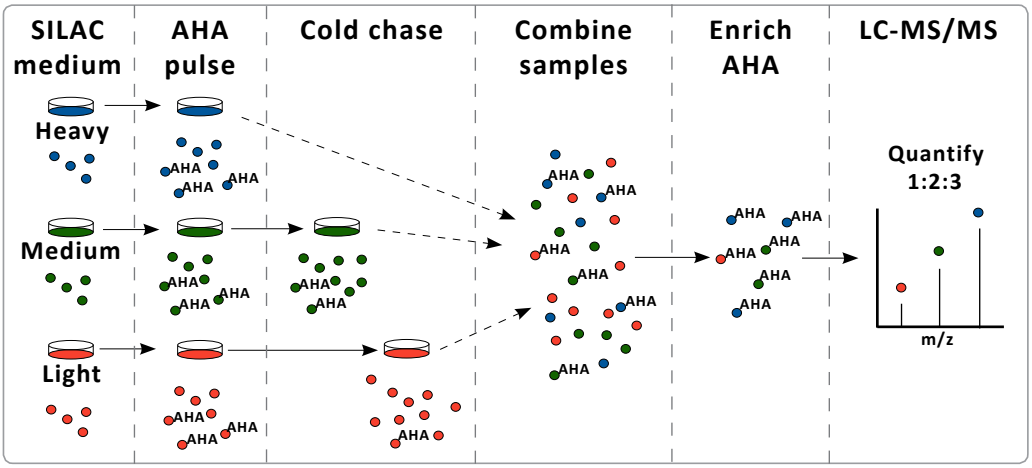
AREAIMOL		http://www.ccp4.ac.uk/
Assembly-prediction		http://github.com/marshlab/assembly-prediction
Bowtie2 v2.1.0		http://bowtie-bio.sourceforge.net/bowtie2
MATLAB R2013a		http://www.mathworks.com/products/matlab/
MATLAB package: Statistics Toolbox		http://www.mathworks.com/products/optimization/
MATLAB package: Optimization Toolbox		http://www.mathworks.com/products/statistics/
MATLAB package: Global Optimization Toolbox		http://www.mathworks.com/products/global-optimization/
MaxQuant v1.5.1.2	(Cox and Mann, 2008)	http://www.biochem.mpg.de/5111795/maxquant

FIGURE 1

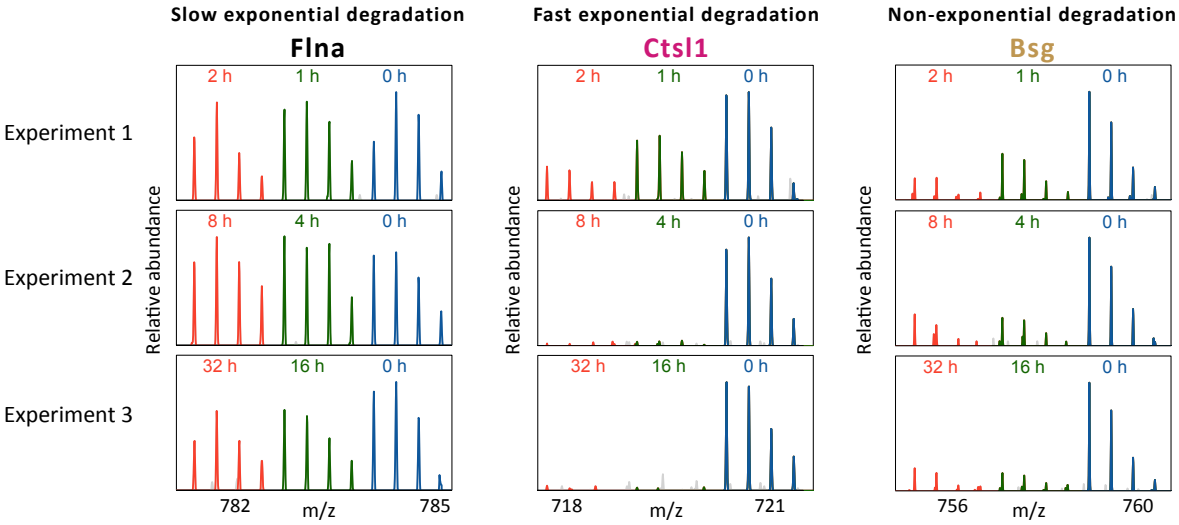
A



B



C



D

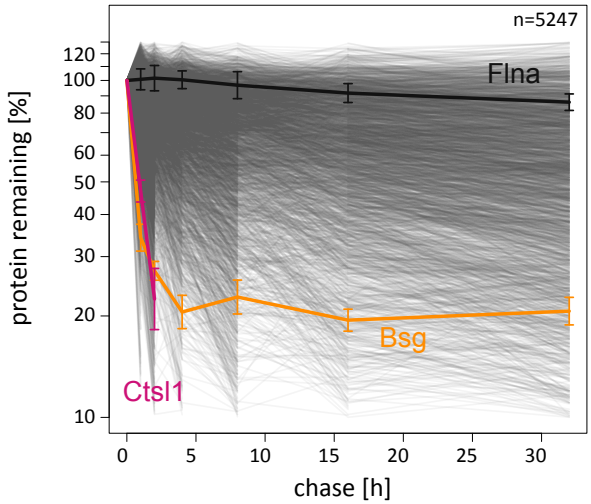


FIGURE 2

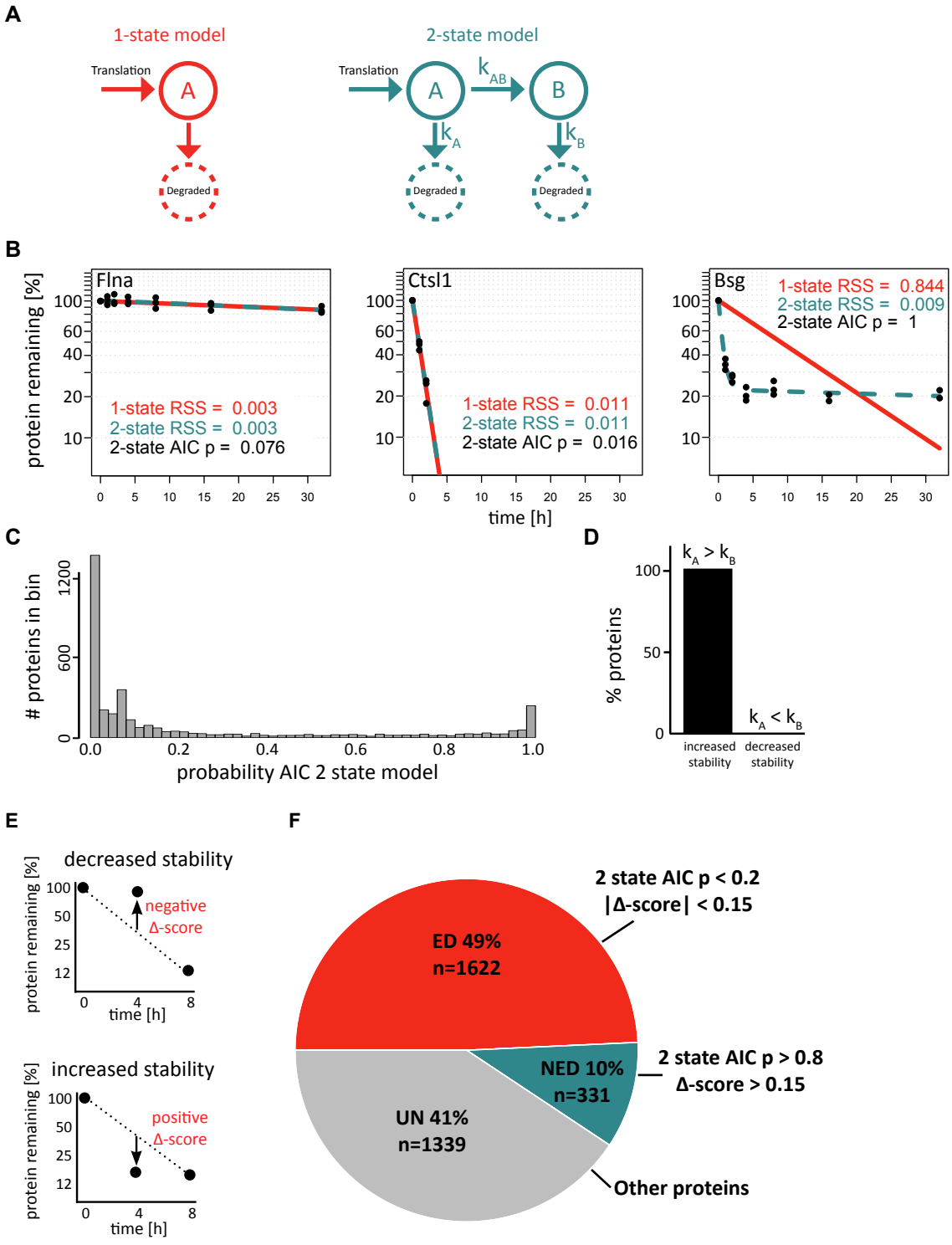


FIGURE 3

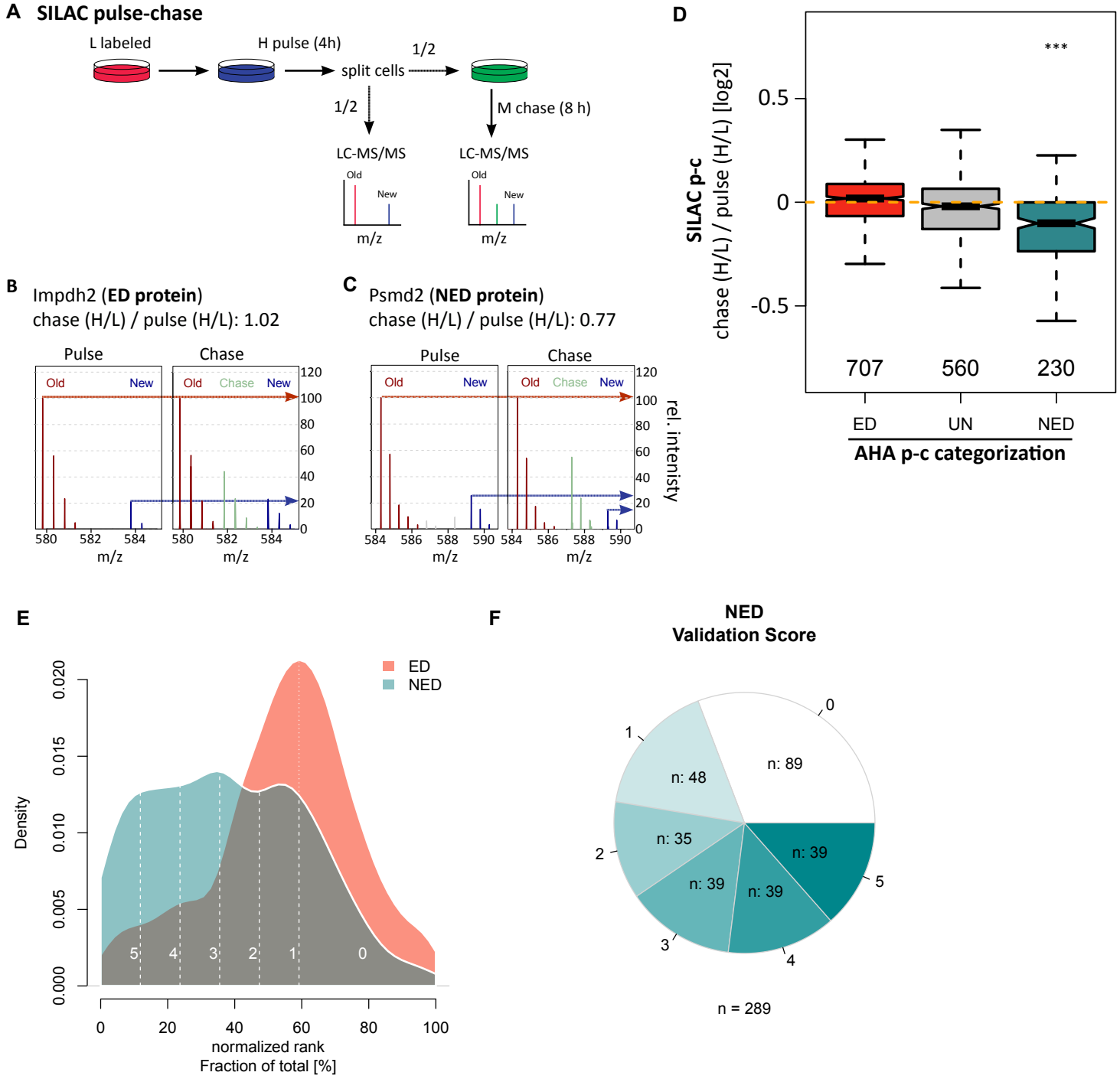
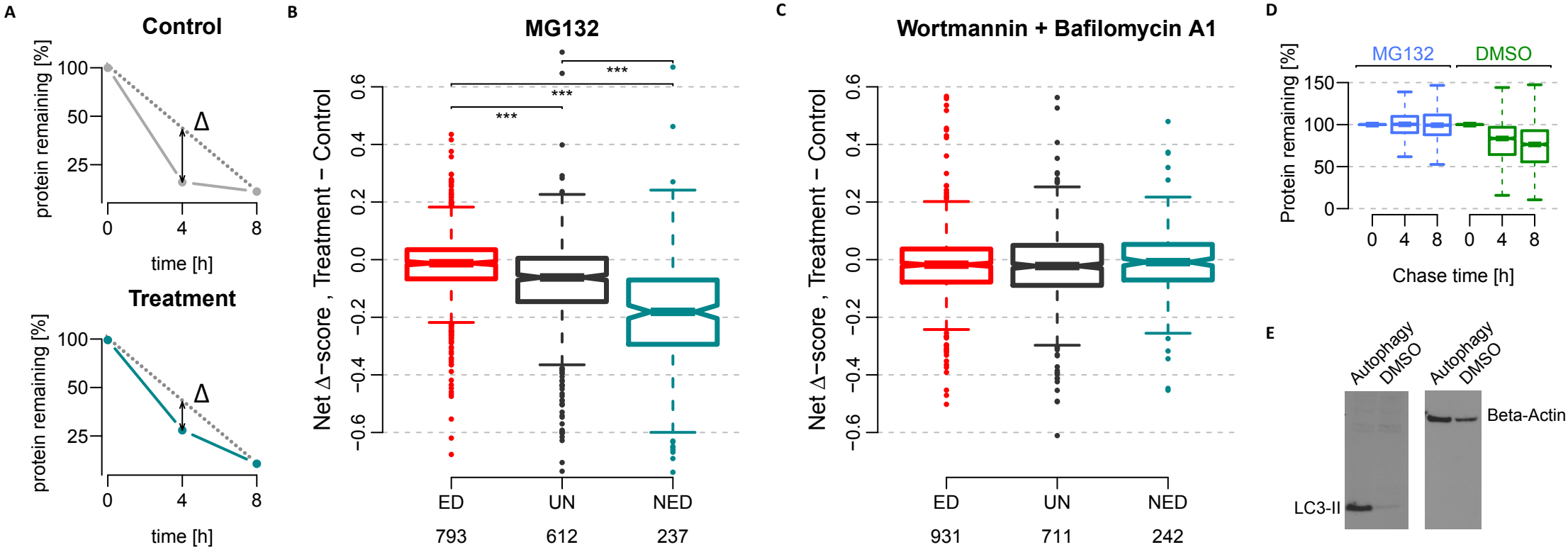


Figure 4

Click here to download Figure Figure 4.pdf

FIGURE 4



Click here to download Figure Figure_5.pdf

FIGURE 5

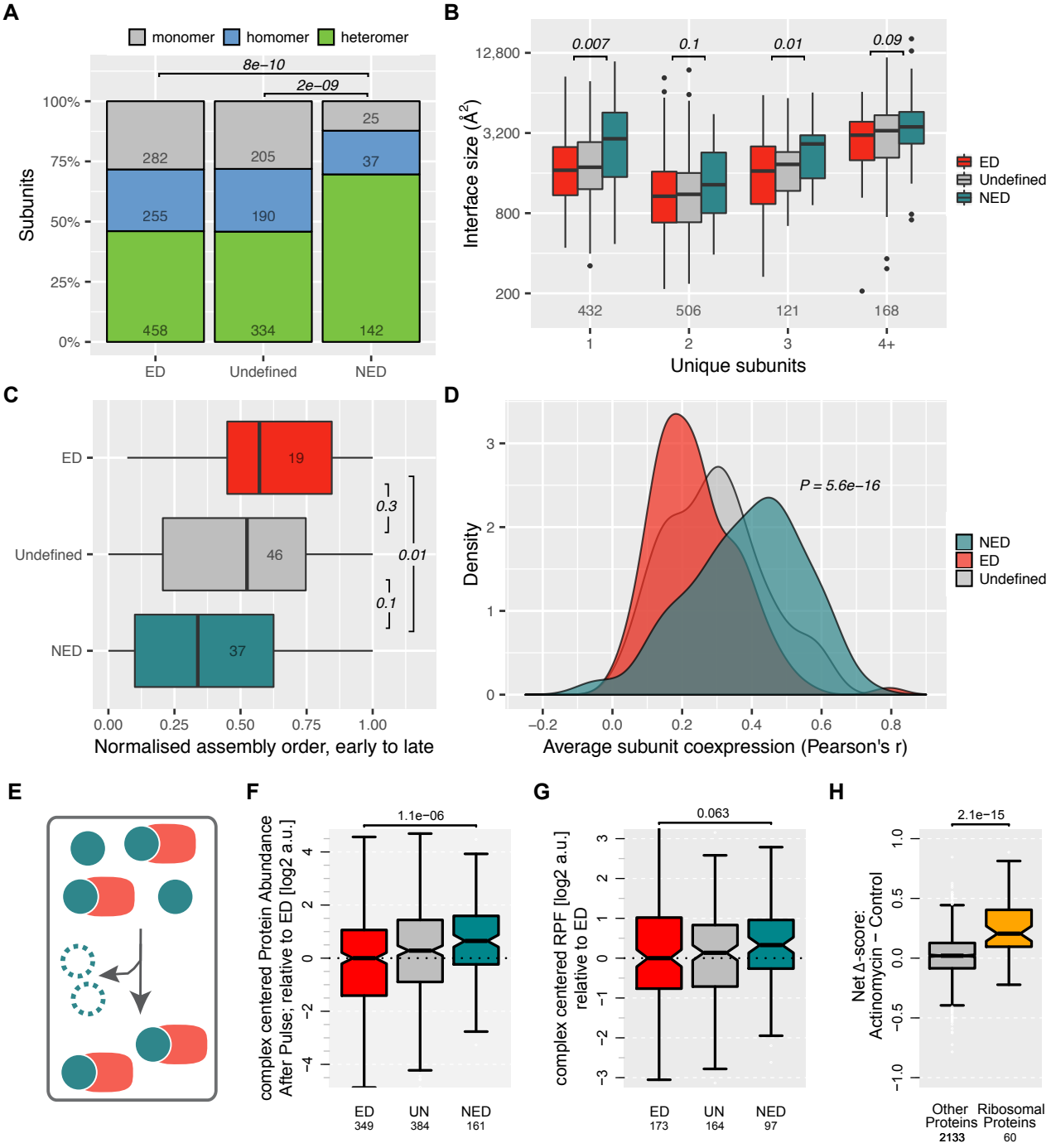
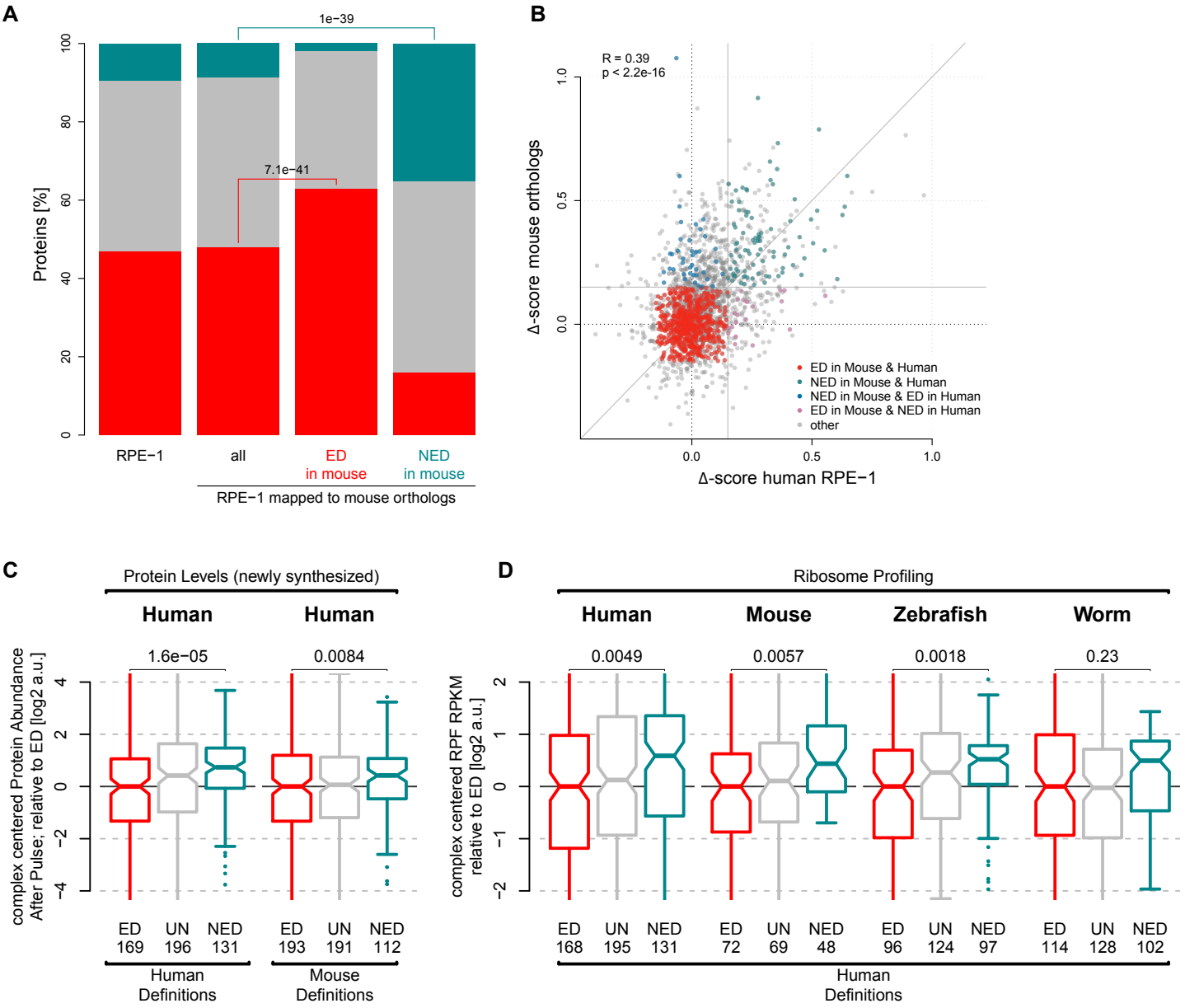
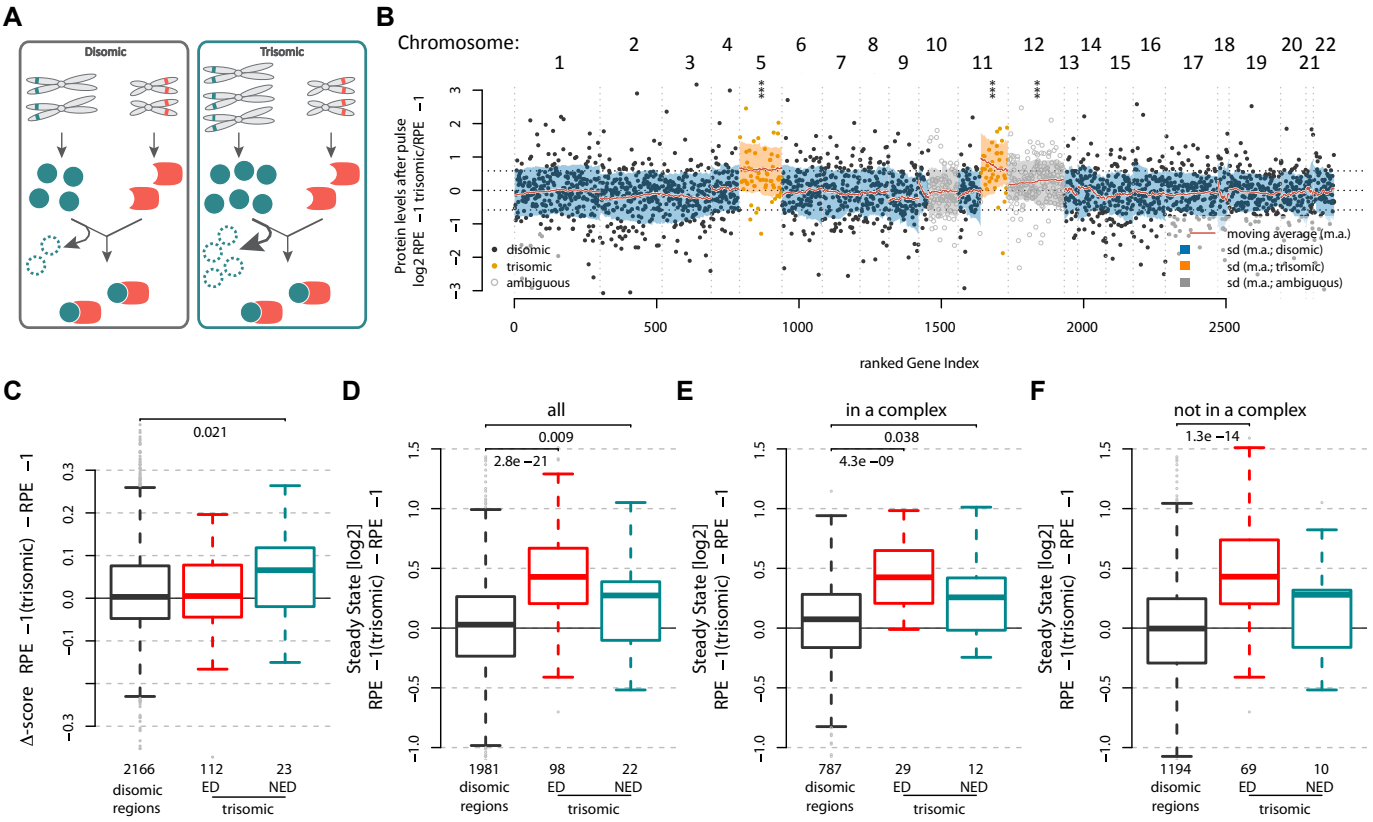


FIGURE 6



Click here to download Figure Figure_7.pdf

FIGURE 7



Supplemental Figure 1
FIGURE S1

Click here to download Supplemental Figure Figure S1.pdf

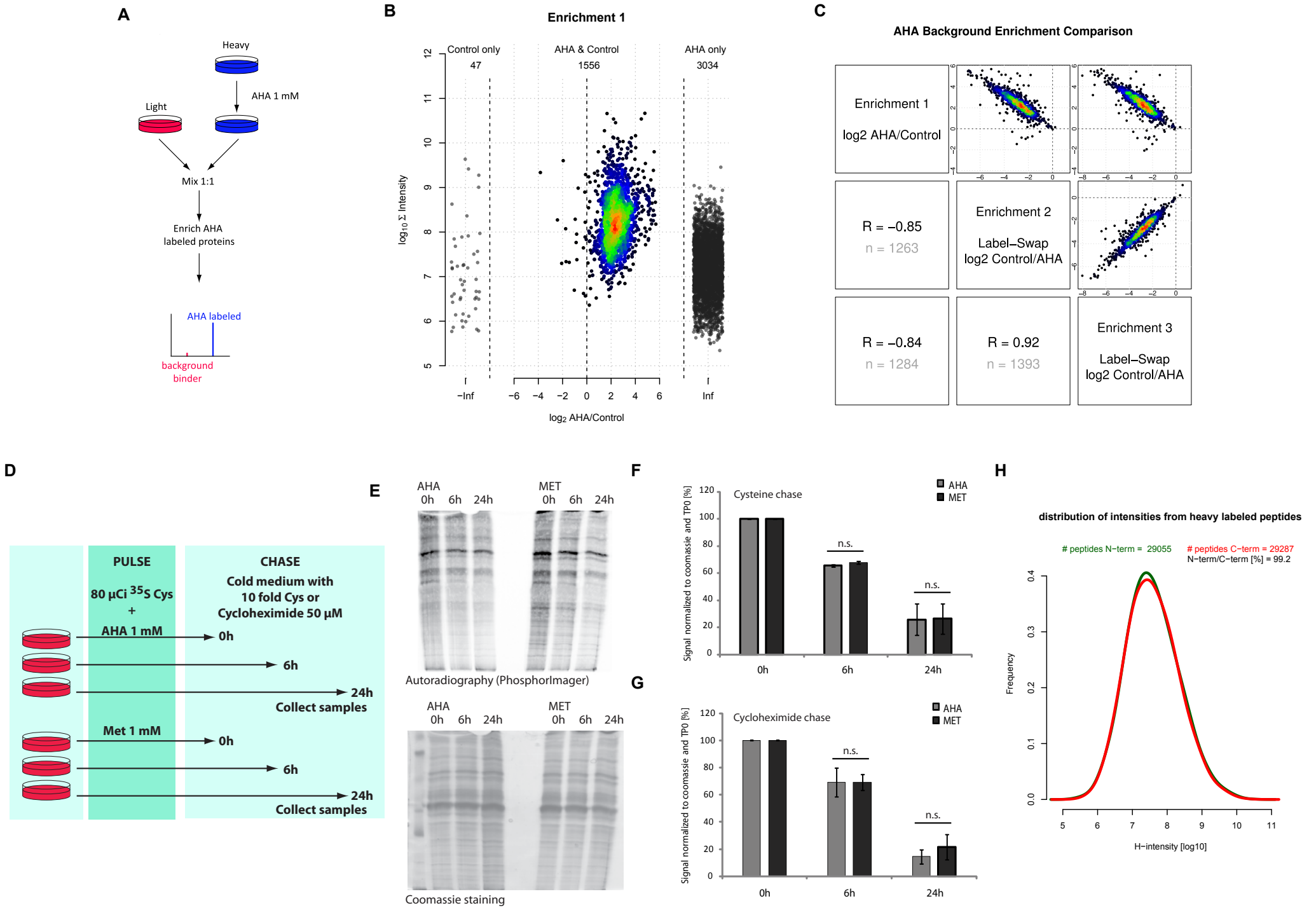
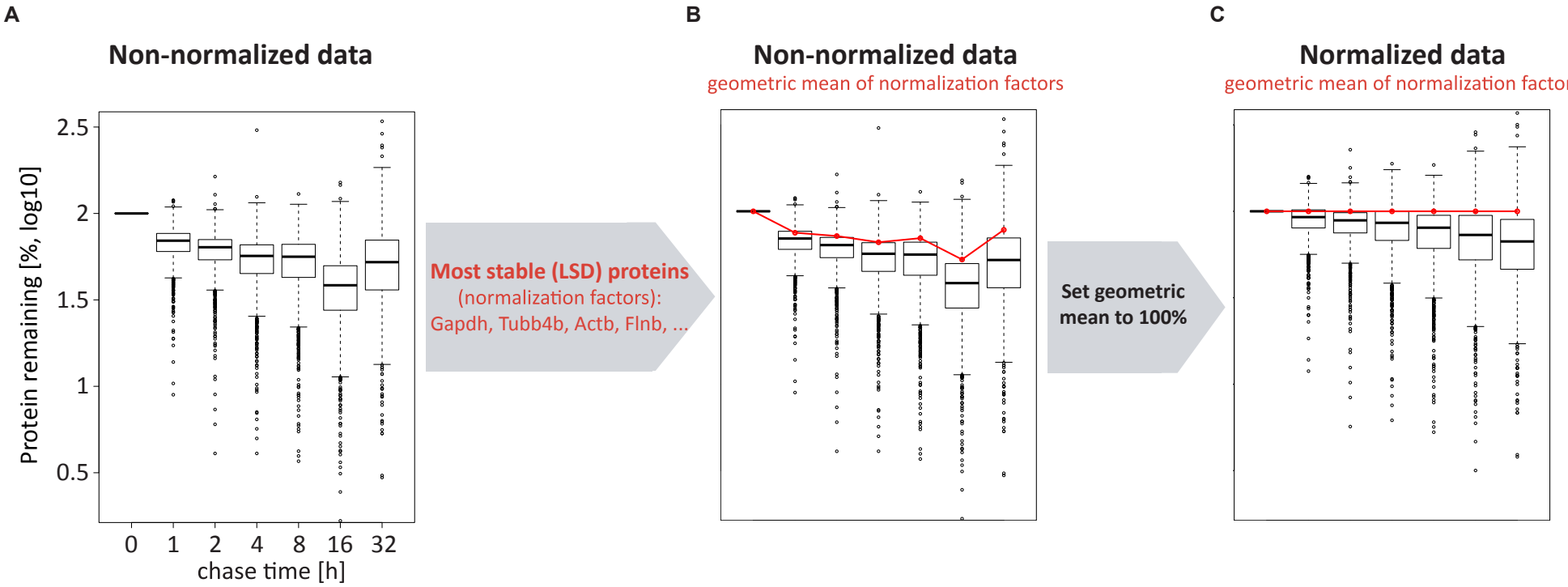


FIGURE S2



Supplemental Figure 3

FIGURE S3

Click here to download Supplemental Figure Figure S3.pdf

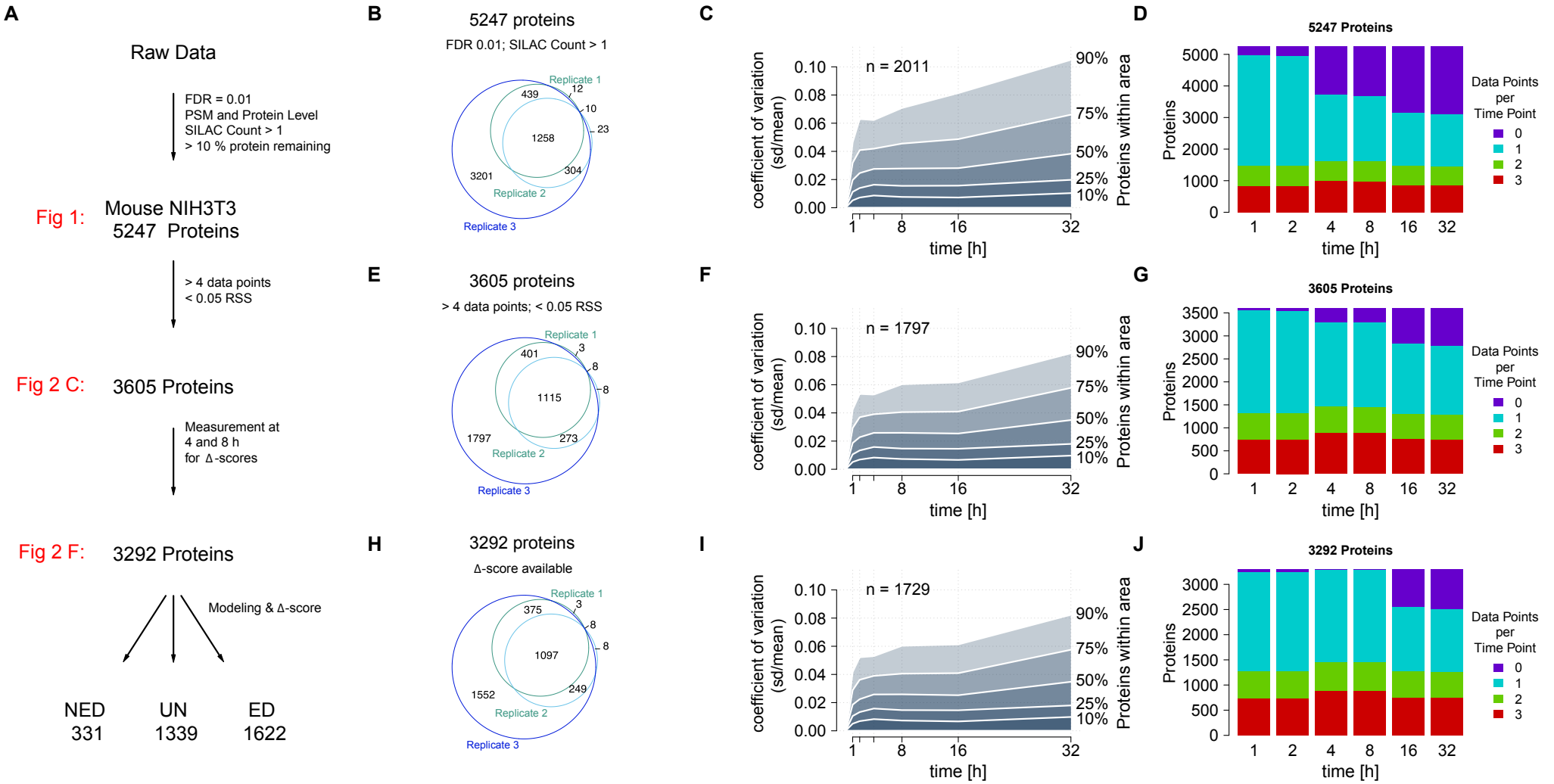


FIGURE S4

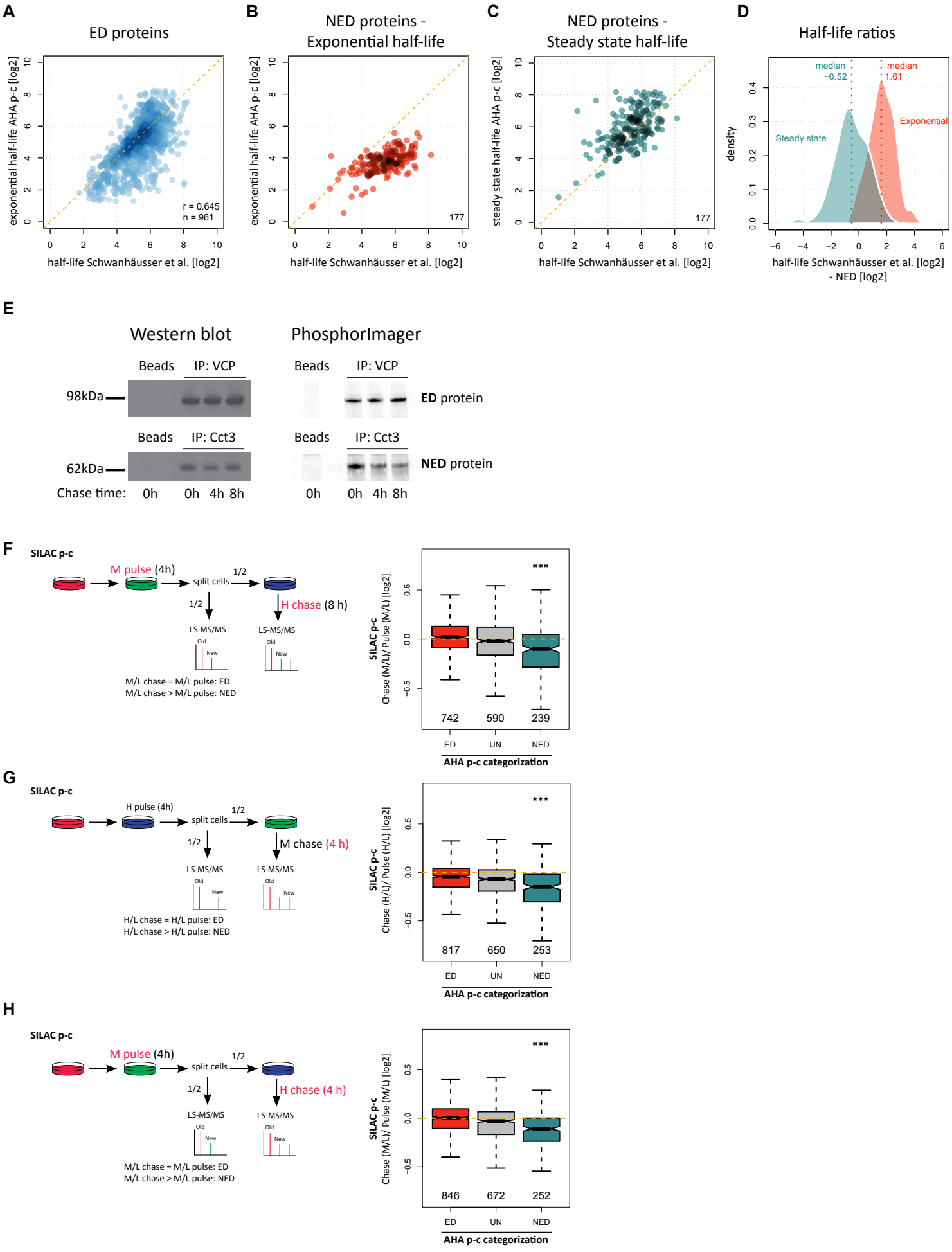


FIGURE S5

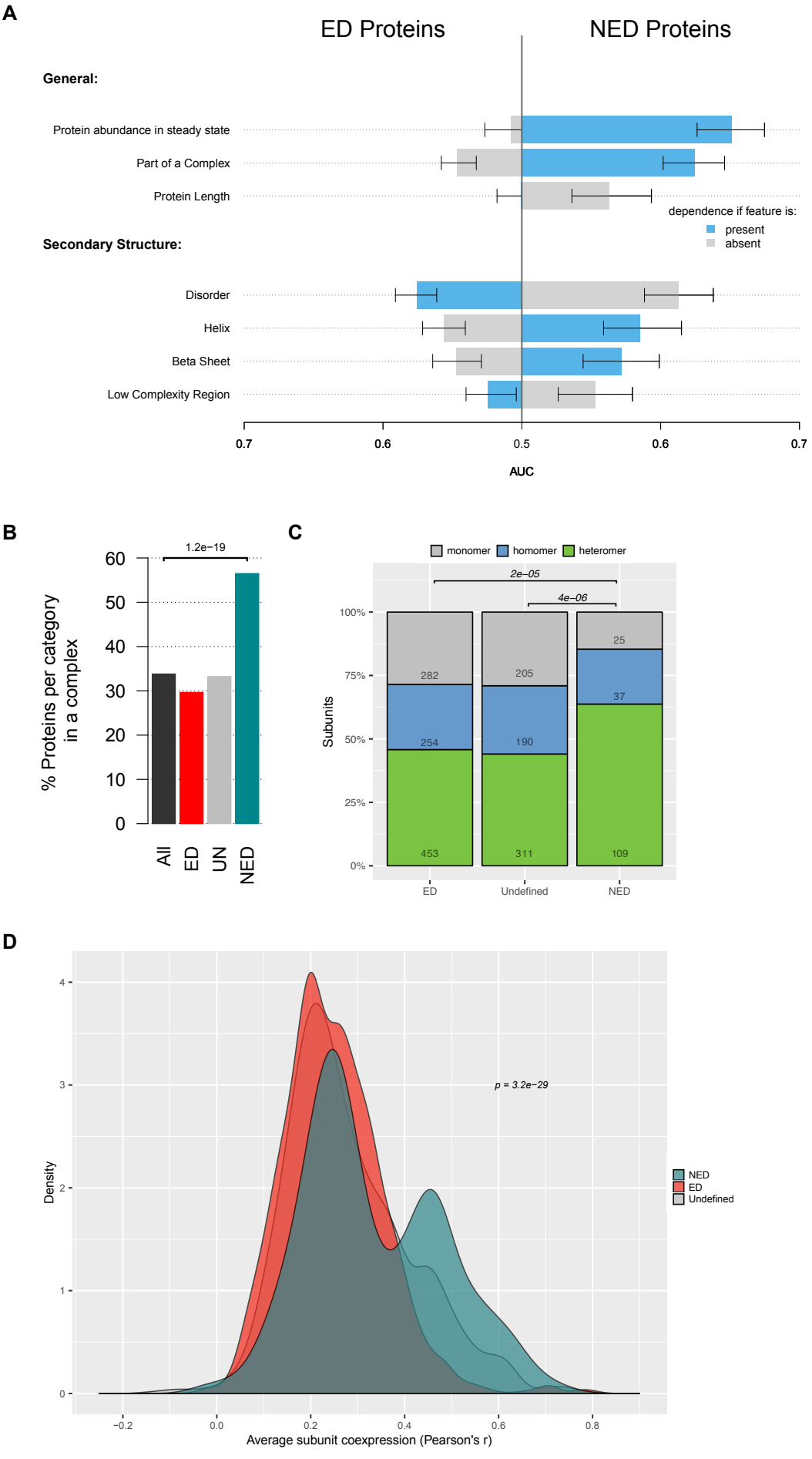
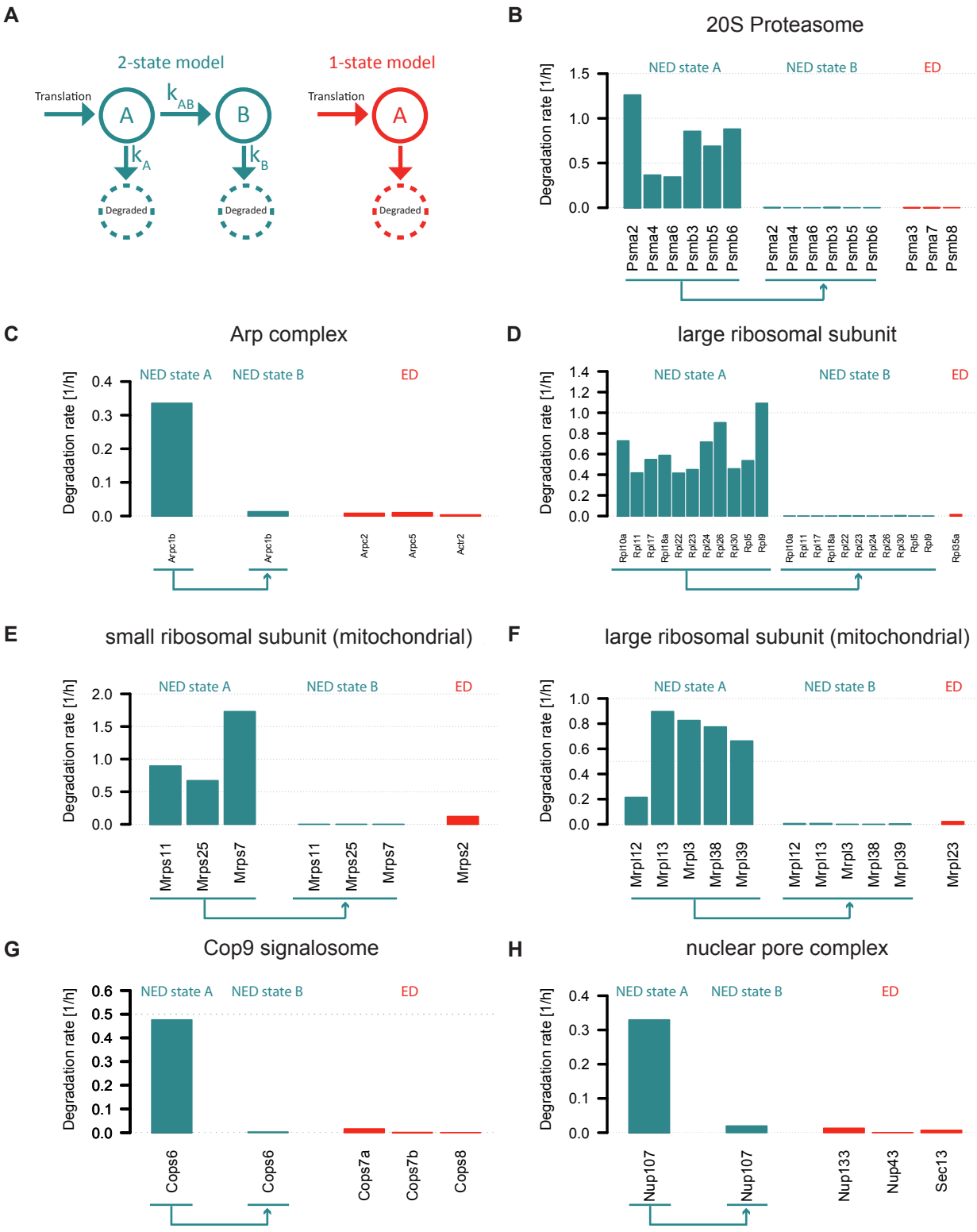


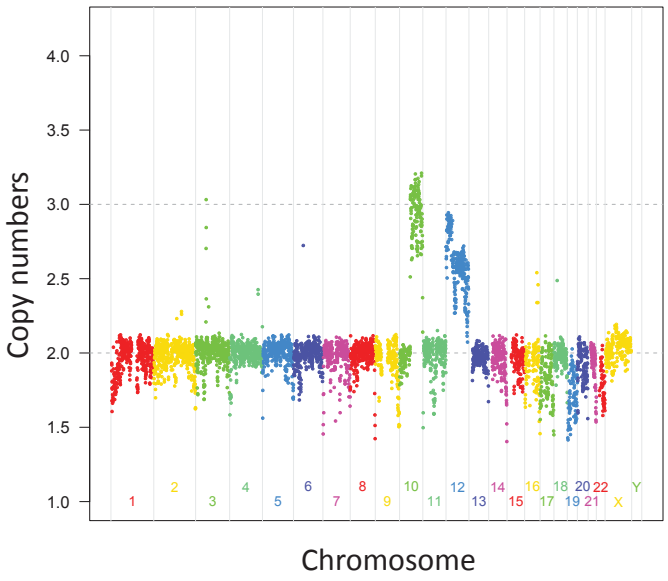
FIGURE S6



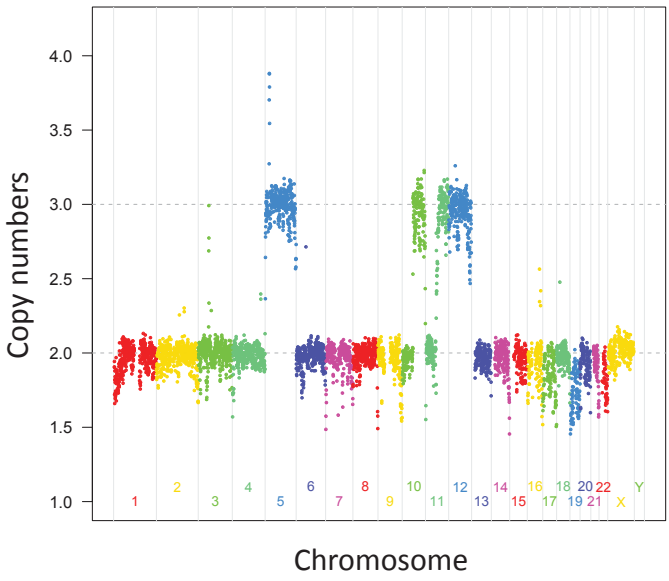
Supplemental Figure 7
FIGURE S7

Click here to download Supplemental Figure Figure S7.pdf

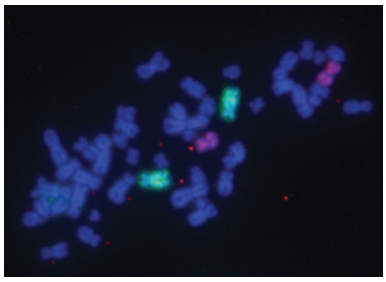
A



B

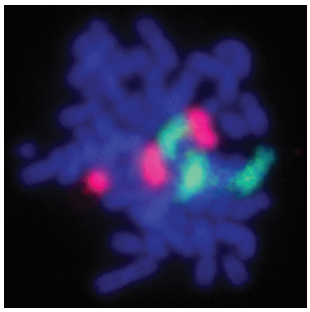


C

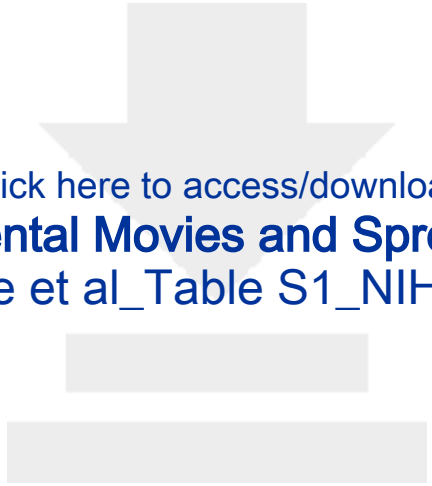


RPE-1
Chrom 11 Chrom 5

D



RPE-1 trisomic
Chrom 11 Chrom 5



Click here to access/download

Supplemental Movies and Spreadsheets

McShane et al_Table S1_NIH3T3.xlsx

Click here to access/download

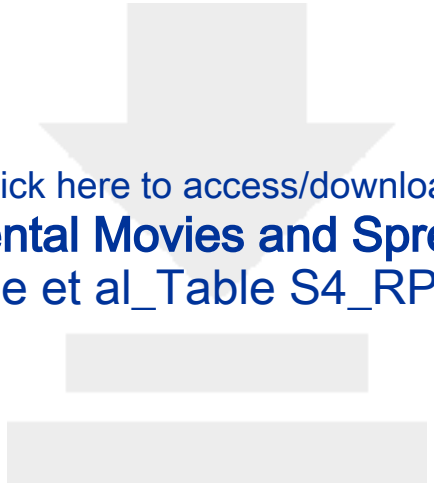
Supplemental Movies and Spreadsheets

McShane et al_Table S2_inhibitors dataset.xlsx

Click here to access/download

Supplemental Movies and Spreadsheets

McShane et al_Table S3_structure dataset.xlsx



Click here to access/download

Supplemental Movies and Spreadsheets

McShane et al_Table S4_RPE-1.xlsx