Global Quantification of Cellular Protein Degradation Kinetics

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m

(Dr. rer. nat.)

im Fach Biologie

eingereicht an der

Lebenswissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

M.Sc. Erik McShane

Präsident der Humboldt-Universität zu Berlin

Prof. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät

Prof. Dr. Bernhard Grimm

Gutachter/innen: 1. Prof. Dr. Matthias Selbach

2. Prof. Dr. Thomas Sommer

3. Prof. Dr. Britta Eickholt

Tag der mündlichen Prüfung: 20/03/2017

for my family

Table of Contents

Abstract

Proteins are thought to be degraded exponentially. That means that newly synthesized proteins have the same probability to be degraded as old proteins. However, evidence has accumulated showing that this is not true in all cases. To analyze this more systematically, we developed a method employing metabolic pulse-labeling by the non-canonical amino acid azidohomoalanine (AHA). AHA enables enrichment of newly synthesized proteins directly after pulse or after chase in AHA-free medium. We used SILAC and shotgun proteomics to quantify how much protein remains after different lengths of chase to create degradation profiles for thousands of proteins. Importantly, these degradation profiles allowed us to detect changes in degradation kinetics as the proteins age. We found that more than 10 % of proteins are non-exponentially degraded (NED). These protein are exclusively stabilized by age. Proteasomal degradation of excess protein complex subunits seems to explain a large fraction of NED. Comparing NED in mouse and human cells, we found that NED is at least partially conserved, seemingly due to cells consistently making too much of certain subunits. These overproduced subunits are on average shorter and more structured than the exponentially degraded proteins within the same complex. Finally, since excess NED proteins are degraded during baseline conditions, we hypothesized that making more of a NED protein would not increase its steady state levels. We employed an aneuploidy cell model and found that indeed NED proteins encoded on trisomic chromosomes did not increase in steady state levels to the same extent as exponentially degraded proteins. Instead, we recorded an increase in initial degradation of these proteins. In summary, we present a method for global pule-chase experiments allowing the detection of age-dependent protein degradation with possible implications for the understanding of aneuploidy and cancer.

Keywords: Azidohomoalanine, SILAC, pulse-chase experiments, non-exponential kinetics, protein degradation, protein complex assembly, shotgun proteomics, mass spectrometry, Aneuploidy

Zusammenfassung

Es wird allgemein angenommen, dass Proteine exponentiell degradiert werden. Das bedeutet, dass neu synthetisierte als auch alte Proteine mit gleicher Wahrscheinlichkeit degradiert werden. Es tauchen jedoch immer mehr Hinweise dafür auf, dass das nicht immer der Fall sein muss. Um diese Fragestellung systematisch anzugehen, haben wir eine Methode zur metabolischen Pulsmarkierung mit der nichtkanonischen Aminosäure Azidohomoalanine (AHA) entwickelt. AHA ermöglicht die Anreicherung von neu synthetisierten Proteinen direkt nach einem Puls oder nach einer „chase“ (Nachverfolgung) Periode in AHA freiem Medium. Wir kombinierten diese Methode mit SILAC und Shotgun Proteomik um zu quantifizieren wieviel Protein nach verschiedenen chase-Perioden übrig bleibt. Damit konnten wir Degradationsprofile für tausende von Proteinen erstellen. Diese erlauben uns Veränderungen in der Degradationskinetik von alternden Proteinen zu ermitteln. Unsere Daten zeigen, dass mehr als 10 % der Proteine nicht exponentiell degradiert werden (NED). Diese Proteine werden mit fortschreitendem Alter ausschließlich stabiler. Proteasomale Degradation von überschüssigen Proteinkomplexuntereinheiten scheint einen Großteil der NEDs zu erklären. Beim Vergleich zwischen murinen und humanen Zellen stellte sich heraus, dass NED teilweise konserviert ist. Das liegt scheinbar daran, dass diese Zellen trotz unterschiedlichem Ursprungs einheitlich bestimmte Untereinheiten überproduzieren. Diese sind im Durchschnitt kürzer und strukturierter als exponentiell degradierte Proteine aus dem gleichen Proteinkomplex. Da überschüssige NED Proteine bereits unter Standardbedingungen degradiert werden, nahmen wir an, dass die zusätzliche Überproduktion eines NED Proteins seine Level im stationären Zustand nicht verändern sollte. Um dies zu zeigen, quantifizierten wir Degradationskinetiken von Proteinen einer aneuploidenZelllinie. Wir fanden, dass NED Proteine, die auf trisomischen Chromosomen codiert sind, nicht in gleichem Maße ihr stationäres Level steigerten wie exponentiell degradierte Proteine. In Übereinstimmung mit unserer Hypothese verzeichneten wir stattdessen eine Zunahme der anfänglichen Degradationsraten dieser NED Proteine. Zusammenfassend ist eine Methode für globale Pulsmarkierung zur Analyse von altersabhängiger Proteindegradation mit möglichen Anwendungsgebieten bei Aneuploidie und Krebs vorgestellt.

Keywords: Azidohomoalanine, SILAC, pulse-chase, Nicht Exponentiell Degradationskinetik, Proteine Degradierung, Proteinekomplex, Shotgun Proteomik, Massenspectrometrie, Aneuploidie

Statement of Contributions

The work presented in this thesis is built on a collaborative effort. Many people were involved at different stages of the project. As a common example, I suggested that we model the number of data points that was required to define a NED protein, Matthias Selbach suggested a boot strap approach which was applied by Celine Sin and finally a figure (Figure 23) was produced by Henrik Zauber. However, as a rule of thumb it's good to know that the main idea (using AHA p-c to investigate non-exponential degradation) came from **Matthias Selbach** and most of the further developments were conjured in collaboration between him, Henrik Zauber and myself with the clear exceptions mentioned below. In addition, most wet lab work was performed by myself, once again with the exceptions below. To avoid confusion I have below outlined the individual contributions. For clarity I have also annotated the figure legends and respective material and methods segments with whom or by whom the work was performed.

Celine Sin and **Angelo Valleriani,** both at the MPI for Colloids and Interfaces, performed and developed the Markov-chain modeling see Figure 20 (script by Celine Sin), Figure 21 and Figure 22 for examples. Based on a suggestion to use the most stable proteins as normalization factors, Celine Sin also created and applied the elegant LSD normalization algorithm (see Figure 18). Both these steps are integral to analyzing the AHA p-c data. This is further exemplified by the fact that most of the data in the supplemental tables 1, 2, 4 and 5 is the direct output from their analysis.

Henrik Zauber performed much of the bioinformatic analysis presented in the thesis. He performed the computational analysis in and created Figure 16, Figure 22-24, Figure 26, Figure 27 E-F, Figure 29 A-D, Figure 35, Figure 37-38, Figure 42-43 and Figure 45. He also developed and implemented the complex centric analysis (Figure 33 B-C, Figure 39 and Figure 44) and distinguished ED and NED related properties using ROC analysis (Figure 30).

Joseph A. Marsh and **Jonathan N. Wells** (University of Edinburgh) performed the complex enrichment and complex assembly analysis and created Figure 31 A and Figure 32.

Neysan Donnelly and **Zuzana Storchova** (MPI for Biochemistry) produced the RPE-1 Cell lines, performed chromosome paintings (Figure 41 B) and sample preparation for whole genome sequencing (Figure 36 and Figure 41).

Xi Wang, **Jingyi Hou** and **Wei Chen** (all MDC) performed whole genome sequencing and corresponding analysis of the RPE-1 cell lines and created Figure 36 and Figure 41 A.

1 Introduction

Protein degradation is a vital biological process which is involved in most cellular functions. In this introduction I will discuss the major concepts related to protein degradation. Known functions of protein degradation and the main molecular mechanisms responsible for it will be introduced. First, I will present a summary of the historical advancement in the field. Second, I will present the open questions this thesis is focused on. Third, I will present our approach to answering these questions.

1.1 A short history of protein degradation

In this historical review I hope to emphasize some facts:

- Protein degradation is not one but many coexisting, highly regulated and complex processes
- Proteins are degraded for a multitude of reasons, ranging from regulatory mechanisms, release of free amino acids during starvation to "quality control" of faulty proteins
- Technical advancements have been key for understanding these biological processes
- Kinetics of protein degradation are more diverse than exponential degradation

1.1.1 Until 1905: The early days

The term *protein,* derived from Greek meaning "of the first rank", was coined by Berzelius in 1838 for a set of important body constituents consisting of a consistent set of chemicals (Hartley, 1951; Mulder, 1838). At the time of Berzelius only four proteins were known: albumin, fibrin, gelatin and casein (Reynolds, 2001; Vickery, 1950). These proteins were thought to exist in both animal and plant bodies but to be made exclusively by plants; animals subsequently needed to eat plants to take up proteins (for review see (Vickery, 1950)). However, what happened with proteins in the animal body was a mystery. How were proteins metabolized and used for energy and catabolic processes?

About one hundred years after the discovery of urea as "the native salt of the body" von Liebig suggested that nitrogen in urine might be taken as a measure of protein destruction (Boerhaave, 1732; Lusk, 1909 ; von Liebig, 1842). This idea was confirmed by von Voit, who measured nitrogen content in meat consumed by a dog and correlated this with the levels of nitrogen in the urine and feces from the dog (Lusk, 1909 ; von Voit, 1860). In addition, polypeptides had been shown to be digested *in vitro* by purified digestive enzymes such as trypsin and pepsin (McCance, 1930). It therefore seemed established that proteins could be degraded within the body.

Nevertheless, where in the body proteins were degraded remained an open question. There existed two distinct hypotheses. The first was put forward by von Voit who suggested that proteins in the blood rarely became a part of cells but were mainly used as an energy source (Folin, 1905b). In contrast, Pflüger proposed that ingested proteins became an integral part of the cells and then degraded (Folin, 1905b; Pflüger, 1893).

1.1.2 1905 - 30: The Folin hypothesis

Folin strongly opposed the intracellular degradation hypothesis because of "*the well-known fact that extremely large quantities of food protein are more or less completely katabolized in the course of a few hours. It is considered incredible that such enormous building up of bioplasm, accompanied by immediate destruction of it, can take place in so short a time*" (Folin, 1905b). Instead he put forward his own hypothesis based on his experimental measurements of urine content of a normal (July 13. Figure 1) and protein starved man (July 20. Figure 1). He found that protein starvation led to a strong change in the composition of urine. For example, the urea content strongly dropped whilst the kreatinin levels where unchanged. "*To explain such changes in the composition of the urine on the basis of protein katabolism, we are forced, it seems to me, to assume that katabolism is not all of one kind. There must be at least two kinds.*" (Folin, 1905b). Folin continues to distinguish between the low endogenous metabolism, for wear and tear, and the larger exogenous metabolism, for energy consumption. Folin's hypothesis became dominant at the time and a consensus idea of protein turnover in mammals had emerged (McCance, 1930).

	JULY 13.	JULY 20.
Volume or urine	1170 c.c.	385 c.c.
Total nitrogen	16.8 gm.	3.60 gm.
Urea-nitrogen	14.70 gm. = 87.5%	2.20 gm. = 61.7%
Ammonia-nitrogen	0.49 gm. = 3.0%	0.42 gm. = 11.3%
Uric acid-nitrogen	0.18 gm. = 1.1%	0.09 gm. = 2.5%
Kreatinin-nitrogen	0.58 gm. = 3.6%	0.60 gm. = 17.2%
Undetermined nitrogen	0.85 gm. = 4.9%	0.27 gm. = 7.3%
Total SO_3	3.64 gm.	0.76 gm.
Inorganic SO_3	3.27 gm. = 90.0%	0.46 gm. = 60.5%
Ethereal SO_3	0.19 gm. = 5.2%	0.10 gm. = 13.2%
Neutral SO_3	0.18 gm. = 4.8%	0.20 gm. = 26.3%

Figure 1. Protein starvation does not uniformly change the composition of urine.
Urine composition of a "*normal person*" on two different diets: milk and egg (July 13) and Starch and cream (July 20). Note that the absolute amount of kreatinin stays constant even though the total nitrogen levels change drastically. Reproduced from (Folin, 1905a).

1.1.3 1930 – 45: Rudolph Schönheimer

In 1936 Schönheimer had fed mice deuterated fatty acids and measured their degradation rates in a study that deployed concepts like “turnover” and “half-life time” of a body constituent for the very first time (Ratner, 1979; Schoenheimer and Rittenberg, 1936). Schönheimer continued his work by deploying enriched heavy nitrogen (^{15}N) isotopes (Urey and Greiff, 1935; Urey, 1937). From these heavy labeled amino acids were synthesized and fed to rats (Schoenheimer and Ratner, 1939). Urine was collected and body tissues were harvested at different time points and the heavy and light nitrogen bound in different molecules were measured on a Bleakney type mass spectrometer (Bleakney, 1929). Analyzing the results it became obvious that Folin’s theory was wrong. 50 % of the administered DL-tyrosine was retained in the body of the rat and even higher values were later measured for the L-isoform (Schoenheimer et al., 1939). Further deploying his techniques, Schönheimer was able to investigate the turnover of individual protein species. He and Ratner followed the synthesis and degradation of antibodies created after active (by vaccination) or passive (by injection of antibodies) immunization of rats (Figure 2) (Heidelberger et al., 1942). The antibodies could then be enriched by the antigen and their turnover measured (Ratner, 1979). The actively immunized antibodies were turned-over and a turnover rate could be calculated (Heidelberger et al., 1942). Schönheimer had thereby opened up the possibility to measure protein degradation kinetics.

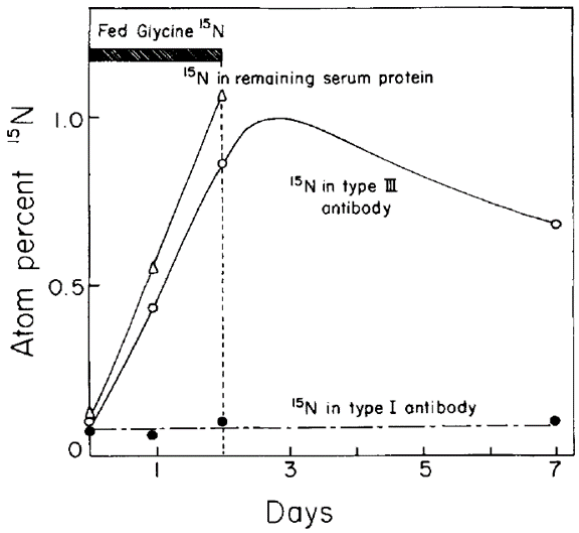

Figure 2. The first turnover measurement for an individual protein.

Abundance of ^{15}N in antibodies from actively (type III) and passively (Type I) immunized rats. Antibodies from passively immunized (i.e. injected with antibodies) rats do not incorporate ^{15}N. Antibodies from actively immunized rats do incorporate ^{15}N. When the heavy isotope glycine diet is removed, the ^{15}N containing antibodies are degraded and a turnover rate can be calculated (not shown). Reproduced with permission from (Ratner, 1979).

1.1.4 1945 – 75: The era of the lysosome and radioactive tracers

1.1.4.1 Long-lived radioactive isotopes open up the field of degradation kinetics

The existence of long-lived radioactive isotopes of atoms contained in organic matter was long questioned (Kamen, 1985). The few radioactive isotopes known all had half-lives too short for most biological experiments e.g. ^{15}O has a half-life of 2 min (See Figure 3 for a discussion on exponential decay). However, when ^{14}C (half-life 5730 years) was discovered in 1939, the radioisotope was quickly employed (Borsook et al., 1950; Ruben, 1941). Similarly, ^{35}S (half-life 88 days) was properly described in 1940 (Levi, 1940). ^{35}S L-Methionine and L-Cysteine could be synthesized in bulk by 1952 (Williams and Dawson, 1952). The cumbersome methods employing stable isotopes by Schönheimer, which required a hand built mass spectrometer, were then quickly abandoned.

Discussion arose surrounding Schönheimer's findings when data showed that exponentially growing bacteria displayed almost no protein degradation when pulsed with radioactive amino acids (Hogness, 1955). The authors suggested that the "dynamic state of body constituents" was merely an effect of protein secretion and cell lysis in the complex mammalian organs. The results gained traction because of the well-controlled model system where cell lysis could be accounted for. However, the controversies were later resolved when results were published demonstrating protein turnover in *E. coli* at stationary state (Mandelstam, 1957).

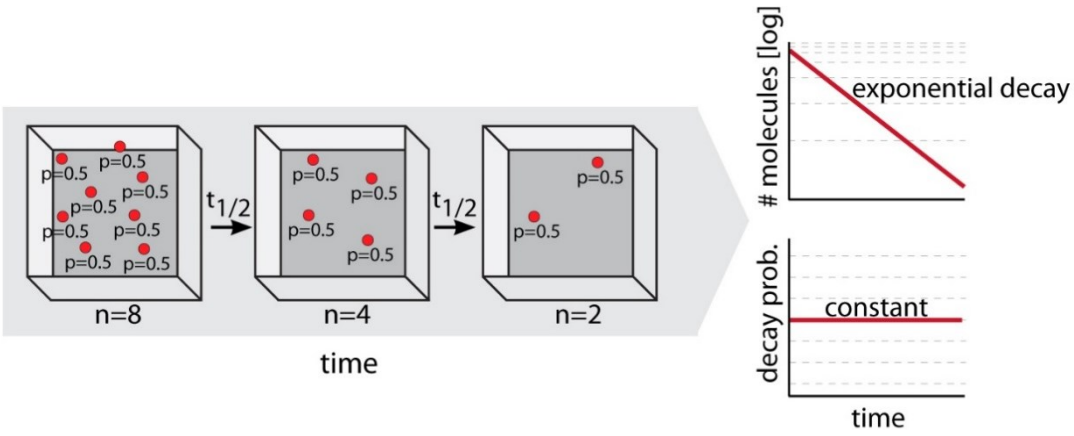

Figure 3. Radioactive isotopes display exponential decay.

By definition the number of radioactive isotopes (red dots) remaining will decrease by half for each half-life time ($t_{1/2}$) that passes. If one plots the number of molecules remaining over time in a semi-exponential plot, it will result in a straight line (upper right panel). However, if one looks at the decay probability for each individual radioactive isotope, it is constant with a 50 % probability over the next half life time. In other words, the decay probability is constant over the age of the molecules (lower right panel). This means that a newly created "young" radioactive isotope has the very same decay probability as an old radioactive isotope.

1.1.4.2 Robert Schimke and the regulation of protein degradation

To study the function of protein degradation Schimke first focused on the final enzyme in the urea cycle, arginase, which converts arginine into urea and ornithine (Schimke, 1964). Schimke applied ^{14}C labeling in rats during different dietary regimes in combination with biochemical and immunological purification of arginase. It became clear that during starvation degradation of the enzyme stopped, increasing the arginase steady state levels, and conversely, after a change from high to low protein diet, the degradation rate of arginase increased. He went on to show that administration of tryptophan increased the synthesis rate and decreased the degradation rate of the enzyme tryptophan hydrolase (Schimke et al., 1965). These two studies led to a paradigm shift because until then protein degradation was thought to be passive and dependent on the inherent stability of the proteins. Importantly, Schimke´s findings created a rationale behind the continued turnover of proteins in the sense that degradation rates could correlate with the ability to adapt to stimuli. For example, the 12 liver proteins with the shortest half-lives were responsible for the rate limiting steps in their respective biochemical pathway (Goldberg and Dice, 1974). They could therefore be quickly up or down-regulated.

1.1.4.3 The discovery of lysosomes by Christian de Duve

The discovery of the lysosome by Christian de Duve was a result of a failed attempt to localize the enzyme glucose-6-phosphatase within the cell (Duve, 1975). De Duve applied the recently developed differential centrifugation technique to homogenized rat liver to distinguish between distinct cellular organelles (Claude, 1946). He found that 95 % of the glucose-6-phosphatase was localized in the microsomal fraction. As a control he assayed for the activity of acid phosphatase in the mitochondrial fraction but found only 10 % of the expected activity based on previous measurements using harsher centrifugation techniques (reviewed in (Bainton, 1981)). However, when the samples where reassessed a few days later, all of the expected activity was recovered. The latent activity of the enzyme was attributed to the breakdown of a membrane that previously prevented the accesses of substrate to the enzyme. Further biochemical analysis revealed that the enzyme was not located in mitochondria but rather in a distinct compartment. A number of proteases with differing substrate specificities but with their highest activity at low pH were also found to be located in this new membrane enclosed compartment (Bainton, 1981). The potential function in cellular degradation processes lead to the naming of the organelles as lysosome (De Duve et al., 1955). The cellular function of lysosomes was quickly investigated (see Figure 4 for a modern understanding of lysosomal function). For example, Werner Straus and colleagues showed that lysosomes could be responsible for the degradation of extracellular proteins taken up by endocytosis (Straus, 1954). De Duve and

others later showed that cells also could degrade internal proteins, a process he termed autophagy (Castro-Obregon, 2010). The first cellular protein degradation machinery had been described.

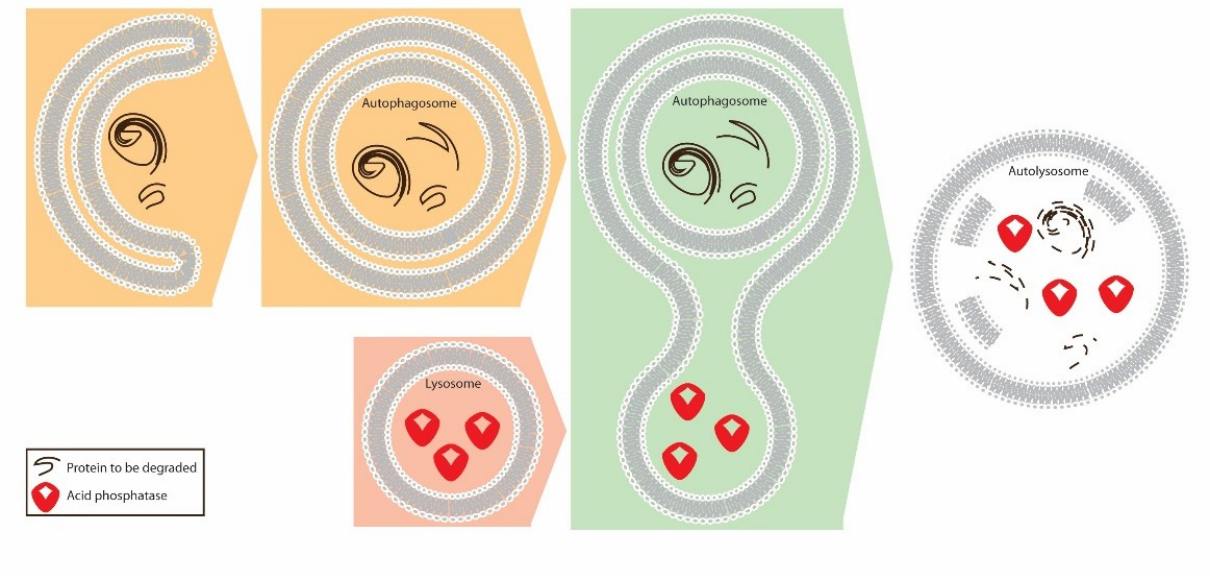

Figure 4. Principles of Autophagy.

Today it's believed that lysosomes mainly play a role in cell maintenance by clearing larger organelles (i.e. mitochondria) especially after cell stimuli (e.g. starvation or growth factor deprivation) and in degradation of endocytosis products (Castro-Obregon, 2010). Autophagy is the process in which autophagosomes (yellow background) engulf intracellular molecules and organelles such a mitochondria (Takeshige et al., 1992). They then fuse with protease containing organelles called lysosomes (red background). The combined organelles, referred to as autolysosomes, are acidified, the proteases are activated and degrade the internalized substrates.

1.1.4.4 Degradation of "abnormal" proteins

"*There must exist strong selective pressure favoring the evolution of a mechanism to degrade abnormal protein, since the intracellular accumulation of inactive or partially active enzymes should be highly deleterious*" – (Goldberg, 1972).

In 1974, Alfred Goldberg and Fred Dice summarized the current understanding of the physiological function of protein degradation (Goldberg and Dice, 1974). They outlined three main purposes for protein degradation:

1. Based on Schimke's findings it was clear that degradation rates were used by the cells to rapidly change the concentration of an enzyme
2. Starvation induced degradation to increase the pool of free amino acids
3. Quality control of abnormal proteins

Goldberg was a strong champion of the third option and he and others had previously shown that anomalous proteins were rapidly degraded in *E. coli* and reticulocytes (Goldberg, 1972; Pine, 1967; Rabinovitz and Fisher, 1964). For example, premature translational termination, by the addition of puromycin, led to unfinished polypeptides rapidly being degraded (Goldberg, 1972). In addition, increasing translation error rate by using an *E. coli* strain with a ribosomal mutation lead to amplified protein degradation (Goldberg, 1972). These and other findings clearly demonstrated that protein degradation could function as a quality control mechanism in cells.

1.1.4.5 Exponential degradation as the dominant model of protein removal

"*Where it has been studied carefully, the breakdown of enzymes obeys first order kinetics, which means that newly synthesized proteins are as likely to be degraded as old ones. In other words, degradation of polypeptides is a random event unlike aging of organisms or red cells.*" – (Goldberg and Dice, 1974).

That quote and similar claims by Schimke laid the foundation to the nowadays common assumption that all proteins are exponentially degraded (Schimke and Doyle, 1970). It is worth noting that at the time two examples of non-exponentially degraded proteins were known. The above mentioned haemoglobin was degraded in concert with the rest of the red blood cell (Shemin, 1946). Red blood cells are disposed of in an age-dependent manner by the liver since they lack most mechanism of self-repair. Similarly, the proteins of the outer rod cells are stable until a certain age when they are displaced and degraded (Figure 5, (Young and Bok, 1969)). Nevertheless, an exponential assumption was to accompany the degradation field for decades to come.

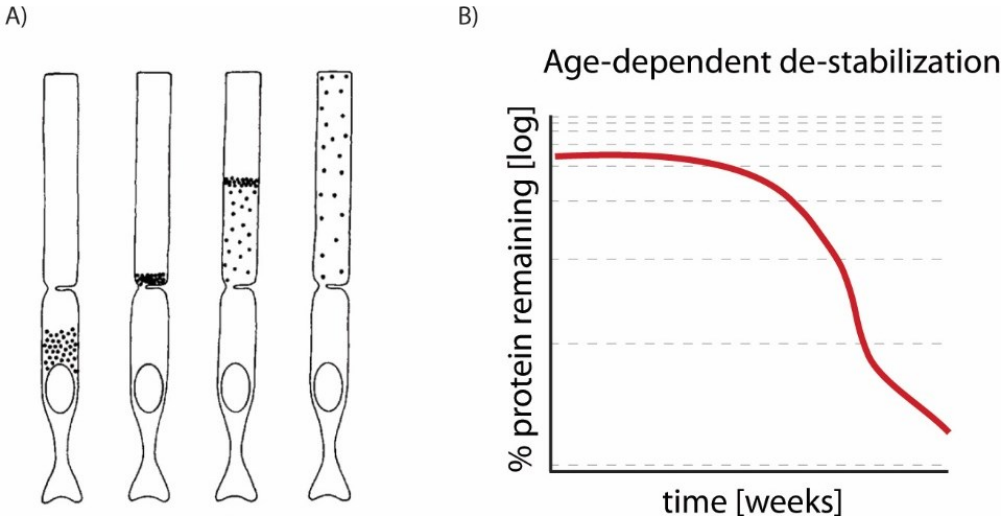

Figure 5. An example of age-dependent destabilization.

A) Illustration of findings from (Young and Bok, 1969). Black dots indicate location of radioactive proteins in rod cells. First rod cell (from the left) shows the localization minutes after injection of L-methionine-^{3}H into frogs. The newly synthesized proteins are then organized in a disc shape (second rod cell from the left). The discs are gradually displaced by newer discs. After about 6 weeks the disc disappears and is phagocytosed by pigmented epithelial cells. B) Graphical representation of the stability of the disc proteins (note the time scale). They show a clear non-exponential profile in which they are first stable and then after ~6 weeks are rapidly degraded. Reproduced with permission from publisher.

1.1.5 1975-2000: the proteasome and spatial distribution of cellular degradation

1.1.5.1 The ubiquitin-proteasome system

Many lines of evidence were accumulating indicating that a second cellular degradation system must exist (for discussion see: (Ciechanover, 2005)). For example:

I. There existed large differences in protein half-lives inconsistent with the slow process of autophagy
II. Lysosomal inhibition did not stop degradation of short-lived or abnormal proteins
III. Most protein degradation required energy in contrast to the exergonic lysosomal processes

For the discovery of the proteasome a certain cell system was required: the reticulocytes. These immature red blood cells lack lysosomes but could nevertheless degrade abnormal proteins (Rabinovitz and Fisher, 1964). Goldberg isolated the system and demonstrated that it degraded proteins *in vitro* in an ATP-dependent manner at neutral pH (Etlinger and Goldberg, 1977). Hershko and Chiechanover further worked out the details and showed by fractionating cell lysates that the system required at least two components (Ciehanover et al., 1978). They focused on the first fraction and demonstrated that it contained a small protein they named APF-1 (ATP-dependent proteolysis factor 1). APF-1 was found to be covalently linked to the protein substrates which were degraded when incubated with the second fraction (Ciechanover et al., 1980b; Hershko et al., 1980). A model was proposed in which APF-1 was conjugated to the substrate and targeted it for degradation. Afterwards APF-1 would be released and reused. APF-1 was soon shown to be the previously described protein ubiquitin (Ciechanover et al., 1980a; Wilkinson et al., 1980). Ubiquitin was shown to be conjugated to other proteins by isopeptide linkage to a lysine (Goldknopf and Busch, 1977; Hunt and Dayhoff, 1977). Soon after the 20S and 26S proteasome were purified and a coherent model of specific protein degradation was proposed (See Figure 6) (Hough et al., 1987; Wilk and Orlowski, 1980).

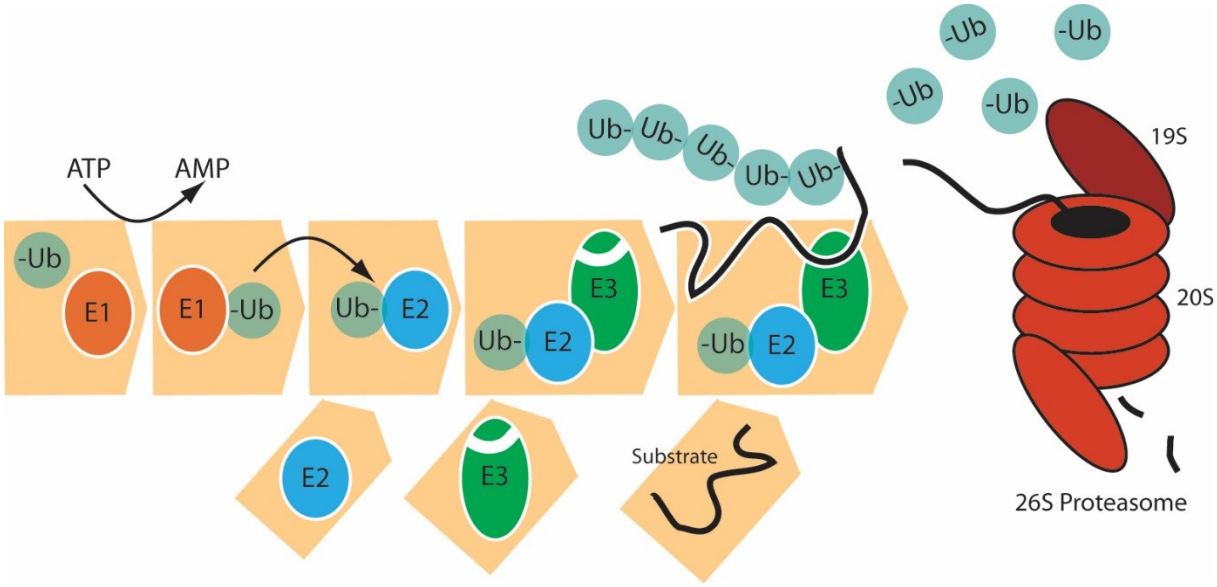

Figure 6. Principles of the ubiquitin-proteasome system.

An E1 ubiquitin-activating enzyme attaches an ubiquitin molecule to an E2 ubiquitin-conjugating enzyme (E2) in an energy-dependent fashion. The E2 then binds to an E3 ubiquitin ligase. The E2/E3 combo then specifically recognizes certain protein substrates and attaches the ubiquitin to a lysine residue. Repeated cycles of ubiquitination take place until a poly-ubiquitin chain has formed. The poly-ubiquitin chain is then recognized by the proteasome and the substrate is guided into the barrel-shaped 20S core particle. Inside the proteasome multiple specific protease activities work on the protein before short peptides are released into the cytosol.

1.1.5.2 ER associated degradation (ERAD)

The endoplasmic reticulum (ER) is a large membrane-enclosed organelle organized in a network-like configuration (Porter et al., 1945). Two forms, rough and smooth, of the ER are usually recognized. The rough ER gets its appearance from the multiple ribosomes associated to the outside of the membrane. Accordingly, the ER is a place where newly synthesized proteins mature, e.g. fold and are post-translationally modified (Meusser et al., 2005). Specifically, those proteins that will be secreted or exposed on the cell surface are co-translationally or translocated directly after translation into the ER (Jan et al., 2014). The proteins are transported via the Golgi apparatus before finally being exported or inserted in the cell membrane.

A quality control system associated with the ER was long predicted because of the numerous examples of proteins being synthesized into the ER and subsequently degraded. One main group of proteins degraded was non-assembled protein complex subunits. For example, subunits of the Asialoglycoprotein receptor (Amara et al., 1989), T-cell receptor (Chen et al., 1988; Lippincott-Schwartz et al., 1988), Gonadotropin hormone dimer (Corless et al., 1987), IgM complex (Dulis et al., 1982), Acetylcholine receptor (Merlie et al., 1982) and Fibrinogen (Plant and Grieninger, 1986). The degradation of many of the examples was shown to be dependent on a posttranslational modification: the addition of a carbohydrate in a processed called glycosylation.

Another ER-contained protein, cystic fibrosis transmembrane conductance regulator (CFTR), helped bring clarity into the degradation mechanism (Jensen et al., 1995; Ward et al., 1995). CFTR that lacked a phenylalanine at position 508 was known to misfold and be thereafter rapidly degraded from the ER. With the recently developed proteasome inhibitors (e.g. MG132 (Goldberg, 2012; Rock et al., 1994) and lactacystin (Fenteany et al., 1995)) it became possible to show that the ER-residing proteins were in fact degraded by the proteasome (Jensen et al., 1995; Ward et al., 1995). However, no proteasomes are located in the ER, so ER-residing substrates needed to be retro-translocated back into the cytosol for destruction (Sommer and Jentsch, 1993). See Figure 7 for a current view on ERAD.

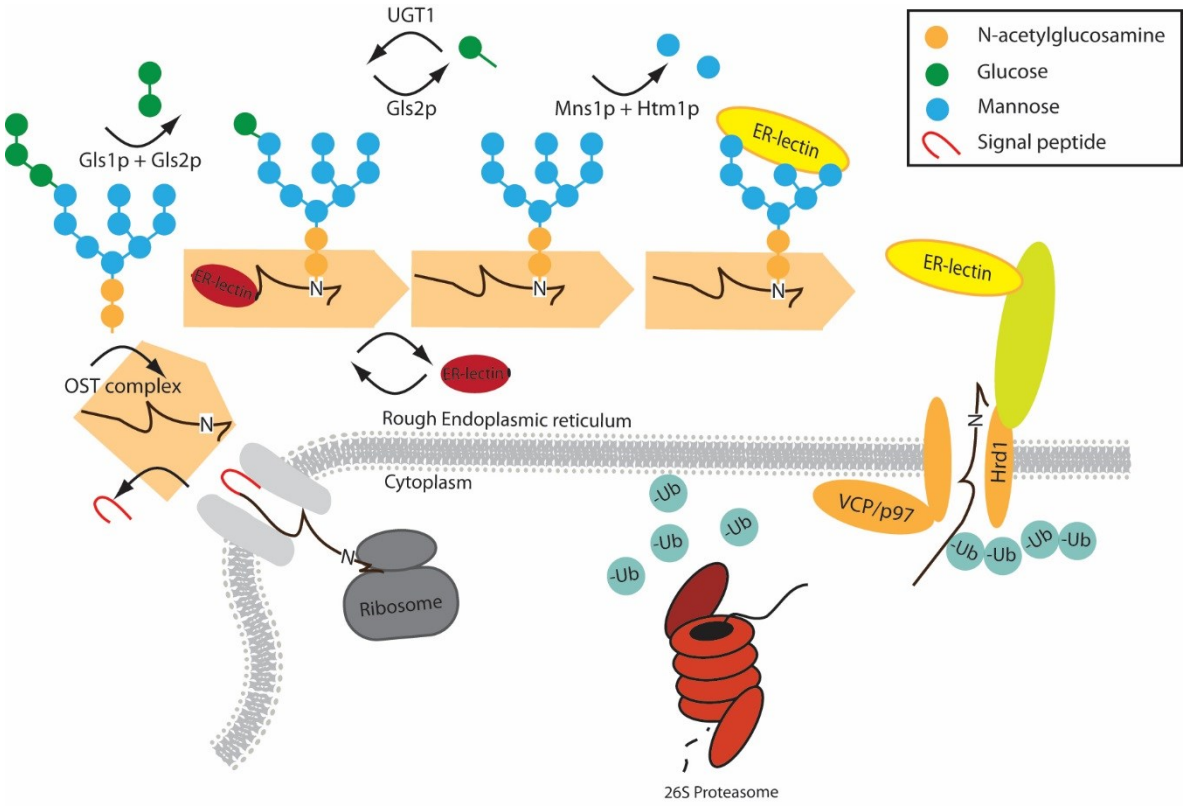

Figure 7. Overview of ER-associated degradation (ERAD).

Newly synthesized proteins are translocated into the ER through a multiprotein channel as unfolded polypeptides (Jan et al., 2014). In the ER the proteins have their ER-targeting signal sequence cleaved off (if present), are folded with the help of heat shock proteins, have disulfide bonds created by oxidoreductases and are modified by covalent attachment of complex sugars to asparagine residues by the oligosaccharyltransferase (Buchberger A, 2010; Mandon et al., 2013). These sugars moieties are further modified leading to a “glyco-code” which is read by lectins, which direct the proteins to ER-exit, refolding or ERAD (Aebi et al., 2010). When ERAD is selected, proteins are exported out of the ER via a translocon while simultaneously being poly-ubiquitinated by ER-associated E3 ubiquitin ligase Hrd1. The poly-ubiquitinated protein is recognized by the AAA-ATP p97/VCP that extracts the protein for subsequent degradation by the cytosolic proteasome (Buchberger A, 2010).

1.1.5.3 Non-exponential degradation and ER associated degradation

Interestingly, many of the endogenous ERAD substrates discussed above, such as the T-cell receptor subunits (Lippincott-Schwartz et al., 1988) and CFTR (Ward and Kopito, 1994), are non-exponentially degraded since they are targeted for degradation mainly before they assemble (TCR) or fold (CFTR) and then become stable when they reach the cell surface. Another interesting example is Basigin (Figure 8). This protein is, similarly to the TCR and CFTR, synthesized into the ER, where it matures, and when assembled is transferred to the cell surface. Interestingly, the unassembled core-glycosylated form has a half-life of about 90 minutes while the cell surface exposed protein is stable (Tyler et al., 2012). This is a clear example of non-exponential degradation kinetics.

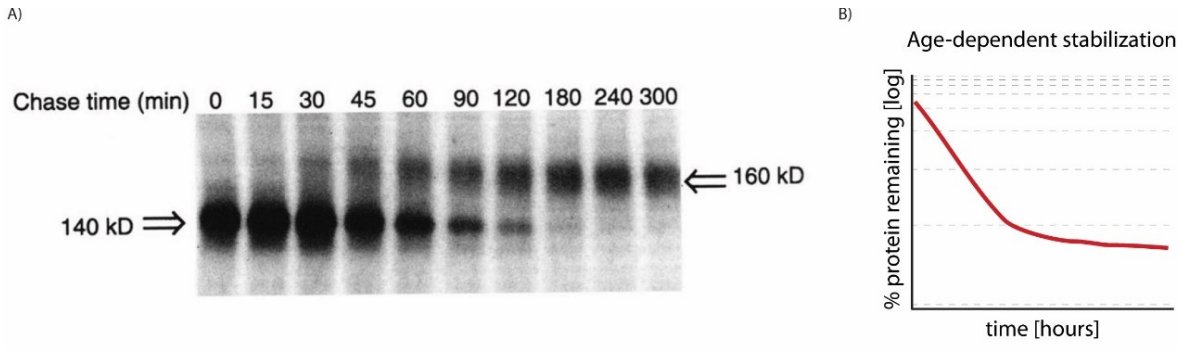

Figure 8. Basigin is a non-exponentially degraded protein.

A) Basigin is synthesized into the ER as a lower molecular weight protein (140 kDa). As time passes, the protein becomes glycosylated in the ER and shows up with a heavier molecular weight (160 kDa). Interestingly, only a fraction of the newly synthesized proteins, ~20 %, survive the transition. The mature protein, however, is strikingly stable. Reproduced (and modified) with permission from (Tyler et al., 2012). B) A schematic plot of what the data in A) would look like in a semi-logarithmic plot. This is an example of a clearly non-exponential profile.

1.1.6 2000 – Current: Global measurements

1.1.6.1 Kinetics of degradation – not so exponential after all

A commonly applied method in the 70s and the 80s was the double labeling technique. For example, cells could be labeled for a longer time period with ^{14}C followed by a shorter labeling period with ^{3}H. The result of this technique was the selective labeling of a long-lived and short-lived protein population (e.g. (Poole and Wibo, 1973)). In 1980, Wheatly and colleagues criticized the interpretation of the two population of proteins: "*The technique of selectively labeling 'short' and 'long-lived' proteins, which implies qualitative differences between them, is more readily interpreted as an artificial polarisation of a declining statistical probability curve of proteolysis with time which is similar for all nascent proteins*" (Wheatley et al., 1980). In other words, they proposed that the suggested two populations of distinct proteins with different life expectancies could equally well be explained by one population of proteins that become more stable over time.

One explanation of why a cell would seemingly wastefully degrade a significant proportion of new proteins was proposed by Yewdell, Anton and Bennink (Yewdell et al., 1996). They proposed that short lived proteins were defective ribosomal products (DRiPs). Peptides from the DRiPs were used for presentation on MHC-I molecules on the cell surface allowing CD8+ T cells to distinguish between self and none-self (Rock and Goldberg, 1999). This mechanism, they claimed, could explain how cells could respond so rapidly to viral infections, before the viral proteins were a significant part of the intracellular protein pool, and also why it was so hard for viruses to avoid cell surface exposure of their proteins. A significant boost to this theory was gained when a paper describing the results of very short ^{35}S Methionine pulses (30 seconds) followed by chases up to 60 minutes using HeLa cells was published (Schubert et al., 2000). The pulses were performed with or without proteasome inhibitors. In the presence of inhibitor roughly 30 % more labeled proteins were recovered. This led to the conclusion that ~30 % of all newly synthesized protein were directly degraded. This number was, however, soon questioned when it was shown that protein synthesis is dependent on proteasome function during amino acid starvation (Vabulas and Hartl, 2005). In other words, part of the increase in radioactive-labeled proteins reported by Schubert et al. could be explained by a more severe amino acid starvation and therefore more efficient loading of tRNA with radio-labeled amino acids, in the presence of proteasome inhibitor. The new investigation limited the amount of proteins rapidly degraded to a few percent on a global level. To date, the question of whether protein turnover can explain the MHC class-I peptidome without assuming degradation of DRiPs is still open (Bassani-Sternberg et al., 2015; Yewdell and Nicchitta, 2006).

In contrast, the existence of non-exponential protein degradation is well backed up. Recently, studies focusing on co-translational ubiquitination and degradation have been published (Duttler et al., 2013; Pechmann et al., 2013; Wang et al., 2013). They concluded that ~15 % (in human) and >1 % (in yeast) of nascent polypeptides are co-translationally ubiquitinated and targeted to the proteasome. Another line of evidence was derived from enriching ubiquitinated peptides in combination with dynamic SILAC (Kim et al., 2011). This study showed that ubiquitinated proteins were consistently younger than the non-ubiquitinated proteins. They further showed that the ubiquitination was increased by blocking the proteasome, strongly suggesting that the young ubiquitinated proteins were targeted for degradation.

1.1.6.2 Global turnover measurements: dynamic SILAC

Although modern ^{15}N labeling and stable isotope labeling by amino acids (SILAC) as developed by the Mann group both share many features with the pioneering work of Schönheimer, there have been some major developments, mainly on the mass spectrometric and computational side (Ong et al., 2002). I will here only discuss the most common approach for proteomics referred to as shotgun proteomics. In shotgun proteomics, digested proteins are detected and quantified by liquid chromatography coupled to tandem mass spectrometry (LS-MS/MS) and then computationally reassembled (Cox and Mann, 2008; Zhang et al., 2013). This method has successfully been employed to measure the abundance of many cellular proteomes with near complete coverage (de Godoy et al., 2008; Nagaraj et al., 2011). Using label free methods it is possible to measure differences between samples with high accuracy (Cox et al., 2014). However, when combined with stable isotope labeling, relative quantification within samples is greatly improved (Ong et al., 2002; Wu et al., 2004). Although ^{15}N labeling is cheap and suitable for measuring protein turnover, the vast amount of nitrogen in proteins and the less than 100 % pure enrichment of heavy nitrogen isotopes make most peptides only partially labeled, impairing both identification and quantification. ^{15}N is therefore mainly used for labeling organisms that synthesize all their own amino acids, like plants, or when large amounts of labeled amino acids would be required, such as when labeling rats (Toyama et al., 2013; Wu et al., 2004).

Dynamic SILAC is based on changing some amino acids in cell culture medium (or food when labeling animals) from light to heavy versions (Doherty et al., 2009; Krüger et al., 2008; Sheean et al., 2014). By taking samples over time and measuring the heavy to light ratio one can calculate the turnover of the proteins, at least if the cell doubling time is known. This type of measurement lead to the ability to acquire global knowledge of turnover rates on a proteome-wide scale (Doherty et al., 2009; Schwanhäusser et al., 2011). One interesting example in this context is the measurement of nuclear

protein dynamics (Andersen et al., 2005; Lam YW, 2007). By combining SILAC pulse labeling with subcellular fractionation, the authors showed that ribosomal proteins are rapidly shuffled into the nucleolus, where they combine with cofactors and rRNA to form the small and large ribosomal subunits. This process was measured to take on average 2 hours before ribosomes were exported out to the cytoplasm. Interestingly, a large part of the newly synthesized proteins were found to be degraded in the nucleus, in stark contrast to the very stable cytoplasmic version of the same proteins (Lam YW, 2007).

1.1.6.3 State of the art - Methodological limitations for protein kinetics research

Dynamic SILAC labeling is limited by the fact that the pre-existing light proteins are still present. That means that if a protein is turned over slowly the relative amount of newly synthesized proteins will be drowned out by the massive signal from the pre-existing light proteins. Thereby, the minimal pulse times for a good signal to noise for most protein are in the range of multiple hours for mammalians cells (Schwanhäusser et al., 2011). In addition, most calculations that are used to calculate turnover are based on an assumption of exponential degradation (Christiano et al.; Hinkson and Elias, 2011; Schwanhäusser et al., 2011). Therefore, the currents methods are sufficient to measure steady state degradation rates and changes of these between cell types or between control and treated cells. However, to enable measurements of changes in degradation kinetics for the same protein as it “ages” would require pulse-chase experiments. It is by following proteins with a limited age-distribution that one can untangle differences in degradation rates. Therefore, the long pulse-times required for dynamic SILAC are not suitable for global kinetic measurements (Hinkson and Elias, 2011; Larance and Lamond, 2015). Other methods, such as global protein stability profiling, bleach-chase or tandem fluorescent timers that rely on fusing protein with fluorescent tags, suffer from similar shortcomings (Eden et al., 2011; Khmelinskii et al., 2012; Yen et al., 2008). The current gold standard of pulse-chase experiments are still performed using radioactive labeling coupled to immunoprecipitation. This methodology does not easily lend itself to multiplexing, which usually limits the analysed proteins to one or two.

1.2 This study

1.3 Are all proteins exponentially degraded?

Reviewing past research it becomes clear that the hypothesis of exponential degradation as put forward by, among others, Schimke does not hold true for all proteins. However, the fraction of the proteome that is non-exponentially degraded (NED) is to date not known. In other words, are the examples known so far just *ad hocs* for specific circumstances or is NED common?

This thesis will try to answer the following questions:

1. How common is NED?
2. Do NED proteins become more or less stable with age?
3. What proteins are NED?
4. What are the molecular mechanisms involved?
5. What are the biological implications of NED?

1.4 How to study NED on a global level

To study kinetics of protein degradation an age-defined population of proteins with a limited age-span is needed. Degradation from this population then needs to be followed over time. If one plots the ”% protein remaining” in semi-log space over time, one would expect the data points to lay on a straight line if the protein is exponentially degraded (e.g. see Figure 3). If the data points are not falling on a straight line, the proteins do not follow first order degradation kinetics (for example Figure 5 and Figure 8). In this thesis we introduce a method that allows for all these types of measurements on a global level.

2 Materials and Methods

2.1 Cell lines

NIH 3T3 mouse fibroblast cells

From ATCC.

Human retinal pigmented epithelial cells (RPE-1)

The cell line RPE-1 immortalized with hTERT (referred to as RPE-1) was a kind gift from Stephen Taylor (University of Manchester, UK) and described in detail in Stingele et al. (2012).

RPE-1 trisomic cells

The derivative cell line RPE-1 5/3 11/3 12/3 (referred to as RPE-1 trisomic) was generated using microcell-mediated chromosome transfer in Stingele et al. (2012). The partial trisomy of chromosome 11 spontaneously occurred and is unique to this study.

Cell culture for the different cells lines is described under the individual experiment sections.

2.2 Experiments using mouse cells

2.2.1 Dynamic SILAC to determine effects of AHA on protein synthesis

NIH 3T3 mouse fibroblast cells were grown in 6-well plates in SILAC DMEM (Life Technologies) complemented with glutamine (Glutamax, Life Technologies) and 1 % Penicillin and Streptomycin (Life Technologies), 10 % dialyzed fetal calf serum (dFCS, Pan-Biotech), light L-arginine (Arg0, Sigma-Aldrich) and light L-lysine (Lys0, Sigma-Aldrich) referred to as Light SILAC DMEM (Ong et al., 2002). When the cells reached 30 % confluency they were starved in Light DMEM depleted of methionine (Biosera Ltd). After 1 h the cells were switched to Heavy SILAC DMEM medium containing heavy L-arginine (Arg10, Sigma-Aldrich) and heavy L-Lysine (Lys8, Sigma-Aldrich) and supplemented with different concentrations of azidohomoalanine (AHA, Anaspec) and/or methionine (Sigma-Aldrich) as shown in the figure. After a 4 h the cells were washed and lysed in modified RIPA buffer (50 mM Tris (pH8), 150 mM NaCl, 0.1 % SDS, 0.5 % sodium deoxycholate, 1 mM EDTA, 1 % NP40) with nuclease (Benzonase, Sigma-Aldrich) and 2-fold protease inhibitor cocktail (Roche) added. Lysates were incubated on ice for 10 minutes before being sonicated in an ice-bath for another 10 min. Proteins were precipitated by Wessel-Flügge precipitation (Wessel and Flugge, 1984). In short, proteins were

precipitated to get rid of SDS by sequentially adding methanol, chloroform and then water before spinning down the samples at 10,000 rcf and finally adding methanol again (Puchades et al., 1999; Sheean et al., 2014). After centrifugation, the retrieved protein pellet was resuspended in 6 M Urea, 2 M Thiourea in 10 mM Hepes (pH8). Proteins were digested with Lysyl endopeptidase (LysC, Wako Chemicals) before being diluted in 50 mM ammonium bicarbonate (ABC) buffer and trypsinated (Invitrogen) overnight. The peptide solution was acidified by addition of TFA before peptides were desalted and stored on StageTips (Rappsilber et al., 2003). StageTips were prepared by putting 3 discs of C18 (3M) in 200 µL pipette tips and activating the material with methanol. Organic solvents were washed away by Buffer A (5 % acetonitrile and 0.1 % formic acid) before peptides were loaded onto the stageTip. Salts were washed away with Buffer A. Peptides were eluted using Buffer B (80 % acetonitrile and 0.1 % formic acid) and acetonitrile was evaporated in a speedvac (Eppendorf). Samples were diluted in Buffer A and peptides were separated on a 15 cm long column with 75 µm inner diameter packed in house with ReproSil-Pur 120 C18-AQ 3µm resin (Dr. Maisch GmbH) using a 3 h linear gradient with 250 nl/min flow rate of increasing concentration of Buffer B in a High Pressure Liquid Chromatography (HPLC) system (ThermoScientific). Peptides were ionized by an electrospray ionization (ESI, ThermoScientific) source and analysed on a LTQ Orbitrap Velos mass spectrometer (ThermoScientific). The mass spectrometer was run in data-dependent mode selecting the top 20 most intense ions in the MS full scans (Orbitrap resolution: 60,000; target value: 1,000,000 ions) for higher energy collision-induced dissociation. The resulting MS/MS spectra from the Orbitrap had a resolution of 17,500 with a target value of 5,000. The resulting raw files were analysed using MaxQuant software version 1.4.1.2 (Cox and Mann, 2008). Default settings were kept except that ”match between runs“ was activated and ”reQuantify“ was turned off. Lys8 and Arg10 were set as labels and oxidation of methionines and n-terminal acetylation were defined as variable modifications. Carbamidomethyl of c-termini was set as fixed modification. The *in silico* digests of the mouse Uniprot database (2013-01) and a database containing common contaminants were done with Trypsin/P. The peptide- and protein-level false discovery rates were set to 1 % and were assessed by searching in parallel a data base containing the reversed mouse database sequences. Plotting was done in R version 2.15.1 (R Foundation for Statistical Computing, Vienna, Austria).

2.2.2 Enrichment efficiency of AHA-containing proteins

NIH 3T3 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 DMEM depleted of methionine and cysteine (Invitrogen) with 10 % dFCS and 1 % penicillin and streptomycin added. Cells were then pulsed for 1 h with 80 µCi final concentration of ^{35}S-Cysteine (Perkin Elmer) in combination

with either 1mM AHA or 1mM methionine. After the radioactive pulse, cells were washed in PBS and scraped in the same and pelleted by centrifugation. Cells were lysed and AHA-containing proteins covalently linked to alkyne beads using the "Click-it protein enrichment" kit (Invitrogen). In brief, cells were lysed in Urea lysis buffer (8 M Urea, 200mM Tris (pH 8), 4 % CHAPS and 1 M NaCl) supplemented with protease inhibitor cocktail (Roche) on ice for 30 min. Half the cell lysate was frozen and kept as input. Alkyne beads were prepared and copper-catalysed click reaction was performed as recommended by the manufacturer (Hou et al., 2015). Beads were washed twice in SDS wash buffer and supernatant was retained. The resulting three fractions: input, supernatant and beads, were added to 5 mL Rotiszint eco plus (Carl Roth) and measured by scintillation counting (Packard). The experiment was repeated twice and was analysed using Excel (Microsoft).

2.2.3 Global ^{35}S cysteine pulse-chase experiments

Confluent NIH 3T3 mouse fibroblasts in 6-well plate wells were washed in pre-warmed PBS before being starved in methionine and cysteine free DMEM as above. 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. Cells were then washed twice in SILAC DMEM before either being directly lysed or chased for 6 or 24 h in cold medium with either 50 µM cycloheximide (Sigma-Aldrich) or 10 fold cysteine (Sigma-Aldrich) added to limit re-incorporation of radioactive amino acids. Finally, cells were scraped and lysed in modified RIPA buffer as above. All samples were frozen at -80 °C before being thawed on ice for 30 min in the presence of Benzonase. 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. SDS-PAGE on a 10 % polyacrylamide gel was used to resolve proteins. In-gel fixation of proteins was accomplished by the addition of 5 % acetic acid and 50 % methanol. Proteins were labeled by Coomassie blue staining (colloidal blue stain kit, Novex). The gel was then vacuum-dried using a gel drying system (Bio Rad). To quantify total loaded protein, dried gels were scanned on a scanner (Canon). The radioisotope signal was measured by exposing the gel to a magnetic photo-stimulable phosphor-plate overnight. The plate was developed by a phosphorimager (Typhoon FLA 9500, GE Healthcare). ImageQuant software (GE Healthcare) was used to quantify radioactive and coomassie images. Each lane was quantified separately and background signal was estimated by quantifying an empty lane. The radioactive signal was then further normalized to total protein input estimated by the Coomassie staining. Three biological replicates with technical triplicates were performed. Statistics and plotting were performed in Excel (Microsoft).

2.2.4 Enrichment efficacy of AHA-labeled proteins

NIH 3T3 mouse fibroblasts were cultured in a 10 cm plate in Heavy SILAC DMEM (life technologies) prepared as above. Cells were washed twice in pre-warmed Heavy SILAC DMEM depleted of methionine (Biosera) before being starved of methionine in the same for 1 h. The starvation was followed by a 2.5 h incubation with 1 mM AHA. The AHA-labeled Heavy cells were washed in ice cold PBS and then scraped in Urea lysis buffer containing two-fold protease inhibitors (Roche) and mixed 1:1 with cells grown in Light SILAC DMEM. The lysate was treated with benzonase on ice for 20 min before being sonicated in an ice bath. A part of the lysate was saved to estimate mixing accuracy (see below). Samples were spun down at 20,000 rcf, the supernatant was transferred to a new tube containing alkyne agarose beads and the click reaction was performed overnight as described above. 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) final concentration. Beads were sequentially washed in SDS buffer, 8 M Urea in 100mM Tris (pH8) and 80 % acetonitrile by centrifugation and decanting supernatant. Proteins were digested “on bead” in 5 % acetonitrile in ABC buffer first 3 h with LysC and then overnight with trypsin. Input samples were Wessel-Flügge precipitated and “on-pellet” digested as above.

Peptides were acidified and stageTipped as above. Eluted peptides were separated using HPLC and a 4 h linear gradient with a 300 nl/min flow rate with increasing buffer B concentration. Peptides were passed over a 2 m long column with a 100 µm inner diameter containing monolithic C18 (kindly provided by Yasushi Ishihama (Kyoto University)). Peptides were ionized using an ESI source and analysed on a Q-Exactive mass spectrometer. The mass spectrometer selected the top 10 most intense ions in the MS full scans for higher energy collision induced dissociation and MS/MS analysis. Full scan parameters: Orbitrap resolution: 70,000; target value: 3,000,000 ions; maximum injection time of 20 ms. MS/MS spectra parameters: Orbitrap resolution: 17,500; target value 1,000,000 ions; maximum ion collection time: 60 ms.

The resulting raw files were analysed as above but with a later version of the Uniprot mouse data base (2014-10) and with deamidation of asparagine and glutamine residues defined as variable modifications.

2.2.5 AHA pulse-chase of mouse fibroblasts

Fully triple SILAC (Light: Lys0, Arg0; Medium: Lys4, Arg6; Heavy: Lys8, Arg10) labeled mouse fibroblast where grown to ~25 % cell density. 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. Cells were starved for methionine as described above and then labeled with 1 mM AHA both for 1 h. After the pulse, Medium and Light cells were washed in PBS followed by SILAC DMEM before being chased in the same medium. The Heavy cells were instead washed in ice cold PBS before being scraped and spun down. Finally, the cell pellet was frozen. After the chase the Medium and Light cells were also scraped and frozen. The pellets were lysed as described for the "enrichment efficacy experiment". In addition, click reaction, washing of beads, denaturation, alkylation and digestion were performed in the same way.

In two out of the three experiments 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 with high pH buffer (20 mM Acetic acid, 20 mM phosphoric acid, 20mM boric acid, pH was adjusted to 11 by titrating with 1M 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). In one experiment the peptides were pre-fractionated by IEF into 12 fractions as described above. In all cases, 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 packed in house with C18 material. Peptides were ionized using an ESI source and analysed on a Q-Exactive with the settings described above.

The acquired raw files were analysed using MaxQuant as for the "enrichment efficacy experiment". 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.

2.2.5.1 AHA pulse-chase data processing

Note to reader: The normalization strategy was developed in collaboration with Celine Sin and Angelo Valleriani (MPI for colloids and interfaces). In addition, the AIC probability and parameter fitting where fully developed and implemented by Celine Sin and Angelo Valleriani. The protein feature analysis was developed by Henrik Zauber (MDC). Large portions of the relevant segments below are re-worked versions of their original texts.

2.2.5.2 Normalization

Normalization is a challenge for experiments measuring abundances. Differences in starting material, labeling efficiency, etc. are all contributors to aberrations in measurements. A common normalization strategy is median-normalization, where the assumption is that the median values should not change throughout a series of experiments. However, for pulse-chase experiments, we expect the median to decrease over time, thus this strategy cannot be used. Instead, our normalization scheme is based on the assumption that there are stable proteins whose amounts decrease little during the time course of the experiment (Figure 18) (Schwanhäusser et al., 2011; Toyama et al., 2013). These proteins would remain unchanged and equal to 100 % remaining throughout the experiment. We can identify these very stable proteins because the Medium/Heavy and Light/Heavy SILAC ratios of these proteins should be among the highest in all experiments. Using this method, we identify the most stable proteins, and then calculate the multiplication 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 %.

To find the most stable proteins, we consider proteins with data at all time points in all replicates to ensure that all normalization candidates are able to contribute to the normalization factor. 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 L/H or M/H signal strength (from smallest to largest, at time t) to the interval (0, 1). Proteins with higher SILAC ratios at each time point will have higher scores. Thus, 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 biological replicate. 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 difference from the maxScore. 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.

2.2.5.3 Parameter fitting

In this study, we consider two simple models: a one-state model (exponential decay) and a two-state model (non-exponential decay). In the 1-state model, proteins are born into state A (see Figure 20). 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. In other words, the 1-state model models exponential decay. In the 2-state model, proteins are born into state A. From state A, the molecule can immediately degrade at the rate of k_A, or it can transition to state B with the rate $k_{A\text{-}B}$. Molecules that reach state B are degraded with rate k_B. I.e. the life expectancy of the molecule depends on the age of the molecule, since the older a molecule is the more likely it is that it has transferred to the second state.

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" and might even be degraded while some molecules are just newly synthesized. Therefore, the length of the pulse must be taken into account for the calculations.

The derivation of the mathematical description consists of two steps. The first is to translate the single molecule dynamics model into 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. Calculation of the response of the system resulting from a pulse has been covered in (Sin, 2016).

In our formalism, the function $\Lambda(t)$ 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:

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

$\Lambda(t_p + \Delta) \propto G(\Delta) - G(t_p + \Delta)$ for the 2-state model

where $G(u) = k_{A-B}(k_{A-B} + k_A)e^{-k_2 0 u} + k_B(k_A - k_B)e^{-(k_{A-B}+k_A)u}$

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.

2.2.5.4 Model selection by the Akaike information criterion

The Akaike Information Criterion indicates the quality of the model for a given set of data (Akaike, 1974; Burnham K.P., 2002). 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, there is the danger of over-fitting – that is the model could be approximating not only the system's dynamics, but also quantities not related to the 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 trade-off 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:

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

2.2.5.5 Delta-score (Δ-score) calculations

If a protein is exponentially degraded, one can draw a straight line in a semi log plot between time point 0 h (100 % protein remaining) and the measured value for another time point and derive the relative protein abundance at any other time point. However, if a measurement for another time point is not on the line, allowing for measurement and quantification errors, the protein is non-

exponentially degraded. We used this fact to estimate the size and direction (increased or decreased stability with age of the molecule) of the non-exponentiality of degraded proteins. We used the median % protein remaining at time point 8 h (tp8) after chase to calculate the expected relative protein abundance at time point 4 h assuming exponential degradation. For this we solved the straight line equation (y = mx + c) for x = 4h, where the intercept c is log(100%), and the slope m is calculated using the value at tp8:

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

Finally, we calculated the distance from the measured value, log(tp4), to the expected value, y(4h):

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

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 taken place by 4 h chase. Thereby we expected to be able to catch age-dependent stabilization (or destabilization) by comparing these two time points. Also, very few proteins had a half-life shorter than 2.5 h and could thereby not be detected at the 8 h time point (See supplemental table 1). In addition, these were almost exclusively exponentially degraded according to the AICp.

2.2.6 Bioinformatic analysis

Note to reader: The protein feature and complex centric analyses were developed by Henrik Zauber (MDC). Protein complex assembly and interface analysis was developed and performed by Jonathan N Wells and Joseph A Marsh (both at the University of Edinburgh). Their original texts have been slightly modified for this purpose.

2.2.6.1 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 listed and not listed as a member of a protein complex as defined in (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. Disordered, Helix and Beta Sheet fractions per protein were taken from secondary structure predictions using the s2d method (Sormanni et al., 2015). 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. Features related to structure were length-normalized. For each feature a ROC-curve was generated and the area under the curve (AUC) 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.

2.2.6.2 N- and C-terminal protein abundance analysis

In silico division of all proteins in the mouse data base, re-matching of all identified peptides and plotting was done using R. Only peptides that were matching uniquely to the protein N- or C-terminus were retained.

2.2.6.3 Estimation of relative protein abundance after pulse

To estimate the protein abundance after the pulse we used intensity based absolute quantification (iBAQ, (Schwanhäusser et al., 2011)). First we calculated iBAQ values for proteins based on the H-intensities (derived from the newly synthesized proteins without chase). Second, we normalized these iBAQ data from the three experiments using the median H-intensities for the LSD-proteins (see normalization above). This allowed the combination of experiments. Finally, we reported the median protein H-iBAQ from the three experiments as the protein abundance.

For a complex-centred analysis of the relative protein abundances after pulse, identified proteins from the mouse dataset were mapped to a filtered version of PDB (see supplemental table 3) using gene names. Protein abundances were normalized in a complex centred manner: First all proteins that mapped to a complex were extracted. Second, stoichiometric differences between subunits within a complex were normalized out by dividing each protein abundance against the listed stoichiometry in the PDB database. Third, abundances of all proteins of a complex were normalized to the average abundance of each complex. The resulting complex-centred abundances were compared between the protein subsets ED, NED and UN. Only proteins from complexes with at least one ED and one NED subunit were considered for the analysis.

2.2.6.4 Protein complex assembly and interface analysis

All protein structures in the Protein Data Bank (2016-02-24) were searched for polypeptide chains with more than 70 % sequence identity to a mouse or human gene. If multiple matches were detected, a single chain was selected by sorting by sequence identity, followed by number of unique subunits in the complex, and finally the number of atoms present in the chain. AREAIMOL was used to calculate pairwise interfaces between all pairs of subunits (Winn et al., 2011). For all complexes, excluding those containing nucleic acid chains, the normalized assembly order was calculated by first predicting the (dis)assembly pathway using all the pairwise interfaces from each heteromeric complex using the *assembly-prediction* package (Marsh et al., 2013; Wells et al., 2016). For subunits represented by multiple copies within a single complex, the mean assembly order of each subunit type was considered. The normalized assembly order was calculated so that the last subunit to assemble has a value of 1, the first has a value of 0, and the average value for all unique subunits in a complex is equal to 0.5.

2.2.7 Serial dilution experiment

Heavy and Light-labeled NIH 3T3 mouse fibroblasts were lysed in RIPA buffer as above. Protein concentrations were measured by Bradford assay using the "Coomassie Plus" kit (Pierce). Heavy proteins were serial diluted into the Light protein solution. The mixed proteins were then precipitated, denatured, alkylated, digested and stageTipped as described above. Each dilution step was analysed using a 2 h gradient of increasing Buffer B concentration on a 15 cm column and measured on a Q-Exactive as described above. The resulting raw files were analysed as for the "enrichment efficiency experiments" but this time once with "reQuantify" once active and once inactive.

2.2.8 SILAC pulse-chase (confirmation experiment)

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 Figure 48). Cells were then washed in PBS before being trypsinated for 2 min at 37 °C. Detached cells were diluted 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 the pellet then frozen. After the Medium chase (see Figure 48 for different chase times), cells were also spun down and frozen. Additional "label-swap" experiments were performed in the same fashion. However, in the label-swap experiments the cells were pulsed with Medium and chased in Heavy amino acids (see Figure 48). Cell pellets were lysed and proteins denatured in 0.2 % SDS, 0.1M DTT and 50 mM ABC buffer (pH8) 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 minutes. Proteins were precipitated by Wessel-Flügge precipitation, digested "on bead" and stageTipped as in the "dynamic SILAC experiment". The peptides were resolved on a 4 m long monolithic column using a 12h gradient of increasing buffer B concentration and a flow rate of 500 nl/min. Peptides were ionized by ESI and analysed on a Q-Exactive Orbitrap. Resulting raw files were analysed on MaxQuant with the same parameter settings as for the AHA p-c experiment. Plotting and statistics were performed using R and figures were modified in Illustrator.

2.2.9 Inhibitor experiments

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

Inhibition of autophagy was monitored by taking a sample for western blotting. Scraped cells were directly lyzed in LDS sample buffer (NuPAGE, Invitrogen) supplemented with DTT. Samples were run on 4-12 % gradient Bis-Tris gel (NuPAGE, Invitrogen) before being blotted onto a PVDF membrane (Immobilon-P, Millipore) using a wet blotting contraption. The membrane was probed against LC3-II (Novus NB100-2220) and was afterwards stripped at 37 °C for 15 min in 2 % SDS and 2 % β-mercaptoethanol in 65 mM Tris Base (pH6.7, Roth) before being re-blotted with an antibody specific for β-actin (Sigma-Aldrich A5441).

Cells for mass spectrometric analysis were scraped, lysed, and had their AHA labeled proteins clicked to alkyne-agarose beads. Proteins were reduced and alkylated before beads were washed. Proteins were digested "on bead" by LysC and then trypsinated overnight all as in the main AHA p-c experiment. Peptide solutions were put on 4mm/1ml C18 columns (Empore, 3M) and washed in buffer A. Peptides were eluted in buffer B and vacuum dried.

Peptides derived from MG132 samples were loaded on a column packed with 2:1 mixture of 3 µm weak anion exchange beads (WAX, PolyLC Inc.) and 3 µm strong cation exchange beads (SCX, PolyLC Inc.) followed by a C18 "trap" segment (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 WAX/SCX material where then separated on a 15 cm column with 2 h gradients of increasing buffer B concentration with a 250 nl/min flow rate. Autophagy inhibited samples were put on SCX tips and washed in no salt buffer, as described above, to minimize polymer contamination. Samples were then eluted with 500mM ammonium acetate before being desalted on stageTips. Eluted samples were separated using an 8 h gradient on a 2,000 mm monolithic column with a flow rate of 300nl/min. Actinomycin D samples were put on SCX tips but manually fractionated by eluting in steps by increasing ammonium acetate concentration in 8, 16, 32 and 500 mM steps before being put back on stageTips (Kulak et al., 2014). Eluted samples were separated on a 15 cm column using 2 h gradients.

Peptides were ionized using an ESI source and analysed 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 analysed using MaxQuant as the standard AHA p-c experiment.

Protein profiles were normalized using values from LSD proteins branded in the AHA p-c mouse dataset. 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 two protein subsets identified as NED, ED and UN.

2.3 Experiments using human RPE-1 and RPE-1 trisomic cells

Note to reader: Chromosome spreads were performed by Neysan Donnelly and Zuzana Storchova (MPI of Biochemistry) and whole genome sequencing and data analysis was performed by Xi Wang, Jingyi Hou and Wei Chen (all MDC). Some of the segments below are modified from their original versions.

2.3.1 Preparation of chromosome spreads and chromosome painting

Cells were grown to 70-80 % confluency before treatment with 50 ng/ml colchicine for 3-5 hours. Subsequently, cells were trypsinated and centrifuged at 250 rcf for 10 min. Pellets were then resuspended in 75mM 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 chromosome 5, 11 and 12 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.

2.3.2 Genomic DNA sequencing and copy number estimation

DNA was isolated using the Blood and Cell Culture DNA kit (Qiagen) according to the manufacturer's recommendations. 1 µg genomic DNA was sheared following the protocol from the SureSelect Target Enrichment System Kit for Illumina Multiplexed Sequencing (Agilent Technologies). 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 1 x 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 100 Kb 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 mapping 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.

2.3.3 AHA p-c of RPE-1 and RPE-1 trisomic cells

RPE-1 and RPE-1 trisomic cells were grown, methionine-starved, AHA-pulsed, chased and lysed as described for the mouse fibroblast. 3 biological replicates were performed per cell line. Experiments were started when cells reach ~30 % confluency to avoid that the cells reached full confluency during the 32 h of chase. 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). Eluted samples were speedVacced and for two of the samples Buffer A was added to 10 µl final volume. Of these, 5 µl was loaded onto a 15 cm column and 5µl onto a 2m monolithic column using a HPLC system. The samples were analysed on a Q-Exactive

orbitrap system, as described above, employing 4 and 6 h linear gradients of increasing Buffer B, for the 15 cm and 2 m column respectively. For one sample the peptides were further SCX fractionated into 2 fractions by 125 mM and 500 mM ammonium acetate as described above. These samples were analysed using 4 h gradients of increasing Buffer B concentration over a 15 cm column. The resulting raw files were analysed using MaxQuant with previously described parameters, with the exception that the Andromeda search engine was matching to the Human Uniprot database (2014-10). Normalization, fitting of models, Δ-score and abundance after pulse calculations were performed as for the mouse fibroblasts.

2.3.4 Steady state protein abundance

Fully Heavy SILAC labeled RPE-1 and Light labeled RPE-1 trisomic cells were grown in 10 cm plates to 70 % confluency. Cells were scraped in ice cold PBS before being spun down at 1000 rcf and PBS decanted. Cell pellets were lysed 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 samples were directly Wessel-Flügge precipitated, resuspended in Urea buffer, “on bead” digested and stageTipped all as described above. Eluted peptides were analysed using a 6 h gradient on a 15 cm column as described above. The Q-Exactive was run with standard settings and raw files were analysed as described for the “AHA p-c 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 where RPE-1 cells were grown in Light SILAC medium and RPE-1 trisomic cells were grown in Heavy SILAC medium.

2.3.5 Bioinformatics of RPE-1 cells

Delta-score (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 the first entry in the gene name column from the proteinGroups.txt tables. 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 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. Overall, 2529 proteins had comparable quantitative information between the RPE-1 and RPE-1 trisomic cells. 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. Trisomic regions included proteins mapping to chromosome 5 and chromosome 11 downstream from position 62650000. Chromosome 12 and 10 starting from nucleotide position 62500000 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. Proteins were further categorized in a complex-centred manner.

3 Results

3.1 Establishment of azidohomoalanine labeling

To study protein degradation kinetics we employed a combination of fully SILAC labeled cells with pulse-chase by the non-natural amino acid azidohomoalanine (AHA) (Figure 9). AHA is metabolically incorporated into newly synthesized proteins in place of methionine (Kiick et al., 2002). In contrast to methionine, AHA contains an azido group which makes it suitable for click chemistry. In principle, any functionality containing an alkyne, e.g. biotin or fluorophores, can be clicked to an AHA-containing protein (Dieterich et al., 2006; tom Dieck et al., 2015). In our approach, we combine AHA-labeled protein with alkyne-bearing agarose beads. By selectively labeling newly synthesized proteins with AHA, one can then efficiently enrich this population and analyse it by mass spectrometry (Dieterich et al., 2006; Eichelbaum and Krijgsveld, 2014; Eichelbaum et al., 2012).

In the AHA pulse chase (AHA p-c) method presented here we label Light, Medium-heavy and Heavy SILAC cell cultures for 1 hour with AHA (Figure 10). These AHA-labeled newly synthesized proteins are the age-defined (0 - 60 minutes old) population that we follow during the chase. One of the cultures we directly collect (0 hour time point) while the other two we chase in medium depleted of AHA ("cold" chase) for different amounts of time. We then combine the samples with different SILAC labels and enrich the newly synthesized proteins. Finally we measure the samples using shotgun proteomics. The SILAC labels allow us to quantify how much of the proteins are left after the different chase times (Figure 10). By combining results from multiple experiments, we can cover many time points.

When establishing AHA p-c, we were mainly concerned with the potential impact of the unnatural amino acids on the cells and the proteins that had incorporated AHA. In this segment I will discuss the influence of AHA labeling on two of the most important biological parameters, namely protein synthesis and degradation. I will also discuss the more technical factor which is the enrichment efficiency of AHA-labeled proteins.

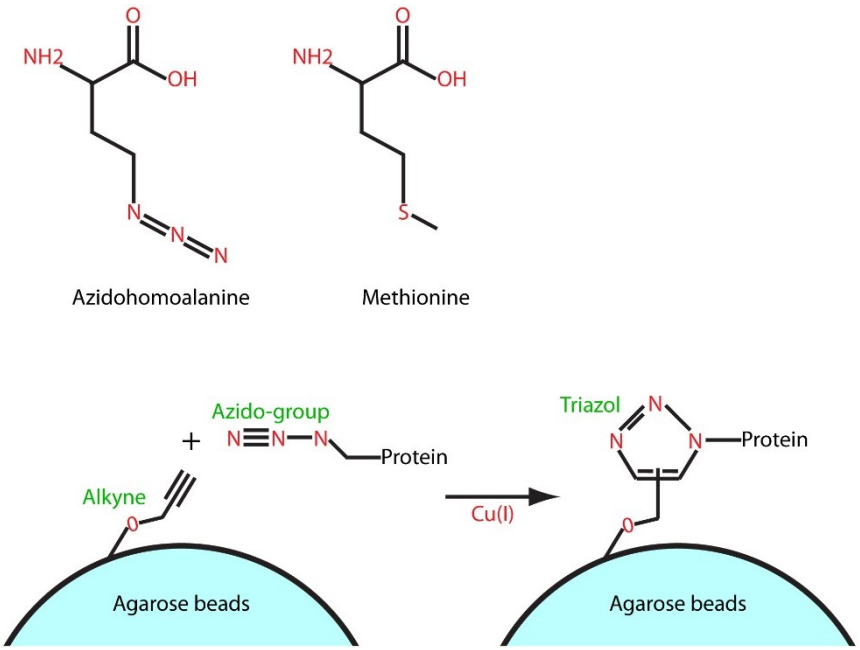

Figure 9. Azidohomoalanine is a "clickable" amino acid.

Upper panel compares the structural formula of azidohomoalanine (AHA) to methionine. AHA is metabolically incorporated into newly synthesized proteins instead of methionine (Kiick et al., 2002). Lower panel displays the principle of the chemical "click" reaction (Eichelbaum et al., 2012). Copper-catalysed 1,3-cycloaddition covalently couples the AHA-containing new proteins to alkyne-bearing agarose beads. The highly stable triazol-group allows for stringent washing of the beads without loss of new proteins. Figure is not to scale.

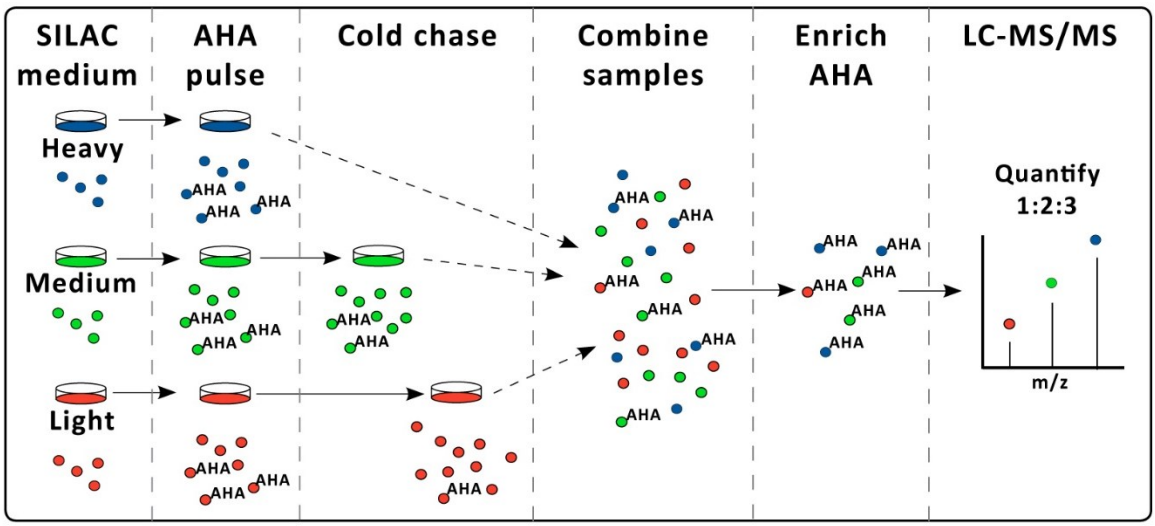

Figure 10. Azidohomoalanine pulse-chase (AHA p-c) experiments.

Fully SILAC labeled methionine starved cells are pulsed with AHA for 1 hour. The Heavy SILAC cells are then directly collected whilst the Medium and Light SILAC cells are chased in medium depleted of AHA but containing methionine (cold chase). After different amounts of chase time, during which the AHA-containing proteins can only degrade, the cells are collected and lysed. The samples are then combined and alkyne-agarose beads and click chemistry are used to enrich the AHA-bearing proteins. The proteins are then digested “on bead” before being analysed by LC-MS/MS. Using the SILAC ratios, the relative amount of protein remaining after chase can be calculated for the identified proteins.

3.1.1 Long AHA-pulses lead to decreased protein synthesis

The first thing we wanted to understand was the impact of AHA-labeling on cells and proteins. One risk of using an artificial amino acid is that it will impact protein synthesis rates. To estimate protein synthesis levels when using AHA, I performed dynamic SILAC experiments in which I pulsed methionine-starved cells with 0.2 mM methionine as a control or different amounts of AHA for 4 h (Figure 11). In parallel with the treatment, I also added Heavy (H) isotope labeled arginine (Arg10) and lysine (Lys8). One can then measure the amount of new Heavy proteins that have been synthesized compared to the old Light (L) population. By comparing the H/L ratio of protein in the methionine-treated control cells to the H/L ratio in the AHA-treated cells, one can estimate the relative amount of new proteins in the different conditions.

The results clearly show a decrease in the amount of proteins synthesized with a median of 67 % and 65 % protein synthesized in experiment 1 (Figure 11) and 2 (n = 168, not shown), respectively. Interestingly, protein synthesis did not increase when titrating with more AHA (Figure 11). This finding is consistent with the fact that the methionyl-tRNA synthase has a 400 fold higher catalytic efficiency when methionine is the substrate as compared to AHA, suggesting that 1 mM AHA already saturates this potentially rate limiting step (Kiick et al., 2002).

Another interesting finding is that when AHA was added in combination with methionine, protein synthesis levels were rescued (Figure 11). This is consistent with previously published data (Bagert et al., 2014). In that study, Bagert and colleagues demonstrated that mixing AHA and methionine in a 1:30 ratio limited the impact of AHA on protein synthesis, even though it also significantly decreased the amount of AHA-containing proteins. Importantly, that the combined AHA and methionine treatment rescues protein synthesis indicates that the addition of AHA *per se* is not toxic for the cell.

A gene ontology analysis (not shown) of the data indicates that proteins involved in protein synthesis and in cell cycle progression are the least synthesized populations when methionine is exchanged for AHA (Huang et al., 2009). A look at only ribosomal proteins confirm this picture, with a median of 31 % (Figure 11, right panel) and 35 % synthesis levels of these proteins in the first and second (n=6, not shown) experiment, respectively. A decrease in synthesis of these populations of proteins could indicate that the cell experiences amino acid starvation on top of the above discussed tRNA loading deficiencies (Efeyan et al., 2012). Methionine is used for many purposes in the cell, such as DNA methylation, and the cell senses and responds to the lack of this and other amino acids (Efeyan et al., 2012). In summary, the addition of AHA does not seem to fully compensate for the methionine depletion and cells respond by decreasing protein synthesis in a specific manner.

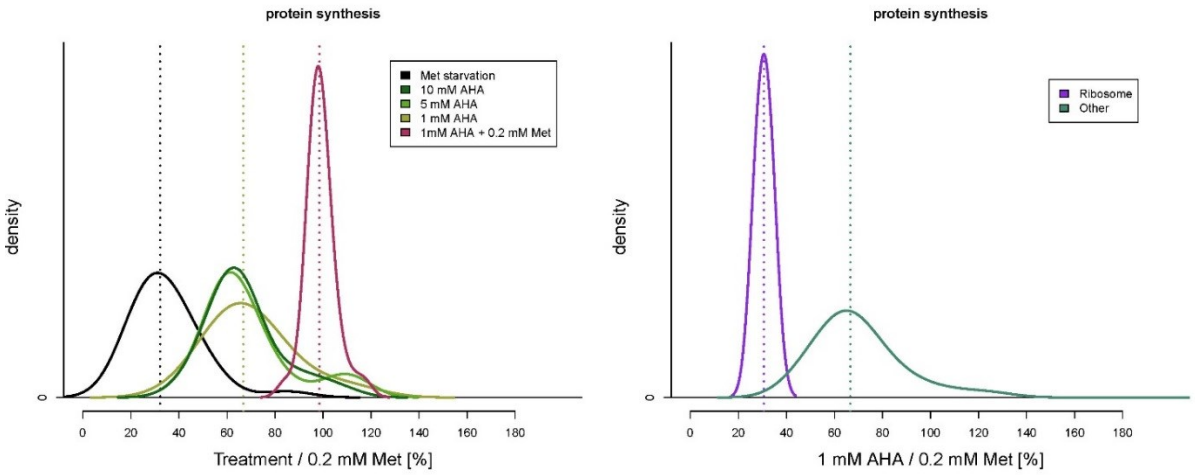

Figure 11. AHA labeling leads to decreased protein synthesis compared to methionine control.

4 h of pulse labeling with different concentrations of AHA or 0.2 mM (standard concentration in DMEM) methionine in combination with Heavy lysine and arginine. A) H/L ratios of methionine treated cells were set as 100 % protein synthesis. Using the ratio of ratios, we estimated the fraction of protein synthesis during the titration of AHA. When using 1mM AHA, proteins were synthesized with a median 67 % efficiency. Increasing the levels of AHA did not rescue synthesis rates. However, adding methionine to the AHA-treated cells rescued protein synthesis (magenta). n = 30, proteins that overlap in all 6 conditions (i.e. the same proteins are plotted for all densities). B) Looking specifically at the ribosomal proteins (purple, n=3), they were more decreased when using 1mM AHA than the average protein (n=73). Shown is 1 representative experiment out of 2.

3.1.2 AHA-labeled proteins can be efficiently enriched

To establish how well we could enrich AHA-labeled proteins, I performed a classical radioactive pulse experiment. I labeled proteins with either methionine or AHA in combination with ^{35}S cysteine for 1 h (Figure 12 A). I then lysed the cells and performed the click chemistry. Finally, I quantified the ^{35}S derived radioactivity in the supernatant and on the beads. I found that I could pull down 70 % of the radioactivity when I labeled the proteins with AHA (Figure 12 B). In contrast, the vast majority of the radioactivity stayed in the supernatant in the methionine-treated cells. Due to recirculation of methionine and the fact that not all proteins contain methionine we would not expect that all proteins would be labeled by both AHA and cysteine, especially after such a short pulse. In conclusion, we can enrich AHA containing proteins with 70 % efficiency after a 1 h pulse.

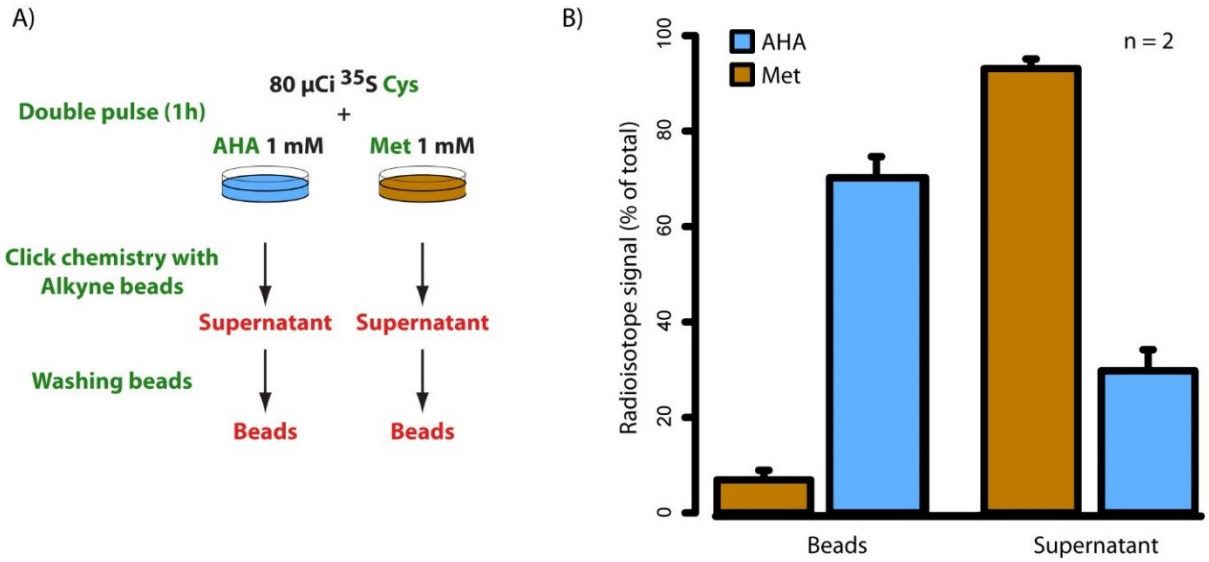

Figure 12. Enrichment efficiency of newly synthesized AHA-containing proteins by click chemistry using alkyne beads. A) Experimental setup: NIH 3T3 cells were labeled in parallel with ^{35}S Cysteine and either methionine (Met) or azidohomoalanine (AHA). Scintillation counter measurements were taken of the supernatant after the click reaction and of AHA-bearing proteins covalently bound to the alkyne-beads (beads). B) The scintillation counts from two experiments were normalized to the total signal (beads + supernatant) to enable direct comparison between the replicates. When AHA was used, the majority ~70 % of radioactively labeled proteins were attached to the beads in comparison to less than 10 % for methionine treated cells. Error bars: standard deviation between the two replicates.

3.1.3 Short AHA-pulses have minor effect on protein synthesis

Because of the large population of pre-existing Light proteins in the dynamic SILAC experiments, I was limited to a relative long pulse of 4 hours to properly quantify the newly synthesized Heavy proteins. This pulse time was significantly longer than the 1 h we planned to use for the AHA p-c experiments. Therefore I additionally performed radioactive pulse experiments to determine the effect of shorter AHA-pulses on protein synthesis. I labeled methionine-starved mouse fibroblasts with ^{35}S cysteine and either 1mM methionine or AHA for 1 h (see time point 0 h in Figure 14 A). I then quantified the incorporated radioactivity as a measure of protein synthesis. A trend towards less proteins synthesis (~90 % of control levels) in the AHA treated cells was visible, although not statistically significant. Nevertheless, this experiment demonstrates that short AHA pulses have a limited impact on protein synthesis.

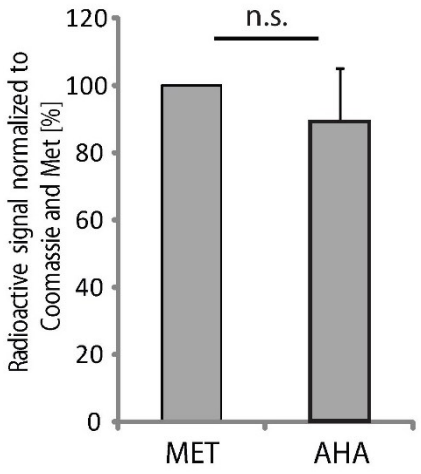

Figure 13. Short pulses with AHA have a minimal impact on protein synthesis.

NIH 3T3 mouse fibroblasts were starved for methionine and cysteine for 45 min before being pulsed with ^{35}S Cysteine in combination with either 1mM of Methionine or azidohomoalanine for 1 h. Proteins in cell lysates were resolved on a SDS-PAGE (see time point 0 h in Figure 14 B). Proteins were fixed in the gel and stained by Coomassie blue. Subsequently, the gel was dried before total loaded protein was estimated by the Coomassie blue stain and the radioactive signal was measured in a phosphorimager. Error bars: standard error, n = 6 biological replicates. n.s, not significant using a two-sided t-test and an alpha = 0.05.

3.1.4 AHA-incorporation has minimal impact on global protein stability

Because the goal of this thesis is to look at protein degradation kinetics, the most important factor to consider is the potential effect of incorporating one or more AHA molecules on a protein's stability. AHA has previously been shown to be non-toxic to sensitive cells, such as neurons, and to whole organisms, such as zebra fish and mice (Calve et al., 2016; Cohen et al., 2013; Dieterich et al., 2007; Hinz et al., 2012; McClatchy et al., 2015; tom Dieck et al., 2015). Nevertheless, there is a risk that AHA stabilizes or destabilizes proteins. To test this I performed radioactive pulse-chase experiments. They were performed by co-labeling proteins in mouse fibroblasts with ^{35}S cysteine and either AHA or methionine (Figure 14 A). I either harvested pulse labeled cells directly or chased them for 6 or 24 h in "cold" medium depleted of radioactive amino acids but in the presence of either 10 fold cysteine and methionine or the protein synthesis inhibitor cycloheximide, which was added to prevent re-incorporation of radiolabeled amino acids. When the three biological experiments from each chase conditions were quantified, I found no significant differences in the levels of newly synthesized proteins after 6 and 24 h of chase (Figure 14). This is consistent with similar studies on AHA and protein stability (Cohen et al., 2013; Howden et al., 2013; tom Dieck et al., 2015). In addition, previous groups have shown that AHA-containing proteins are not more ubiquitinated than normal proteins (Dieterich et al., 2006). AHA has even been used previously for pulse chase experiments (Chai et al., 2016; Cohen et al., 2013; Mirigian et al., 2014; tom Dieck et al., 2015). In conclusion, consistent with previous literature, we find no evidence that AHA has global effect on protein degradation.

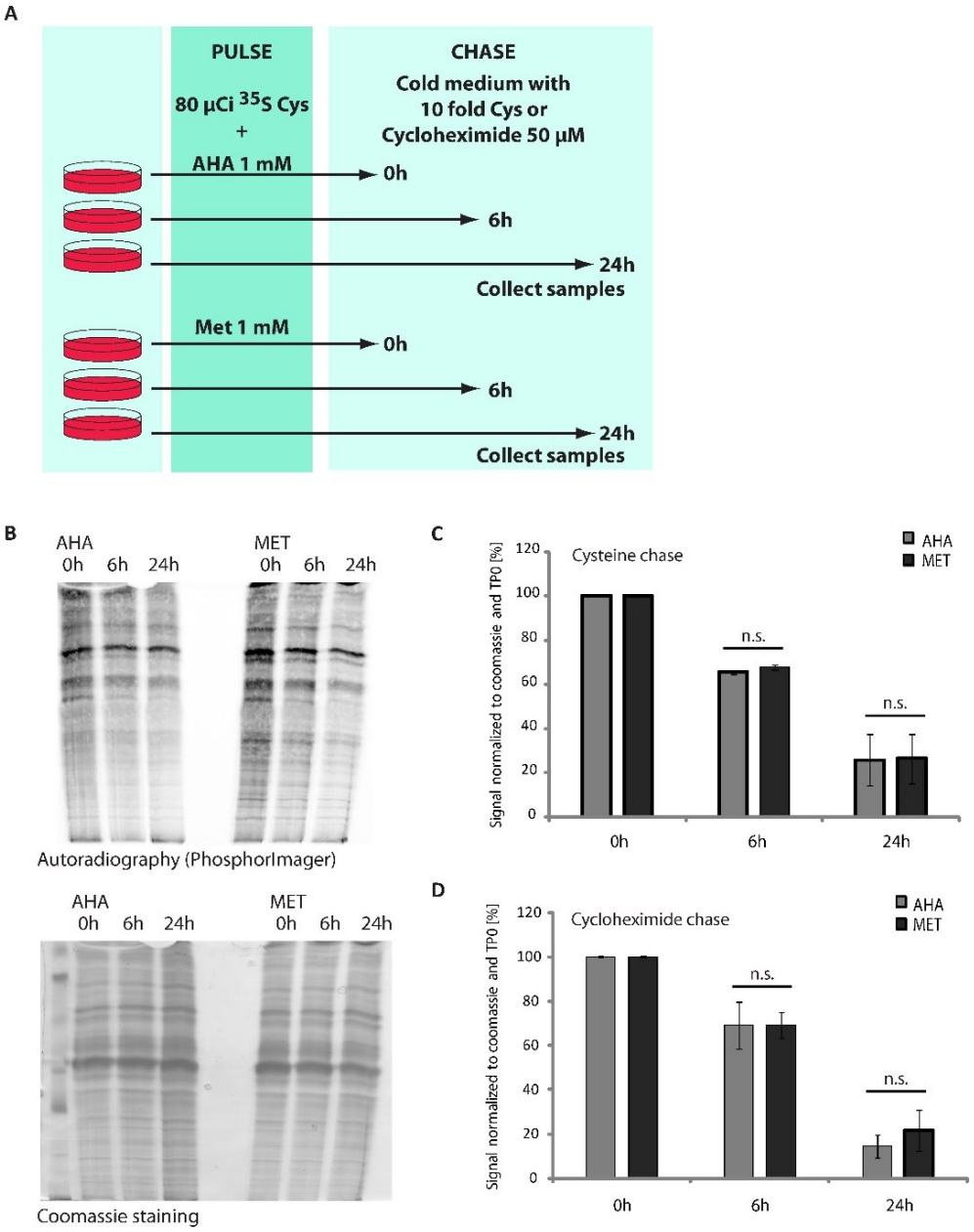

Figure 14. AHA does not affect global protein degradation.

A) Experimental design. Cells were co-labeled with 80µCi ^{35}S Cysteine and either 1mM AHA or 1mM methionine (MET) for 1 h. The cells were then washed and chased in medium depleted of ^{35}S cysteine but with 10 fold cysteine or cycloheximide added to prevent re-incorporation of radiolabeled amino acids. B) Radioactive proteins were probed on a dried SDS-PAGE gel. Figure shows one replicate from the cysteine chase experiment. C) and D) quantification of the cysteine chase (n=3 biological replicates) and cycloheximide chase (n=3) experiments, respectively. No significant differences in abundance after pulse were detected. Error bars: standard error. n.s, not significant using a two-sided t-test and an alpha= 0.05.

3.2 Proteome-wide analysis of protein degradation kinetics

After establishing that short pulses with AHA have a very limited impact on protein synthesis and degradation and that I can efficiently enrich the AHA-containing proteins, I set out to perform the AHA p-c experiments. In this segment I will discuss the AHA p-c method, including data normalization, cut-off strategy, fitting of mathematical models and finally the categorization of non-exponentially degraded proteins. In addition, I will discuss some important controls, such as the estimation of the impact of background binding and SILAC ratio compression on the NED categorization.

3.2.1 AHA pulse-chase experiments using mouse fibroblasts

I set out to perform the AHA p-c experiments in NIH 3T3 mouse fibroblasts as outlined in Figure 10. I completed three experiments in which the Heavy cell population was always harvested immediately after the pulse and thus served as the common reference time point 0 h. I then chased proteins for 1 and 2 h, 4 and 8 h and finally 16 and 32 h, respectively. Figure 15 A shows the data for three peptides belonging to the exponentially degraded proteins Filamin alpha (slow exponential kinetics) and Cathepsin L1 (fast exponential kinetics) and the NED protein Basigin (Tyler et al., 2012). As is apparent in Figure 15, Basigin levels rapidly decline until 2 h of chase after which they stay relatively stable over the 32 h of chase. We performed 3 biological replicates and in total identified 5247 proteins. The profiles for all the proteins are displayed in Figure 15 B and the three profiles from the examples in Figure 15 A are highlighted. More than 2000 proteins showed up in more than one biological replicate and could thus be used to assess reproducibility (Figure 16). Overall, the reproducibility was good with a coefficient of variation below 10 % for 90 % of the proteins at time point 32 h (Figure 16).

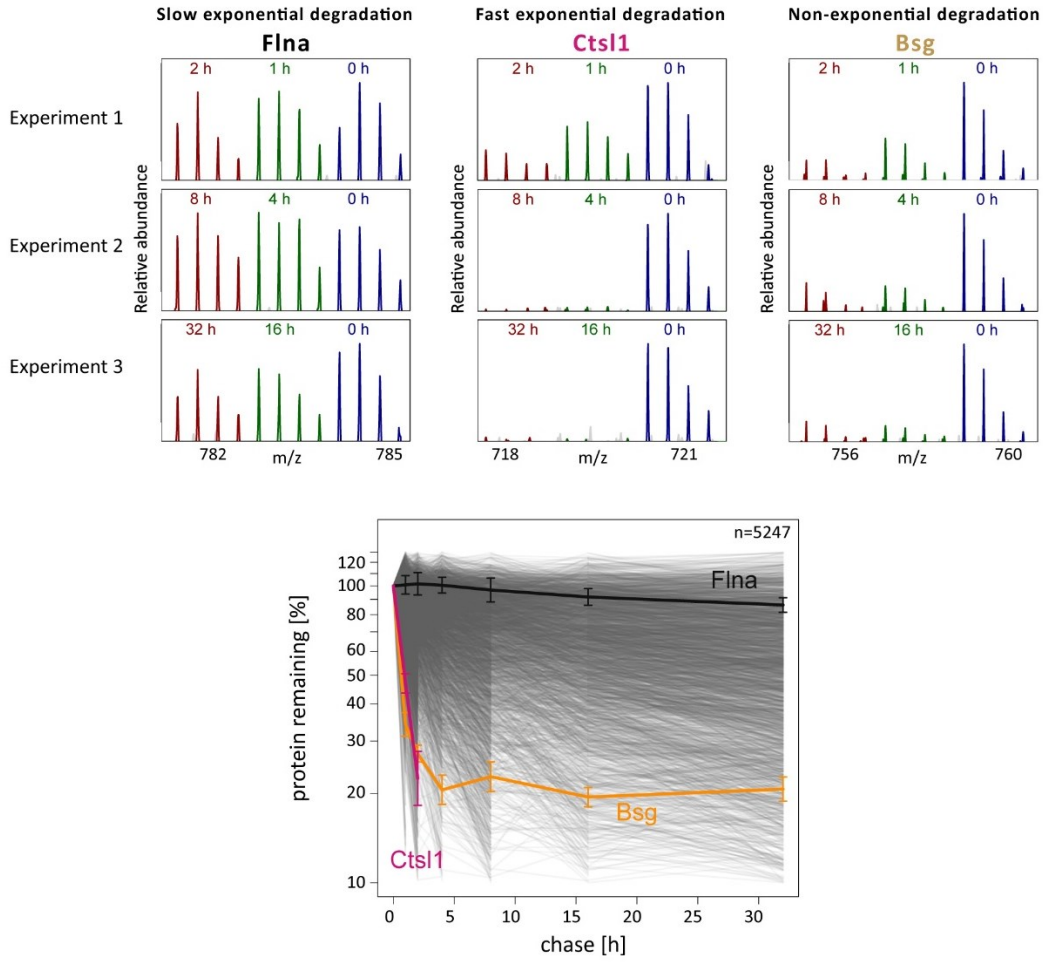

Figure 15. AHA p-c of mouse fibroblasts.

Upper panel shows example MS-spectra from the three experiments for three proteins. The three experiments were combined using the time point 0 h Heavy peaks (blue in plot) as a common reference point. Filamin alpha (Flna, represented by the peptide "ASGPGLNTTGVPASLPVEFTIDAK") is relatively stable over the 32 h chase, while Cathepsin L1 (Ctsl1,"NLDHGVLLVGYGYEGTDSNK") is rapidly degraded; both show exponential kinetics. Basigin (Bsg, "VLQEDTLPDLHTK"), a previously described NED protein, shows non-exponential kinetics also in AHA p-c. B) All the profiles from the 5247 quantified proteins in the three replicates (See **Figure 16** A). The median values from the replicates were used for plotting. The three example profiles from A) are highlighted. For the highlighted profiles standard deviations are displayed.

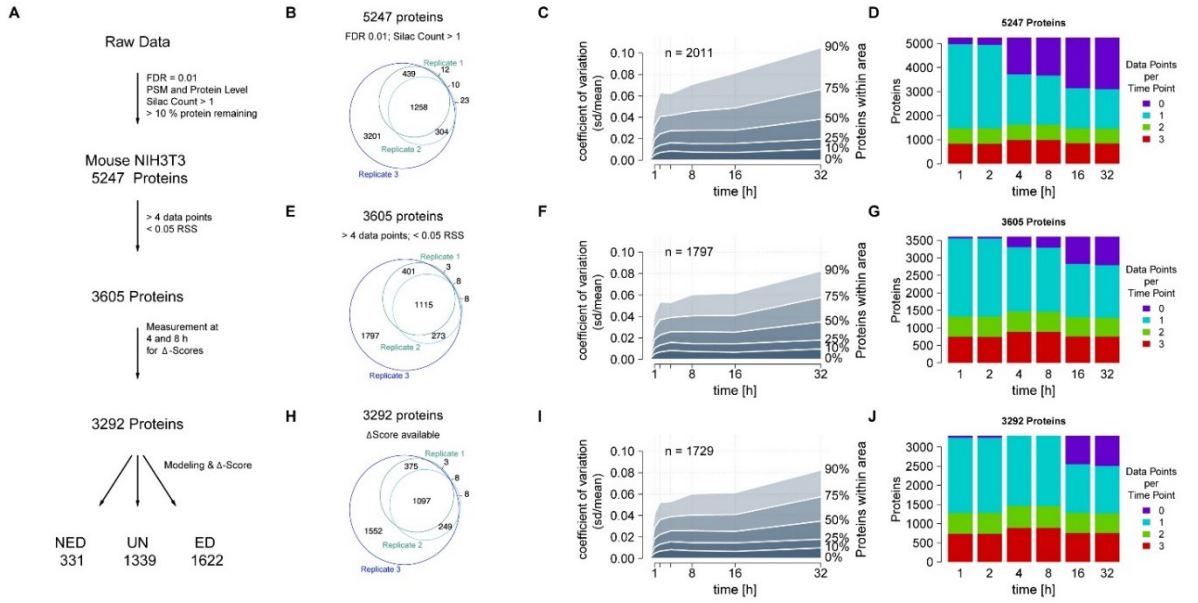

Figure 16. Overview of data processing and quality.

A) Data filtering steps. 5247 proteins quantified with a FDR on both the protein and peptide level and using the 10 % cut-off rule (see below) were further filtered for the number of data points, residual sum of squares from the model fits (RSS, see below), measurements for Δ-scores (see below) resulting in 3292 protein profiles used for downstream analysis. B, E and H) Venn diagrams displaying the number of proteins overlapping in the three biological replicates after each filtering step. C, F and I) show the coefficient of variation (CV) after each filtering step for all overlapping proteins in B, E and H). The fraction of proteins in the shaded areas has a CV equal or smaller at each time point. D, G, J) The fraction of proteins covered by a fraction of data points at each time point. Figure generated by Henrik Zauber, MDC.

3.2.2 AHA labeling does not induce detectable premature translational termination

Unnatural amino acids could interfere with protein synthesis by interrupting the ribosome and inducing premature protein synthesis termination. Since we had measurements for thousands of newly synthesized proteins (the Heavy population, time point 0 h), I could now measure the occurrence of premature translational termination. If there were consistent ribosome fall-off, we would expect to find more peptides mapping to the beginning of the proteins. To investigate this possibility, I divided *in silico* all proteins in half and re-matched all identified peptides to either the N- or C-terminal end of the proteins. Peptides matching to the middle of the *in silico* divided proteins were discarded. We found no bias towards N-terminal peptides, indicating that no major ribosome fall-off was detectable (Figure 17). Furthermore, we found no difference in the distribution of the MS-intensities, which roughly equals peptide abundance, between the N- and C-terminal peptides (Figure 17). However, this does not fully exclude that premature translation termination takes place and that the resulting truncated proteins are degraded too fast for us to detect the difference between N- and C-terminal peptides.

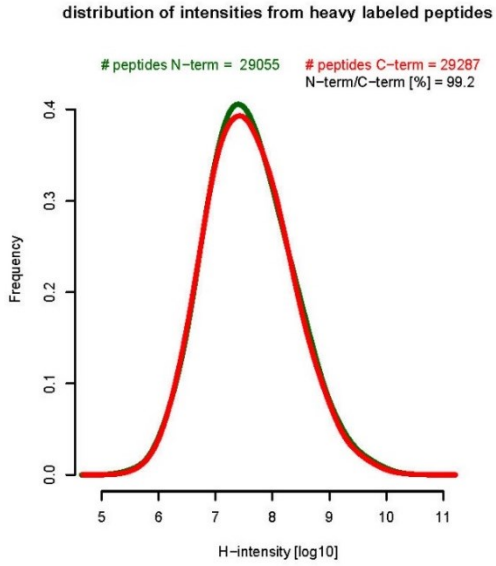

Figure 17. Distribution of intensities from newly synthesized Heavy peptides.

The frequency distribution of Heavy peptides measured in the AHA p-c experiment is similar for N-terminal (green line) and C-terminal (red) peptides. There is also no difference in the number of peptides that matched to the beginning (N-term, n=29055) or end (C-term, n=29287) of proteins. This suggest that AHA does not induce premature translational termination as this would skew the proteins mapping to the end of the protein to be less abundant, as judged by number of identifications and their intensities.

3.2.3 Normalization of AHA p-c data

Multiple errors can accumulate during the creation and handling of proteomic samples, e.g. differences in labeling efficacy, starting material, sample mixing and so forth. Therefore, most proteomic experiments require some form of normalization procedure. Mostly, the normalization strategy is based on an assumption that some proteins are always equally abundant in the cells (e.g. GAPH as a loading control in a western blot) or that most proteins are not affected by a treatment, for example, median normalization in a standard SILAC experiment. However, in the case of AHA p-c, none of these assumption can reasonably be expected to hold true. Instead, we based our normalization strategy on the assumption that the most stable proteins in our data are not degraded during the chase (Figure 18). We selected the 200 most stable proteins (least square deviation proteins, LSD proteins) based on their abundance at later time points during the chase (see Material and Methods). It is worth noting that results are relatively insensitive to the numbers of LSD proteins as long as they stay above 30 and below 200 (data not shown). In addition, since the whole population is shifted, the relative degradation profiles do not change much, even though the absolute values derived from the calculation, such as half-lives, do shift somewhat (data not shown). We then calculated the geometric mean for LSD-proteins to find the multiplicative factor that would shift the geometric mean to “100 % protein remaining” for this population. We then shifted the whole population accordingly. Each time point was normalized separately as shown in Figure 18.

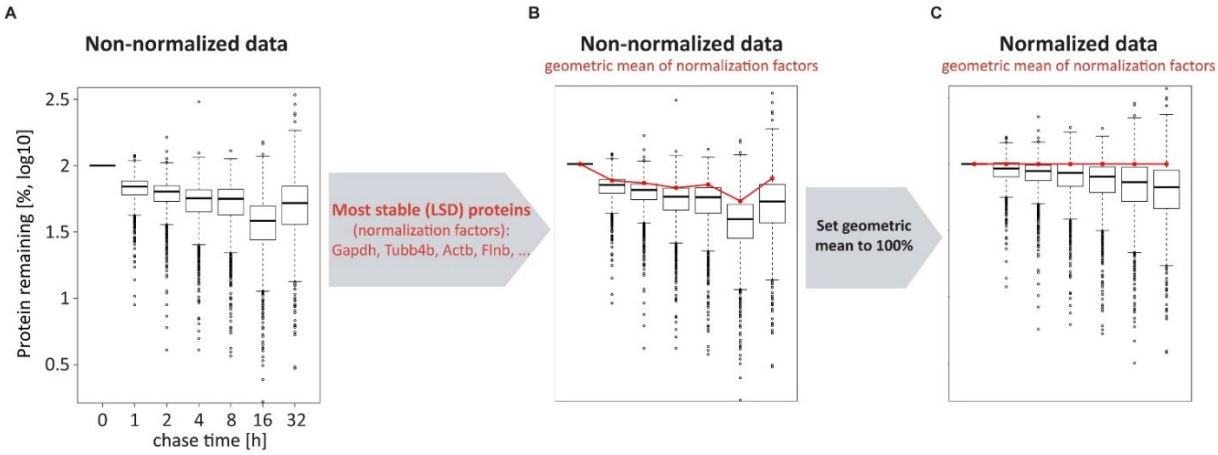

Figure 18. Normalization strategy.

In A) the distribution of non-normalized data from one replicate is displayed. From the data in all replicates we identified the most stable (least square deviating, LSD) proteins (see material and methods). Example LSD proteins are highlighted in the grey arrow. The geometric mean for this population is showed in red in B). Note that all time points get an individual normalization value by shifting the time point specific geometric mean of the LSD proteins to 100 % remaining. C) The data after normalization. In collaboration with Celine Sin and Angelo Valleriani, MPI.

3.2.4 SILAC ratio compression and cut-off strategy

A problem in SILAC-based strategies for quantification is ratio compression. Ratio compression is the underestimation of ratios that occurs when the true abundance ratios are very large. This could lead to false positive classification of NED proteins. For example, imagine that SILAC ratios larger than 1:5 would always be reported at 1:5. Then proteins that were rapidly degraded would seem to stabilize at around 20 % protein remaining (i.e. a SILAC ratio of 1:5). The first thing I did to investigate this potential problem was to perform a serial dilution experiment (Figure 19). I diluted cell lysate from Heavy labeled mouse NIH 3T3 cells into Light cell lysate. I then performed the standard shotgun proteomics work flow (See material and methods). I analysed the resulting raw files twice using the MaxQuant software once with the reQuantify option active and once with reQuantify off. The reQuantify feature allows for detection of more extreme SILAC ratios that were not detected by the first round of analysis by the software (Cox et al., 2009). With reQuantify active, the SILAC ratios stabilized at ~1/7 and the number of quantified proteins stayed constant at around 1000. However, when reQuantify was deactivated the SILAC ratios stabilized at a much lower level ~1/31 (equivalent to 3 % protein remaining) and the number of quantified proteins dropped sharply. Consequently, we performed all analysis of the AHA p-c experiments with reQuantify turned off to minimize false positive stabilization of protein profiles (Cox et al., 2009). In addition, we applied a stringent cut-off to the AHA p-c data in which all data points below “10 % protein remaining” were change to NA (see also Figure 16). This further limits the effect of artificial stabilization of protein profiles by simply excluding the data in the potentially affected range (around 3 % protein remaining).

In summary, ratio compression could have an impact on the AHA p-c data by “artificially stabilizing” degradation profiles of rapidly degraded proteins. We resolve this issue by excluding data points in the potentially affected range (< 10 % protein remaining) and by analysing the data in a conservative manner (with reQuantify disabled).

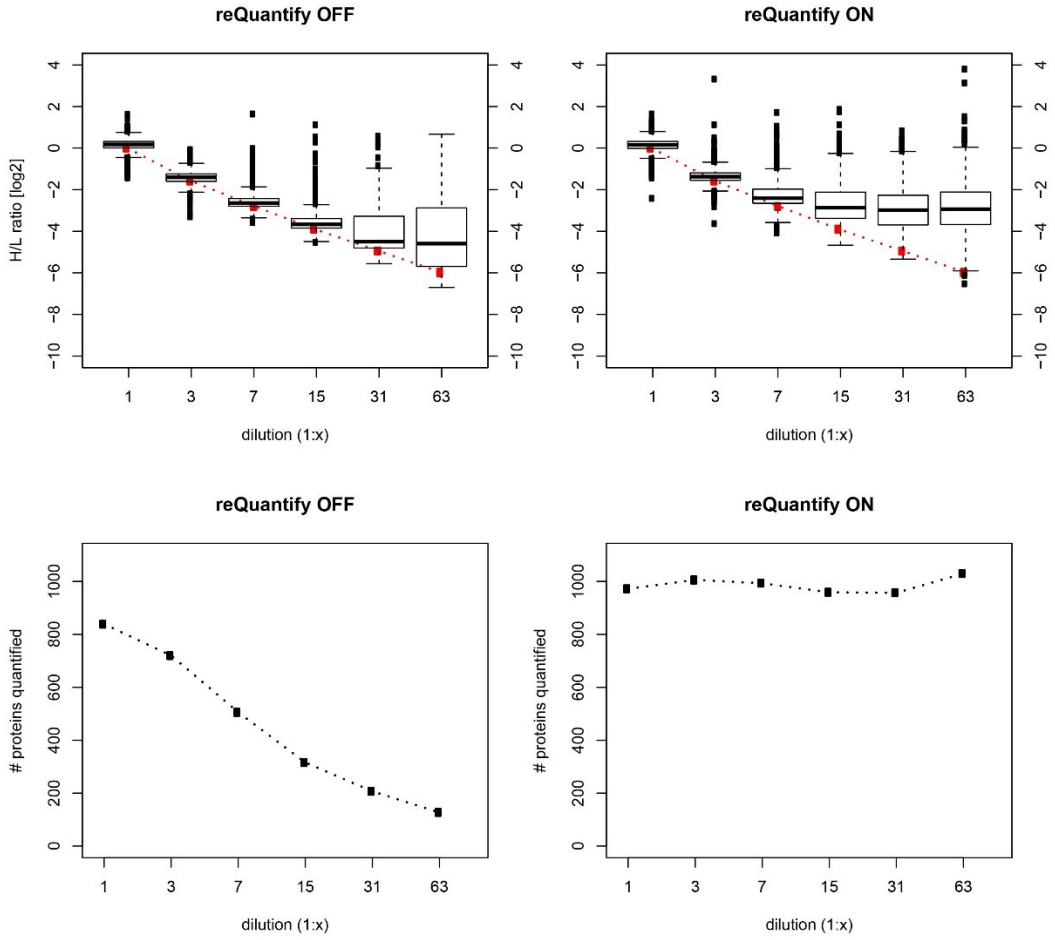

Figure 19. Serial dilution experiment to assess ability to measure large SILAC ratios.

Heavy labeled SILAC NIH 3T3 cell lysate was serially diluted into Light SILAC labeled NIH 3T3 cell lysate. In total 6 dilution steps were performed and analysed separately using LC-MS/MS. Raw files were analysed twice by the MaxQuant software, once with reQuantify parameter turned on ("ON" in figure) and once with reQuantify turned off ("OFF"). Upper panels show the H/L ratios when reQuantify is on or off. Red line indicates the expected values and black boxplots the distribution of measured values. With reQuantify on the ratios stabilize around 1/7, while when reQuantify is off the ratios reach ~1/31. The lower panels show the number of quantified (>1 SILAC count) proteins. It becomes clear that the number of quantified proteins drops rapidly when the ratios are high, limiting the number of proteins wrongly quantified only when reQuantify is off.

3.3 Mathematic modeling reveals non-exponential degradation

To systematically test if each individual profile represents a non-exponential or exponential case we employed Markov-chain based modelling approach. Celine Sin and Angelo Valleriani, both at the Max Planck Institute for Colloids and Interfaces, developed two mathematical models (Figure 20). In the first scenario, a protein is synthesized into a single state A (1-state model) from which they would be degraded with a certain degradation rate k_A. This would model exponentially degraded proteins. In the alternative 2-state model, proteins are synthesized into the state A but can now either be degraded with rate k_A or be transferred to a second state B with the transfer rate k_{AB}. Importantly, the second state has a distinct degradation rate k_B from the first state. This model would fit a non-exponential scenario.

All 5247 degradation profiles were fitted with each model and the number of data points-normalized residual sum of squares (RSS/n) was calculated (Figure 16). A probability for which model best explains the data was calculated using the Akaike information criterion (AICp, (Akaike, 1974)). In principle, the RSS/n was compared for both model fits and when they were similar the simplest model was selected to explain the data. This is because the fact that adding more parameters (the 1-state model has 1 parameter whilst the 2-state models has three) will always result in an at least equal RSS/n (see examples in Figure 20). Therefore, the AIC is used to weight the 2-state model so that it is not used to explain data that could similarly well be explained by the simpler exponential fit. In other words, the AIC probability is a conservative approach to categorize NED profiles.

We calculated AIC probabilities for all the profiles and retained proteins which had at least 5 data points (see Figure 16 and discussion below). The AIC only determines which of the tested models better explains the data with the least number of parameters. However, if none of the tested models explain the data well, the AIC would not be able to tell. We thereby excluded all profiles where the RSS/n was larger than 0.05 to exclude profiles that could not be explained well with either model (Figure 16). This resulted in 3605 proteins with an AIC probability (Figure 21 A). Out of these, 509 had a 2-state AICp higher than 0.8. We thereby conclude that a significant number of proteins are not exponentially degraded under our experimental conditions.

We also reasoned that the degradation rates from state A and B could be used to determine if the proteins were becoming more or less stable over time. By simply comparing the first degradation rate, k_A, with the second degradation rate, k_B, we found that k_A was always higher than k_B. Thus, all 2-state proteins become more stable as they age (Figure 21B).

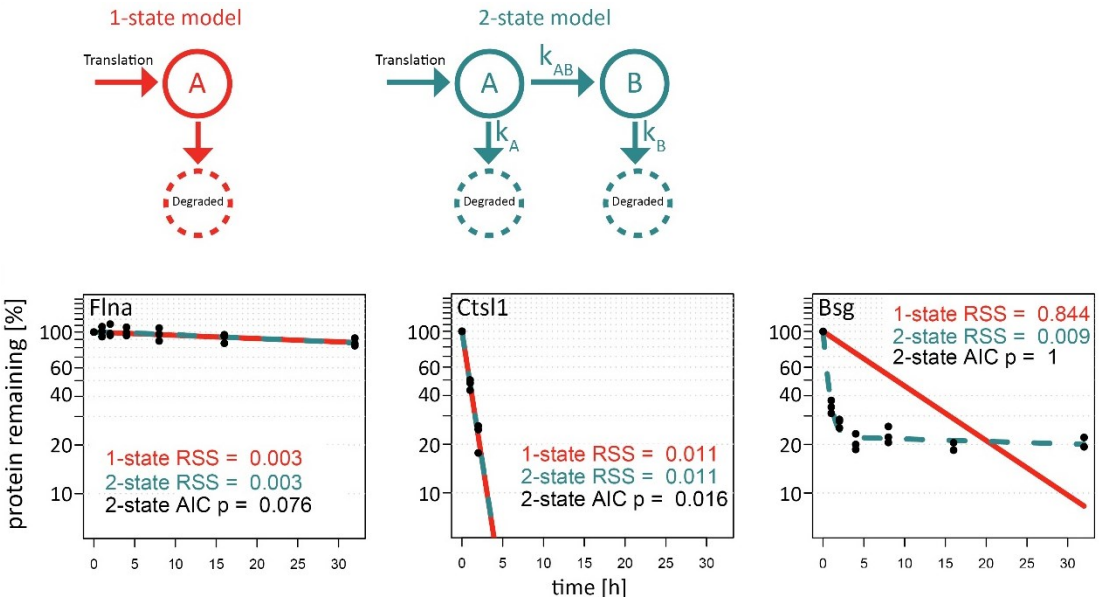

Figure 20. Markov chain modelling to distinguish between exponential and non-exponentially degraded proteins.

Upper panel displays the two models used for determining exponential or non-exponential degradation. The 1-state model, in which a protein is synthesized into a state A and from where it can only be degraded, contains only one parameter. The 2-state model contains three parameters: two distinct degradation rates k_A and k_B and one transfer rate between the two states A and B (k_{AB}). Note that the flow of the molecule is unidirectional and memoryless. The bottom panel shows the RSS/n for the two models when applied to the three example profiles from Figure 15. For Flna and Ctsl1 the two fits are almost identical and the AIC probability selects the simpler 1-state model. For Basigin the 2-state model has a much better fit and consequently the AIC recommends the 2-state model. In collaboration with Celine Sin and Angelo Valleriani, MPI.

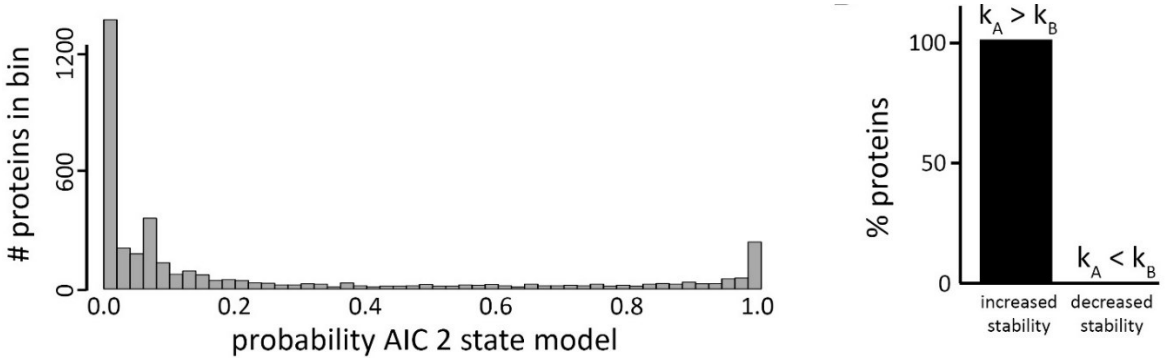

Figure 21. Distribution of AIC probabilities for the 2-state model.

Histogram (left panel) displays the distribution of 2-state AIC probabilities (n = 3605). 509 proteins have a large probability (2-state AIC probability > 0.8). These proteins are exclusively stabilized by age as judged by their degradation rates in state A and B (right panel). In collaboration with Celine Sin and Angelo Valleriani, MPI.

3.3.1.1 Δ-scores quantify the size of NED

The AIC probability only uses the RSS of the fits to determine which model better explains the data. It does not take into account the size of the non-exponentiality of degradation. To measure this we developed a simple score, the delta-score (Δ-score). The Δ-score is calculated using only the 0, 4 and 8 h time points. We drew a straight line between the 0 h time point and the 8 h time point in a semi-log plot (Figure 22 A). If the protein is exponentially degraded, we would expect that the 4 h time point measurement would land on the straight line. We thereby calculated the distance between the measurement and the straight line. This value is the Δ-score. A low Δ-score, a small distance between the straight line and the measurement, indicates that the protein is exponentially degraded. A large negative value would indicate that the protein is age-dependently destabilized and a large positive value that the protein would be age-dependently stabilized.

We selected the 4 and 8 h time points for the Δ-score calculations because the vast majority of the initial degradation was found to be completed before 4 hours of chase (e.g. see Basigin in Figure 15). In addition, very few protein were degraded so rapidly that they lacked data for the 8 h time point. Finally, these rapidly degraded proteins were almost exclusively exponentially degraded (See supplemental table 1). Therefore, by using the 4 and 8 hour time points we capture the NED of most proteins.

3.3.1.2 NED categorization

Employing the Δ-score and the AIC probability, we defined NED proteins as proteins that had a 2-state AICp higher than 0.8 and Δ-score larger than 0.15. This ensures that NED proteins not only have non-exponential degradation profile but also that they are strongly non-exponentially degraded. Conversely, ED proteins were defined by having a 2-state AIC probability lower than 0.2 and an absolute Δ-score lower than 0.15. Proteins falling outside these definitions we simply refer to as undefined (UN) from here on. It's worth noting that UN is a purely technical definition and most undefined profiles would probably be categorized either as ED or NED if more data would be gathered. In any case, the NED population becomes more robustly defined when taking into account both the probability and size of the effect.

Using these definitions we found 331 proteins (10 %) that show strong NED while 1622 (49 %) display exponential degradation (Figure 22).

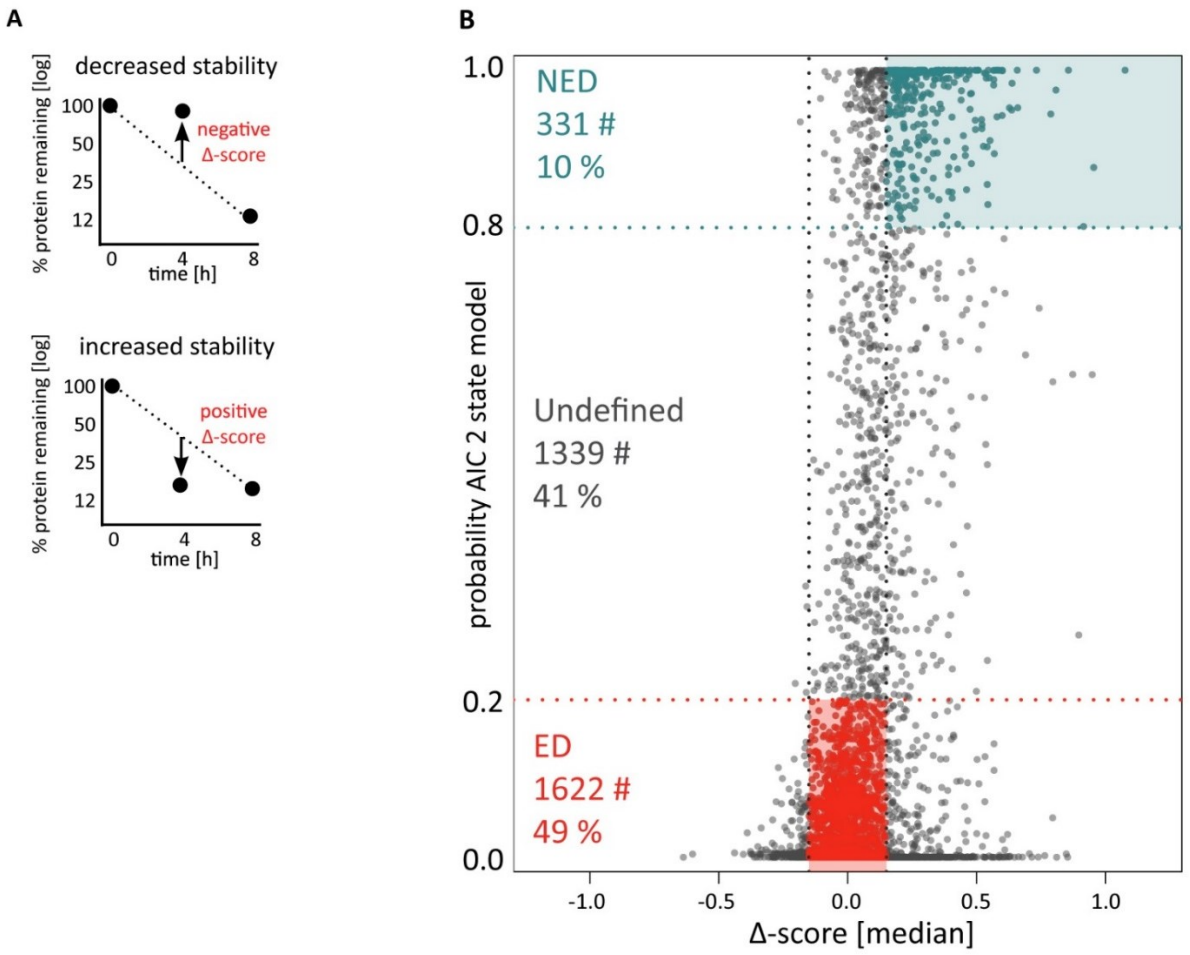

Figure 22. 10 % of all proteins are strongly non-exponentially degraded (NED).

A) Δ-scores are calculated by drawing a straight line between time point 0 h and 8 h. The distance from the measured 4 h time point and the straight line is then calculated. A large distance, or Δ-score, indicates non-exponential degradation since the definition of exponentiality is the fact that the measurements should fall on a straight line in a semi-log plot. B) Using the Δ-score and the 2-state AIC probability, we defined three categories of proteins NED, ED and undefined (UN). NED proteins have a probability for the 2-state model > 0.8 and a Δ-score > 0.15. ED proteins, in contrast, are defined by having a 2-state AIC p < 0.2 and an absolute Δ-score < 0.15. The remaining proteins are considered undefined. This should be considered a technical definition and probably does not reflect biology. Collection of more data points would most likely lead to the proteins within this category to be defined either as NED or ED. In collaboration with Henrik Zauber (MDC), Celine Sin and Angelo Valleriani (MPI).

3.3.1.3 Number of data points required

Another question that arose was the number of data points that are required to make a distinction between the two models. To investigate this we used a bootstrapping approach. We took the data from two profiles, one of an ED and one of a NED protein (Figure 23). We then randomly sampled a number of the data points from each profile. We performed 500 replicates for each number of data points selected, i.e. the boot strappings were performed by grabbing 4 random measurements from the Uba2 profile 500 times and then grabbing 5 time points another 500 times and so forth. We found that when using 5 data points we recalled the vast majority of Basigin profiles as either undefined or NED. By using seven data points we were almost 100 % correct. The ED protein was much less sensitive to the number of data points. This finding reflects the conservative approach we took when identifying NED proteins. In summary, a NED protein with 5 data points measured will be correctly classified many more times than it would be incorrectly recalled as an ED protein. We therefore used profiles with a minimum of five data points for the NED categorization (see also Figure 16).

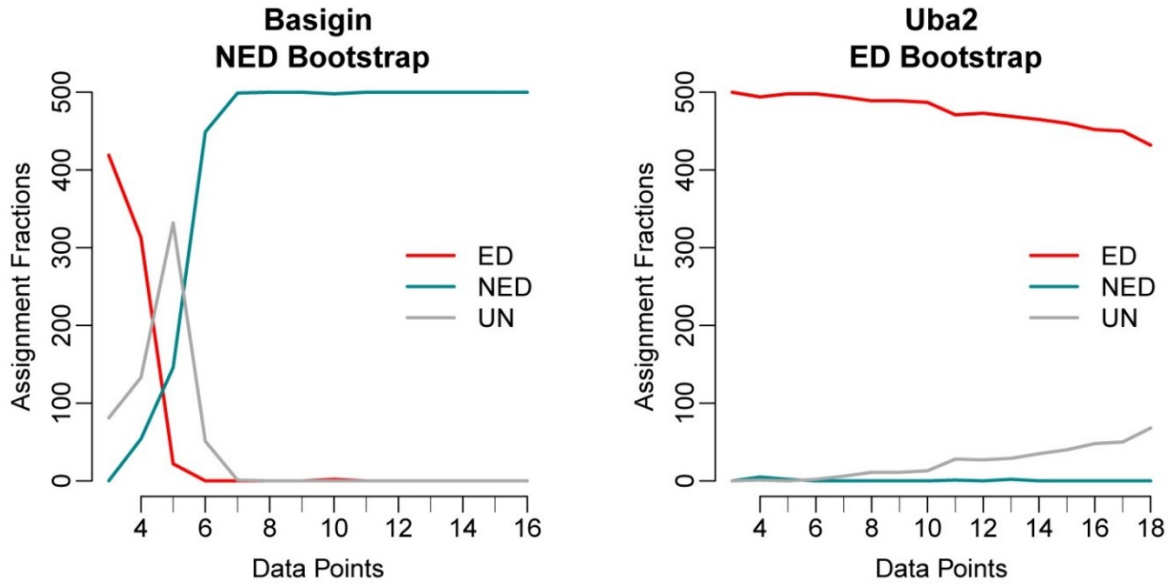

Figure 23. Impact of the number of data points on our classification.

We used the experimental data for a NED protein (Basigin, left) and an ED protein (Uba2, right). From these data, we generated sub-samples of different sizes (that is, "Data Points") using bootstrapping (500 profiles per sample size). We then applied our AIC-based modelling, computed Δ-scores and applied our cut-offs. The number of profiles that were classified as NED, ED and UN is shown as a function of the number of data points. Plot by Henrik Zauber (MDC) and analysis by Celine Sin (MPI).

3.4 Background binding has a minimal impact on NED categorization

3.4.1.1 Measurement of background binding

A potential technical issue could be unspecific binding of proteins to the alkyne-beads. If proteins would display high background binding, this would once again introduce false positive NED profiles by artificially stabilizing rapidly degrading proteins. Imagine a protein that has very high background binding levels. In this case we might not be able to detect degradation at all since the large population of old proteins would mask the signal from the enriched AHA-labeled newly synthesized proteins.

To estimate the amount of background binding I performed a control experiment in which I AHA-labeled a population of fully Heavy SILAC labeled cells (Figure 24). I then mixed these cells 1:1 with unlabeled Light cells. I enriched the AHA containing proteins and performed shot gun proteomics. If the enrichment was specific, we would expect to only find Heavy peptides. In contrast, background binders should show up in the light channel. I detected 4673 proteins in the first replicate. Out of these, 3034 were detected only in the Heavy channel (infinity ratios). 47 showed up only in the Light channel and contained mainly proteins, such as serum albumin, that have an extracellular origin and proteins with very few identified peptides, both groups probably representing contaminants. 1556 proteins showed up in both channels with a median ratio larger than 11. Therefore, the vast majority of protein are below the conservative 10 % cut-off, limiting the potential impact of background binding on NED categorization.

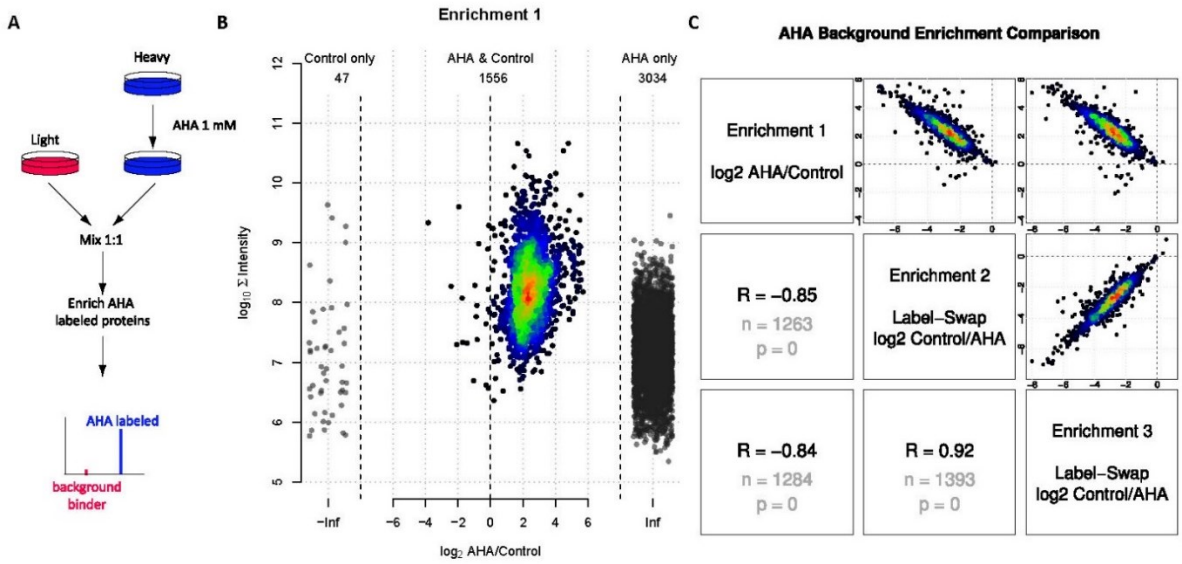

Figure 24. Enrichment specificity of AHA bearing proteins.

The experimental design is shown in A). Fully Heavy SILAC labeled NIH 3T3 cells were pulsed for 2.5 h with AHA. After the pulse the cells were mixed 1:1 with fully Light labeled cells which had not been in contact with AHA. The standard click-reaction and enrichment procedure was performed and the resulting peptides were analysed by LC-MS/MS. By measuring the H/L ratio we can estimate how much of the signal comes from proteins that potentially can contain AHA (Heavy channel) compared to background binders or other contaminants (Light channel). The results from one such experiment are displayed in B). As expected, most proteins (n = 3034) show up only in the Heavy channel. Very few protein (n = 47) show up only in the Light channel. 1556 more abundant proteins, as judged by their high summed intensity, show up in both channels. C) The potential background binding population was found to be highly reproducible with many of the same proteins showing up with a similar ratio in all three experiments analysed. The median ratio for the three experiments was > 11. Analysis in collaboration with Henrik Zauber, MDC.

3.4.1.2 Subtraction of background signal has limited impact on NED categorization

Since we had defined the NED, ED and UN populations, we decided to further test the impact of potential background binding on our categorization. Even though the majority of proteins did not show any background binding in the assay (Figure 24), the proteins that did display background binding showed highly reproducible values (Figure 24 C). For the group of proteins that showed background binding, we used the data from Figure 24 and calculated a protein-specific background binding value (i.e. the % of signal that was potentially derived from background binding). We then simply subtracted this value from the AHA p-c experiment and rescaled the values. With this new data, we re-performed all the AICp and Δ-score calculations. Finally, we compared these results with the previous results using the 10 % cut-off (Figure 25). We found that very few proteins were affected by the protein-specific background subtraction and therefore decided to use the simpler 10 % cut-off strategy.

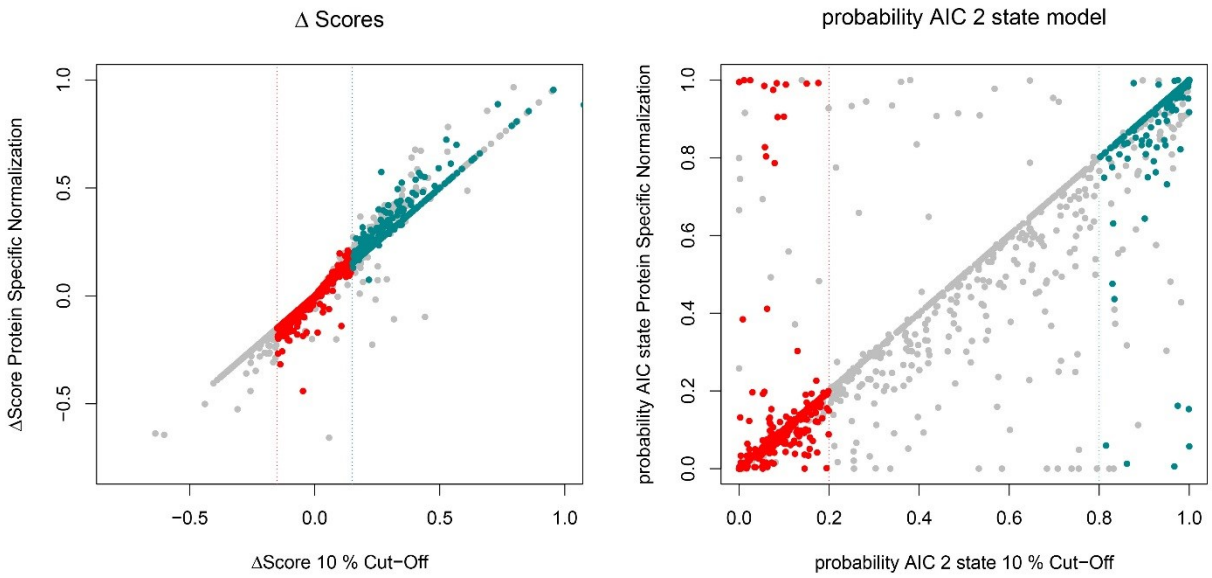

Figure 25. Impact of different strategies to deal with background binding.

The plots show the Δ-scores (left) and AIC probability (right) using either the 10 % cut-off strategy (x-axes) or subtraction of the protein-specific background (y-axes). Protein-specific backgrounds were derived from the enrichment experiment shown in Figure 24 and subtracted from the experimental data. If no ratio was measured in Figure 24, the standard 10 % cut-off was applied also to these proteins. With the new values, the standard modeling and Δ-score calculations were performed. Plot and analysis by Henrik Zauber (MDC), Celine Sin and Angelo Valleriani (MPI).

3.5 Confirmation of NED categorization

When we had defined NED and ED proteins by AHA p-c, we sought to confirm our definitions with alternative methods. I will discuss in this segment the orthogonal methods we used to confirm our categorization and the results gained.

3.5.1 AHA p-c provides half-lives similar to previously published

A good starting point when confirming one's findings is to check that the data matches published literature. The first thing we did was to see if the half-lives we had calculated correlate with a previously published data set using the same cell line (Schwanhäusser et al., 2011). The previously published half-lives were quantified using a dynamic SILAC approach which is very different from the AHA p-c approach. We first compared the half-lives from the exponentially degraded proteins to Schwanhäusser et al. and found a good correlation (r = 0.645, Figure 26). It was also comforting to see that the absolute values of the half-lives reported were similar. It should be noted that half-lives larger than 300 h were excluded since we cannot accurately quantify them in our approach with the longest chase time being 32 hours.

For the NED proteins we applied two different approaches. First, we compared the exponential half-lives from the 1-state model fit (Figure 26 B). Second, we calculated steady state half-lives based on the 2-state model (Figure 26 C). These half-lives are based on the time where half the molecules would be left starting at steady state rather than after the 1 h pulse. Finally, we calculated the difference in half-lives between the two approaches and Schwanhäusser et al (Figure 26 D). We found that, when using an exponential assumption for the NED proteins, the reported half-lives were shorter than the previously published ones. This is consistent with the fact that the linear regression does not take into account that NED proteins stabilize with age. In contrast, the steady state half-lives using the 2-state model resulted in a much smaller deviation from Schwanhäusser et al. In summary, using the 2-state model gives better half-life correlation for the NED proteins.

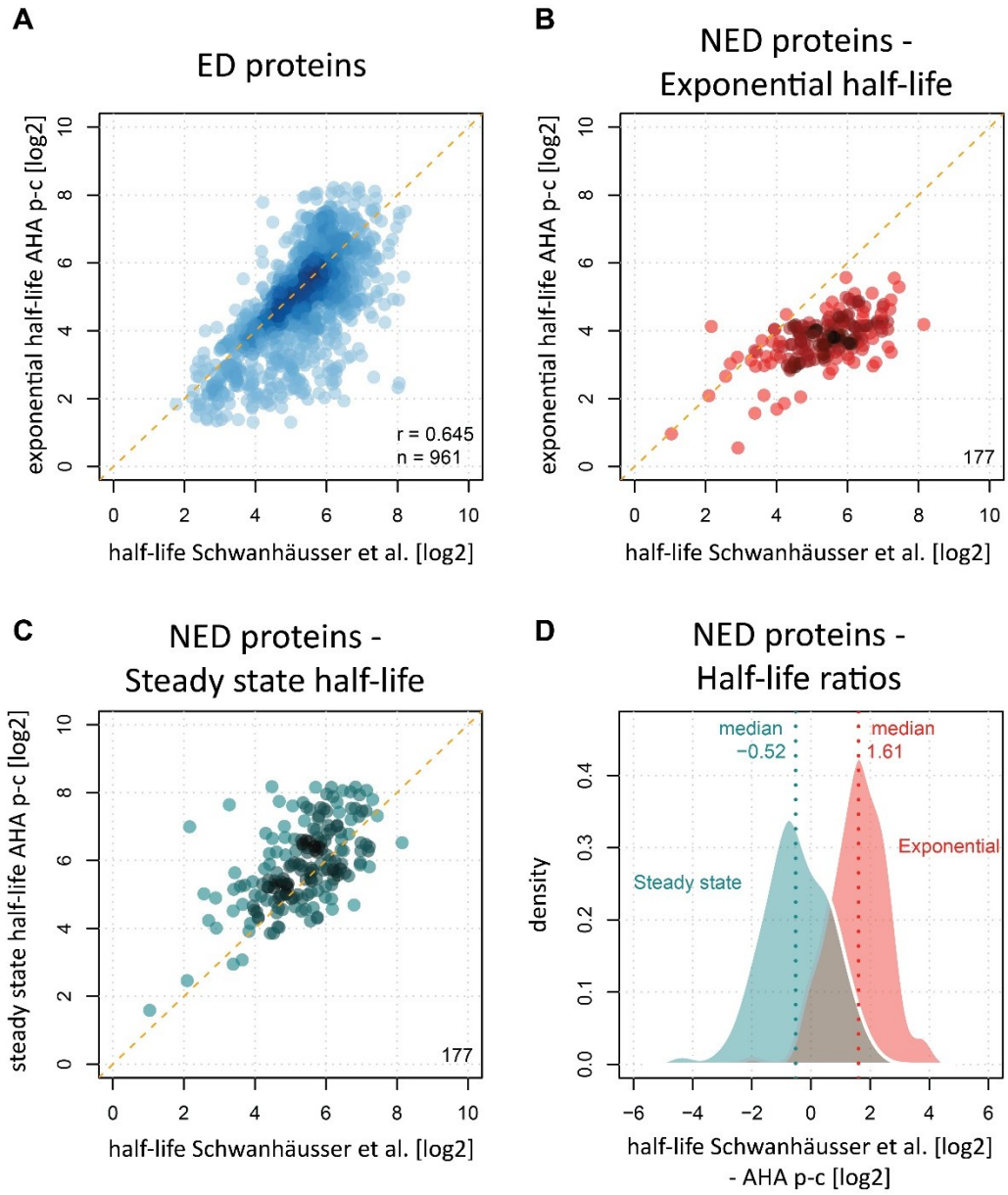

Figure 26. Half-life comparison between AHA p-c and Schwanhäusser et al.

Half-lives derived from AHA p-c were compared to half-lives derived from dynamic SILAC measurements in the same cell line (Schwanhäusser et al., 2011). A) ED proteins have similar half-lives according to both methods. B) Using an exponential fit when calculating half-lives for NED proteins systematically underestimates the half-lives of these proteins. C) Taking the stabilization of NED proteins into account by calculating a steady state half-life (see material and methods) gives more similar half-lives to dynamic SILAC measurements. D) A direct comparison between half-lives derived from an exponential fit and steady state calculations show that steady state half-lives which take the 2-state model into account give more accurate half-lives. In all panels, half-lives > 300 h were excluded because they cannot be accurately calculated with a maximum chase time of 32 h in AHA p-c. Figure in collaboration with Henrik Zauber, MDC.

3.5.2 SILAC p-c confirms AHA p-c categorization

Since comparing half-lives is only an indirect test of the non-exponential decay kinetics, we sought to directly test this. To this means we developed a method, SILAC pulse-chase (SILAC p-c), which does not depend on artificial amino acids and enrichment of newly synthesized proteins. This would limit artifacts caused by background binders and impacts of incorporating an artificial amino acid. In the SILAC p-c I used fully Light labeled mouse fibroblasts and pulsed them for 4 hours in Heavy SILAC DMEM (Figure 27). I then split the cells in two and directly froze one population. The other population I switched to Medium-heavy SILAC amino acids and chased the cells for 8 h before harvesting them. I then quantified the Heavy/Light ratio before and after the chase. I ignored the Medium-heavy population synthesized during the chase. If the new (Heavy) proteins have the same degradation rate as the Light pre-existing proteins the H/L ratio should be the same before and after the Medium-Heavy pulse. Any changes in the ratio of ratios would indicate NED. For example, the ED protein Lmph2 showed no difference in the abundance of new (Heavy) and old (Light) proteins after the pulse (ratio of ratios 1.02, Figure 27 B). In contrast, the NED protein Psmd2 has less new protein remaining after the pulse, indicating that this population is degraded faster than the old population (ratio of ratios 0.77, Figure 27 C). When looking at all the proteins quantified in one experiment by SILAC p-c (Figure 27 D), I saw a significant decrease in the newly synthesized protein specifically in the NED population as judged by the ratio of ratios.

In total, I performed 4 experiments with different chase times and labels swaps (see Figure 48 for the other 3 experiments). To systematically compare between the different experiments we ranked all the proteins in each experiment from the lowest to highest ratio of ratios. We then plotted the normalized rank (Figure 27 E) for all the protein categorized as either ED or NED by AHA p-c. We found that the NED population was clearly shifted to the lowest ratio of ratios (i.e. having the lowest ranks and shifted to the left in the plot). Interestingly, we found a few ED proteins also among the lowest ranked ratio of ratios potentially consisting of incorrectly classified NED proteins. This is to be expected because we are very conservative with our NED categorization and some NED protein could be wrongly classified (Figure 23). From this plot we devised a NED validation score. In total 289 out of the 331 NED proteins acquired a validation score due to different coverage in the different experiments. We considered 200 of the 289 proteins at least partially validated with a validation score of one or larger. Importantly, the long pulse times in SILAC p-c impose a strong limit to the sensitivity of this assay. Therefore, a lack of confirmation by SILAC p-c should not be interpreted as a false positive in the AHA p-c assay but solely as unconfirmed.

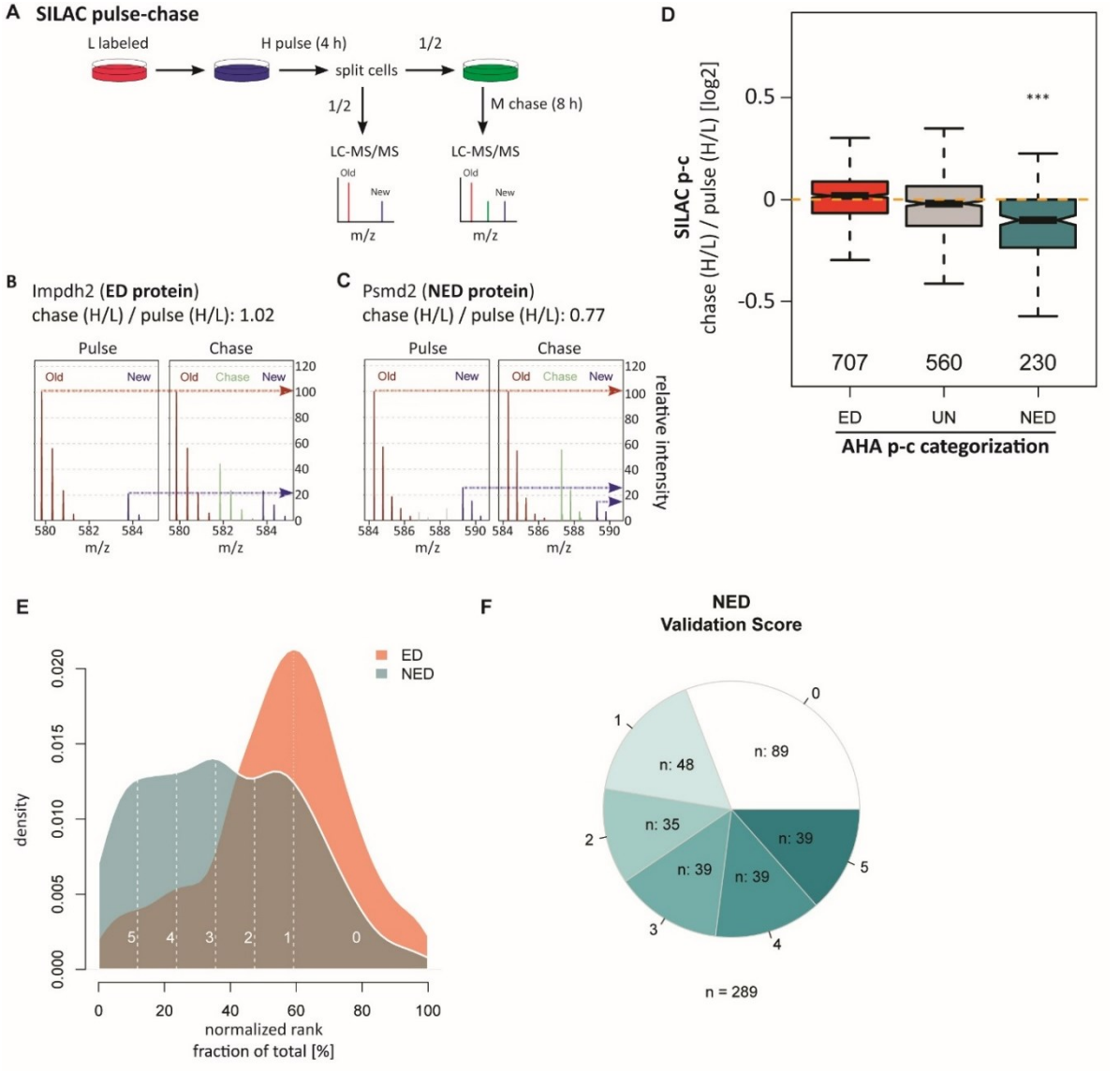

Figure 27. SILAC pulse-chase (p-c) experiments confirm the NED categorization from AHA p-c.

A) SILAC p-c is based on switching Light cells to Heavy SILAC amino acids. After the Heavy-pulse cells are split and one half directly frozen. The other half is chased in Medium-heavy amino acids. By comparing the H/L ratio before and after the Medium-heavy-chase one can determine if the degradation rate for the old (Light) and new (Heavy) proteins are the same. B and C) MS1 spectra from one ED and one NED protein according to AHA p-c before and after the chase. The ED protein Impdh2 has the same H/L ratio before and after the pulse. For the NED protein Psmd2 the new proteins have been degraded faster than the Light old proteins. D) Showing the ratio of ratios for SILAC p-c for the NED, ED and UN protein from AHA p-c. NED proteins have a significantly lower ratio of ratios according to a one-sided Wilcoxon sum-rank test comparing the populations to the total (***: $p < 0.0001$, alpha 0.05). E) Using the normalized rank from the displayed experiment and the three experiments in Figure 48 we assigned a validation score for all NED protein. F) Summary of the distribution of validation scores. Figure in collaboration with Henrik Zauber, MDC.

3.5.3 Radioactive pulse-chase coupled to immunoprecipitation confirms AHA p-c profiles

Finally, we sought to confirm two profiles using the gold standard radioactive pulse-chase coupled to immunoprecipitation assay (RIPA). We pulse labeled mouse fibroblasts for 1 h with radioactive cysteine and methionine and then chased the cells for 0, 4 or 8 hours. From the same cell lysates we immunoprecipitated one ED and one NED protein. We quantified the input using western blot and the radioactive signal using a phosphorimager. VCP, an exponentially degraded protein with a half-life of more than 100 h in AHA p-c, was found to be very stable also using RIPA. In contrast, Cct3, a NED protein according to AHA p-c, showed the expected rapid decay followed by stabilization at 4 h. We conclude that we can confirm some profiles from AHA p-c using the gold standard radio-immunoprecipitation assay.

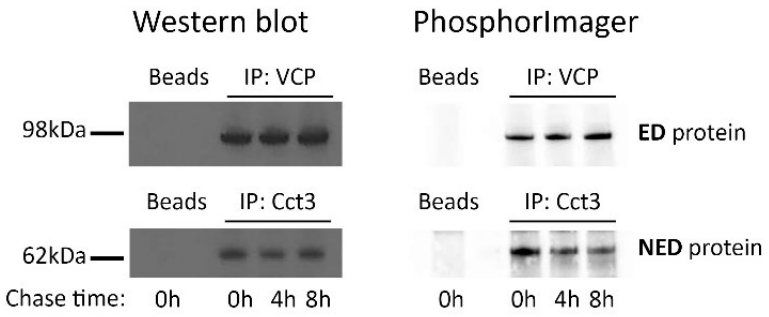

Figure 28. Radioactive pulse chase of VCP and Cct3.

NIH 3T3 cells were pulsed with ^{35}S Cysteine and methionine for 1 h before being washed and then chased for 4 or 8 h in medium containing 10-fold methionine and cysteine. Cell lysates were split in two and incubated with antibodies against the ED protein VCP or the NED protein Cct3. The proteins were subsequently enriched before being separated by SDS-PAGE. Proteins were western blotted and the membrane was probed with the same primary antibodies as for the IP ("Western blot" in figure). The same membrane was also used to expose a magnetic phosphor plate that was then read using a phosphorimager. The stable ED protein according to AHA p-c showed no sign of degradation during the 8 h of chase. The NED protein Cct3, on the other hand, rapidly decayed to be stabilized by the 4 h time point with very little subsequent degradation. Shown is one representative experiment out of two.

3.6 The ubiquitin proteasome system is responsible for some of the NED

Next we sought to uncover the molecular mechanism underlying non-exponential degradation. To this means we employed AHA p-c experiments in the presence of small molecule inhibitors or vector controls (DMSO) (Figure 29 A). This time we only chased the cells for 4 and 8 hours to be able to calculate Δ-scores. We then looked for changes in the Δ-scores after drug treatment compared to the control (Figure 29 A).

Treatment with the proteasome inhibitor MG132 clearly decreased the Δ-score of most NED proteins (Figure 29 B and D). In contrast, blocking autophagy had no significant impact on the NED of proteins (Figure 29 C and E). This would clearly implicate the ubiquitin proteasome system (UPS) as a main degradative pathway responsible for NED. This is consistent with previous findings that newly synthesized proteins are mainly degraded by the proteasome (Duttler et al., 2013; Kim et al., 2011; Wang et al., 2013).

Not all proteins were strongly affected by the proteasome inhibition. We therefore performed a GO-term enrichment analysis for the 20 % of NED proteins (n = 47) which were least impacted by proteasome inhibition (Figure 49). We found a significant enrichment of mitochondrial inner membrane proteins (n = 13, 3.5 fold enrichment) and mitochondrial proteins (n = 20, 2 fold enrichment). This is consistent with the fact that inner mitochondrial proteins are not accessible to the proteasome and are thought to be mainly degraded by proteases residing in mitochondria (Tatsuta and Langer, 2008).

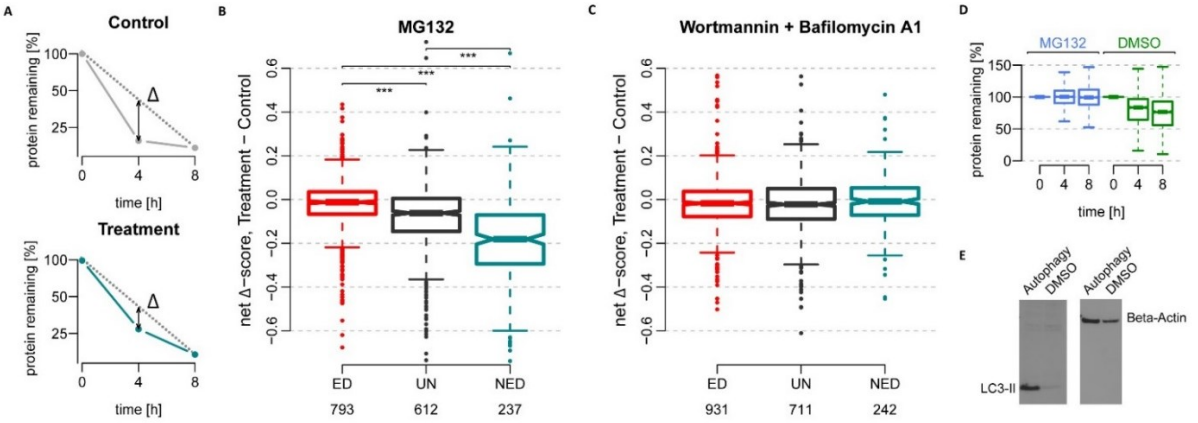

Figure 29. Blocking the proteasome but not autophagy decreases NED.

A) Δ-score were calculated by measuring only the 4 and 8 h time points after treatment with small molecule inhibitors or control (DMSO). By calculating the net Δ-score (i.e. Treatment Δ-score – Control Δ-score) changes in NED can be measured, revealing the molecular mechanism behind NED. B) Inhibition of the proteasome using 20 µM MG132 during the chase had a strong and significant effect by decreasing the amount of NED, as judged by net Δ-score compared to DMSO-treated control cells. The effect was most apparent for the NED proteins. NED, ED and UN are defined from the main AHA p-c experiment. C) Blocking autophagy by the addition of 250 nM of the PI3K inhibitor Wortmannin and 500 nM of the v-ATPase inhibitor Bafilomycin A1 had no significant effect on NED. D) To control for the efficiency of the proteasome inhibitor we monitored the degradation of all proteins as judged by the % protein remaining after the different chase times. MG132 clearly stabilized the vast majority of proteins, in contrast to DMSO. E) Autophagy blockage was monitored by taking a sample of the cells after 8 h chase. The cell lysates were probed by western blotting against β-actin as a loading control and for the autophagy marker LC3-II. The cytosolic protein LC3-I is normally conjugated to phosphatidylethanolamine to form LC-II that is recruited to membranes of autophagosomes (Tanida et al., 2008). LC3-II is then degraded inside the lysosomes. Accumulation of LC3-II is thus a good marker of inhibited autophagy. LC3-II is only accumulating when the inhibitors are applied (“Autophagy” in plot) but not when DMSO alone is added. p-values are derived from a one-sided Wilcoxon sum-rank test. ***: $p < 0.001$. Figure in collaboration with Henrik Zauber, MDC.

3.7 Some NED can be explained by stabilization via complex formation

Since we could confirm the NED of many protein and we had found that the proteasome was mainly responsible for their degradation, we next wanted to know what made a protein NED to begin with. In this section I will discuss our approaches to determine what makes NED proteins different from ED proteins.

3.7.1 NED proteins are abundant, structured and part of multiprotein complexes

The first thing we did was to ask if NED proteins have anything in common. We made receiver operating characteristics (ROC) curves asking how well we could predict a protein being NED or ED dependent on different features. By quantifying the area under the curve (AUC) in ROC-plots one can measure if features are a good predictor of a category. An AUC of 0.5 equal random distribution of the feature (i.e. no predictive power of the feature on the categorization being tested) while an AUC higher than 0.5 indicates a link between the corresponding feature and the applied category. We tested a number of features (some displayed in Figure 30) and found a number of features which positively predict being a NED protein: high protein abundance at steady state, being part of a complex, being structured (high abundance of the secondary structure: helixes and β-sheets) and short (lack of length in this case). ED proteins, on the other hand, were enriched in disordered and low complexity regions.

The fact that NED proteins are more abundant than ED proteins could be caused by a technical bias because more abundant proteins are also measured more frequently. Being covered by more data points correlates positively with being correctly classified as NED (Figure 23). Therefore, some lower abundant NED proteins are probably wrongly classified as undefined. However, we did most downstream analyses also in a binned manner where we binned proteins based on their cellular abundance at steady state and found that the findings below hold true even in this case (not shown). We conclude that NED and ED proteins represent significantly different protein population in the cells.

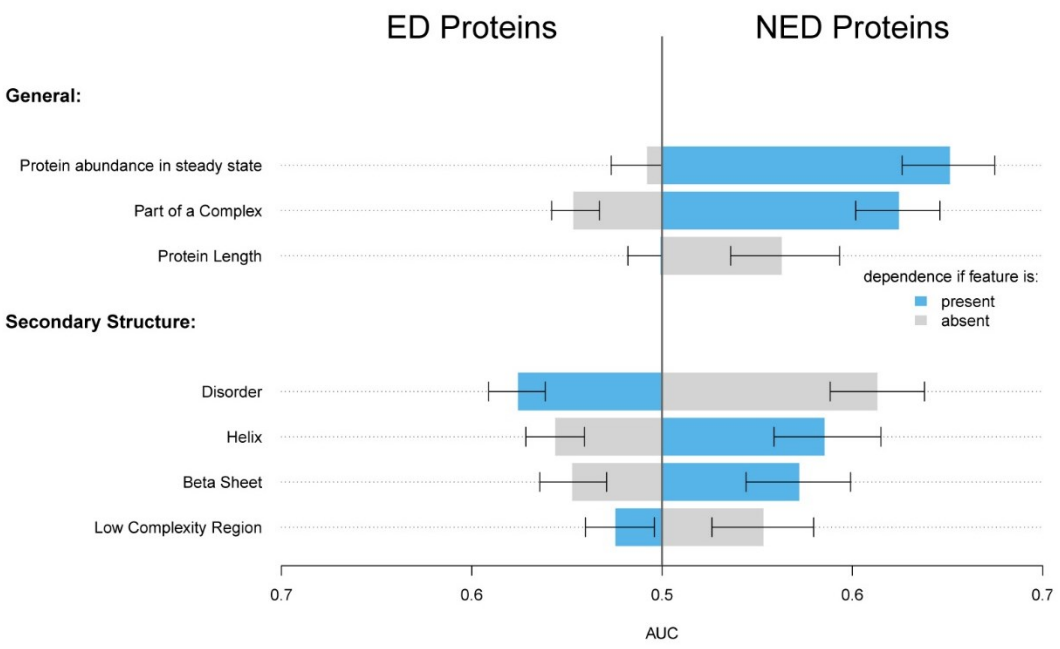

Figure 30. Predicting NED and ED using different features.

The ability of different protein features to predict ED or NED was tested by measuring the area under the curve (AUC) of receiver operating characteristics (ROC) curves. Features can either be present (blue) or absent (grey). For example the secondary structure "disorder" is grey for NED proteins indicating that NED proteins tend to lack disorder (i.e. they're more structured) compared to ED and UN proteins. Error bars indicate 90 % confidence intervals and are derived from repeating 200 bootstraps. Figure and analysis courtesy of Henrik Zauber, MDC.

3.7.2 NED proteins are enriched in multi-protein complexes

To confirm the finding from the ROC curves we performed enrichment analysis of the NED proteins in two protein complex data bases: PDB (Berman et al., 2000) and CORUM (Ruepp et al., 2008) (Figure 31). We found a highly significant enrichment in both cases. It was also found that NED proteins are preferentially enriched in heteromeric complexes with large number of unique subunits (p = 2.2e-16, analysis performed by Joseph Marsh and Jonathan Wells, both at the University of Edinburgh). As a control, the PDB analysis was also performed when excluding the ribosome and still found to be significant (p = 2.5e-5 against ED, figure not shown). This was done since ribosomes are both very abundant and contain exceptionally many subunits and can thereby easily skew the analysis.

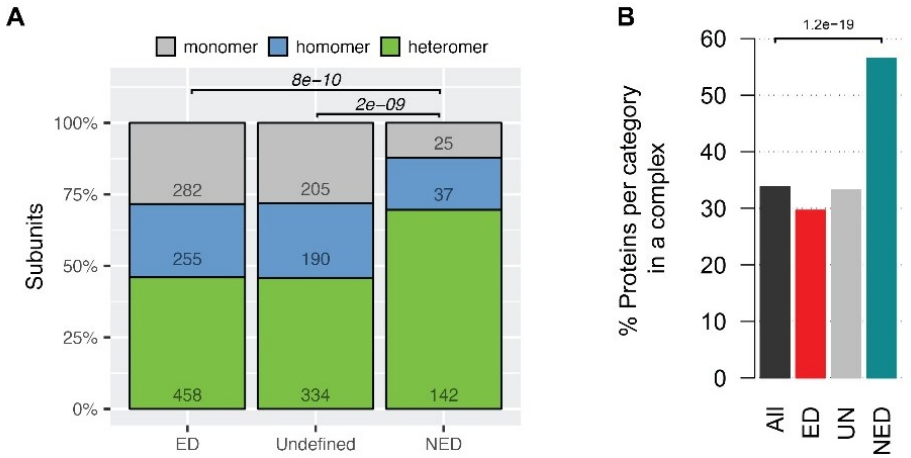

Figure 31. NED proteins are enriched in multiprotein complexes.

NED proteins are enriched in multiprotein complexes both when mapping the categories to A) a structural data base (PDB (Berman et al., 2000)) or B) a manually curated protein complex data base (CORUM, (Ruepp et al., 2008)). p-values are derived from A) Fisher's exact test testing heteromeric compared to monomeric or homomeric subunits and B) hypergeometric test testing enrichments relative to the total population ("All" in figure). Panel A) courtesy of Joseph A. Marsh and Jonathan N. Wells (University of Edinburgh). Panel B) courtesy of Henrik Zauber, MDC.

3.7.3 NED proteins have larger interfaces within complexes and assemble earlier

Even though we found a strong enrichment for NED proteins in multiprotein complexes, these also contain ED proteins. We therefore asked if there are differences between the two populations within complexes. We found that NED proteins tend on average to have a larger interface towards other proteins in a complex (Figure 32). This would indicate that NED proteins are more buried inside a complex than their ED counterparts. Since complexes have a determined evolutionary conserved pathway to assemble, we then looked into whether NED proteins tended to assemble in a specific manner (Marsh et al., 2013). We found that NED proteins tend to assemble earlier than ED proteins in the same complex. These two findings would indicate that NED proteins are “core” members in complexes, making up a “founding” set of proteins on which the ED proteins attach. To further test this hypothesis we looked at mRNA co-expression data and found that NED proteins are strongly co-expressed with other NED proteins whilst ED proteins are only weakly co-expressed with the other members of the same complex (data not shown), supporting the idea that NED proteins are indeed core members of complexes.

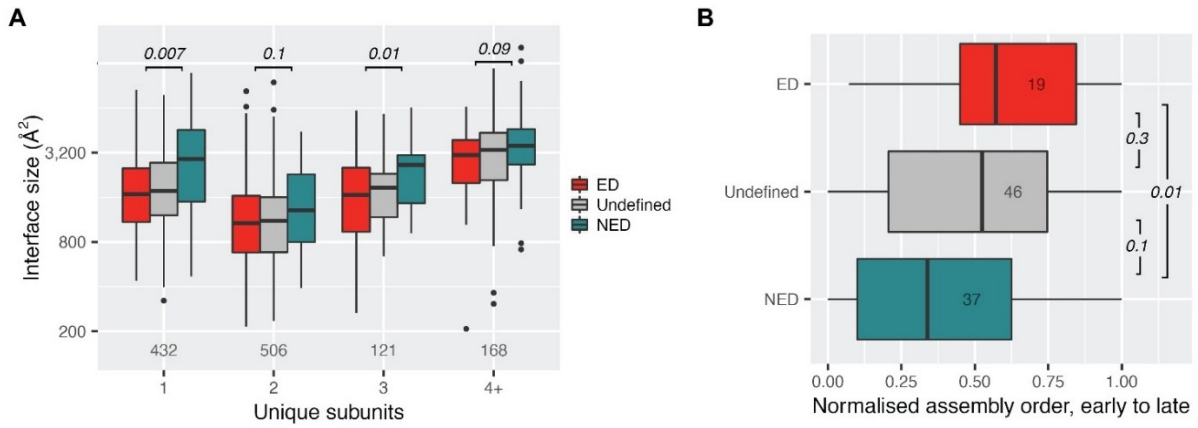

Figure 32. NED proteins have larger interfaces in complexes and assemble earlier.

A) NED proteins have larger interfaces than ED proteins in complexes. The interface size between the protein and other proteins in the same complex was calculated in square Ångström based on the crystal structure for the complex. Complexes are binned for the number of unique subunits to counter the fact that NED proteins are especially enriched in complexes with a large number of subunits. The number of proteins in each bin is annotated at the bottom of the panel. B) Protein complexes are assembled in an ordered manner (Marsh et al., 2013). Proteins were ranked from 0, the first protein in the complex, to 1, the last protein to assemble in the complex. Then, the normalized assembly order was calculated for the three different categories of proteins. Subunit counts are given in each boxplot. All p-values are derived from Wilcoxon rank-sum tests. Analysis courtesy of Joseph A. Marsh and Jonathan N. Wells (University of Edinburgh).

3.7.4 NED proteins are produced in super-stoichiometric amounts

The idea that complex formation could stabilize protein subunits has been around for a long time (Goldberg, 2003; Kuriyama and Omura, 1971). Several individual examples support this hypothesis (Blikstad et al., 1983; Johnson et al., 1998; Lam YW, 2007; Minami et al., 1987; Shemorry et al., 2013; Toyama et al., 2013). Complex formation could explain some cases of non-exponential degradation: if NED proteins were produced in excess relative to ED proteins, the degradation of the non-assembled subunits would lead to non-exponential decay (Figure 33 A). Based on the fact that complexes contain both NED and ED proteins, we hypothesized that NED proteins must be made in super stoichiometric amounts relative to the ED proteins in the same complex. We therefore analysed the abundance of newly synthesized proteins directly after the pulse. We then looked in a complex-centric manner to see if the NED proteins were synthesized in excess (Figure 33 B). We found that this was indeed the case. The prediction also held true (although not statistically significant) when we looked at an independent ribosome profiling dataset (Subtelny et al., 2014) (Figure 33 C). In conclusion, two different methods support the simple model that NED proteins are made in super-stoichiometric amounts relative to ED proteins in the same complex.

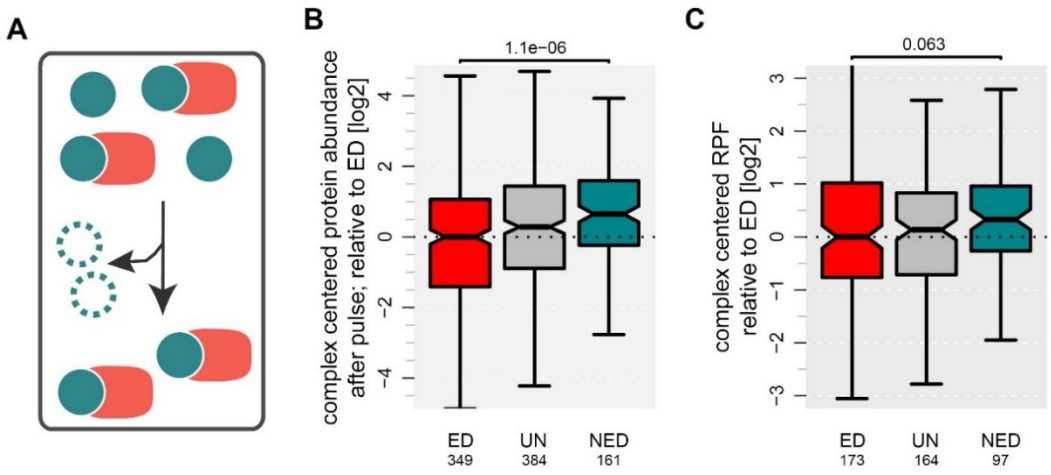

Figure 33. NED proteins are over synthesized compared to other proteins in the same complex.

A) A simple model can explain NED of proteins in complexes. NED proteins (turquoise in figure) are synthesized in super-stoichiometric amounts compared to the ED proteins (red) in the same complex. Excess NED proteins that do not make it into the complex are then degraded. B) Normalized iBAQ values were calculated for all proteins based on the Heavy SILAC label in the AHA p-c experiment (i.e. the abundance after pulse). The relative abundance after pulse was then calculated in a complex-centric manner. NED proteins were found to have a higher relative abundance than ED proteins in the same complex. C) The same as in B but this time ribosome profiling data from the same cell line was used (Subtelny et al., 2014). p-values are based on a one-sided Wilcoxon rank-sum tests (ED vs. NED). Analysis courtesy of Henrik Zauber, MDC.

3.7.5 NED and ED degradation rates support the stabilization by complex formation model

Our hypothesis states that NED proteins that are part of multiprotein complexes are synthesized in super-stoichiometric amounts and that the proteins that do not make it into a complex are degraded. This hypothesis allows another prediction. NED proteins have two degradation rates, k_A and k_B, representing the degradation rates from the two states A and B (Figure 34 A). State A in our hypothesis represents the free (or partially assembled) subunit and state B represents the subunit within a fully assembled complex. The ED proteins, on the other hand, only have one state and one exponential degradation rate, which would represent being in a complex. We can therefore predict that the second state of the NED proteins should have a degradation rate similar to the degradation rate of the only state for the ED proteins since both would be the degradation rate of the complex. To test this we plotted the degradation rate for state A and B for the NED protein and the exponential degradation rate for the ED protein for a number of complexes (Figure 34 B-H). We found that the degradation rate k_B was in all cases more similar to the ED proteins exponential degradation rate than the k_A for the NED proteins. In summary, our finding that the degradation rates for the second state of NED proteins are similar to the exponential degradation rates of the ED proteins in the same complex further strengthen the proposed model.

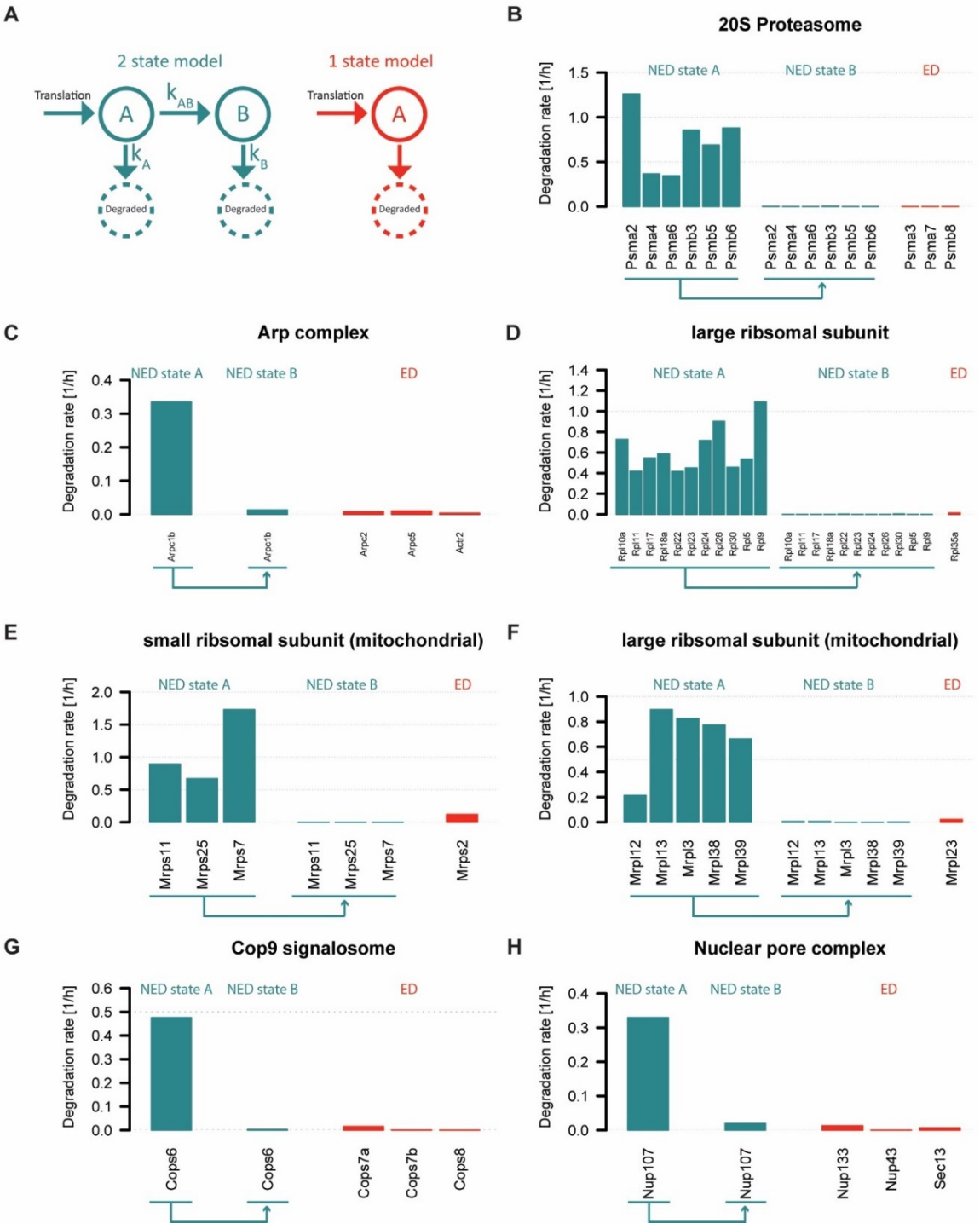

Figure 34. Degradation rates for NED proteins tend to be similar to the degradation rates of ED proteins in the same complex.

A) The two models used to determine degradation profiles. NED proteins are represented by the 2-state model and ED proteins by the 1-state model. B-H) Degradation rates for the NED state A (k_A) and state B (k_B) in addition to the exponential degradation rate of the ED proteins within the annotated complexes are displayed. In the case of NED the same protein is always plotted twice (indicated by the arrow) before and after the transition to the second state. The ED proteins only exist in one state (in our model complexed) and are therefore only plotted once.

3.7.6 Prevention of ribosome assembly increases NED of ribosomal subunits

To more directly test our observation we performed an experiment with the same setup as for the degradation inhibitors. We used the small molecule inhibitor Actinomycin D, which at low doses inhibits primarily ribosomal RNA (rRNA) synthesis by RNA Polymerase 1 (Bensaude, 2011; Jao and Salic, 2008). Blocking synthesis of rRNA subsequently prevents ribosome assembly (Andersen et al., 2005; Girard et al., 1964; Warner, 1966; Warner, 1977). We reasoned that an increased competition for assembly would increase the pool of free unassembled subunits, thereby increasing the measured NED. Our prediction held true and we found increased Δ-scores specifically for ribosomal proteins after treatment (Figure 35), consistent with previous data (Lam YW, 2007; Warner, 1977). It is worth noting that the apparent high stability of ribosomal proteins and other proteins in complexes using dynamic SILAC or cycloheximide chase methods is not in conflict with these findings. Rather, it can be well explained by our model since we find that at steady state the vast majority of proteins are in state B, in other words, in the complexed stable form (Figure 50).

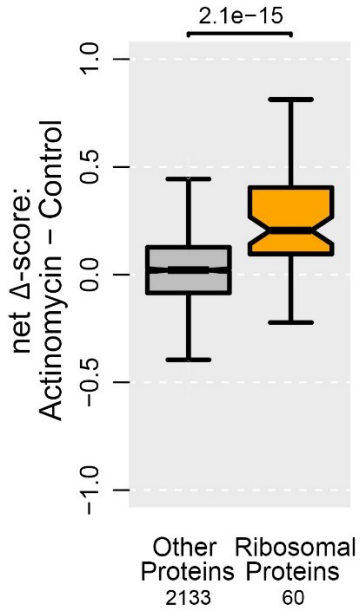

Figure 35. Actinomycin D treatment prevents ribosome assembly and increases NED of ribosomal proteins.
The AHA p-c experiment was performed as in Figure 29. NIH 3T3 cells were treated with either 100 nM Actinomycin D or DMSO (Control) during the pulse and chase. Net Δ-scores were calculated as in Figure 29. Ribosomal proteins have a large increase in Δ-scores indicating that prevention of complex assembly increases the number of proteins in the first state with a higher degradation rate. Statistics are based on a one-sided Wilcoxon rank-sum test. Figure courtesy of Henrik Zauber, MDC.

3.8 NED is an evolutionarily conserved property

One worry with our finding that super stoichiometric production of proteins in complexes cause non-exponential degradation is the model system we used for our initial findings. The NIH 3T3 mouse fibroblasts are mainly tetraploid but have a complex karyotype with frequent further amplification and deletions (Leibiger et al., 2013). In total, less than 2 % of the NIH 3T3 cell genome is disomic. The over-synthesis could therefore be an artifact of our model system in which amplified regions cause increased protein levels, inducing the NED of proteins that are uncoordinatedly produced. We therefore set out to perform AHA p-c experiments in a disomic cell line with limited karyotype issues.

3.8.1 Whole genome sequencing of RPE-1 cells

We turned to the disomic human retinal pigmented epithelial cell line (RPE-1) (Stingele et al., 2012). To confirm that this cell line indeed is disomic we performed low coverage whole genome sequencing (Figure 36). We found that, except for a partial trisomy of chromosome 10 and clonal expansion of a population with trisomy of chromosome 12, the cell line was mainly disomic. We therefore continued our investigation using this cell line.

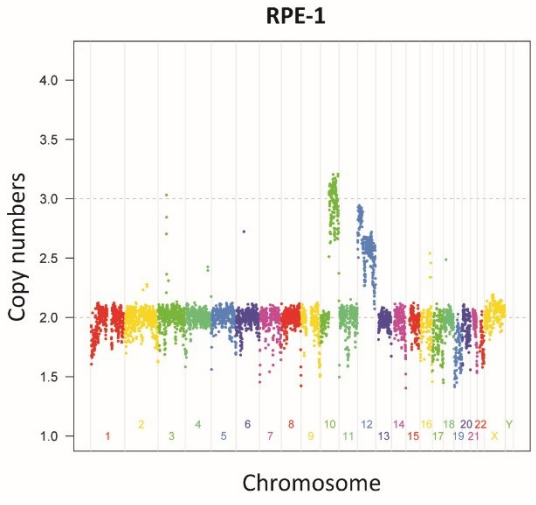

Figure 36. Low coverage whole genome sequencing of the RPE-1 cell line reveals a mainly disomic karyotype. A library of RPE-1 cell genomic DNA was created and sequenced with a depth of about 30 million reads (see material and methods). Displayed are the genomic copy numbers over the chromosomes. Sequencing and figure courtesy of Xi Wang, Jingyi Hou and Wei Chen (MDC). Cell line and sample preparation Neysan Donnelly and Zuzana Storchova (MPI for Biochemistry).

3.8.2 AHA p-c of human RPE-1 cells

After confirming the mainly disomic nature of the RPE-1 cells, we set out to perform AHA p-c experiments. We performed three biological replicates and used the same data processing pipeline as for the mouse fibroblast cells (Figure 37). All in all, the human numbers were similar to the mouse fibroblasts, with 1474 ED (47 %), 1363 UN and 296 NED (8 %) out of the 3133 profiles passing all of our quality criteria. Taken together, AHA p-c experiments of a human disomic cell line show that the number of NED proteins is not mainly dependent on the chromosomal aberration of the mouse fibroblasts cells but rather a consistent phenomenon.

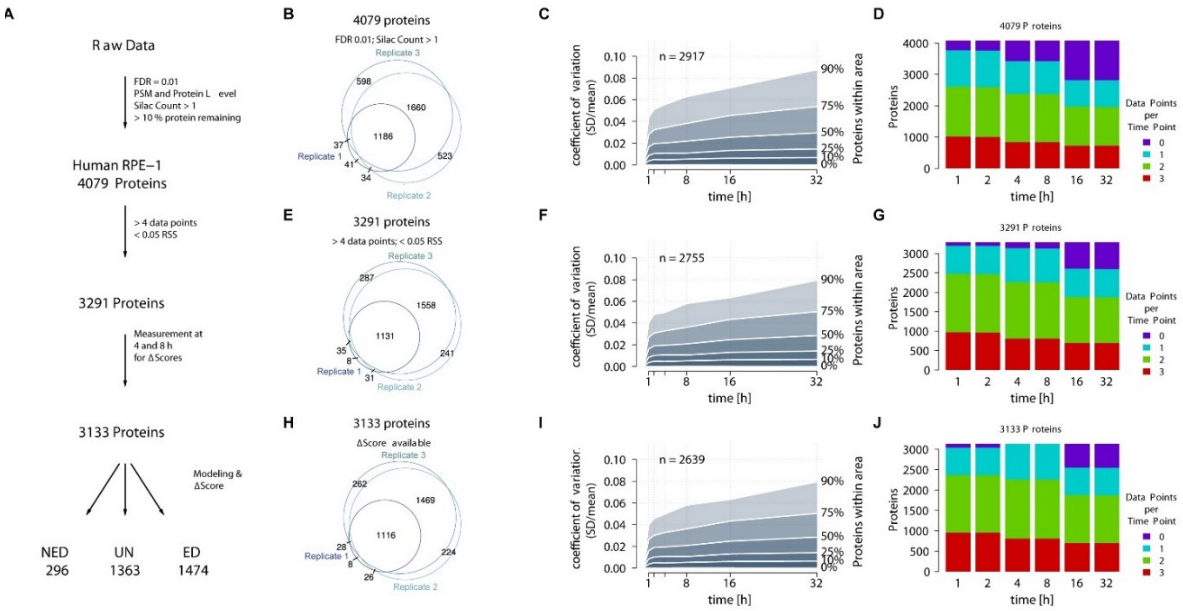

Figure 37. Overview of data processing and quality for AHA p-c experiments using RPE-1 cells.
As in **Figure 16**. Analysis courtesy of Henrik Zauber, MDC.

3.8.3 NED is conserved between mouse and human

Since we now had data on non-exponential degradation from two different species, we realized that we could find out the level of evolutionary conservation of NED. First, we mapped the mouse orthologs to their human counterparts. We then asked if there was an enrichment for human ED proteins among the mouse ED proteins and of human NED proteins among mouse NED proteins (Figure 38 A). In both cases we found a very significant enrichment.

We then turned to the Δ-scores to ask if there was a correlation between the human Δ-scores and the Δ-scores of their mouse orthologs (Figure 38 B). Once again we found a highly significant correlation. From these data we conclude that NED is at least partially evolutionarily conserved between mouse and human and also consistent between different cell types.

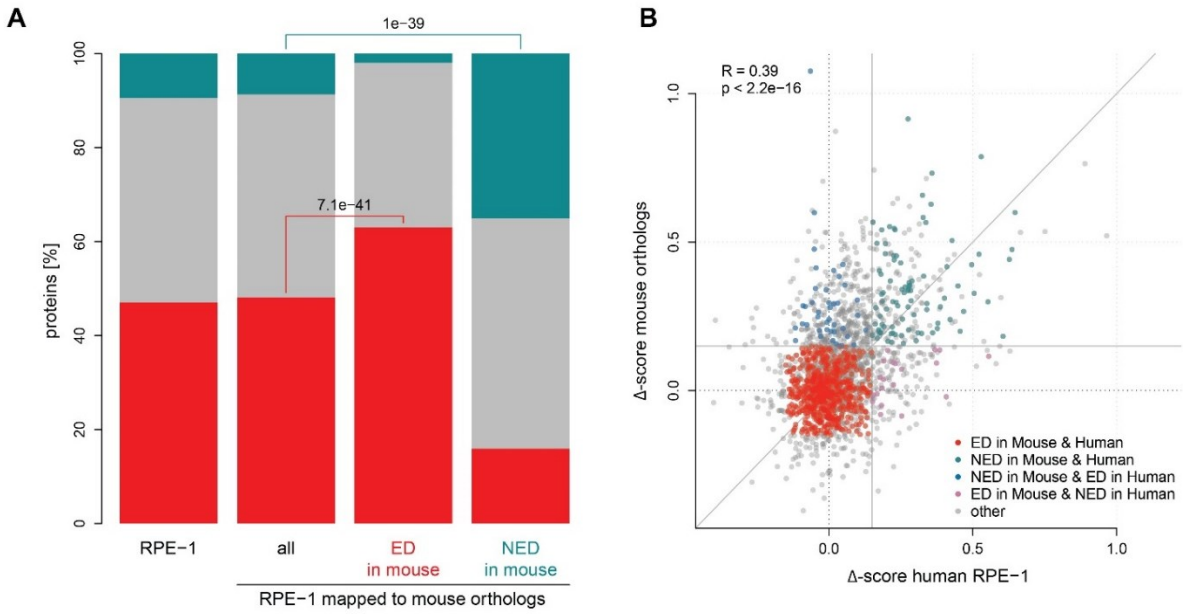

Figure 38. NED is conserved between human and mouse.

Human proteins were mapped to their mouse orthologs. A) Displaying the fraction of proteins being categorized as NED, ED or UN in human. From the left; all the protein as categorized in Figure 37, second the same categories when looking only at the portion of the human proteins mappable to their mouse orthologs. Third, using only the proteins defined as ED from mouse and finally using the NED proteins from mouse. p-values are based on a hypergeometric test. B) Significant correlation between human and mouse ortholog Δ-scores. Proteins were mapped as in A) but this time Δ-scores were plotted. R is Pearson's correlation coefficient. Analysis and figure courtesy of Henrik Zauber, MDC.

3.8.4 Super-stoichiometric synthesis of NED proteins is evolutionarily conserved

Finding an evolutionary conservation of NED between mouse and human cells, we turned back to the hypothesis regarding over-synthesis of NED members of protein complexes. We wanted to see if the over-synthesis also was conserved as this would strongly support the model we put forward. We therefore once again performed the protein-centric analysis as shown in Figure 33. Looking at the protein abundance after pulse in the human cells we again found that NED proteins were synthesized in higher amounts relative to the ED proteins in the same complex (Figure 39 A). This held true also when we used the mouse definitions mapped back to the human orthologs, indicating that the same proteins are over-synthesized in mouse and human. To further test the hypothesis and to exclude any potential mass spectrometry-derived artifacts, we turned to previously published ribosome profiling datasets and mapped the human definitions to them (Figure 39 B). We found an evolutionarily conserved over-synthesis of NED proteins in mouse and zebrafish and the same trend, although not significant, also for worm. From these two analyses we conclude that our model in which NED proteins are over-synthesized relative to the ED proteins of the same complex is evolutionarily conserved.

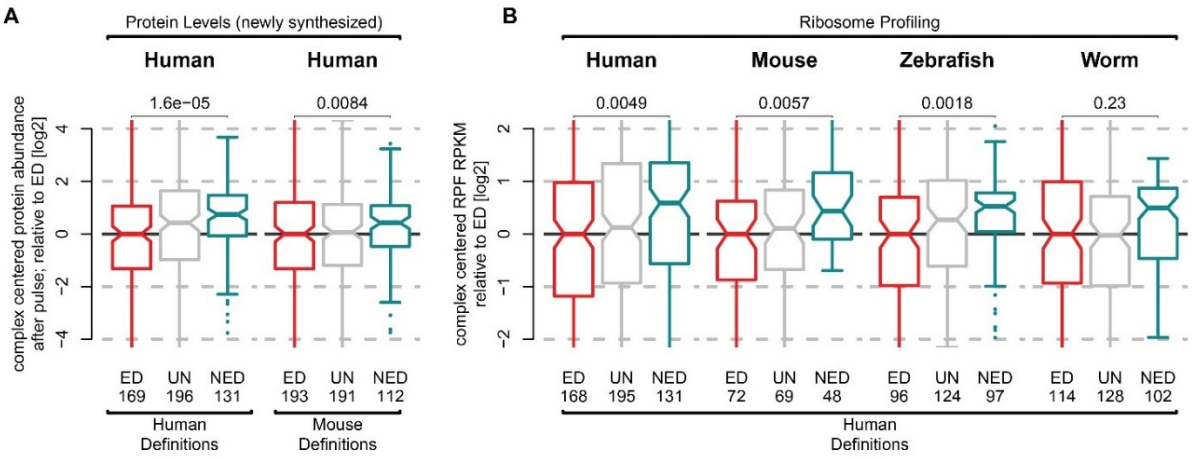

Figure 39. Super stoichiometric synthesis of NED protein is evolutionarily conserved.

A) The analysis of the AHA p-c experiments using human RPE-1 cells was performed as in **Figure 33**. NED proteins are over-synthesized in human cell using either the human or mouse categorization of NED, ED and UN. B) As in A) but using ribosome profiling (RPF) data from human (Liu et al., 2013), mouse (Shalgi et al., 2013), zebrafish (Chew et al., 2013) and worm (Nedialkova and Leidel, 2015). In all cases the NED population has more RPF reads (see material and methods). p-values are derived from a one-sided Wilcoxon sum-rank test. Analysis and figure courtesy of Henrik Zauber, MDC.

3.9 NED predicts attenuation in aneuploidy

Learning that NED proteins are produced in super-stoichiometric amounts we wondered what would happen if one modified the levels of newly synthesized NED proteins. We would predict that increasing synthesis levels of a NED protein would not increase the steady state level of that protein since the excess proteins would be degraded (Figure 40). Importantly, the model we propose would allow one to make predictions regarding steady state protein levels in aneuploidies. Aneuploidy is a frequent cause for birth defects such as mental retardation and spontaneous abortions and is one of the hallmarks of cancer (Hassold et al., 2007; Santaguida and Amon, 2015). Interestingly, for a number of proteins derived from amplified genes in aneuploid cells steady state levels do not increase as much as expected (Dephoure et al., 2014; Torres et al., 2007). This is the case even though transcript levels are elevated for these proteins, indicating posttranscriptional regulation. The concept of protein abundances regressing to diploid levels in aneuploidy cells has been termed attenuation. Attenuation has been noted to be more common for proteins in multi-subunit complexes (Dephoure et al., 2014; Geiger et al., 2010; Stingele et al., 2012).

We would expect that attenuation can be predicted using our acquired degradation profiles. To test if our hypothesis is correct and if we have a better predictive ability than the simpler hypothesis that proteins in multiprotein complexes are attenuated we turned to an aneuploidy cell model. We acquired RPE-1 cells that had an extra copy of chromosome 5 generated by microcell-mediated chromosome transfer, referred to as RPE-1 trisomic cells below (Stingele et al., 2012). This model had the added benefit that we could use the parental RPE-1 cell line as a direct control.

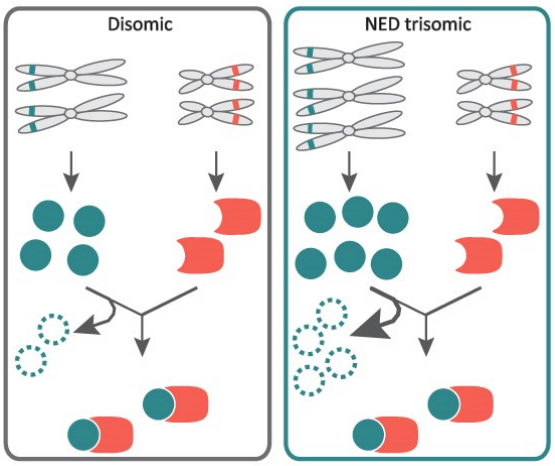

Figure 40. The simple super stoichiometric synthesis model allows for testable predictions.
We predict that increased synthesis of a NED protein caused by aneuploidy would not lead to increased protein steady state levels.

3.9.1 RPE-1 trisomic cells: sequencing and chromosome paints

We first sought to characterize the trisomic cell line and once again performed low coverage whole genome sequencing (Figure 41). We found that, in contrast to the parental cell line, the trisomic cell line had an extra copy of chromosome 5 and also a partial trisomy of chromosome 11. The partial trisomies as well as the trisomy of chromosome 12 were all acquired spontaneously by the cell lines. To confirm the sequencing results we performed chromosome painting (see material and methods) and found that the trisomy of chromosome 5 and a translocation event coupling part of chromosome 11 to an unidentified chromosome were unique to the trisomic cell line (Figure 41). Since chromosome 12 and parts of chromosome 10 were amplified in both the trisomic and parental cell lines, they were ignored in all downstream analysis.

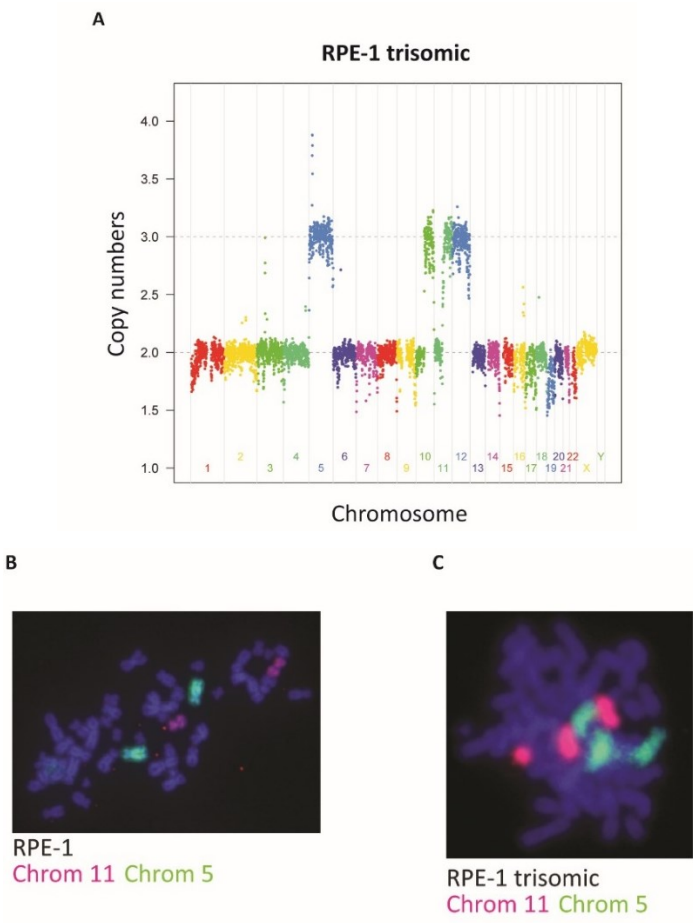

Figure 41. Low coverage whole genome sequencing and chromosome paints of the RPE-1 trisomic cell.

A) Genome sequencing was performed as in Figure 36. B) Chromosome paints reveal a trisomy of chromosome 5 and parts of chromosome 11 that have been amplified and fused to an unknown chromosome in the RPE-1 trisomic cell line. Cell line, sample preparation and chromosome paints by Neysan Donnelly and Zuzana Storchova (MPI for Biochemistry). Sequencing and plot courtesy of Xi Wang, Jingyi Hou and Wei Chen (MDC).

3.9.2 Genomic amplification leads to increased levels of newly synthesized proteins

Next we performed AHA p-c experiments with three biological replicates of the RPE- trisomic cell line. We then calculated the abundance of newly synthesized proteins after the pulse in these cells (Figure 42). We found that proteins encoded on the trisomic regions were significantly more abundant than their disomic counter parts after the pulse. Indicating an increased protein synthesis of protein encoded on trisomic regions. These finding confirm the whole genome sequencing analysis and show that the genomic alteration translates into changes in levels of newly synthesized proteins.

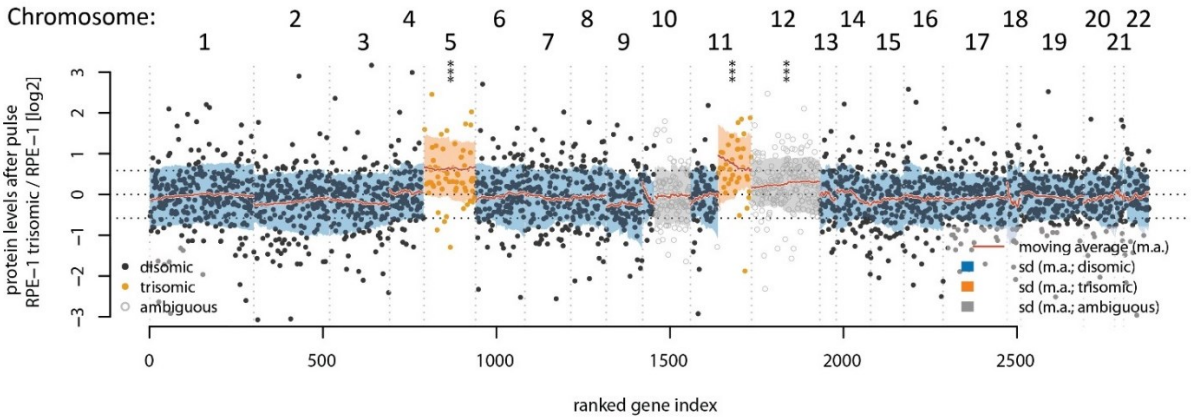

Figure 42. Genomic amplification leads to increase in newly synthesized protein levels.

iBAQ values for proteins after the pulse (Heavy intensities, time point 0 h) were calculated for the RPE-1 and RPE-1 trisomic cell lines. Then log2 fold changes were calculated for each protein. The proteins are ranked according to their chromosomal position. Trisomic regions are highlighted in orange, disomic regions in blue and grey indicates ambiguous regions with amplifications in both the parental and trisomic cells. The moving average is highlighted as an orange line. Regions with significantly different fold changes compared to the parental cells are shown with asterisks. ***: $p < 0.0001$ according to a one-sided Wilcoxon sum-rank test (alpha = 0.05). Analysis and figure courtesy of Henrik Zauber, MDC.

3.9.3 Over-synthesis of NED proteins lead to increased NED

Our model predicted that over-synthesized NED protein would become more NED if the abundance of newly synthesized proteins increased. To test this we calculated Δ-scores for the proteins in the RPE-1 trisomic cell line and compared them to the parental cell line (Figure 43). We found that there was a significant increase in the Δ-score of NED protein encoded on trisomic chromosomes. In contrast, ED proteins showed no increase in non-exponentiality when over-expressed. This would strongly indicate that NED but not ED proteins increase their initial degradation when over-expressed.

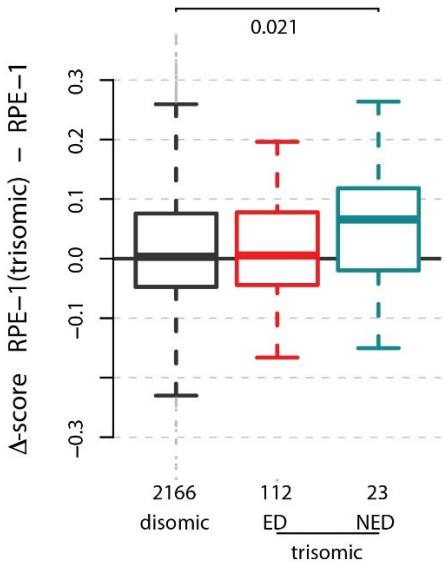

Figure 43. Δ-scores of NED proteins increase after genomic amplification of the respective genes.
Net Δ-scores are displayed comparing NED and ED proteins encoded on trisomic regions (as defined in Figure 42) and disomic proteins. NED proteins show a significant increase in Δ-scores when derived from genomically amplified regions as determined by a one-sided Wilcoxon sum-rank test. In contrast, ED proteins encoded on the same regions show no increase in NED as determined by Δ-scores. Analysis and figure courtesy of Henrik Zauber, MDC.

3.9.4 NED predict attenuation at steady state in trisomic cells

Since protein level attenuation in aneuploidy has previously been mainly explained by complex formation, we wanted to see if NED would be a better predictor of attenuation. We therefore performed a standard SILAC experiment to measure the relative protein levels RPE-1 and RPE-1 trisomic cells at steady state. We then mapped our NED categorization to the quantified proteins (Figure 44). We found that ED proteins encoded on trisomic regions were increased at steady state. NED proteins however, were found at lower levels than ED proteins. This held true also when we compared proteins defined as being in a complex or proteins not in a complex. We conclude that NED is a better predictor of attenuation than belonging to a protein complex in RPE-1 trisomic cells.

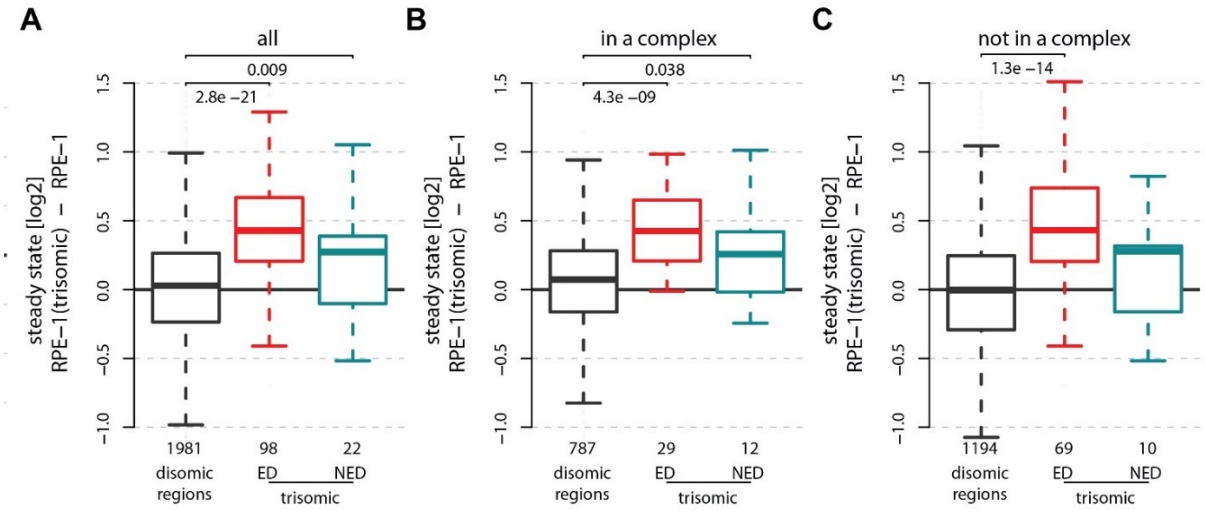

Figure 44. NED proteins are more attenuated than ED proteins in trisomic RPE-1 cells.

Displayed are steady state protein fold changes comparing RPE-1 and RPE-1 trisomic cells. ED proteins are more elevated than NED proteins when looking at A) all proteins, B) proteins in a complex and C) proteins not in the CORUM database. Statistics are from a one-sided Wilcoxon sum-rank test. Analysis and figure courtesy of Henrik Zauber, MDC.

4 Discussion

I will here summarize the main findings and try to put them in perspective. I will take some time to discuss the technical limits of AHA p-c because an understanding of the shortcomings of the method is crucial in interpreting the data. Finally, I will discuss the biological implications of our findings.

4.1 Summary of findings

To evaluate changes in protein degradation kinetics as proteins age we developed a global pule-chase method. We combined SILAC labeling with pulse and chase by the non-canonical amino acid azidohomoalanine. Using LC-MS/MS we then determined the degradation profiles for thousands of proteins. Applying Markov chain-based modeling we determined the propensity of different proteins to show NED. Our major finding is that NED is common with at least 10 % of proteins showing this type of decay pattern (see Figure 45). This means that the common assumption of exponential degradation is not always correct. We further found that NED is mainly executed by ubiquitin proteasome system. In addition, a great deal of NED could potentially be explained by the super-stoichiometric production of protein complex subunits. In our simple model, the excess subunits that do not make it into a complex are then degraded. Furthermore, within the same complex different proteins show NED and ED with the NED proteins tending to be shorter and more structured. Interestingly, the same proteins are on average overproduced in different species, indicating evolutionary conservation. Finally, we found that further increasing the levels of newly synthesized NED proteins did not translate to increased steady state levels of these proteins. Thereby, we can use our NED categorizations to predict attenuation in aneuploidy.

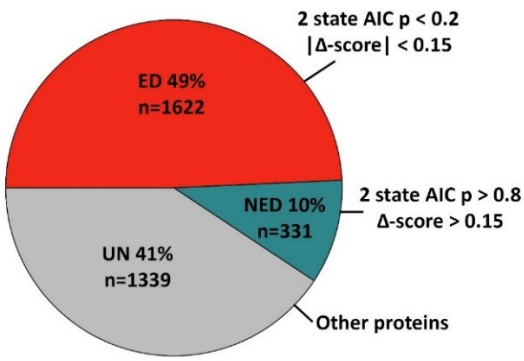

Figure 45. Summary of NED categorization.

Applying both a cut-off of the AIC probability and the Δ-score we defined 10 % of proteins in NIH3T3 mouse fibroblasts as NED.

4.2 Limitations of the method

Even though AHA p-c allows a novel understanding of cell biology, it is important to keep in mind its limitations. I will below discuss the major potential caveats when applying AHA p-c

4.2.1 Azidohomoalanine is an unnatural amino acid

Azidohomoalanine is an unnatural amino acid and the cell distinguishes between it and methionine (Kiick et al., 2002). This is in sharp contrast to stable isotope encoded amino acids which are virtually indistinguishable from standard amino acids for the cells.

One effect of AHA-labeling is decreased protein synthesis (Figure 11) (Bagert et al., 2014). Cells seems to respond to AHA-labeling by an amino acid starvation response e.g. by decreasing biogenesis of ribosomes. The effect is not caused by incorporation of AHA itself because when methionine and AHA are used concurrently, the effects on protein synthesis are gone, even though AHA has been incorporated into proteins (Bagert et al., 2014). Another complicating factor is that titrating in higher concentrations of AHA does not further increase protein synthesis (Figure 11 A). These issues indicate that:

- AHA does not fully compensate for the lack of methionine in the cells
    - Perhaps reflecting a methionine sensor in the cell incapable of recognizing AHA (Efeyan et al., 2012)
    - Or the lack of other functions of methionine in cells such as methionine being the starting material for DNA methylation
- there exist a rate limiting process during the AHA labeling
    - Possibly the loading of methionyl-tRNA (Kiick et al., 2002).

One strategy we deployed to alleviate the effects on proteins synthesis was to limit the pulse time to 1 hour. This minimizes the starvation sensed by the cells to a total of 2 h when including the 1 h methionine depletion before the AHA-pulse. When using a 1 h AHA-pulse we reach roughly 90 % of the expected protein synthesis during the pulse (Figure 13).

Another issue is that we cannot fully exclude some impact of AHA-incorporation on protein stability, even though the radioactive experiments, RIPA experiment, half-life experiment and the current literature exclude any major impact on most proteins surveyed (Cohen et al., 2013; Howden et al.,

2013; tom Dieck et al., 2015). I will in the following segment discuss a special case of potential issues of AHA-labeling on protein degradation kinetics.

4.2.1.1 Exchanging the N-terminal methionine with AHA

That AHA substitutes methionine in proteins comes with many benefits such as the fact that methionine is a frequently used amino acid. However, it also creates its own set of unique challenges. Chiefly, the main difference between methionine and other amino acids is that the first amino acid incorporated into newly synthesized proteins is the initiator methionine (encoded by the start codon AUG). Moreover, the amino acids at the amino terminus of proteins are thought to be an important factor that determines protein stability. Since AHA is incorporated instead of methionine, the potential impact of this artificial amino acid on protein degradation needs to be considered in this context.

The effect of N-terminal amino acids on the half-life of a protein in yeast was reported by the Varshavsky lab (Bachmair et al., 1986), where they created a ubiquitin-β-gal fusion expression vector. Since ubiquitin is commonly expressed as a fusion protein, there exist cellular proteases responsible for the separation of ubiquitin with its co-synthesized neighbor (Baker et al., 1992). Varshavsky took advantage of this and created 20 different ubiquitin-β-gal mutants only differing in the N-terminal amino acid. Ubiquitin was cleaved off *in vivo* and consequently β-gal with different N-terminal residues was produced. Strikingly, the different mutants had very different half-lives, ranging from >20 h for methionine to 2 minutes for arginine. From this finding the lab proposed the “N-end rule”, which states that the N-terminal amino acid regulates the half-life of a protein. The N-end rule was later refined (and complicated) by the Ac/N-end rule (for N-terminally acetylated eukaryotic proteins), Arg/N-end rule (for non-acetylated proteins) and the mitochondrial N-end rule (Shemorry et al., 2013; Varshavsky, 2011).

Most cellular proteins are at least partially N-terminally acetylated in human cells (Arnesen et al., 2009; Harper and Bennett, 2016). Interestingly, the Varshavsky lab demonstrated that Cog1 and Hcn1 were both degraded after an acetylated N-terminal methionine functioned as a degradation signal, also known as a “degron” (Shemorry et al., 2013). However, when the proteins were in their complexed form in the “conserved oligomeric Golgi complex” and “anaphase-promoting complex” respectively, the degron was shielded by the interaction and the proteins were stabilized. Theoretically, this molecular mechanism could explain many cases of NED. However, we still see NED even when it can be assumed that many N-terminal methionines have been replaced by AHA (Bagert et al., 2014). To further elucidate the relevance of the Ac/N-end rule in NED, it would be relevant to find out how many of the N-terminal methionines have been replaced. It would also be interesting to

knock-out Doa10, the E3 ubiquitin ligase responsible for targeting the Ac/N-end substrates for destruction (Shemorry et al., 2013).

It could also be speculated that some NED detected by our method is an artifact caused by the creation of a population of proteins which lack the destabilizing N-terminal methionine acetylation and therefore are stabilized by the N-terminal AHA-incorporation. In this case, two lines of evidence – the SILAC p-c experiment and the fact that we report similar to previously published protein half-lives – point to a limited impact of replacing N-terminal methionines with AHA. This is in line with the fact that the relevance of the Ac/N-end rule has recently been questioned (Boisvert et al., 2012).

Finally, the full removal of the N-terminal methionine is a common co-translational modifications (Varland et al., 2015). Two further issues related to N-terminal AHA-incorporation are instigated by the excision of the initiator methionine:

- First, the excision of the N-terminal amino acid is less effective when methionine has been replaced by AHA (Bradshaw et al., 1998; Wiltschi et al., 2009). This could lead to the creation of two populations one with and one without the second amino acid after the initial amino acid exposed.
- Second, another scenario is when a protein species only incorporate one AHA molecule on average (e.g. very short proteins during suboptimal labeling conditions). Imagine that a protein contains two methionines which of one is the initiator methionine. If half of these proteins contain an AHA only at the initiator position and the other half an internal AHA it would look like half with the initiator-AHA display rapid degradation kinetics due to the quick removal of the said initiator-AHA. This is because when we lose the only AHA we could use to enrich the protein with it appears like it has been degraded in our assay.

However, as discussed above, the fact that we see NED in SILAC p-c experiments suggests a limited impact of such artefacts. In addition, we decreased NED by inhibiting the proteasome with MG132. This compound should not inhibit the methionine aminopeptidases 1 and 2, which are responsible for the N-terminal methionine excision. The strong effect of MG132 limits the possibility that we have wrongly categorized proteins as NED in cases where the initiator-AHA or a short signal peptide is cleaved off. In addition, with a one hour long pulse most co-translation effects are probably too fast for our assay to detect.

In summary, most N-terminal methionine related issues are probably controlled for with the SILAC p-c experiments and the half-life analyses but we cannot fully exclude all potential artefacts caused by these mechanisms.

4.2.2 Background binding to the alkyne-beads

As with all enrichment based methods, there is a risk that background binding could distort the results. This would be especially problematic in our assay because background binding could mimic non-exponential degradation profiles for rapidly degraded proteins. This is because when we have consistent background binding, consider for example a scenario where 20 % of the signal is background, a protein cannot go below this value. If a protein would be fully degraded the signal would only drop to 20 % (background level) and then remain at this level which we would interpret as a biologically relevant stabilization of this protein. This potential "artificial stabilization" of proteins is of concern and below I will discuss our strategies for controlling for such events.

Our control experiments indicate that we have a surprisingly consistent background binding (Figure 24 C). Interestingly, we found that background binders were enriched for more hydrophobic amino acids and contained more reactive amino acids like cysteine and serine (data not shown). A potential model for background binding is that more hydrophobic proteins interact more with the beads and that this allows for improper covalent linking of reactive amino acids to the alkyne moiety. In any case, the consistent background binding allowed us to control for the impact of background binding. We found that after subtracting background levels from the control experiment from the proteins in the AHA p-c experiment we still categorized the NED protein as NED proteins (Figure 25).

Our main approach to handling background binding (and SILAC ratio compression) was to exclude data points in the "danger zone" below 10 %. This is because the background binding control experiments indicated that the vast majority of proteins do not show background binding above this limit (Figure 24). It is also worth noting that most NED proteins stabilize much above this limit (median 47 and 41 % protein remaining at the 4 and 8 h time point respectively) and that very few proteins display background binding close to these numbers.

In addition, if background binding is causing artificial stabilization of NED proteins, the later time points would not reflect biology but rather only reflect background binding. Therefore, we calculated half-lives assuming exponential degradation using data derived only from the first 4 h for NED proteins (not shown). This would model a scenario where the protein is rapidly degraded followed by artificial stabilization. Half-lives computed from only the early time points correlated very poorly and were much shorter than previously published half-lives. This observation indicates that proteins do indeed stabilize at later time points. Thus, the stabilization of NED proteins in the AHA p-c data does not seem to be due to background binding.

All in all, the fact that most NED proteins could be confirmed by SILAC p-c a method not dependent on enrichment – limits the impact of the enrichment process on creating false positives (Figure 27). Nevertheless, not all NED proteins could be confirmed by SILAC p-c because either they were not covered or did not show NED in the SILAC p-c experiments. The SILAC p-c experiments are based on very long pulses and thereby suffer from a lower sensitivity than AHA p-c. Therefore, a lack of confirmation by SILAC p-c does not mean a false positive NED protein but just that it is unconfirmed. Further research is needed to uncover the false positive rate of AHA p-c in categorizing NED proteins.

4.2.3 Assumption related to normalization procedure

To alleviate issues regarding different labeling efficacy, starting material and pipetting errors, we needed to normalize the data. All normalization procedures depend on assumptions. In our case, the assumption is that certain proteins are fully stable. Although this assumption is probably not fully realistic, there are many literature examples of proteins with extremely long half-lives e.g. (Toyama et al., 2013). We found that using between 30 – 200 LSD proteins gave very consistent normalization factors, indicating that many proteins in our assay are very stable. In fact, our LSD normalization proteins contain many known stable proteins commonly used for input normalization, such as β-actin and GAPDH. The fact that the LSD-protein based normalization yields half-lives similar to those in previous publications is also comforting (Figure 26).

Importantly, the degradation profiles *per se* are quite independent of the normalization factors, i.e. changing the number of LSD proteins increased or decreased the average reported half-life but had a very limited impact on the NED/ED categorization (data not shown). This is because the degradation profile of a protein is relative to the degradation profile of all other proteins in the same experiment. Shifting all data points by a multiplicative normalization factor does not shift the relative position of the data points.

4.2.4 Compromises regarding pulse and chase length

Perhaps the most important limitation of AHA p-c is the pulse-length. The pulse length is a compromise between the:

- Depth of the analysis
    - Shorter pulses decrease the number of identified proteins drastically
- Amount of background binding
    - The ratio of labeled to unlabeled proteins increases with pulse length
- Kinetics we can resolve

- Shorter pulses allow a more detailed resolution of degradation kinetics. For example, most co-translational degradation can be assumed to be undetected in our assay since after the 1 h long pulse most co-translated proteins are already degraded and we would not be able to detect a drop in protein levels (Duttler et al., 2013).

Conversely, the chase length also limits the analysis. We limited the longest chase time point to 32 hours. This was done to avoid the cells reaching full confluency during the experiments, which can be expected to change the degradation kinetics, since the cells go from exponential growth mode to stationary phase. Therefore, we cannot say anything about kinetic changes taking place after 32 h of chase. For example, we are not able to identify proteins that become destabilized after that time (Young and Bok, 1969). It is worth noting that most of the initial degradation from state A has taken place at around 1-4 h, indicating that chasing much longer than 8 h to capture protein stabilization is probably not necessary for the majority of NED proteins. Therefore, we also limited the analysis to 4 and 8 h time points in the inhibitor-treated cells.

4.2.5 Limitations related to the modeling approach

The more time points that are sampled, the more advanced models can be applied without risking over fitting the data (Burnham K.P., 2002). For example, studies looking at mRNA degradation commonly use more sampled time points and more complicated multi step models (Deneke et al., 2013; Sin et al., 2016). In this study, we limited the number of time points to 7 because of the triple SILAC approach, in which we needed to add one experiment and more machine time with each two time points added (time point 0 h is always present), quickly increasing the amount of work needed. It is interesting that these two simple models could fit the data quite well. In fact, very few proteins were excluded by the lack of fit to either model (not shown).

Finally, regarding the model, we interpret the model outcome as an ageing of individual proteins. However, the two state model applied does not distinguish between that scenario and a scenario in which proteins are “born” into two different states and that these two states have different degradation rates (Aalen, 1994). One such scenario could be that proteins are either mistranslated or correctly translated, where the mistranslated proteins would have a higher degradation rate.

4.2.6 Abundance bias

The more data points we collected, the more certain we can be when categorizing a NED protein. This means that more abundant and thereby better covered proteins are more likely to correctly be categorized as NED. Until we have full coverage of the proteome with AHA p-c this bias will to some

extent, remain. To minimize the impact of the abundance bias we performed most analyses in a protein complex-centric manner. Since most complex subunits are expressed at similar levels, this bias is alleviated.

4.3 Biological implications of NED degradation

Taking the caveats discussed above into account we can turn to the biological function of NED. I will here discuss the historic reasons for the general assumption of exponential degradation still remaining to date. I will then turn to the reasons for proteins to be degraded by non-exponential kinetics. Finally, I will discuss the implications of this in the context of aneuploidy and in cancer.

4.3.1 Where does the assumption of exponential degradation come from?

There are a number of reasons why exponential degradation became the way to describe protein degradation. Probably the main reason is the fact that it is a simple and convenient term to discuss turnover. However, it should be noted that at the time when exponential degradation became the paradigm the possibility to detect NED was limited. For example, all the proteins referred to in Robert Schimke's seminal 1970 review paper "Control of enzyme levels in animals", with the exception of ferritin, were enzymes (Schimke and Doyle, 1970). From this publication, which is probably the best summary of known half-lives at the time, it is clear that:

1. There are only a handful of enzymes with a described half-life.
2. The vast majority of listed proteins are active as monomers or homomeric complexes. In other words, they are not members of heteromeric complexes, which are the main component of the NED population. Consistently, the proteins with one exception are ED or undefined also in this study. Interestingly, the only exception, ferritin, was actually shown to display NED in the presence of iron (see Figure 46 (Drysdale and Munro, 1966)).
3. Most kinetic analysis methods (e.g. Figure 47) were not suitable for detection of NED. Even though many studies used gold standard radioactive pulse-chase methodology the use of whole animals, imprecise enrichment techniques and suboptimal time points severely limited the ability to find anything other than ED.

Most of the above points can be explained by a detection bias. When the first enzyme, Urease, was crystalized and showed to be a protein, it opened up a whole new line of research (Sumner, 1926). Mainly, it revolutionized the enrichment of proteins. Even though enrichment still required strenuous

biochemical separation methods, it was now possible to estimate the enrichment factor by testing for enzyme activity (e.g. (Archibald, 1945; Price et al., 1962; Schimke, 1964)).

This, however, led to that only enzymes with known chemical reactions could be monitored. Even after the advent of immunological enrichment methods the detection bias remained because the only way to create the antibodies to begin with was to immunize animals with enriched protein e.g. known enzymes. The low numbers of measured half-lives and the monomeric/homomeric nature of the measured proteins are both derived from the requirement of known biochemistry for each enzyme and to the limitation in enrichment techniques. When these requirements were relaxed the variation in degradation profiles emerged.

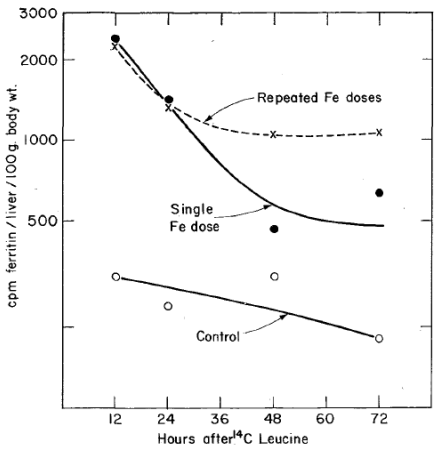

Figure 46. Turnover of rat liver ferritin after injection of Leucine ^{14}C from (Drysdale and Munro, 1966).
Baseline Ferritin (Control) or higher levels of ferritin were induced by either single or repeated doses of 400 µg iron. After animals were sacrificed at indicated times, the specific radioactivity of ferritin (enriched by chromatographic isolation) was measured. The control ferritin showed exponential degradation, whilst the addition of iron stabilized a population of ferritin, leading to a non-exponential behaviour. The authors conclude "The injected iron might act at the time of protein synthesis, prevent breakdown of nascent ferritin by stabilization of an unstable precursor". Reproduced with permission from publisher.

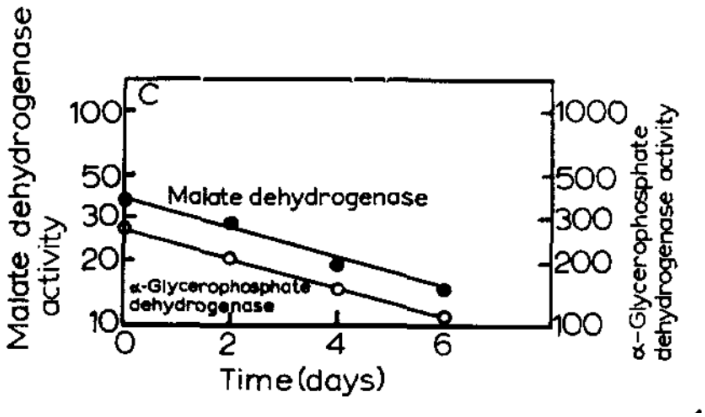

Figure 47. A common way to estimate protein turnover was by enzyme kinetics.

Example from (Tarentino et al., 1966), where rats where treated with thyroid hormone (T4) for 10 days to induce maximum levels of malate dehydrogenase and α-glycerophosphate dehydrogenase. At day 0 the animals are taken off thyroid hormone and are instead treated with ethionine, which prevents induction of the enzymes. Rats where sacrificed at indicated days and liver homogenates where prepared and enzyme activity was measured. Although a clear exponential fit can be seen, it's quite obvious that this method is not optimal for detecting non-exponential degradation because the "chase" starts with the whole steady state population of the proteins and not a younger age-defined population. In other words, the "pulse" is much too long to detect age-dependent changes in degradation kinetics. Reproduced with permission from publisher.

4.3.2 Why are proteins non-exponentially degradaded?

As mentioned in the introduction, few individual proteins were known to show non-exponential degradation kinetics before this thesis. The two proteins mentioned, Basigin and CFTR, are both substrates of ER-associated protein degradation (ERAD) (Tyler et al., 2012; Ward et al., 1995). However, these proteins are not representative for the majority of NED proteins reported here because improper folding seems to be the mechanism behind their initial destabilization and not lack complex formation. I will thereby focus the following discussion on NED of excess protein complex subunits.

4.3.2.1 NED by proteins in complexes

We found that one major mechanism behind NED is stabilization of free subunits by complex formation. This mode of protein stabilization was suggested already in the 1970s (Goldberg, 2003; Kuriyama and Omura, 1971). Examples of proteins stabilized by interactions are common e.g.: α-spectrin (Blikstad et al., 1983), Matα2 (Johnson et al., 1998), survival of motor neuron (SMN) proteins (Burnett et al., 2009), RI-alpha (Amieux et al., 1997), N-methyl-D-aspartate receptor subunit 1 (Huh and Wenthold, 1999), Connexin 56 (Berthoud et al., 2000), tubulins (Katz et al., 1990), keratins (Kulesh et al., 1989), hemoglobin alpha (Shaeffer, 1988), Map2 (Okabe and Hirokawa, 1989), ribosomal proteins (elBaradi et al., 1986; Lam YW, 2007; Maicas et al., 1988), histones (Quivy and Almouzni, 2003) and T cell antigen receptor subunits (Minami et al., 1987). Equally common are examples of proteins destabilized by the removal of its interaction partners, e.g.: PP2A subunits (Silverstein et al., 2002), Oligosaccharyl transferase subunits (Mueller et al., 2015), nicalin-nodal modulator membrane complex subunits (Dettmer et al., 2010; Haffner et al., 2007), Fatty acid synthase subunit (Scazzari et al., 2015; Schuller et al., 1992), five friends of methylated Chtop complex subunits (Fanis et al., 2012), Protein subunits of the electron transportation chain (Pearce and Sherman, 1995; Stiburek et al., 2012), APC/C complex subunits (Clark and Spector, 2015) and presenilin (Edbauer et al., 2002). Interestingly, most examples above come from either over expression or knock down experiments. Very few studies look at base line super-stoichiometric synthesis of complex subunits. The few exception include the T-cell receptor, the ribosome and histone 3 (Gunjan and Verreault, 2003; Lam YW, 2007; Minami et al., 1987; Quivy and Almouzni, 2003).

This work is to our knowledge the first study to produce a systematic overview of steady state over production of complex subunits and the subsequent degradation of the non-complexed proteins.

4.3.2.2 Mechanisms of degradation of excess complex subunits

It seems clear from our data and from previously published examples that the cell needs to be able to distinguish between complexed and free subunits and to degrade the latter with faster kinetics. The previously discussed N-end rule is one proposed mechanism in which a degradation signal (or "degron") is hidden by interaction (Shemorry et al., 2013). Recently, a conserved mechanism for targeting of excess ribosomal proteins was uncovered (Sung et al., 2016). In here the yeast E3 ubiquitin ligase Tom1 (Huwe1 in humans) recognizes stretches of positively charged amino acids in ribosomal proteins that are covered in the holoenzyme. Similarly, uncomplexed histone 3 is recognized by Rad53 which further directs the protein for degradation via the ubiquitin proteasome system (Gunjan and Verreault, 2003). Yet another ubiquitin ligase recognizes Fas2 when its interaction partner Fas1 in the fatty acid synthase is missing (Scazzari et al., 2015). In addition, multiple different chaperones work on unassembled proteins (Harper and Bennett, 2016).

Another interesting example is the mitochondrial degradation pathways. This is especially true since we find many proteins belonging to mitochondrial inner membrane, such as members of the electron transportation chain among the NED proteins. Many of these proteins we found to be insensitive to proteasome inhibition (Supplemental table 2 and Figure 49). Proteins in the mitochondrial matrix are degraded by two different proteases, ClpXP and Lon, both highly homologous to their bacterial counterparts (Bross et al., 1995; Kang et al., 2005; Suzuki et al., 1994). In the inner mitochondrial membrane, the location of the electron transportation chain, no less than three independent proteases are located. The hetero oligomeric matrix directed m-AAA protease (m-AAA), which consists of paraplegin and *ATPase family gene 3-like 2* (AFG3L2) and the intermembrane space pointed, i-AAA protease, is a homo-oligomeric complex composed of *Yeast Mitochondrial escape protein 1-Like 1* (YME1L1) (Baker and Haynes, 2011). Both the i-AAA and m-AAA are believed to function in protein quality control by recognizing and degrading unfolded proteins or free subunits of the electron transportation chain (Arlt et al., 1996; Korbel et al., 2004; Pajic et al., 1994; Stiburek et al., 2012). In the inter-membrane space the homotrimeric serine protease *high temperature requiredA2* (HtrA2) resides and degrades a number of substrates (Faccio et al., 2000; Johnson and Kaplitt, 2009). Finally, proteins in the outer membrane are thought to be extracted and degraded by the ubiquitin proteasome system in a manner partially analogous to ERAD (Baker and Haynes, 2011).

In conclusion, many different mechanisms are involved in handling excess protein subunits, although the proteasome seems to play the dominant role in the cytosol and nucleus. Future work will be needed to untangle all the molecular pathways involved in NED in general and in NED of proteins in complexes in particular.

4.3.2.3 Why are excess subunits produced only to be degraded?

Why do we find so much NED in proteins involved in multiprotein complexes? There are multiple and not mutually exclusive answers to this question. The first thing to realize is that most complexes contain both ED and NED proteins.

1. The first possible explanation for the excess protein synthesis is that cells are just not very good at coordinating the amount of protein they need. However, orthologs tend to remain ED or NED in different species (Figure 38) indicating that the process of becoming a NED protein is not fully random. In a random scenario we would expect the proteins that are NED to change over an evolutionary time frame.
2. One possible explanation is that NED proteins accumulate more errors during translation or are harder to fold. In that case they need to be made in excess to compensate for the fact that many proteins never reach a functional state. We therefore looked at the differences between ED and NED proteins within the same complexes. We found that NED proteins on average are more structured and shorter. Conversely ED proteins are longer and less structured. Hence, mistranslation is not a likely explanation. However, misfolding of free NED subunits cannot by excluded as a limiting factor in the amount of available proteins to make a complex.
3. Alternatively, NED proteins could be over-synthesized to ensure that the long and unstructured ED proteins always have an interaction partner and do not aggregate causing disease (Kim et al., 2016). Certain proteins are toxic when overexpressed (Gelperin et al., 2005; Sopko et al., 2006). The toxicity is thought to be caused by increased amount of interaction-prone molecules, which increase the number of improper interactions made in a cell (Marcotte and Tsechansky, 2009; Vavouri et al., 2009). In multicellular organism's aggregation prevention might be worth the extra energy expenditure of synthesizing proteins only to degrade them. In contrast, studies in *E. coli* have shown that bacteria produce proteins in well-coordinated stoichiometry (Li et al., 2014). This would indicate that single cell organism that do not suffer as much from aggregation related disorders coordinate the proteins levels without degradation of excess subunits.
4. Another potential mechanism is that ED proteins function as rate-limiting factors in complex formation. The cell would then only need to upregulate the synthesis of one protein and not all the members of a complex to rapidly increase the levels of a complex. For example, a recent study demonstrated that the steady state levels of the CCT complex in different mouse strains could be explain by the variation of one subunit Cct6a (Chick et al., 2016). The authors suggested that degradation of the other super-stoichiometric CCT subunits was the most

probable explanation of the differences in the steady state levels of the complex. We have at this point only indirect evidence in support of this mechanism in the finding that ED proteins are more tissue-specifically expressed than NED proteins. This could mean that NED proteins in a complex are generally highly expressed but that ED proteins determine the tissue specific abundance of the final complex.

5. The super-stoichiometric synthesis levels of NED proteins might also help complex formation (Marsh et al., 2013; Matalon et al., 2014; Veitia, 2016). For example, proteins with lower affinity interactions might need to be synthesized at higher levels for their interactions to be sufficiently strong. Similarly, super-stoichiometric subunits might first interact among themselves before they interact with lower abundant peripheral subunits and in this way support an ordered assembly of a complexes.

Which, if any, of the suggested mechanism are at play we cannot say at this point. In any case, it has been estimated that 1 in 3 proteasomes are unengaged at steady state, meaning that cells have plenty of room to take care of excess, misfolded or otherwise unwanted proteins (Harper and Bennett, 2016).

One interesting idea is that the mechanism at play is most likely not a quality control instrument. We have no data suggesting a qualitative difference between the NED population that gets degraded and the NED population that is stabilized. Not unlikely is the incorporation of a protein into a complex a purely random event. Therefore, the term quality control does not apply to degradation of excess protein subunits in contrast to, for example, the clearance of misfolded Basigin or CFTR.

In summary, multiple mechanisms can explain why cells produce super-stoichiometric amounts of certain complex subunits. More research is needed to resolve which mechanism or mechanisms that mainly responsible.

4.3.3 NED in aneuploidy

Another major finding in this study is the fact that proteins defined as NED are also more likely to be attenuated at the protein level in aneuploid cells. The concept of attenuation was derived from studies of haploid yeast in which individual chromosomes were added to a disomic state (Dephoure et al., 2014; Torres et al., 2007). It was found that transcript levels doubled with the increase in genomic information in the diploid yeast. However, in 20 % of cases protein levels did not follow but stabilized at 1.6 fold increase (Dephoure et al., 2014). This effect was most strongly seen for proteins in protein complexes and seemed to be dependent on protein degradation, as judged by blockage of the UPS. Importantly, the attenuation of proteins derived from genes encoded on supernumerary chromosome

can also be seen in human cancer cells with genomic aberrations or with artificial tri- and tetrasomies (Geiger et al., 2010; Stingele et al., 2012).

Our finding that NED predicts attenuation in our triploid cells lines further extends this observation. First, we found that a number of proteins are “overexpressed” already at baseline, i.e. without chromosomal aberrations. Second, mainly the proteins that were super-stoichiometrically produced to begin with were attenuated in the triploid cell line (Figure 44). That means that belonging to a protein complex is not the main predictor of attenuation but rather being NED.

5 Outlook

5.1 Potential impact of the work

Hopefully, this work will dispel the common assumption of exponential degradation of proteins. This assumption is frequently applied to degradation modelling, e.g. Schwanhäusser et al. (2011). An exponential model will fit data based on degradation of a steady state (i.e. not age-defined) population without too much error because most NED proteins at steady state occur only in the second state and are therefore mainly impacted by the second degradation rate (Figure 50). However, an exponential model severely underestimates the importance of protein degradation in establishing the proteome. Only when applying more careful kinetic measurements such as AHA p-c the full impact of degradation on the proteome can be understood. For example, the apparent poor correlation between transcript and protein abundances could potentially improve if one took into account the initial degradation rate of NED proteins (Vogel and Marcotte, 2012). E.g. NED could be responsible for cases of lower than expected protein abundance based on transcript levels.

Importantly, the method described here should be applicable in answering many open questions in the field (see below). One particular area where AHA p-c should allow important progress is in the functional definition of different chaperones, E3 ubiquitin ligases and proteases involved in protein quality control and removal of excess protein subunits. For example, one could perform a knock-out screen of potential PQC proteins in combination with AHA p-c to look for substrates with increased NED.

Finally, the implication of NED proteins in establishing the steady state proteome during aneuploidy should increase the understanding of disease phenotypes in developmental syndromes and cancers depending on large genetic alterations.

5.1.1 Open questions

This work generates many questions. For example:

- Orthologs display similar degradation kinetics in different species, which strongly points to a biological function of NED. In addition, differences in molecular features between ED and NED proteins further indicates functionality. However, the question why proteins are NED is still open (see discussion for different hypotheses).

- Two pools of NED proteins exist with one having a higher degradation rate than the other. How does the cell distinguish between the two?
    - Are uncovered “degrons” such as N-terminal acetylation or stretches of positively charged amino acids responsible for the recognition of all uncomplexed subunits?
    - Furthermore, which are the folding-factors, ubiquitin ligases and proteases involved? Of particular interest are the mitochondrial proteases since a substantial fraction of NED proteins reside in mitochondria and do not respond proteasome inhibition (Figure 49).
- What is the function of ED proteins in complexes?
    - Are they rate limiting subunits? Would increasing their abundance increase steady state levels of their respective complexes?
    - Are ED proteins “toxic” when overexpressed? Do NED protein function as chaperones for ED proteins?
- What is the role of NED in buffering differences in transcript levels between different species, organs and in diseases? For example, it is known that most of the genetic and transcriptome differences between humans and chimps are buffered at the protein levels (Khan et al., 2013). To what extent is NED responsible for these kinds of observations? Studies looking at protein quantitative trait locis (pQTLs) indicate that post translational mechanisms are indeed responsible for much of the transcriptome buffering (Battle et al., 2015; Chick et al., 2016).
- How detrimental are NED proteins for cells during Aneuploidy?
    - Does the fact that excess NED proteins are not lingering in cells make them less harmful than ED protein that accumulate? The fact that “overexpression” of individual toxic protein in aneuploidy does not seem to cause disease would argue against ED proteins as the main driver of disease phenotype (Donnelly and Storchová, 2014; Torres et al., 2007).
    - Alternatively, is the extra burden on the cellular PQC systems by the extra NED proteins destined for degradation one of the drivers of disease phenotypes? General PQC stress in aneuploidy implicates NED proteins in aneuploidy related disease phenotypes (Donnelly and Storchová, 2014; Torres et al., 2007).

6 Supplementary Information

6.1 Abbreviations

Abbreviation	Expansion
ABC	Ammonium BiCarbonate
AHA	L-Azidohomoalanine
AIC	Akaike Information Criterion
ATP	Adenosine Triphosphate
AUC	Area Under the Curve
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
DAPI	4',6-Diamidino-2-Phenylindole
dFCS	dialyzed Fetal Calf Serum
DMEM	Dulbecco's Modified Eagle's Medium
DNA	Deoxyribonucleic Acid
DRiPs	Defective Ribosomal Products
DTT	Dithiothreitol
ED	Exponentially Degraded
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic Reticulum
ERAD	Endoplasmic Reticulum- Associated Protein Degradation
ESI	Electrospray Ionization
FITC	Fluorescein Isothiocyanate
HPLC	High-Performance Liquid Chromatography
iBAQ	Intensity Based Absolute Quantification
LC-MS/MS	Liquid Chromatography Tandem Mass Spectrometry
LDS	Lithium Dodecyl Sulfate
LSD	Lowest Score Deviation
LTQ	Linear Ion Trap Quadrupole
LysC	Lysyl Endopeptidase
MDC	Max Delbrück Center for Molecular Medicine

Met	Methionine
MHC-I	Major Histocompatibility Complex I
MPI	Max Planck Institute
NED	Non-Exponentially Degraded
PBS	Phosphate-Buffered Saline
p-c	Pulse-Chase
PQC	Protein Quality Control
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristics
rRNA	Ribosomal Ribonucleic Acid
SAX	Strong Anion Exchange
SCX	Strong Cation Exchange
SILAC	Stable Isotope Labeling by Amino Acids in Cell Culture
StageTip	Stop and Go Extraction Tip
TAMRA	5-Carboxytetramethylrhodamine
TCR	T-Cell Receptor
TFA	Triflouroacetic Acid
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer Ribonucleic Acid
UN	Undefined
UPS	Ubiquitin proteasome system
WAX	Weak Anion Exchange

6.2 Supplementary figures

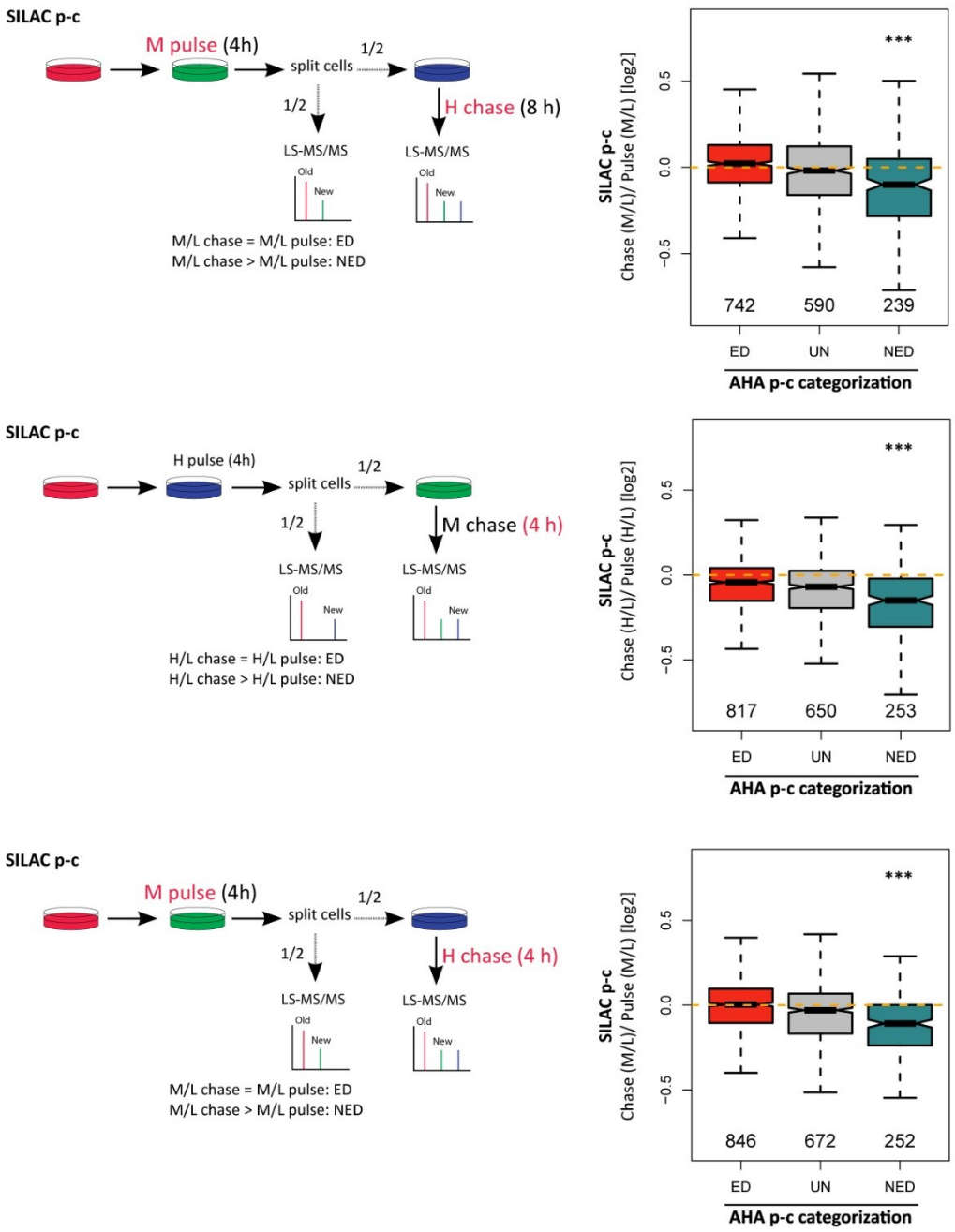

Figure 48. SILAC pulse-chase experiments as in Figure 27.

Differences between Figure 27 and the three experiments displayed are highlighted in red.

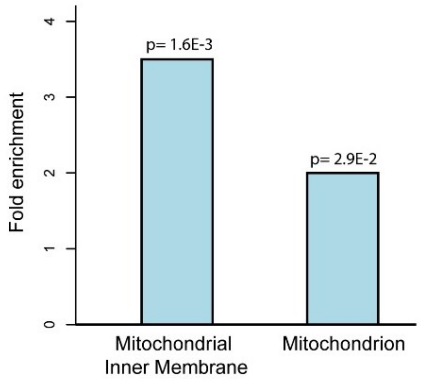

Figure 49. NED proteins insensitive to proteasome inhibition are enriched for mitochondrial proteins. Taking the 20 % of NED proteins with the smallest response to MG132 in Figure 29 B (n = 47) and looking for enrichment of the GO-terms subcategory subcellular location using the DAVID tool (Huang et al., 2009) we found an enrichment for proteins in the mitochondrial inner membrane using all NED proteins in Figure 29 B as background (n = 236). The fold enrichment of the two most significant hits are displayed. p-values are based on a modified Fisher exact test and have been corrected for multiple hypothesis testing using Benjamini-Hochberg as reported by the DAVID tool (Huang et al., 2009).

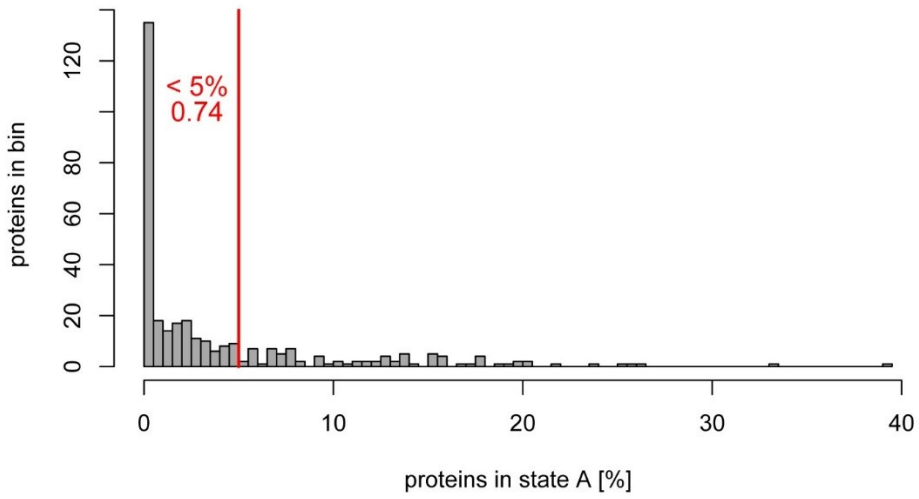

Figure 50. Distribution of NED proteins between state A and B at steady state.

Histogram showing the % of NED proteins in state A at steady state. All NED proteins are mainly in the second state at steady state. 74 % of NED proteins have less than 5 % of molecules in state A at steady state. That would mean that, at steady state, most NED proteins that are stabilized by complex formation are found complexed.

6.3 Supplementary tables

Supplemental tables 1-5 can be downloaded from:

http://www.cell.com/action/showImagesData?pii=S0092-8674%2816%2931248-X

6.4 Raw-data access

6.4.1 Mass spectrometric data

Mass spectrometric data has been uploaded to the ProteomeXchange data base with the following accession (Vizcaíno et al., 2014):

Mouse: PXD004929

Human: PXD004915

6.4.2 Genomic sequencing data

Genomic sequencing data has been uploaded to the NCBI BioProject with the following accession:

Human: PRJNA339199

7 Acknowledgements

"If I have seen further it is by standing on the shoulders of Giants" – Isaac Newton

The giants whose shoulders I have been standing on are very real. The first person I would like to express gratitude to is my always supportive PhD advisor Prof. Matthias Selbach. He accepted with stoic calm my many failures (for example he did not, literarily or figuratively, explode when I mixed up the lab stock of SILAC amino acids wasting a few thousand euros …) and encouraged the (few) successes that showed up along the road. If I become half the scientist he is I would count myself lucky.

I am indebted to my "doktorvater" Prof. Thomas Sommer whom has encouraged me throughout the PhD and has help by lending me the benches and minds of his lab.

I would also like to thank the further members of my PhD committee Prof. Britta Eickholt, Prof. Hanspeter Herzel and Prof. Achim Leutz for taking the time to read my dissertation and for participating in my defence.

Many entertaining evenings/nights where spent in the presence of Prof. Angelo Valleriani and Celine Sin (and Rico!) discussing science and everything else over pizza and beer. If we would have ended up with no results at all I would still have counted this as a great collaboration. Wei Chen, Xi Wang and Jingyi Hou (and Sunny) have all been there for urgent sequencing needs and great discussions! It was a pleasure to work with Prof. Joe Marsh and Jon Wells. Their fantastic analyses and cleaver minds help to crystalize many ideas. Prof. Zuzana Storchova and Neysan Donnelly's friendly and active support during this project made a great difference. Prof. Thomas Langer and Simon Tröder were invaluable during my one year detour (?) into mitochondrial protein degradation.

Henrik Zauber is a legend. Without his work this story would have had a very different outcome. I would also like to thank Martha and Christian for all technical help! Christian has helped me with pretty much everything during my years in the lab (including my taxes). Eva Kärgel saved me (and probably the MDC) from a nuclear melt down. Joao accepted the overwhelming challenge of proof reading my thesis. Without him this thesis would have been plagued with far more spelling mistakes and misaligned figures!

A special thank you goes to Matthias Sury. He guided me through my first year(s) in the Selbach lab and I will always be grateful for his support!

I am the last lab member to bridge all the Selbach lab people and as such I have many people to thank. I'm forever grateful to the old guys that introduced me to the lab and taught me everything I know about proteomics (and half the things I know about having fun): Björn, Olivia, Jimmy, Fabian, Marie, Murphy, and Florian. The new guys whom I taught far less in return: Boris, Katrina, Kamila, Katrin, Djordje, Michal, Koshi-博士, Tommaso, Damir, Carlos and Matthias. Also a special shout out to some of the most memorable short time members: Murat, Anna-Laura, Piotr and Jonas!

I would also like to thank Gunnar Dittmar for fruitful discussions and his lab for fantastic times especially Oliver, Tamara, Patrick, Daniel and Evelyn.

Petra and Sabine have made my life at the MDC so much easier. I am grateful for all they have done for me (that includes at one point arranging an extra month's salary!).

During my time at the MDC I also had the privilege to collaborate on some fantastic projects with some fantastic people. They made my time extra enjoyable: Maria, Carmen, Rebecca, Anup and Linda.

Finally, my family has supported me throughout the years.

The latest family member Luna (perhaps influenced by the Selbach lab baby boom?) made my last year more focused and distracted at once.

My wife Eugenia's never ending understanding and backing made this work possible I owe her everything. Thank you BG.

I dedicate this thesis to them.

8 Selbständigkeitserklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe.

Berlin, den

Erik McShane

9 References

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

Aebi, M., Bernasconi, R., Clerc, S., and Molinari, M. (2010). N-glycan structures: recognition and processing in the ER. Trends Biochem Sci *35*, 74-82.

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

Amara, J.F., Lederkremer, G., and Lodish, H.F. (1989). Intracellular degradation of unassembled asialoglycoprotein receptor subunits: a pre-Golgi, nonlysosomal endoproteolytic cleavage. The Journal of cell biology *109*, 3315-3324.

Amieux, P.S., Cummings, D.E., Motamed, K., Brandon, E.P., Wailes, L.A., Le, K., Idzerda, R.L., and McKnight, G.S. (1997). Compensatory regulation of RIalpha protein levels in protein kinase A mutant mice. J Biol Chem *272*, 3993-3998.

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.

Archibald, R. (1945). Colorimetric determination of Urea. J Biol Chem, 507.

Arlt, H., Tauer, R., Feldmann, H., Neupert, W., and Langer, T. (1996). The YTA10–12 Complex, an AAA Protease with Chaperone-like Activity in the Inner Membrane of Mitochondria. Cell *85*, 875-885.

Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J.E., Vandekerckhove, J., Lillehaug, J.R., Sherman, F*., et al.* (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. Proceedings of the National Academy of Sciences *106*, 8157-8162.

Bachmair, A., Finley, D., and Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. Science *234*, 179-186.

Bagert, J.D., Xie, Y.J., Sweredoski, M.J., Qi, Y., Hess, S., Schuman, E.M., and Tirrell, D.A. (2014). Quantitative, time-resolved proteomic analysis by combining bioorthogonal noncanonical amino acid tagging and pulsed stable isotope labeling by amino acids in cell culture. Mol Cell Proteomics *13*, 1352-1358.

Bainton, D.F. (1981). The discovery of lysosomes. The Journal of cell biology *91*, 66s-76s.

Baker, B.M., and Haynes, C.M. (2011). Mitochondrial protein quality control during biogenesis and aging. Trends Biochem Sci *36*, 254-261.

Baker, R.T., Tobias, J.W., and Varshavsky, A. (1992). Ubiquitin-specific proteases of Saccharomyces cerevisiae. Cloning of UBP2 and UBP3, and functional analysis of the UBP gene family. J Biol Chem *267*, 23364-23375.

Bassani-Sternberg, M., Pletscher-Frankild, S., Jensen, L.J., and Mann, M. (2015). Mass Spectrometry of Human Leukocyte Antigen Class I Peptidomes Reveals Strong Effects of Protein Abundance and Turnover on Antigen Presentation. Molecular & Cellular Proteomics : MCP *14*, 658-673.

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.

Bensaude, O. (2011). Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? Transcription *2*, 103-108.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic acids research *28*, 235-242.

Berthoud, V.M., Tadros, P.N., and Beyer, E.C. (2000). Connexin and Gap Junction Degradation. Methods *20*, 180-187.

Bleakney, W. (1929). A new method of positive ray analysis and its applications to the measurement of ionization potentials in mercury vapor. Physical review *34*, 157-160.

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.

Boerhaave, H. (1732). Elementa Chemicae, Vol 2 (Leipzig: Caspar Fritsch).

Boisvert, F.M., Ahmad, Y., Gierlinski, M., Charriere, F., Lamont, D., Scott, M., Barton, G., and Lamond, A.I. (2012). A quantitative spatial proteomics analysis of proteome turnover in human cells. Mol Cell Proteomics *11*, M111.011429.

Borsook, H., Deasy, C.L., Haagen-Smit, A.J., Keighley, G., and Lowy, P.H. (1950). METABOLISM OF C14-LABELED GLYCINE, l-HISTIDINE, l-LEUCINE, AND l-LYSINE. Journal of Biological Chemistry *187*, 839-848.

Bradshaw, R.A., Brickey, W.W., and Walker, K.W. (1998). N-terminal processing: the methionine aminopeptidase and N alpha-acetyl transferase families. Trends Biochem Sci *23*, 263-267.

Bross, P., Andresen, B.S., Knudsen, I., Kruse, T.A., and Gregersen, N. (1995). Human ClpP protease: cDNA sequence, tissue-specific expression and chromosomal assignment of the gene. FEBS letters *377*, 249-252.

Buchberger A, B.B., Sommer T (2010). Protein Quality Control in the Cytosol and the Endoplasmic Reticulum: Brothers in Arms. Molecular Cell *40*, 238-252.

Burnett, B.G., Munoz, E., Tandon, A., Kwon, D.Y., Sumner, C.J., and Fischbeck, K.H. (2009). Regulation of SMN protein stability. Molecular and cellular biology *29*, 1107-1115.

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

Calve, S., Witten, A.J., Ocken, A.R., and Kinzer-Ursem, T.L. (2016). Incorporation of non-canonical amino acids into the developing murine proteome. Scientific reports *6*, 32377.

Castro-Obregon, S. (2010). The Discovery of Lysosomes and Autophagy. Nature Education *3*, 49.

Chai, Q., Webb, S.R., Wang, Z., Dutch, R.E., and Wei, Y. (2016). Study of the degradation of a multidrug transporter using a non-radioactive pulse chase method. Analytical and bioanalytical chemistry.

Chen, C., Bonifacino, J.S., Yuan, L.C., and Klausner, R.D. (1988). Selective degradation of T cell antigen receptor chains retained in a pre-Golgi compartment. The Journal of cell biology *107*, 2149-2161.

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.

Chick, J.M., Munger, S.C., Simecek, P., Huttlin, E.L., Choi, K., Gatti, D.M., Raghupathy, N., Svenson, K.L., Churchill, G.A., and Gygi, S.P. (2016). Defining the consequences of genetic variation on a proteome-wide scale. Nature *534*, 500-505.

Christiano, R., Nagaraj, N., Fröhlich, F., and Walther, Tobias C. Global Proteome Turnover Analyses of the Yeasts S. cerevisiae and S. pombe. Cell reports *9*, 1959-1965.

Ciechanover, A. (2005). Proteolysis: from the lysosome to ubiquitin and the proteasome. Nature reviews Molecular cell biology *6*, 79-87.

Ciechanover, A., Elias, S., Heller, H., Ferber, S., and Hershko, A. (1980a). Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes. J Biol Chem *255*, 7525-7528.

Ciechanover, A., Heller, H., Elias, S., Haas, A.L., and Hershko, A. (1980b). ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. Proc Natl Acad Sci U S A *77*, 1365-1368.

Ciehanover, A., Hod, Y., and Hershko, A. (1978). A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. Biochemical and Biophysical Research Communications *81*, 1100-1105.

Clark, E., and Spector, D.H. (2015). Studies on the Contribution of Human Cytomegalovirus UL21a and UL97 to Viral Growth and Inactivation of the Anaphase-Promoting Complex/Cyclosome (APC/C) E3

Ubiquitin Ligase Reveal a Unique Cellular Mechanism for Downmodulation of the APC/C Subunits APC1, APC4, and APC5. Journal of Virology *89*, 6928-6939.

Claude, A. (1946). FRACTIONATION OF MAMMALIAN LIVER CELLS BY DIFFERENTIAL CENTRIFUGATION : II. EXPERIMENTAL PROCEDURES AND RESULTS. The Journal of experimental medicine *84*, 61-89.

Cohen, L.D., Zuchman, R., Sorokina, O., Müller, 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.

Corless, C.L., Matzuk, M.M., Ramabhadran, T.V., Krichevsky, A., and Boime, I. (1987). Gonadotropin beta subunits determine the rate of assembly and the oligosaccharide processing of hormone dimer in transfected cells. The Journal of cell biology *104*, 1173-1181.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol Cell Proteomics *13*, 2513-2526.

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.

Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J.V., and Mann, M. (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nat Protocols *4*, 698-705.

De Duve, C., Pressman, B.C., Gianetto, R., Wattiaux, R., and Appelmans, F. (1955). Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. The Biochemical journal *60*, 604-617.

de Godoy, L.M.F., Olsen, J.V., Cox, J., Nielsen, M.L., Hubner, N.C., Frohlich, F., Walther, T.C., and Mann, M. (2008). Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. Nature *455*, 1251-1254.

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.

Dettmer, U., Kuhn, P.-H., Abou-Ajram, C., Lichtenthaler, S.F., Krüger, M., Kremmer, E., Haass, C., and Haffner, C. (2010). Transmembrane Protein 147 (TMEM147) Is a Novel Component of the Nicalin-NOMO Protein Complex. The Journal of Biological Chemistry *285*, 26174-26181.

Dieterich, D.C., Lee, J.J., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2007). Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. Nat Protoc *2*, 532-540.

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. Journal of Proteome Research *8*, 104-112.

Donnelly, N., and Storchová, Z. (2014). Dynamic karyotype, dynamic proteome: buffering the effects of aneuploidy. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1843*, 473-481.

Drysdale, J.W., and Munro, H.N. (1966). Regulation of Synthesis and Turnover of Ferritin in Rat Liver. Journal of Biological Chemistry *241*, 3630-3637.

Dulis, B.H., Kloppel, T.M., Grey, H.M., and Kubo, R.T. (1982). Regulation of catabolism of IgM heavy chains in a B lymphoma cell line. J Biol Chem *257*, 4369-4374.

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

Duve, C. (1975). Exploring cells with a centrifuge. Science *189*, 186-194.

Edbauer, D., Winkler, E., Haass, C., and Steiner, H. (2002). Presenilin and nicastrin regulate each other and determine amyloid beta-peptide production via complex formation. Proc Natl Acad Sci U S A *99*, 8666-8671.

Eden, E., Geva-Zatorsky, N., Issaeva, I., Cohen, A., Dekel, E., Danon, T., Cohen, L., Mayo, A., and Alon, U. (2011). Proteome Half-Life Dynamics in Living Human Cells. Science *331*, 764-768.

Efeyan, A., Zoncu, R., and Sabatini, D.M. (2012). Amino acids and mTORC1: from lysosomes to disease. Trends in molecular medicine *18*, 524-533.

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.

elBaradi, T.T., van der Sande, C.A., Mager, W.H., Raue, H.A., and Planta, R.J. (1986). The cellular level of yeast ribosomal protein L25 is controlled principally by rapid degradation of excess protein. Current genetics *10*, 733-739.

Etlinger, J.D., and Goldberg, A.L. (1977). A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. Proceedings of the National Academy of Sciences of the United States of America *74*, 54-58.

Faccio, L., Fusco, C., Chen, A., Martinotti, S., Bonventre, J.V., and Zervos, A.S. (2000). Characterization of a novel human serine protease that has extensive homology to bacterial heat shock endoprotease HtrA and is regulated by kidney ischemia. J Biol Chem *275*, 2581-2588.

Fanis, P., Gillemans, N., Aghajanirefah, A., Pourfarzad, F., Demmers, J., Esteghamat, F., Vadlamudi, R.K., Grosveld, F., Philipsen, S., and van Dijk, T.B. (2012). Five friends of methylated chromatin target of protein-arginine-methyltransferase[prmt]-1 (chtop), a complex linking arginine methylation to desumoylation. Mol Cell Proteomics *11*, 1263-1273.

Fenteany, G., Standaert, R., Lane, W., Choi, S., Corey, E., and Schreiber, S. (1995). Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. Science *268*, 726-731.

Folin, O. (1905a). LAWS GOVERNING THE CHEMICAL COMPOSITION OF URINE. Am J, Physiol *13*, 66.

Folin, O. (1905b). A THEORY OF PROTEIN METABOLISM. Am J, Physiol *13*, 117.

Geiger, T., Cox, J., and Mann, M. (2010). Proteomic changes resulting from gene copy number variations in cancer cells. PLoS genetics *6*, e1001090.

Gelperin, D.M., White, M.A., Wilkinson, M.L., Kon, Y., Kung, L.A., Wise, K.J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H.*, et al.* (2005). Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. Genes & development *19*, 2816-2826.

Girard, M., Penman, S., and Darnell, J.E. (1964). THE EFFECT OF ACTINOMYCIN ON RIBOSOME FORMATION IN HELA CELLS. Proc Natl Acad Sci U S A *51*, 205-211.

Goldberg, A. (1972). Degradation of Abnormal Proteins in Escherichia coli. Proc Natl Acad Sci U S A *69*, 422-426.

Goldberg, A.L. (2003). Protein degradation and protection against misfolded or damaged proteins. Nature *426*, 895-899.

Goldberg, A.L. (2012). Development of proteasome inhibitors as research tools and cancer drugs. The Journal of cell biology *199*, 583-588.

Goldberg, A.L., and Dice, J.F. (1974). Intracellular protein degradation in mammalian and bacterial cells. Annu Rev Biochem *43*, 835-869.

Goldknopf, I.L., and Busch, H. (1977). Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. Proceedings of the National Academy of Sciences of the United States of America *74*, 864-868.

Gunjan, A., and Verreault, A. (2003). A Rad53 Kinase-Dependent Surveillance Mechanism that Regulates Histone Protein Levels in S. cerevisiae. Cell *115*, 537-549.

Haffner, C., Dettmer, U., Weiler, T., and Haass, C. (2007). The Nicastrin-like protein Nicalin regulates assembly and stability of the Nicalin-nodal modulator (NOMO) membrane protein complex. J Biol Chem *282*, 10632-10638.

Harper, J.W., and Bennett, E.J. (2016). Proteome complexity and the forces that drive proteome imbalance. Nature *537*, 328-338.

Hartley, H. (1951). Origin of the word 'protein'. Nature *168*, 244.

Hassold, T., Hall, H., and Hunt, P. (2007). The origin of human aneuploidy: where we have been, where we are going. Human molecular genetics *16 Spec No. 2*, R203-208.

Heidelberger, M., Treffers, H.P., Schoenheimer, R., Ratner, S., and Rittenberg, D. (1942). BEHAVIOR OF ANTIBODY PROTEIN TOWARD DIETARY NITROGEN IN ACTIVE AND PASSIVE IMMUNITY. Journal of Biological Chemistry *144*, 555-562.

Hershko, A., Ciechanover, A., Heller, H., Haas, A.L., and Rose, I.A. (1980). Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. Proceedings of the National Academy of Sciences of the United States of America *77*, 1783-1786.

Hinkson, I.V., and Elias, J.E. (2011). The dynamic state of protein turnover: It's about time. Trends Cell Biol *21*, 293-303.

Hinz, F.I., Dieterich, D.C., Tirrell, D.A., and Schuman, E.M. (2012). Non-canonical amino acid labeling in vivo to visualize and affinity purify newly synthesized proteins in larval zebrafish. ACS Chem Neurosci *3*, 40-49.

Hogness, D., Melvin, C., Monod, M. (1955). Studies on the induced synthesis of β-galactosidase in Escherichia coli: The kinetics and mechanism of sulfur incorporationOriginal Research. Biochimica et Biophysica Acta *16*, 99-116.

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.

Hough, R., Pratt, G., and Rechsteiner, M. (1987). Purification of two high molecular weight proteases from rabbit reticulocyte lysate. J Biol Chem *262*, 8303-8313.

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.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic acids research *37*, 1-13.

Huh, K.H., and Wenthold, R.J. (1999). Turnover analysis of glutamate receptors identifies a rapidly degraded pool of the N-methyl-D-aspartate receptor subunit, NR1, in cultured cerebellar granule cells. J Biol Chem *274*, 151-157.

Hunt, L.T., and Dayhoff, M.O. (1977). Amino-terminal sequence identity of ubiquitin and the nonhistone component of nuclear protein A24. Biochem Biophys Res Commun *74*, 650-655.

Jan, C.H., Williams, C.C., and Weissman, J.S. (2014). Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. Science *346*, 1257521.

Jao, C.Y., and Salic, A. (2008). Exploring RNA transcription and turnover in vivo by using click chemistry. Proc Natl Acad Sci U S A *105*, 15779-15784.

Jensen, T.J., Loo, M.A., Pind, S., Williams, D.B., Goldberg, A.L., and Riordan, J.R. (1995). Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. Cell *83*, 129-135.

Johnson, F., and Kaplitt, M.G. (2009). Novel Mitochondrial Substrates of Omi Indicate a New Regulatory Role in Neurodegenerative Disorders. PLoS ONE *4*, e7100.

Johnson, P.R., Swanson, R., Rakhilina, L., and Hochstrasser, M. (1998). Degradation signal masking by heterodimerization of MATalpha2 and MATa1 blocks their mutual destruction by the ubiquitin-proteasome pathway. Cell *94*, 217-227.

Kamen, M.D. (1985). Radiant science, dark politics

A memoir of the nuclear age (Berkely and Los Angeles, California: University of California Press).

Kang, S.G., Dimitrova, M.N., Ortega, J., Ginsburg, A., and Maurizi, M.R. (2005). Human mitochondrial ClpP is a stable heptamer that assembles into a tetradecamer in the presence of ClpX. J Biol Chem *280*, 35424-35432.

Katz, W., Weinstein, B., and Solomon, F. (1990). Regulation of tubulin levels and microtubule assembly in Saccharomyces cerevisiae: consequences of altered tubulin gene copy number. Molecular and cellular biology *10*, 5286-5294.

Khan, Z., Ford, M.J., Cusanovich, D.A., Mitrano, A., Pritchard, J.K., and Gilad, Y. (2013). Primate Transcript and Protein Expression Levels Evolve Under Compensatory Selection Pressures. Science *342*, 1100-1104.

Khmelinskii, A., Keller, P.J., Bartosik, A., Meurer, M., Barry, J.D., Mardin, B.R., Kaufmann, A., Trautmann, S., Wachsmuth, M., Pereira, G.*, et al.* (2012). Tandem fluorescent protein timers for in vivo analysis of protein dynamics. Nat Biotechnol *30*, 708-714.

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.

Kim, Y.E., Hosp, F., Frottin, F., Ge, H., Mann, M., Hayer-Hartl, M., and Hartl, F.U. (2016). Soluble Oligomers of PolyQ-Expanded Huntingtin Target a Multiplicity of Key Cellular Factors. Mol Cell *63*, 951-964.

Korbel, D., Wurth, S., Kaser, M., and Langer, T. (2004). Membrane protein turnover by the m-AAA protease in mitochondria depends on the transmembrane domains of its subunits. EMBO reports *5*, 698-703.

Krüger, M., Moser, M., Ussar, S., Thievessen, I., Luber, C.A., Forner, F., Schmidt, S., Zanivan, S., Fässler, R., and Mann, M. (2008). SILAC Mouse for Quantitative Proteomics Uncovers Kindlin-3 as an Essential Factor for Red Blood Cell Function. Cell *134*, 353-364.

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.

Kulesh, D.A., Cecena, G., Darmon, Y.M., Vasseur, M., and Oshima, R.G. (1989). Posttranslational regulation of keratins: degradation of mouse and human keratins 18 and 8. Molecular and cellular biology *9*, 1553-1565.

Kuriyama, Y., and Omura, T. (1971). Different turnover behavior of phenobarbital-induced and normal NADPH-cytochrome c reductases in rat liver microsomes. Journal of biochemistry *69*, 659-669.

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., and Lamond, A.I. (2015). Multidimensional proteomics for cell biology. Nature reviews Molecular cell biology *16*, 269-280.

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.

Levi, H. (1940). Radio-sulphur. Nature *145*, 588.

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.

Lippincott-Schwartz, J., Bonifacino, J.S., Yuan, L.C., and Klausner, R.D. (1988). Degradation from the endoplasmic reticulum: disposing of newly synthesized proteins. Cell *54*, 209-220.

Liu, B., Han, Y., and Qian, S.B. (2013). Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. Mol Cell *49*, 453-463.

Lusk, G. (1909). The Elements Of The Science Of Nutrition, 3 edn (Philadelphia and London: W. B. Saunders company).

Maicas, E., Pluthero, F.G., and Friesen, J.D. (1988). The accumulation of three yeast ribosomal proteins under conditions of excess mRNA is determined primarily by fast protein decay. Molecular and cellular biology *8*, 169-175.

Mandelstam, J. (1957). Turnover of Protein in Starved Bacteria and its Relationship to the Induced Synthesis of Enzyme. Nature *179*, 1179-1181.

Mandon, E.C., Trueman, S.F., and Gilmore, R. (2013). Protein Translocation across the Rough Endoplasmic Reticulum. Cold Spring Harbor Perspectives in Biology *5*.

Marcotte, E.M., and Tsechansky, M. (2009). Disorder, Promiscuity, and Toxic Partnerships. Cell *138*, 16-18.

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.

McCance, R. (1930). The chemistry of the degradation of protein nitrogen. Physiological reviews *10*, 1-34.

McClatchy, D.B., Ma, Y., Liu, C., Stein, B.D., Martinez-Bartolome, S., Vasquez, D., Hellberg, K., Shaw, R.J., and Yates, J.R., 3rd (2015). Pulsed Azidohomoalanine Labeling in Mammals (PALM) Detects Changes in Liver-Specific LKB1 Knockout Mice. J Proteome Res *14*, 4815-4822.

Merlie, J.P., Sebbane, R., Tzartos, S., and Lindstrom, J. (1982). Inhibition of glycosylation with tunicamycin blocks assembly of newly synthesized acetylcholine receptor subunits in muscle cells. J Biol Chem *257*, 2694-2701.

Meusser, B., Hirsch, C., Jarosch, E., and Sommer, T. (2005). ERAD: the long road to destruction. Nat Cell Biol *7*, 766-772.

Minami, Y., Weissman, A.M., Samelson, L.E., and Klausner, R.D. (1987). Building a multichain receptor: synthesis, degradation, and assembly of the T-cell antigen receptor. Proc Natl Acad Sci U S A *84*, 2688-2692.

Mirigian, L.S., Makareeva, E., and Leikin, S. (2014). Pulse-chase analysis of procollagen biosynthesis by azidohomoalanine labeling. Connective tissue research *55*, 403-410.

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.

Mueller, S., Wahlander, A., Selevsek, N., Otto, C., Ngwa, E.M., Poljak, K., Frey, A.D., Aebi, M., and Gauss, R. (2015). Protein degradation corrects for imbalanced subunit stoichiometry in OST complex assembly. Mol Biol Cell *26*, 2596-2608.

Mulder, G. (1838). SUR LA COMPOSITION DE QUELQUES SUBSTANCES ANIMALES. Bulletin des Sciences Physiques et Naturelles en Néerlande.

Nagaraj, N., Wisniewski, J.R., Geiger, T., Cox, J., Kircher, M., Kelso, J., Pääbo, S., and Mann, M. (2011). Deep proteome and transcriptome mapping of a human cancer cell line. Molecular systems biology *7*, 548-548.

Nedialkova, D.D., and Leidel, S.A. (2015). Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. Cell *161*, 1606-1618.

Okabe, S., and Hirokawa, N. (1989). Rapid turnover of microtubule-associated protein MAP2 in the axon revealed by microinjection of biotinylated MAP2 into cultured neurons. Proc Natl Acad Sci U S A *86*, 4127-4131.

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.

Pajic, A., Tauer, R., Feldmann, H., Neupert, W., and Langer, T. (1994). Yta10p is required for the ATP-dependent degradation of polypeptides in the inner membrane of mitochondria. FEBS letters *353*, 201-206.

Pearce, D.A., and Sherman, F. (1995). Degradation of cytochrome oxidase subunits in mutants of yeast lacking cytochrome c and suppression of the degradation by mutation of yme1. J Biol Chem *270*, 20879-20882.

Pechmann, S., Willmund, F., and Frydman, J. (2013). The ribosome as a hub for protein quality control. Mol Cell *49*, 411-421.

Pflüger, E. (1893). Archiv für die gesammte physiologie des menschen und der thiere (Bonn: verlag von Emil Strauss).

Pine, M.J. (1967). Response of intracellular proteolysis to alteration of bacterial protein and the implications in metabolic regulation. Journal of bacteriology *93*, 1527-1533.

Plant, P.W., and Grieninger, G. (1986). Noncoordinate synthesis of the fibrinogen subunits in hepatocytes cultured under hormone-deficient conditions. J Biol Chem *261*, 2331-2336.

Poole, B., and Wibo, M. (1973). Protein degradation in cultured cells. The effect of fresh medium, fluoride, and iodoacetate on the digestion of cellular protein of rat fibroblasts. J Biol Chem *248*, 6221-6226.

Porter, K.R., Claude, A., and Fullam, E.F. (1945). A STUDY OF TISSUE CULTURE CELLS BY ELECTRON MICROSCOPY: METHODS AND PRELIMINARY OBSERVATIONS. The Journal of experimental medicine *81*, 233-246.

Price, V.E., Sterling, W.R., Tarantola, V.A., Hartley, R.W., and Rechcigl, M. (1962). The Kinetics of Catalase Synthesis and Destruction in Vivo. Journal of Biological Chemistry *237*, 3468-3475.

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.

Quivy, J.P., and Almouzni, G. (2003). Rad53: a controller ensuring the fine-tuning of histone levels. Cell *115*, 508-510.

Rabinovitz, M., and Fisher, J.M. (1964). Characteristics of the inhibition of hemoglobin synthesis in rabbit reticulocytes by threo-α-amino-β-chlorobutyric acid. Biochimica et Biophysica Acta (BBA) - Specialized Section on Nucleic Acids and Related Subjects *91*, 313-322.

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.

Ratner, S. (1979). The dynamic state of body proteins. Annals of the New York Academy of Sciences *325*, 188-209.

Reynolds, C.T.a.J. (2001). Natures's robots, A history of proteins (Great Clarendon Street, Oxford, Great Britain: Oxford University press).

Rock, K.L., and Goldberg, A.L. (1999). Degradation of cell proteins and the generation of MHC class I-presented peptides. Annual review of immunology *17*, 739-779.

Rock, K.L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A.L. (1994). Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. Cell *78*, 761-771.

Ruben, S.a.K., M.D. (1941). Long-Lived Radioactive Carbon: C14. Physical Review *59*, 349-354.

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.

Scazzari, M., Amm, I., and Wolf, D.H. (2015). Quality control of a cytoplasmic protein complex: chaperone motors and the ubiquitin-proteasome system govern the fate of orphan fatty acid synthase subunit Fas2 of yeast. J Biol Chem *290*, 4677-4687.

Schimke, R.T. (1964). The Importance of Both Synthesis and Degradation in the Control of Arginase Levels in Rat Liver. Journal of Biological Chemistry *239*, 3808-3817.

Schimke, R.T., and Doyle, D. (1970). Control of enzyme levels in animal tissues. Annu Rev Biochem *39*, 929-976.

Schimke, R.T., Sweeney, E.W., and Berlin, C.M. (1965). The Roles of Synthesis and Degradation in the Control of Rat Liver Tryptophan Pyrrolase. Journal of Biological Chemistry *240*, 322-331.

Schoenheimer, R., and Ratner, S. (1939). STUDIES IN PROTEIN METABOLISM: III. SYNTHESIS OF AMINO ACIDS CONTAINING ISOTOPIC NITROGEN. Journal of Biological Chemistry *127*, 301-313.

Schoenheimer, R., Ratner, S., and Rittenberg, D. (1939). STUDIES IN PROTEIN METABOLISM: VII. THE METABOLISM OF TYROSINE. Journal of Biological Chemistry *127*, 333-344.

Schoenheimer, R., and Rittenberg, D. (1936). DEUTERIUM AS AN INDICATOR IN THE STUDY OF INTERMEDIARY METABOLISM: VI. SYNTHESIS AND DESTRUCTION OF FATTY ACIDS IN THE ORGANISM. Journal of Biological Chemistry *114*, 381-396.

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.

Schuller, H.J., Fortsch, B., Rautenstrauss, B., Wolf, D.H., and Schweizer, E. (1992). Differential proteolytic sensitivity of yeast fatty acid synthetase subunits alpha and beta contributing to a balanced ratio of both fatty acid synthetase components. European journal of biochemistry / FEBS *203*, 607-614.

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.

Shaeffer, J.R. (1988). ATP-dependent proteolysis of hemoglobin alpha chains in beta-thalassemic hemolysates is ubiquitin-dependent. J Biol Chem *263*, 13663-13669.

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.

Shemin, D.R., D. (1946). The Life Span of a Human Red Blood Cell. Journal of Biological Chemistry *166*, 627.

Shemorry, A., Hwang, C.S., and Varshavsky, A. (2013). Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. Mol Cell *50*, 540-551.

Silverstein, A.M., Barrow, C.A., Davis, A.J., and Mumby, M.C. (2002). Actions of PP2A on the MAP kinase pathway and apoptosis are mediated by distinct regulatory subunits. Proc Natl Acad Sci U S A *99*, 4221-4226.

Sin, C., Chiarugi, D., and Valleriani, A. (2016). Degradation Parameters from Pulse-Chase Experiments. PLoS ONE *11*, e0155028.

Sin, C.C., D.; Valleriani, A. (2016). Degradation Parameters from Pulse-Chase Experiments. Under review.

Sommer, T., and Jentsch, S. (1993). A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. Nature *365*, 176-179.

Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S.G., Cyert, M., Hughes, T.R.*, et al.* (2006). Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell *21*, 319-330.

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.

Stiburek, L., Cesnekova, J., Kostkova, O., Fornuskova, D., Vinsova, K., Wenchich, L., Houstek, J., and Zeman, J. (2012). YME1L controls the accumulation of respiratory chain subunits and is required for apoptotic resistance, cristae morphogenesis, and cell proliferation. Mol Biol Cell *23*, 1010-1023.

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.

Straus, W. (1954). Isolation and biochemical properties of droplets from the cells of rat kidney. J Biol Chem *207*, 745-755.

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.

Sumner, J.B. (1926). THE ISOLATION AND CRYSTALLIZATION OF THE

ENZYME UREASE. . J Biol Chem *69*, 435.

Sung, M.-K., Porras-Yakushi, T.R., Reitsma, J.M., Huber, F.M., Sweredoski, M.J., Hoelz, A., Hess, S., and Deshaies, R.J. (2016). A conserved quality-control pathway that mediates degradation of unassembled ribosomal proteins. eLife *5*, e19105.

Suzuki, C., Suda, K., Wang, N., and Schatz, G. (1994). Requirement for the yeast gene LON in intramitochondrial proteolysis and maintenance of respiration. Science *264*, 273-276.

Takeshige, K., Baba, M., Tsuboi, S., Noda, T., and Ohsumi, Y. (1992). Autophagy in yeast demonstrated with proteinase-deficient mutants and conditions for its induction. The Journal of cell biology *119*, 301-311.

Tanida, I., Ueno, T., and Kominami, E. (2008). LC3 and Autophagy. Methods in molecular biology (Clifton, NJ) *445*, 77-88.

Tarentino, A.L., Richert, D.A., and Westerfeld, W.W. (1966). The concurrent induction of hepatic alpha-glycerophosphate dehydrogenase and malate dehydrogenase by thyroid hormone. Biochim Biophys Acta *124*, 205-309.

Tatsuta, T., and Langer, T. (2008). Quality control of mitochondria: protection against neurodegeneration and ageing. The EMBO Journal *27*, 306-314.

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.

Torres, E.M., Sokolsky, T., Tucker, C.M., Chan, L.Y., Boselli, M., Dunham, M.J., and Amon, A. (2007). Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science *317*, 916-924.

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.

Urey, H.C., and Greiff, L.J. (1935). Isotopic Exchange Equilibria. Journal of the American Chemical Society *57*, 321-327.

Urey, H.C.H.J.R.T.H.G.a.F.M. (1937). Goncentration of N15 by Chemical Methods. The Journal of Ghemical Physics *5*, 856.

Vabulas, R.M., and Hartl, F.U. (2005). Protein synthesis upon acute nutrient restriction relies on proteasome function. Science *310*, 1960-1963.

Varland, S., Osberg, C., and Arnesen, T. (2015). N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. PROTEOMICS *15*, 2385-2401.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. Protein science : a publication of the Protein Society *20*, 1298-1345.

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.

Veitia, R.A. (2016). A Fresh Look at 'Aging' Proteins. Trends in Biochemical Sciences.

Vickery, H.B. (1950). The Origin of the Word Protein. The Yale Journal of Biology and Medicine *22*, 387-393.

Vizcaíno, J.A., Deutsch, E.W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J.A., Sun, Z., Farrah, T., and Bandeira, N. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nature biotechnology *32*, 223-226.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat Rev Genet *13*, 227-232.

von Liebig, J. (1842). Die organische Chemie in ihrer Anwendung auf Physiologie und Pathologie (Braunschweig: Verlag von Friedrich Vieweg und Sohn).

von Voit, C. (1860). Die Gesetze der Ernährung des Fleischfressers durch neue Untersuchungen festgestellt (Leipzig and Heidelberg: Wintersche).

Wang, F., Durfee, L.A., and Huibregtse, J.M. (2013). A cotranslational ubiquitination pathway for quality control of misfolded proteins. Mol Cell *50*, 368-378.

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.

Ward, C.L., Omura, S., and Kopito, R.R. (1995). Degradation of CFTR by the ubiquitin-proteasome pathway. Cell *83*, 121-127.

Warner, J.R. (1966). The assembly of ribosomes in HeLa cells. Journal of Molecular Biology *19*, 383-398.

Warner, R. (1977). In the absence of ribosomal RNA synthesis, the ribosomal proteins of HeLa Cells are synthesized normally and degraded rapidly. Journal of Molecular Biology *115*, 315-333.

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.

Wessel, D., and Flugge, U.I. (1984). A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Analytical biochemistry *138*, 141-143.

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.

Wilk, S., and Orlowski, M. (1980). Cation-sensitive neutral endopeptidase: isolation and specificity of the bovine pituitary enzyme. Journal of neurochemistry *35*, 1172-1182.

Wilkinson, K.D., Urban, M.K., and Haas, A.L. (1980). Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. J Biol Chem *255*, 7529-7532.

Williams, R.B., and Dawson, R.M. (1952). The biosynthesis of L-cystine and L-methionine labelled with radioactive sulphur (35S). The Biochemical journal *52*, 314-317.

Wiltschi, B., Merkel, L., and Budisa, N. (2009). Fine Tuning the N-Terminal Residue Excision with Methionine Analogues. Chembiochem *10*, 217-220.

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.

Wu, C.C., MacCoss, M.J., Howell, K.E., Matthews, D.E., and Yates, J.R., 3rd (2004). Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. Analytical chemistry *76*, 4951-4959.

Yen, H.-C.S., Xu, Q., Chou, D.M., Zhao, Z., and Elledge, S.J. (2008). Global Protein Stability Profiling in Mammalian Cells. Science *322*, 918-923.

Yewdell, J.W., Anton, L.C., and Bennink, J.R. (1996). Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? Journal of immunology (Baltimore, Md : 1950) *157*, 1823-1826.

Yewdell, J.W., and Nicchitta, C.V. (2006). The DRiP hypothesis decennial: support, controversy, refinement and extension. Trends Immunol *27*, 368-373.

Young, R.W., and Bok, D. (1969). PARTICIPATION OF THE RETINAL PIGMENT EPITHELIUM IN THE ROD OUTER SEGMENT RENEWAL PROCESS. The Journal of cell biology *42*, 392-403.

Zhang, Y., Fonslow, B.R., Shan, B., Baek, M.-C., and Yates, J.R. (2013). Protein Analysis by Shotgun/Bottom-up Proteomics. Chemical reviews *113*, 2343-2394.