Mapping quality control systems for mislocalized proteins

Dissertation

zur Erlangung des Grades

Doktor der Naturwissenschaften

am Fachbereich Biologie

der Johannes Gutenberg-Universität Mainz

vorgelegt von

Zhaoyan Li

geboren am 05.04.1992

in Beijing, China

Mainz, Mai 2023

Dekan: Prof. Dr. Eckhard Thines

1. Gutachter: [redacted]

2. Gutachter: [redacted]

Tag der mündlichen Prüfung: 15.06.2023

1 TABLE OF CONTENTS

2 LIST OF PUBLICATIONS

Kong KYE*, Fischer B*, Meurer M*, Kats I, **Li Z**, Rühle F, Barry JD, Kirrmaier D, Chevyreva V, San Luis BJ, Costanzo M, Huber W, Andrews BJ, Boone C, Knop M and Khmelinskii A (2021) **Timer-based proteomic profiling of the ubiquitin-proteasome system reveals a substrate receptor of the GID ubiquitin ligase.** *Mol Cell*, 81:2460-2476.e11

Li Z*, Wang Y*, Li Y*, Yin W, Mo L, Qian X, Zhang Y, Wang G, Bu F, Zhang Z, Ren X, Zhu B, Niu C, Xiao W and Zhang W (2018) **Ube2s stabilizes β-Catenin through K11-linked polyubiquitination to promote mesendoderm specification and colorectal cancer development.** *Cell Death and Disease*, 9:456

Wang J*, Zhang Y*, Hou J, Qian X, Zhang H, Zhang Z, Li M, Wang R, Liao K, Wang Y, **Li Z**, Zhong D, Wan P, Dong L, Liu F, Wang X, Wan Y, Xiao W and Zhang W (2016) **Ube2s regulates Sox2 stability and mouse ES cell maintenance.** *Cell Death Differ*, 23(3) 393-404

3 SUMMARY

The functional state of a cell is ultimately determined by the state of its proteome. Maintaining a balanced proteome requires not only correct protein synthesis and folding, but also proper trafficking of proteins to different organelles and assembly of multi-protein complexes with precise stoichiometry. Proteins that fail to reach their native compartments or assemble into native complexes are defined as mislocalized proteins. Protein mislocalization is a constitutive problem caused by intrinsic inefficiencies in cellular processes and increases with aging. Aberrant accumulation of mislocalized proteins brings harmful effects on normal cellular functions and is associated with a wide range of diseases, such as neurodegenerative diseases and some types of cancer. Therefore, eukaryotic cells have evolved protein quality control systems to prevent the abnormal accumulation of mislocalized proteins. Consequently, due to their low abundance, detecting mislocalized proteins that are subject to quality control can be challenging under non-perturbed conditions.

Here we sought to address this issue in the budding yeast *Saccharomyces cerevisiae*. By modeling protein mislocalization through overexpression of individual proteins combined with the tandem fluorescent protein timer assay, we identified about 15% of the proteome that is attenuated specifically upon overexpression. This set of attenuated proteins is strongly enriched for subunits of protein complexes. Extensive validation experiments proved the reproducibility of the high-throughput screening results. Moreover, cell culture conditions and the large size of the C-terminal tFT tag did not widely affect the protein attenuation phenotypes.

Protein abundance can be reduced by either deceleration of protein synthesis or acceleration of protein degradation. Taking advantage of the tFT readout, whereby the redFP/greenFP fluorescence ratio reports on protein degradation kinetics, we found that at least 43.5% of attenuated proteins are also destabilized. Further experiments on several randomly chosen destabilized proteins showed that they are mainly targeted for degradation by the ubiquitin-proteasome system. In order to know which E3 ligases mediate their degradation, we tested how protein abundance is affected by mutations in key E3s involved in protein quality control. Our results show that 20% of attenuated and destabilized proteins accumulate in the *ubr1*Δ, *san1*Δ or *tom1*Δ mutants, suggesting these E3s have important roles in elimination of mislocalized proteins.

This work provides a systematic survey of mislocalized proteins that are subject to quality control and suggests Ubr1, San1, and Tom1 as key quality control factors responsible for removal of mislocalized proteins. Our work paves the way for further dissecting features and mechanisms involved in the recognition of mislocalized proteins by quality control machinery.

4 ZUSAMMENFASSUNG

Der funktionelle Zustand einer Zelle wird letztlich durch den Zustand ihres Proteoms bestimmt. Die Aufrechterhaltung eines ausgewogenen Proteoms erfordert nicht nur eine korrekte Proteinsynthese und -faltung, sondern auch den ordnungsgemäßen Transport von Proteinen zu verschiedenen Organellen und den Aufbau von Multiproteinkomplexen mit präziser Stöchiometrie. Proteine, die ihre nativen Kompartimente nicht erreichen oder sich nicht zu nativen Komplexen zusammenfügen, werden als fehllokalisierte Proteine bezeichnet. Die Fehllokalisierung von Proteinen ist ein konstitutives Problem, das durch intrinsische Ineffizienzen in zellulären Prozessen verursacht wird und mit dem Alterungsprozess zunimmt. Die Anhäufung fehlgelagerter Proteine hat schädliche Auswirkungen auf die normale Zellfunktion und wird mit einer Vielzahl von Krankheiten in Verbindung gebracht, z.B. mit neurodegenerativen Erkrankungen und einigen Krebsarten. Daher haben eukaryotische Zellen Systeme zur Qualitätskontrolle von Proteinen entwickelt, um die abnormale Anhäufung fehlgelagerter Proteine zu verhindern. Aufgrund ihrer geringen Häufigkeit kann der Nachweis fehllokalisierter Proteine, die der Qualitätskontrolle unterliegen, unter ungestörten Bedingungen schwierig sein.

Wir haben versucht, dieses Problem in der Hefe Saccharomyces cerevisiae zu lösen. Durch die Modellierung der Proteinfehllokalisierung durch Überexpression einzelner Proteine in Kombination mit dem Tandem-Fluoreszenzprotein-Timer-Assay konnten wir etwa 15 % des Proteoms identifizieren, die bei Überexpression spezifisch abgeschwächt werden. Diese Gruppe von abgeschwächten Proteinen ist stark angereichert mit Untereinheiten von Proteinkomplexen. Umfangreiche Validierungsexperimente bewiesen die Reproduzierbarkeit der Ergebnisse des Hochdurchsatz-Screenings. Darüber hinaus hatten die Zellkulturbedingungen und die große Größe des C-terminalen tFT-Tags keinen großen Einfluss auf die Phänotypen der Proteinabschwächung.

Die Proteinhäufigkeit kann entweder durch eine Verlangsamung der Proteinsynthese oder durch eine Beschleunigung des Proteinabbaus verringert werden. Unter Ausnutzung des tFT-Readouts, bei dem das Verhältnis von rotemFP zu grünemFP über die Kinetik des Proteinabbaus Auskunft gibt, fanden wir heraus, dass mindestens 43,5 % der abgeschwächten Proteine auch destabilisiert sind. Weitere Experimente an mehreren zufällig ausgewählten destabilisierten Proteinen zeigten, dass sie hauptsächlich durch das Ubiquitin-Proteasom-System abgebaut werden. Um herauszufinden, welche E3-Ligasen ihren Abbau vermitteln, haben wir getestet, wie die Proteinhäufigkeit durch Mutationen in wichtigen E3s, die an der Qualitätskontrolle von Proteinen beteiligt sind, beeinflusst wird. Unsere Ergebnisse zeigen, dass sich 20 % der abgeschwächten und destabilisierten Proteine in den ubr1Δ-, san1Δ- oder tom1Δ-Mutanten akkumulieren, was darauf hindeutet, dass diese E3s eine wichtige Rolle bei der Beseitigung fehllokalisierter Proteine spielen.

Diese Arbeit bietet einen systematischen Überblick über fehllokalisierte Proteine, die der Qualitätskontrolle unterliegen, und legt nahe, dass Ubr1, San1 und Tom1 als wichtige

Qualitätskontrollfaktoren für die Beseitigung fehllokalisierter Proteine verantwortlich sind. Unsere Arbeit ebnet den Weg für die weitere Untersuchung von Merkmalen und Mechanismen, die an der Erkennung fehllokalisierter Proteine durch Qualitätskontrollmechanismen beteiligt sind.

5 INTRODUCTION

5.1 PROTEIN HOMEOSTASIS

Of the about 6000 different proteins in the budding yeast *Saccharomyces cerevisiae*, nearly all are initially synthesized by ribosomes that are free in the cytosol, and ~3000 are targeted to various cell organelles[1]. Because correct localization is important for normal cellular function, extensive efforts have been aimed at understanding how newly made proteins are targeted to the appropriate subcellular localization[2,3,4].

5.1.1 PROTEIN TARGETING PATHWAYS TO DIFFERENT ORGANELLES

5.1.1.1 PROTEIN TARGETING TO THE NUCLEUS

The nucleus is a membrane-enclosed organelle and functions as the control center of the cell because it contains the genetic material. The sorting of nuclear proteins from cytoplasm depends on nuclear localization signals (NLS). They are generally short peptide motifs and locates at anywhere within the protein sequence[2]. The classical NLSs consist of one or two clusters of basic amino acids comprised primarily of lysine (K) and arginine (R) residues[5]. For example, the NLS of simian virus 40 (SV40) large T antigen (PKKKRKV) contains five consecutive positively charged amino acids[6]. In the importin α/β pathway, classical NLSs are recognized by the importin-α subunit. Then, the importin-β transports complexes of cargo through the nuclear pore complex (NPC)[7].

Although the importin α/β pathway plays a dominant role in the nuclear import, many proteins can bypass the requirement for the importin-α subunit and bind directly to the transport receptor (karyopherin-β family proteins, Kapβs) through non-classical NLSs[8]. Unlike the classical NLSs, non-classical NLSs cannot be sufficiently described by a consensus sequence. Instead, they are described by a collection of biophysical properties such as length, charge and hydrophobicity. Several proteins with similar characteristics within their NLSs are recognized by one or two of the 14 members of the karyopherin-β family[9]. For example, Kap123 recognition motifs typically contain more than 25 amino acids and are enriched in basic amino acids, especially lysine. It is the major importer of many ribosomal proteins as well as histones H3 and H4[9,10].

Over the last two decades, a large number of NLSs have been identified and the mechanisms by which importin recognizes and transports proteins into the nucleus have been well studied. However, due to the complexity of human proteome, seven isoforms of importin α are found in mammalian cells[11]. Certain isoforms are preferentially present for certain cellular types and states. Take the well-characterized SV40 NLS as an example; it has been widely used in a variety of experiments such as CRISPR-mediated genome editing to direct recombinant proteins to the nucleus. However, its import efficiency is suboptimal in neurons[12]. Recent studies have selected uncommon NLSs or generated NLS nucleotide variants to obtain high

nuclear import dynamics, which helps diversify the implementation of NLSs in basic research and therapeutics[12,13].

5.1.1.2 PROTEIN TARGETING TO THE ENDOPLASMIC RETICULUM (ER)

The endoplasmic reticulum (ER) is a large, dynamic structure, where the vast majority of membrane proteins fold and assemble before traveling to their final destinations. Membrane protein targeting to the ER is generally mediated by the targeting sequence, which can either be a cleavable peptide, typically at the N terminus, or the first transmembrane domain (TMD) anywhere within the polypeptide[3].

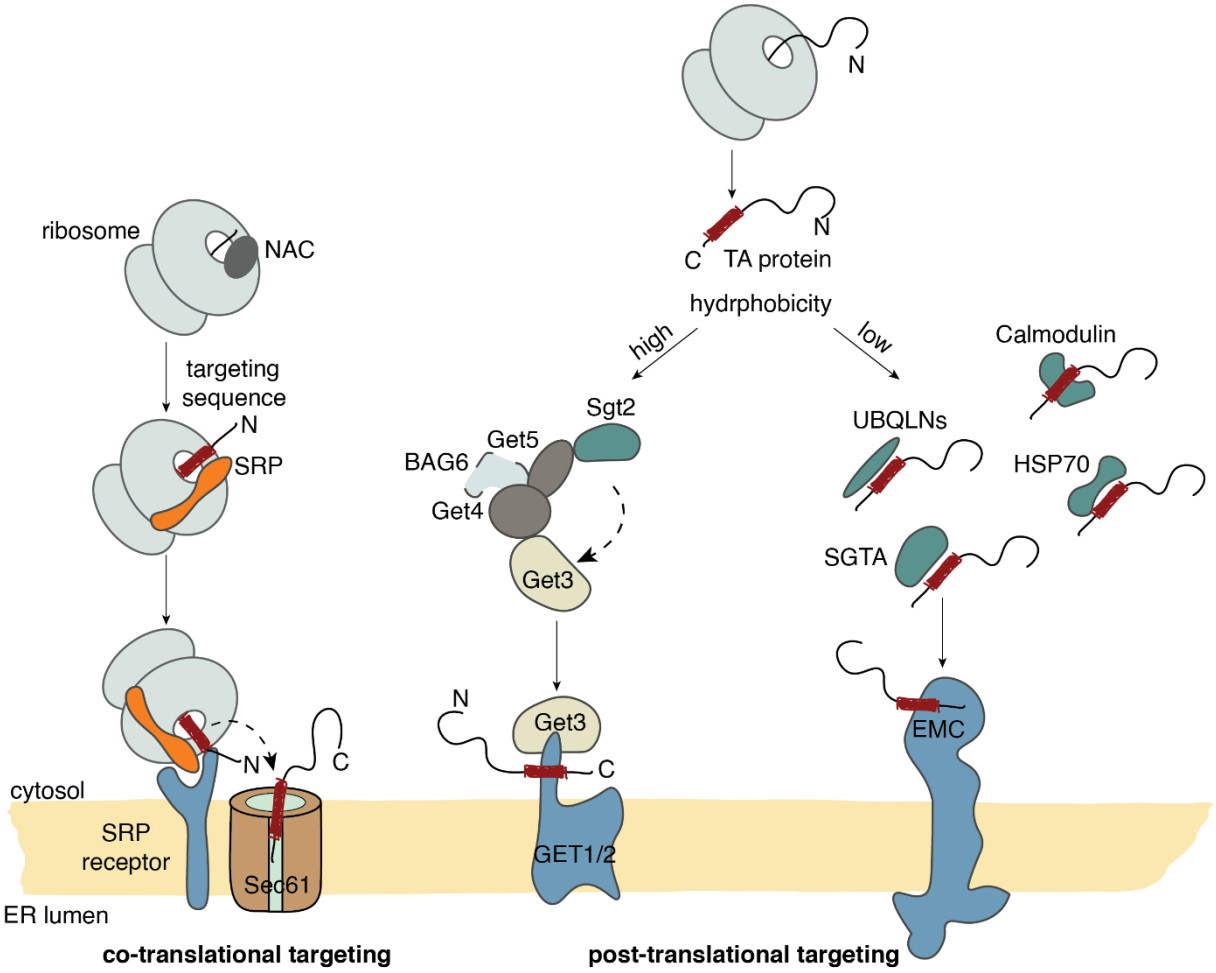


Figure 1. Protein targeting to the ER.

During co-translational targeting, the nascent polypeptide-associated complex (NAC) prevents SRP from binding to the ribosome. Once the hydrophobic targeting sequence is released from the ribosome, NAC is replaced by SRP. SRP then interacts with the SRP receptor on the ER membrane, which guides the ribosome to the ER surface where the translation is coupled with the translocation of the nascent polypeptide into the ER lumen.

Post-translational targeting of TA proteins to the ER membrane. TA proteins with the TMDs of high hydrophobicity are captured by chaperone Sgt2 with the help of the Get4/5 complex (which in mammals also contains BAG6). From Sgt2 TA proteins are transferred to Get3. Get3 interacts with the Get1/2 complex. TMDs of low hydrophobicity are kept soluble in the cytosol by cycles of binding and release from any of several chaperones. Upon the release from the chaperon, the TMD engages the cytosolic domain of the EMC.

C, carboxy terminus; N, amino terminus.

Most proteins target to the ER co-translationally, where the signal recognition particle (SRP) recognizes hydrophobic targeting sequences at the mouth of the ribosome exit tunnel and binds to the ER-localized SRP receptor[14]. Next, the targeting sequence of membrane proteins engages Sec61 translocan, which leads to the lateral gate of Sec61 opening[15,16]. From this point, the polypeptide downstream of the signal peptide is to be pulled into the pore of Sec61[17]. Further translational elongation results in translocation[3] (Figure 1). Due to the hydrophobicity, signal peptide diffuses into the membrane and is cleaved by the signal peptidase[18].

However, if the targeting sequence locates at the very C-terminus, it is only exposed after translation has terminated. Proteins must be targeted to the ER post-translationally in an SRP-independent manner[19]. This group of proteins is termed as "tail-anchored (TA) membrane proteins". The TMDs of TA proteins differ widely in hydrophobicity, a key feature that determines their mechanisms of targeting[3]. Those of high hydrophobicity are targeted by the guided entry of TA protein (GET) pathway[20]. These proteins are initially captured near the ribosome surface by the chaperone Sgt2 (SGTA in mammals) in the cytosol and transferred to ATPase Get3 (TRC40), with the aid of the Get4-Get5 complex (in mammals also contains BAG6). TA proteins that engage Get3 are targeted to the ER membrane via a receptor composed of Get1 and Get2. After TA proteins insertion into the ER, Get3 is recycled to the cytosol and initiate a new round of targeting[21] (Figure 1).

If the TMD is of low hydrophobicity, it is captured by cytosolic chaperones, like the 70 kDa heat shock proteins (HSP70), ubiquilin family proteins (UBQLNs) to keep soluble[22]. Then the TMD engage the ER membrane complex (EMC) for their biogenesis[23]. Protein targeting to the ER is a comparatively mature area. It has been known that high-fidelity targeting not only relies on the selective recognition by TMD recognition factors but also prompt clearance if targeting fails. Thus, various protein quality control factors have been identified which will be discussed in detail in the later section.

5.1.1.3 PROTEIN TARGETING TO MITOCHONDRIA

Mitochondria are crucial for numerous cellular processes and are known as the powerhouse of the cell since they produce the bulk of cellular ATP. These organelles are surrounded by a double-membrane system, consisting of the inner and outer membranes. Yeast mitochondria contain about 1000 different proteins[24]. Of these, 99% are synthesized as precursors on cytosolic ribosomes and have to be imported into mitochondria[25].

So far, the targeting steps occurring in the cytosol are only understood in part. Translocation into the mitochondrial membranes has been well studied. The translocase of the outer membrane (TOM complex) imports most of the precursor proteins into mitochondria[25]. The central component of TOM complex is Tom40, which forms the channel for the translocation of unfolded or partially folded proteins across the outer membrane[26,27]. The peripheral receptors Tom20 and Tom70 recognize precursor proteins or associated chaperones and then

deliver precursors to the central receptor Tom22[28,29]. From here, the precursors are inserted into the Tom40 channel (Figure 2).

Next, three principal pathways direct proteins to their intramitochondrial destinations. First, the translocase of inner mitochondrial membrane 23 (TIM23) and its associated motor (PAM) transport precursor proteins with a cleavable presequence into the inner membrane or matrix (Figure 2, I)[28,30]. The reaction cycle of the TIM23 complex utilizes ATP and the membrane potential across the inner membrane generated by the activity of the respiratory chain[31]. Upon import, the mitochondrial processing peptidase (MPP) removes the presequence. The presequence pathway represents the major import route into mitochondria[25]. Second, the carrier translocase (TIM22 complex) inserts proteins with multiple transmembrane segments like carrier proteins into the inner membrane (Figure 2, II)[32]. Third, the mitochondrial intermembrane space import and assembly (MIA) machinery transports cysteine-rich proteins into the intermembrane space (Figure 2, III)[33].

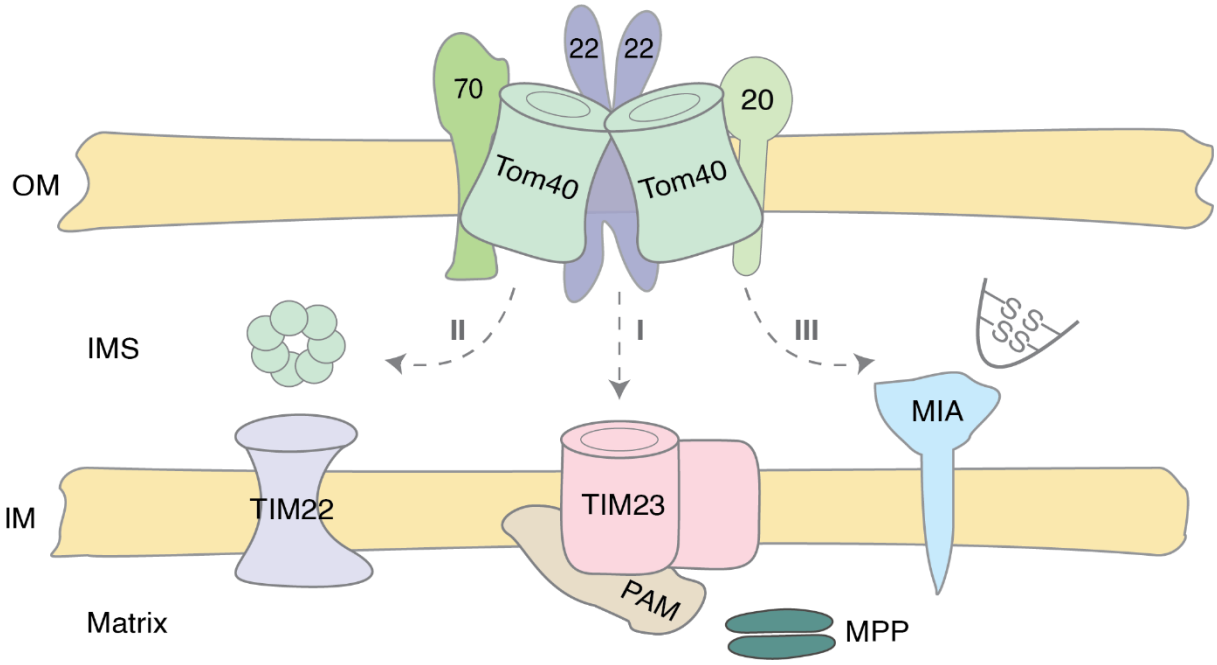


Figure 2. Mitochondrial protein import pathways.
The TOM complex forms the entry gate for most mitochondrial precursor proteins. After passage of the TOM channel, several protein translocases sort the proteins into the mitochondrial subcompartments. (I) The presequence translocase (TIM23 complex) transports proteins with a cleavable presequence into the inner membrane (IM) and mitochondrial matrix. The PAM drives protein transport into the matrix in an ATP-dependent manner. Upon import, the mitochondrial MPP removes the presequence. (II) The carrier translocase (TIM22 complex) inserts precursor proteins with multiple transmembrane domains into the inner membrane. The small TIM proteins transport these precursor proteins from the TOM complex to the TIM22 complex. (III) The mitochondrial intermembrane space import and assembly (MIA) machinery drives via oxidative folding the transport of proteins with cysteine-rich motifs into the intermembrane space (IMS).

Proteins targeting to mitochondria can occur pre-, co- or post- translationally[4]. Proximity-specific ribosome profiling revealed that mRNA of several mitochondrial proteins, particularly inner membrane proteins, are enriched at the mitochondrial surface[34]. Such mRNA targeting is mediated by the member of the Pumilio-homology domain family protein 3 (Puf3)[35] (Figure 3). Although most studies reported that Puf3 is localized to the cytosol, several findings show

that a portion of Puf3 co-localizes with mitochondria and specifically binds to the 3' untranslated regions of mRNAs[36,37,38]. However, the role of Puf3 in mRNA targeting is still being debated, as it is unclear whether translation can be initiated specifically on the mitochondrial surface.

According to the current view, up to one-third of the mitochondrial proteins target to mitochondria co-translationally[25]. At the ribosome exit tunnel, the nascent polypeptide associated complex (NAC) plays an important role in the correct targeting of precursor proteins to mitochondria[39,40]. NAC binds to the outer membrane proteins OM14 and Sam37, which could facilitate protein import into mitochondria[41] (Figure 3). Moreover, deletion of NAC in *C. elegans* results in proteins mislocalized to the ER[42]. Although it is not clear which cluster of mitochondrial proteins are translated specifically on mitochondria-bound ribosomes, it suggests that co-translational import is one of the mechanisms to promote mitochondrial targeting specificity.

Actually, the majority of precursor proteins are targeting to mitochondria in a post-translational manner, although they need to spend more time in the aqueous cytosol. After synthesis, chaperones of heat shock protein families Hsp70 and Hsp90 bind to the precursor proteins, which keep them in an unfolded, import-competent state[43]. Hsp70 is a high abundant chaperone that function in different processes in the cytosol, including protein folding, targeting, disaggregation and protein turnover[44]. It can bind not only to the receptor of mitochondrial outer membrane (OM), but also to the TMD of ER-destined proteins. How to keep the specificity for mitochondrial targeting?

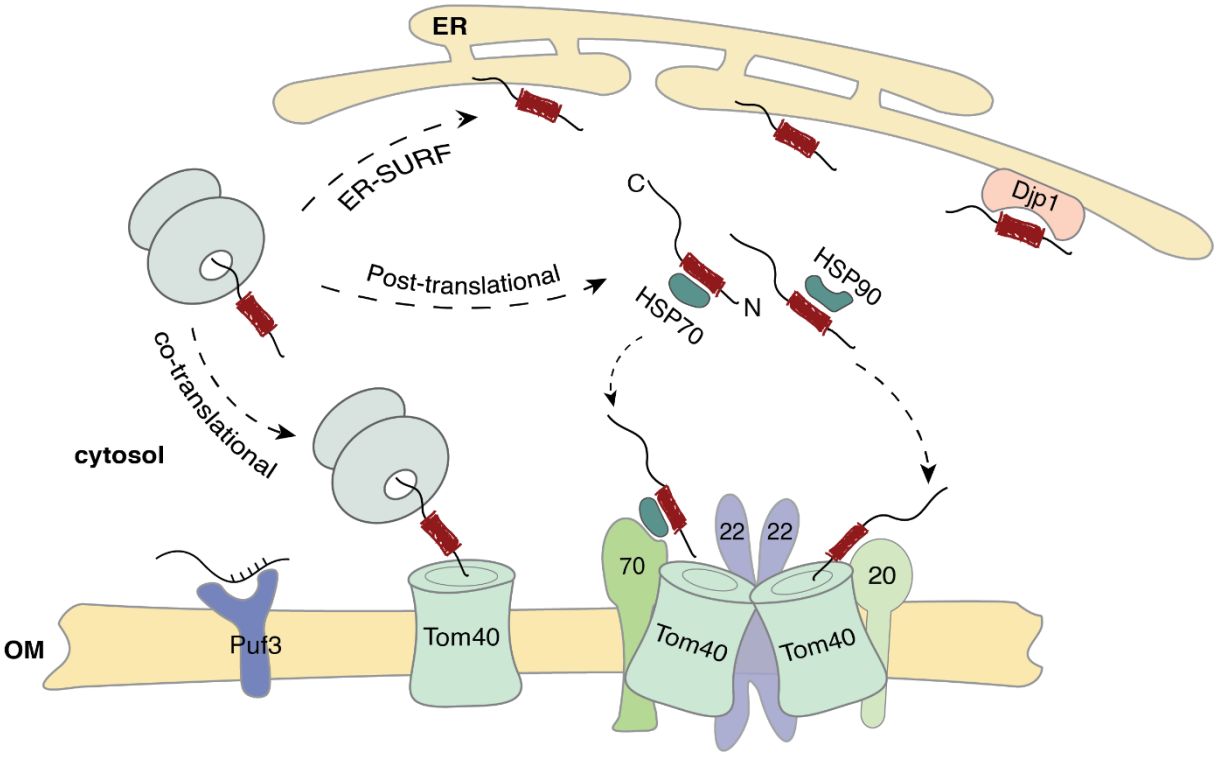


Figure 3. Protein targeting to the mitochondria.

After mRNAs exit the nucleus, they can be targeted to the mitochondrial surface via RNA-binding proteins Puf3. If translation is initiated in the cytosol, ribosome- nascent chain complexes (RNCs) can be targeted to the OM via targeting sequences (co-translational). If the precursor synthesis was completed in the cytosol, cytosolic factors can bind the precursor, keep it in an unfolded, import-competent, state, and facilitate its transfer to mitochondrial receptor proteins (Post-translational). Alternatively, some precursors can be first targeted to the ER and then transferred to mitochondria via the ER surface-mediated protein targeting (ER-SURF) pathway.

Some proteins contain mitochondrial targeting sequences (MTS) that direct the transport of a protein to the mitochondria. They are usually short peptides, about 15–70 amino acids long, bearing positively charged basic residues. Canonical MTSs are located at the N-terminal within the protein sequences, which are recognized by Tom20 receptor at mitochondria OM and cleaved by MPP after import[45]. However, numerous multispanning proteins have non-cleavable and internal MTS. Their recognition relies on Tom70 receptor (Figure 3). Several studies revealed that the recognition spectrum of Tom70 is much larger than previously thought, compromising hydrophobic proteins without MTS and a docking site for Hsp70[43,46]. This reflects different sorting routes for mitochondrial proteins. A recent study found that mitochondrial proteins can be transported from the ER surface and then reroutes to mitochondria with the aid of the ER-localized chaperone Djp1[47] (Figure 4).

5.1.2 PRODUCTION AND ASSEMBLY OF PROTEIN COMPLEXES

In addition to ensuring that proteins are correctly sorted into their native compartments, more than half of all proteins require the assembly of multi-protein complexes with precise stoichiometry[48]. In prokaryotic organisms, different subunits of the same complex are often co-translated from a single mRNA molecule (operon) with a production rate proportional to their required stoichiometry[49]. However, operons are relatively uncommon in eukaryotic organisms. The genome complexity of eukaryotes is also much higher than prokaryotes. How do eukaryotes maintain relative proportions of protein complex components?

Ribosome profiling data has shown that subunits are produced in the correct relative proportions for vast majority of protein complexes in budding yeast. This proportional synthesis also applies to large complexes in higher eukaryotes. Later on, proteomic studies have revealed that actually around 10%-15% of protein complex subunits are made in excess and subsequently degraded[50]. Therefore, eukaryotes use a combination of proportional synthesis and post-translational degradation to achieve precise proportions of different subunits within complexes.

Following protein synthesis, the formation of complexes can occur either co-translationally, where a protein interacts with a partner as it emerges from the ribosomal exit tunnel, or through the formation of complexes between existing proteins within the cell[51].

5.2 PROTEIN MISLOCALIZATION

5.2.1 MISTARGETING

Protein targeting is a process in which a targeting factor recognizes a targeting signal within the newly synthesized protein to guide it to the appropriate organelle. However, due to the intrinsic limitations of specific binding, a certain degree of inefficiency is unavoidable. In addition, targeting pathways face the challenge that targeting signals are characterized by certain biophysical features rather than distinct amino acid sequences. The similarity or overlapping of these biophysical characteristics among different organelles imposes a major challenge of achieving high-fidelity targeting. Studies in mammalian cells have shown that the efficiency of protein trafficking to ER via SRP pathway ranges between 95% and 60%, which indicates that even with the best signal there is still constitutive low-level mislocalization[52]. This problem also exits for TA proteins. When certain mitochondrial OM proteins are overexpressed, they are mislocalized to ER via a Get3-dependent manner[53]. These studies demonstrate that intrinsic mistargeting occurs even under normal circumstances.

The failure rate of protein targeting increases when targeting signals or factors are lost or disrupted. Missense mutations within the NLS motif are capable of abrogating protein nuclear localization. Nuclear-localized SV40 large T antigen was detected in the cytosol when the third amino acid lysine of the NLS was mutated to threonine (PK**K**KRKV→PK**T**KRKV)[54]. Moreover, some ER-destinated proteins are mistargeted to mitochondria in the absence of SRP[55]. Additionally, large-scale protein mislocalization can occur when cells are exposed to environmental stresses. For example, acute ER stress attenuates translocation of secretory and membrane proteins[56]. Oxidative stress inhibits the mitochondrial import capacity, leading to the accumulation of precursor proteins in the cytosol[57].

5.2.2 UNASSEMBLED SUBUNITS OF MULTI-PROTEIN COMPLEXES

Despite the fact that vast majority of protein complexes in budding yeast are produced with defined stoichiometry, 10%-15% of protein complex subunits are produced inherently in excess and cannot assemble into complexes[50]. While this may appear to be a waste of resources, there can be advantages to producing extra subunits. For example, some complexes use a shared scaffold or catalytic core with exchangeable subunits that provide substrate specificity in order to respond to a potential rapid change in its environment. Therefore, even under normal circumstances some subunits fail to assemble into complexes.

Moreover, negative feedback regulation in general is uncommon in yeast. This conclusion is based on the observation that when one of the two genomic copies of 730 different genes was eliminated, more than 80% of the tested genes exhibited a 50% reduction in protein level[58]. When transient imbalances in gene copy number during DNA replication or fluctuations in gene expression occur, the absence of negative feedback regulation may result in the production of complex subunits that lack their binding partners.

Aneuploidy is a condition of having one or more extra chromosomes that is not an exact multiple of the haploid complement. This condition leads to a dramatic alteration in gene dosage, which in turn causes corresponding changes in protein levels produced by genes located on the aneuploid chromosomes. As a result, dramatic alterations in cellular protein composition could significantly increase the burden of unassembled subunits of multi-protein complexes.

5.3 PROTEIN MISLOCALIZATION AND DISEASES

Accurate protein targeting to their native compartments, or their precise assembly with other proteins in the correct proportions, is an indispensable process for proper protein function within the cell. However, when a protein is mislocalized, it may prevent the protein from carrying out its intended function, which can lead to diseases or other deleterious effects on the cell. Mutations in targeting signals, which result in proteins failing to reach their proper subcellular compartments, are associated with a wide range of diseases. Table 1 provides several examples of mutations in signal peptides or NLSs that cause diseases.

Table 1. Genetic alterations in localization signals associated with human genetic disorders

Gene	Disease	Mechanism	Mislocalization
CASR	Familial hypocalciuric-hypercalcemia; Neonatal hyperparahyroidism	Mutation of signal peptide[59]	Impaired protein transportation to cellular membrane
DSPP	Dentinogenesis Imperfecta	Mutation of signal peptide[60]	Failure of protein translocation into the ER
SRY	Swyer syndrome	Mutation of NLS[61]	Loss of nuclear localization
SHOX	Léri–Weill dyschondrosteosis	Mutation of NLS[62]	Cytoplasmic retention
FOXP2	Speech–language disorder	Mutation of NLS[63]	Loss of nuclear localization
AGT	Primary hyperoxaluria type 1	Polymorphism and/or mutation[64]	Mitochondrial mislocalization

Moreover, mutations in the complex interface that impair assembly of the intact complex can also cause diseases. The most common set of diseases involving the defect of complex assembly are the Thalassemias, which is caused by the reduced level of mature hemoglobin[65]. Hemoglobin consists of α- and β-globin. Mutations in the α-β interface impede the hemoglobin formation, which leads to hemolytic anemia that is related to the Thalassemias[66].

Proteins that aberrantly accumulate in the wrong compartment are also related to various diseases. Abnormal accumulation of certain proteins leads to the formation of toxic

aggregates, which can damage cells and potentially lead to cell death. Formation of amyloid-like aggregates in the central nervous system is a hallmark of many neurodegenerative diseases such as Alzheimer's and Parkinson's disease. Table S2 lists some examples of aberrant accumulation of mislocalized proteins that cause neurodegenerative diseases.

Table 2. Aberrant accumulation of mislocalized proteins cause human diseases

Gene	Disease	Mechanism	Mislocalization
APP	Alzheimer's disease	Accumulation of beta-amyloid plaques[67]	Outside of neurons in the brain
MAPT	Alzheimer's disease	Accumulation of tau protein[67]	From the axons and dendrites of neurons to the cell body
TARDBP	Amyotrophic lateral sclerosis (ALS)	Accumulation of TDP-43[68]	Cytoplasmic retention
HNRNPA1	ALS	Accumulation of hnRNPA1[69]	From the nucleus to the cytoplasm
FOXP2	ALS	Accumulation of FUS[70]	From the nucleus to the cytoplasm

5.4 THE UBIQUITIN-PROTEASOME SYSTEM ENSURING PROTEIN HOMEOSTASIS

In order to minimize the detrimental impact of mislocalized proteins, eukaryotic cells have evolved quality control systems to handle these proteins. One of the key mechanisms is protein degradation via the ubiquitin-proteasome system (UPS).

The UPS is a complex biochemical pathway that plays a critical role in the regulation of cellular processes, including protein turnover, cellular signaling, and the control of gene expression[71]. Ubiquitination is the process by which ubiquitin is covalently attached to a target protein. This process is carried out by a cascade of enzymes, including the E1 ubiquitin-activating enzyme, the E2 ubiquitin-conjugating enzyme, and the E3 ubiquitin ligase. The E1 enzyme activates the ubiquitin molecule and transfers it to the E2 enzyme. The E3 ligase facilitates the transfer of ubiquitin from the E2 enzyme to the target protein, forming an isopeptide bond between the C-terminal glycine residue of ubiquitin and a lysine residue on the target protein[72] (Figure 4).

Specificity in the UPS is primarily determined by E3s[73]. There are ~100 putative E3s in yeast, all falling into two major classes: HECT (homologous to E6-AP C-terminus) and RING (really interesting new gene) E3s[71]. RING and HECT domain E3s follow distinct mechanisms to catalyze ubiquitylation. HECT domain E3s contain an active site cysteine within the HECT domain, which forms a thioester with ubiquitin received from an E2 prior to its transfer to the substrate[74]. RING E3s do not form thioester intermediates; they instead facilitate the transfer

of ubiquitin by bringing the charged E2~Ub in close proximity to the acceptor lysine in the substrate[75].

Ubiquitin can tag a protein as a single molecule through mono-ubiquitination or as a polyubiquitin chain by attaching ubiquitin molecules on top of each other[72]. Ubiquitin attach at different lysine residues, generating M1, K6, K11, K27, K33, K48, and K63 linkages[76]. The type of linkage can affect the ultimate fate of the target protein. K48-linked polyubiquitination are typically recognized by the proteasome and targets substrates for proteasomal degradation[71]. Ubiquitin can be removed from substrates by deubiquitinating enzymes (DUBs), preventing its degradation or altering its function or localization within the cell[77]. During the last decade, various components of the UPS have been identified as quality control factors of mislocalized proteins, which will be described in the following section.

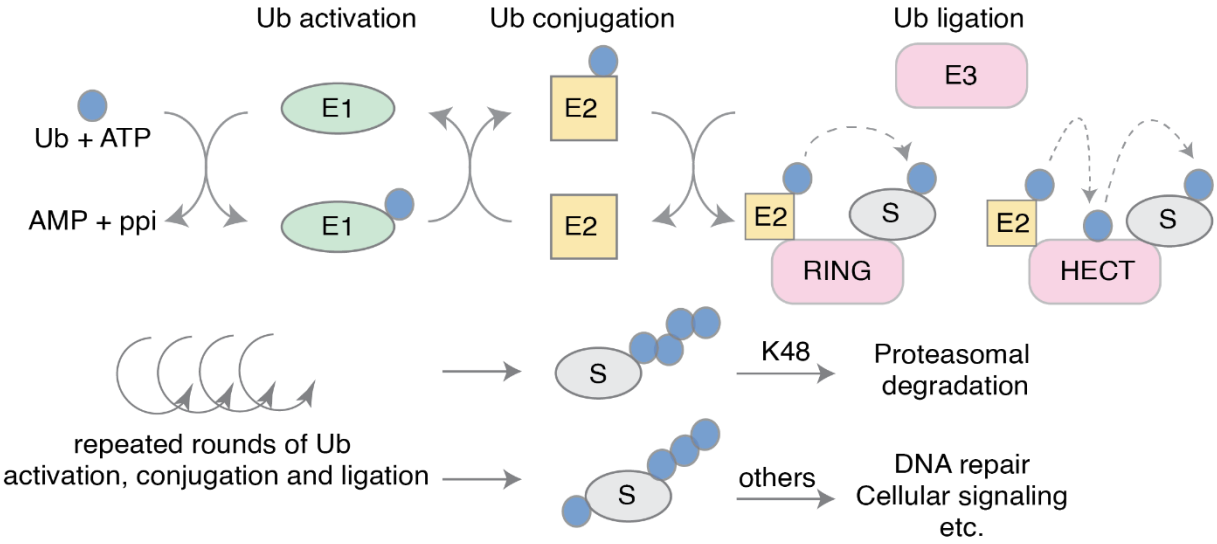


Figure 4: Protein ubiquitylation.
Activation of Ub requires the hydrolysis of ATP by an Ub-activating enzyme (E1), which is followed by the covalent conjugation to an Ub-conjugating enzyme (E2) and the ligation by an Ub-ligase (E3) to the substrate. The process can be repeated, which results in the formation of different homotypic or heterotypic Ub chains. Chains of the K48 linkage-type are associated with proteasomal degradation of the substrate protein. Other chain types have been reported to be involved in other biological processes such as DNA repair or cellular signaling. S: Substrate.

5.5 QUALITY CONTROL OF MISLOCALIZED PROTINS

5.5.1 PROTEIN QUALITY CONTROL IN THE CYTOSOL

As mentioned earlier, many secretory or mitochondrial proteins require chaperones to maintain them in an unfolded, import-competent state until they reach their destination organelles[43]. Therefore, it is a simple fact that these proteins do not fold properly in the cytosol. Due to mislocalized proteins that are subject to quality control are not easily detected under normal conditions, previous studies in yeast used Δss-CPY* as a model mislocalized substrate[78] to dissect quality control mechanisms. This model was created by removing the ER targeting signal from the misfolded vacuolar protease carboxypeptidase Y (CPY*), which restricts the protein to the cytosol[79]. Genetic screening identified the E3 ligase Ubr1 as a

necessary factor for the degradation of Δss-CPY*[78]. Subsequently, additional misfolded/mislocalized substrates of Ubr1 were discovered[80], broadening the scope of Ubr1's role as a quality control factor in the cytosol.

Ubr1 is a single-subunit RING-type E3 ligase and catalyzes K48-linked ubiquitination[81]. It has been well studied in the Arg/N-degron pathway[82]. This pathway targets proteins for degradation through recognition of the N-terminal residues termed N-degrons[83]. In *S.cerevisiae*, Ubr1 contains two types of substrate-binding sites: the type-1 site recognizes substrates starting with basic residues, while the type-2 site is specific for bulky hydrophobic residues[84]. These binding sites allow Ubr1 to recognize and target specific proteins for degradation via the Arg/N-degron pathway. Additionally, functioning as a quality control factor in the cytosol, Ubr1 works together with chaperone proteins such as Ssa1 and Sse1 to recognize and ubiquitinate substrates[85]. Therefore, the recognition mechanism of Ubr1 in this process is distinct from that in the Arg/N-degron pathway.

The studies in yeast uses model substrates lacking the signal sequence to identify quality control factors in the cytosol. In mammalian cells, different quality control factors have been identified by performing functional reconstitution assay. Radiolabeled secretory or membrane proteins are synthesized in rabbit reticulocyte lysate. The absence of the ER or mitochondria confines proteins to the cytosol.

Using this system, it was discovered that Bag6 is a component of the cytosolic quality control pathway[86]. Bag6 acts as a co-chaperone and works together with other E3 ligases to transfer ubiquitin to substrate proteins. When protein targeting fails, Bag6 recognizes the exposed targeting signal of membrane proteins in the cytosol and recruits the ubiquitin ligase RNF126 to facilitate substrate ubiquitination and subsequent proteasome degradation[87]. In addition to mediating ubiquitination, Bag6 also plays an important role in the GET pathway, where it facilitates the targeting of TA proteins to ER[88]. At this point, Bag6 is at the center of a triage reaction that routes proteins towards targeting (in the case of TA proteins) or degradation (for other membrane proteins)[86].

In addition to Bag6, proteins of the Ubiquilin family (UBQLNs) also function as key cytosolic quality control factors for membrane proteins. However, unlike Bag6, UBQLNs engage substrates dynamically. During an initial period, UBQLNs bind to the targeting signal of membrane proteins, which provides an opportunity for them to engage with the insertion machinery. Over time, as would be the case if import fails, UBQLNs recruit a yet-unidentified E3 ligase to mediate substrate degradation[89]. The dual functions of UBQLNs also reveal the dual role of targeting signals in membrane protein biosynthesis and quality control.

Current research on cytosolic quality control factors is not limited to mislocalized ER- and mitochondria-destined proteins. Dedicated factors have also been identified that recognize unassembled proteins of multisubunit complexes in the cytosol for degradation. To date, the Ac/N-degron pathway is the best characterized quality control pathway for unassembled cytosolic subunits. In S. cerevisiae, more than 60% of proteins are irreversibly Nα-terminally

acetylated (Nt-acetylated) by Nt-acetylases[90]. The Ac/N-degron pathway targets proteins for degradation through their Nt-acetylated residues by Doa10 and Not4[90]. Several multi-protein complexes containing N-acetylated subunits are stable in the intact complex, because their N termini are shielded by other subunits. Failure to assemble into a complex (or complex disassembly) exposes the N-terminal degron, resulting in subunits degradation[91]. Examples include the Cog1 subunit of the yeast COG complex and the Hcn1 subunit of the *S. pombe* APC/C complex.

While many substrates may carry degrons, for those that do not have a predicted unstable terminal, which quality control factors are responsible and what is the recognition mechanisms? Recent study has identified the hybrid E2/E3 enzyme UBE2O as an important quality control factors that recognizes the exposed assembly interface, which is normally shielded in the intact complex, and mediate unassembled subunits for degradation[92]. For example, during adult hemoglobin assembly, α-globin uses α -haemoglobin stabilizing protein (AHSP) to shield the interface that is eventually occupied by β-globin. Mutations in α-globin impairs the interaction with AHSP (and β-globin). Exposed assembly interface is recognized and ubiquitinated by UBE2O. The key elements for UBE2O recognition are juxtaposed basic and hydrophobic patches, a feature also seen on the surface of nascent ribosomal proteins. Therefore, in addition to α-globin, some unassembled ribosomal proteins are also targeted by UBE2O for degradation before they import into nuclear[92]. The ubiquitination of a subset of substrates by UBE2O requires chaperone NAP1L1 as an adapter[93].

In separate studies, unassembled ribosomal proteins are recognized and ubiquitinated by the HECT domain E3 ligase Tom1 in yeast[94]. The ubiquitination sites of Tom1 suggest that it probably interacts with basic regions that would ordinarily interact with rRNA in nucleus[94], suggesting that the exposed interface plays a crucial role in the recognition of unassembled ribosomal subunits by Tom1. In mammals, the closest homolog of Tom1 is HUWE1. Proteomic analysis identified 72 potential HUWE1 substrates (but no ribosomal proteins), almost all of which are known subunits of multi-protein complexes[95]. Most of these substrates operate in the nucleus, yet HUWE1 appears to function in the cytosol. Thus, HUWE1 may identify proteins that fail localization, assembly, or both. The molecular features shared by HUWE1/Tom1 clients and the basis of their recognition remain to be addressed.

Furthermore, a recent study has shown that not only individual unassembled subunits but also the assembly intermediates can be targeted for degradation[96]. During complex assembly, unassembled subunits of a complex are held temporary by an assembly factor, until it is displaced by the appropriate interacting subunits. Over time, if the assembly factor retains, as would be the case of failed assembly, the assembly intermediates can be recognized as abnormal and targeted for degradation. An example of this is the proteasome assembly process. The 26S proteasome consists of the catalytic 20S proteasome and the 19S regulatory particle, both of which are composed of a set of multiple distinct subunits[97]. During proteasome assembly, the assembly factor PAAF1 binds to the regulatory subunit PSMC5 to promote its incorporation into the 19S regulatory particle of the proteasome. However, if

PSMC5 fails to incorporate properly, it can remain associated with PAAF1 and other assembly intermediates for an extended period of time. Eventually, this abnormal complex can be recognized by the HECT E3 ligase HERC1 and targeted for degradation[96].

5.5.2 PROTEIN QUALITY CONTROL AT THE ER

Misfolded or improperly targeted proteins to the ER can be removed by the ER-associated degradation (ERAD) pathway. The ERAD pathway involves a complex series of steps, which involve recognizing misfolded or mislocalized proteins, ubiquitinating them through E3 ubiquitin ligases, extracting them from the ER membrane via AAA ATPase Cdc48, retrotranslocating them back into the cytosol, and ultimately degrading them through the 26S proteasome[98]. There are several different E3 ligases involved in ERAD, each of which recognizes specific types of substrates. Hrd1 is a multi-pass transmembrane protein that functions as a key component of the Hrd1 complex[99]. This complex, including Ubc1 and Ubc7, the E2 ubiquitin-conjugating enzymes, is responsible for recognizing exposed hydrophobic and charged residues of the misfolded proteins from the ER lumen and then retrotranslocating to the cytosol for subsequent degradation[100]. Doa10 is another E3 ubiquitin ligase that functions in the ERAD pathway. It localizes throughout the ER[98]. Therefore, the substrate spectrum of Doa10 is quite broad, including misfolded, mistargeted TA proteins, or unassembled subunits[101,102,103]. Additionally, the Asi1/3 complex is responsible for removal of misfolded or mislocalized proteins at the inner nuclear membrane[104,105,106].

5.5.3 PROTEIN QUALITY CONTROL AT THE MITOCHONDRIA

Protein import into mitochondria can be compromised when translocating proteins are stuck in the protein translocases. A recent study revealed that Ubx2, a protein involved in the ERAD pathway, is also present in the OM of mitochondria. In this location, it recruits the AAA ATPase Cdc48 to extract precursor proteins that have become arrested in the TOM channel, facilitating their degradation by the proteasome. This pathway mechanically resembles the ERAD. Thus, it is termed as mitochondrial protein translocation-associated degradation (mitoTAD) pathway.

Moreover, defects in the targeting pathways can lead to mislocalization of ER- or peroxisome resident proteins to the outer mitochondrial membrane. The membrane embedded AAA ATPase Msp1 (ATAD1 in mammals) extracts mislocalized TA proteins from mitochondria outer membrane and facilitates their transfer to the ER, where they are ubiquitylated by the membrane bound E3-ubiquitin ligase Doa10[101,107]. Subsequently, Cdc48 delivers these proteins for proteasomal degradation.

5.6 AIM OF THIS STUDY

Eukaryotic cells have evolved protein quality control systems to prevent the abnormal accumulation of mislocalized proteins. Consequently, due to their low abundance, detecting mislocalized proteins that are subject to quality control can be challenging under non-perturbed conditions. Recent studies have mainly used artificial substrates to investigate the mechanisms that control the mislocalization of proteins. Several factors have been identified, which are a significant discovery. However, for the majority of the proteome, the systems responsible for quality control of mislocalized proteins remain unknown.

The main goal of this study is to systematically examine the quality control mechanisms of mislocalized proteins. To achieve this goal, we established a platform to promote protein mislocalization and used the tFT as a reporter to track the fate of individual proteins.

Our first objective was to identify potential substrates that are subject to quality control when proteins are mislocalized. When proteins are mislocalized, they are eliminated by the protein quality control systems, which leads to a reduction in protein levels.

This decrease in protein abundance can be caused by deceleration of protein synthesis or acceleration of protein degradation. Our second objective was to identify the subset of proteins that are being degraded upon protein mislocalization.

Next, by focusing on this subset, we aimed to determine which known quality control factors are primarily responsible for the removal of mislocalized proteins.

6.1.1 MODELING PROTEIN MISLOCALIZATION THROUGH OVEREXPRESSION COMBINED WITH THE TANDEM FLUORESCENT PROTEIN TIMER ASSAY

Due to limiting amounts of endogenous binding partners, proteins produced in excess will either fail to assemble into native complexes or end up in the wrong compartment. Based on this assumption, we established a platform to promote protein mislocalization through overexpression of individual proteins. In order to systematically identify potential substrates of mislocalized proteins, we monitored the abundance and turnover of each protein by the tandem fluorescent protein timer (tFT) expressed from the genomic locus when the same protein without the tFT tag is overproduced from a multicopy plasmid (Figure 5a).

We used a plasmid carrying the *YGR045C* gene, which encodes an uncharacterized and putative protein of unknown function, as a comparison to simulate the native condition. Upon overexpression, if the protein is not subject to quality control, its abundance or turnover should be similar to the native condition. However, if the amount of the timer-tagged protein is reduced, we consider this protein is a potential substrate of mislocalized proteins (Figure 5b).

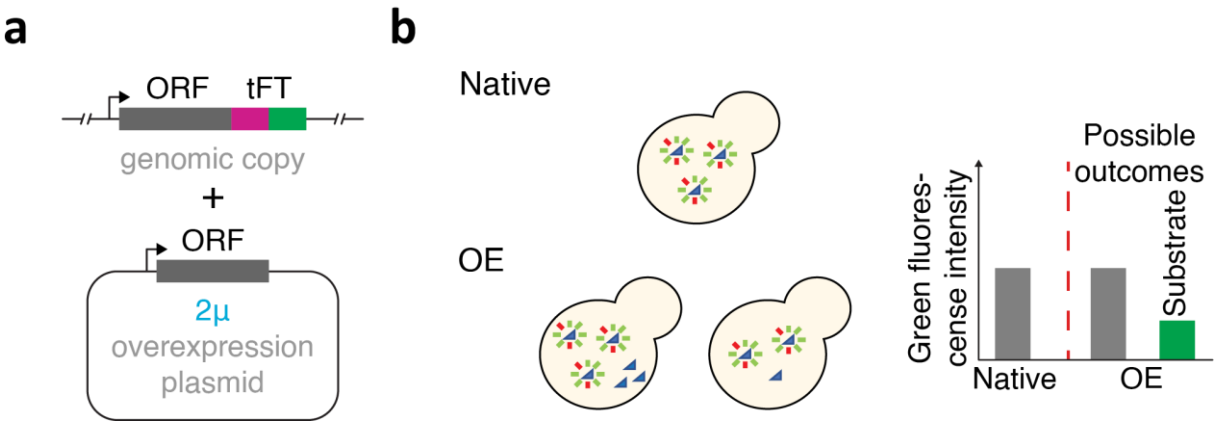


Figure 5. (a) Cartoon of modeling protein mislocalization through overexpression combined with the tandem fluorescent timer to follow the fate of each protein. (b) A tFT strain carrying a control plasmid gives the fluorescence level of the native condition (Native). The same tFT strain carrying the corresponding untagged overexpression plasmid gives the fluorescence level of the overexpression condition (OE) with two possible outcomes. Compared to the native condition, proteins that show reduced green fluorescence intensities upon overexpression are potential substrates of quality control.

7 RESULTS

7.1 PROOF OF PRINCIPLE

A previous study revealed that ribosomal proteins produced in stoichiometric excess over other ribosomal subunits fail to assemble into ribosomes and are rapidly degraded by the ubiquitin-proteasome system[108]. In order to verify the setup of overexpression combined with tFT assay for dissecting quality control mechanisms of mislocalized proteins, I used one of the ribosomal subunits, Rpl26b, as proof of principle.

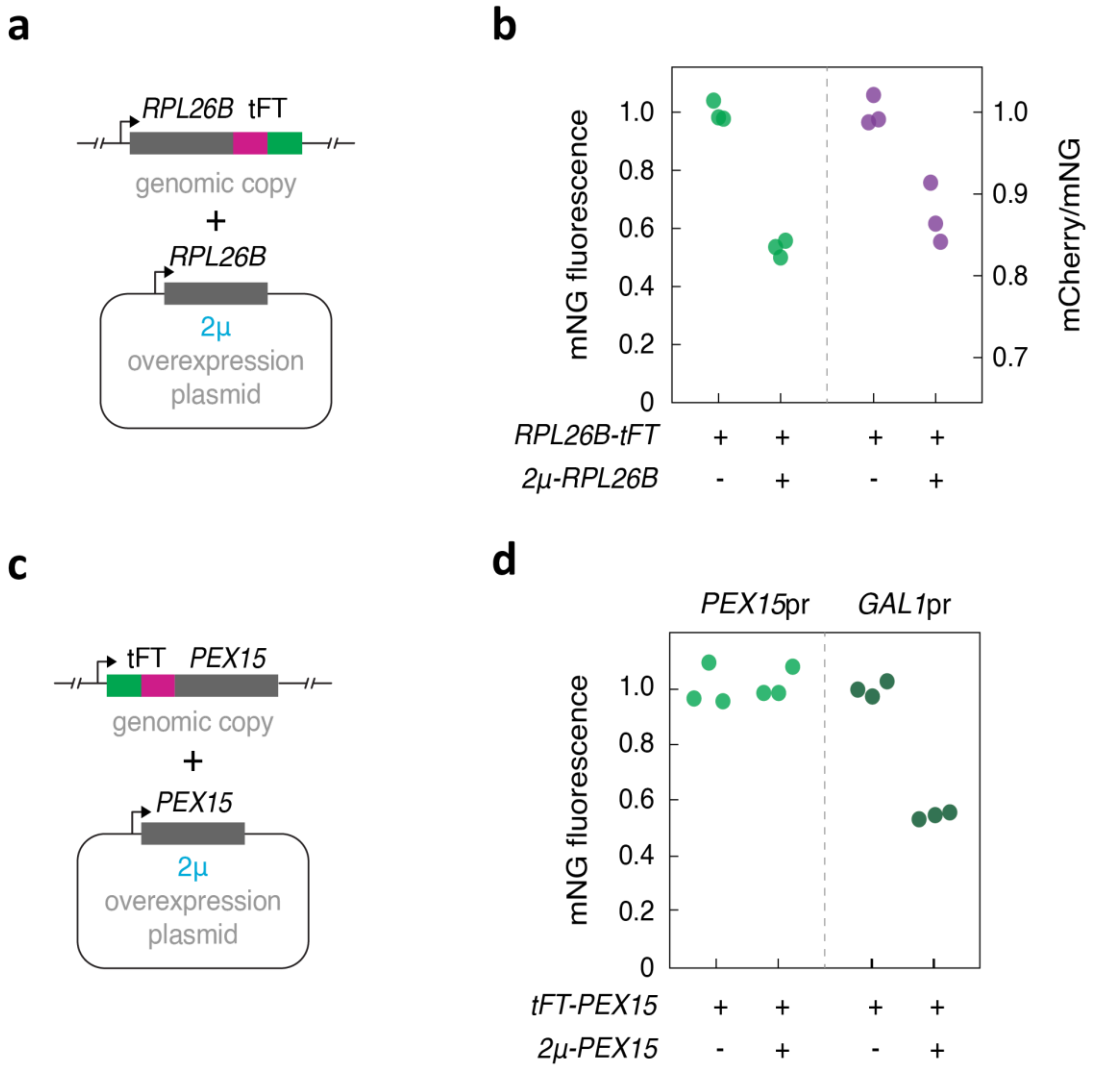


Figure 6. (a) Cartoon of Rpl26b overexpression combined with the tFT assay. (b) The mNG intensity (green dots) and ratio (purple dots) showed attenuation when the Rpl26b protein expressed under the regulation of its native promoter and terminator from 2μ plasmids. Relative protein abundance and turnover were normalized to the native condition, in which the *RPL26B-tFT* strain was transformed with the plasmid carrying the uncharacterized gene *YGR045C*. (c) Cartoon of Pex15 overexpression combined with the tFT assay. (d) Fluorescence measurements of *tFT-PEX15* strains. Left: the expression of *PEX15* was regulated by its native promoter and terminator from 2μ plasmids. Right: the expression of *PEX15* was controlled by the *GAL1* promoter and terminator from 2μ plasmids. Cells were grown in galactose-containing medium for 4h30min to induce Pex15 protein expression. Relative protein abundance were normalized to the native condition.

Rpl26b was endogenously tagged at the C-terminus with the mCherry-mNeonGreen (mNG) tFT. *RPL26B* plasmids were isolated from the Molecular Barcoded Yeast ORF 2.0 (MoBY2.0) library, in which each gene was cloned into a 2micron (2µ) high-copy plasmid backbone (p5476) and controlled by its native promoter and terminator[109]. By measuring mCherry and mNG fluorescence intensities, we found *RPL26B-tFT* stain carrying the corresponding overexpression plasmid exhibit attenuated protein abundance and stability (Figure 6a and b). This result demonstrates that ectopically expressed proteins from high-copy plasmids result in failure of proteins to assemble into native complexes. Since unassembled ribosomal subunits are targeted for protein degradation, attenuation in fluorescence intensities detected from the genomic locus proves our assumption that quality control systems cannot distinguish the tFT-tagged endogenous protein from the non-tagged exogenous protein is true.

Another protein used as proof of principle is Pex15 which is localized to peroxisomal membrane in wild-type cells. Studies have shown that Pex15 mistargeted to the mitochondrial outer membrane is extracted by the AAA-ATPase family protein Msp1 and eventually degraded by E3 ubiquitin ligase Doa10[101,107]. Pex15 is a tail-anchor (TA) protein. Its targeting sequence is located at the C-terminus. In this case, the C-terminal tFT tag would impair protein targeting to the correct organelle. Therefore, we decided to endogenously tag Pex15 with tFT at the N-terminus and then introduced the overexpression plasmids (Figure 6c). By measuring fluorescence intensities, we observed that ectopically expressed Pex15 regulated by *GAL1* promoter but not its native promoter results in attenuation of endogenous protein levels (Figure 6d), indicating that a higher amount of Pex15 is required to saturate the peroxisomal targeting machinery and leads to protein mislocalization. This result suggests that the extent of overexpression necessary to exceed the amounts of binding partners and promote mislocalization can vary between proteins. Meanwhile, this result further proves that mislocalized proteins that are subject to quality control can be identified by attenuated levels of endogenously tFT-tagged protein.

7.2 SCREEN FOR MISLOCALIZED PROTEINS THAT ARE SUBJECT TO QUALITY CONTROL

7.2.1 IDENTIFICATION OF POTENTIAL SUBSTRATES OF QUALITY CONTROL

The proof-of-principle experiments revealed overexpression combined with the tFT assay can be used as a tool for identification of mislocalized proteins that are subject to quality control. The next aim was to apply this method across the yeast proteome.

We first constructed a genome-wide library by the C-terminal SWAp-Tag (C-SWAT) approach. In the C-SWAT library, an acceptor module was inserted into the endogenous chromosomal loci before the stop codon of each open reading frame (ORF) in *Saccharomyces cerevisiae*[110]. Next, a donor plasmid containing mCherry-mNG tFT tag was introduced into the C-SWAT strain. Conditional expression of the I-SceI endonuclease induced double-strand breaks (DSBs) in the acceptor module and the donor plasmid. DSB repair by homologous recombination led to replacement of the acceptor module by the tFT tag. By automated genetic crossing the C-

SWAT library with a donor strain carrying the donor plasmid, each ORF was endogenously tagged with mCherry-mNG at the C-terminus in a high-throughput manner (Figure 7a).

To take into account that the degree of overexpression necessary to induce mislocalization can vary between proteins, we used two sets of 2μ-based overexpression constructs. The MoBY2.0 library comprises 4981 open reading frames (ORFs) of *Saccharomyces cerevisiae* and the expression of each gene was controlled by its native promoter and terminator. While in the MORF (moveable ORF) library, each ORF was regulated by the *GAL1* promoter[111]. The levels of overexpression are higher from the MORF library than those in the MoBY2.0 library.

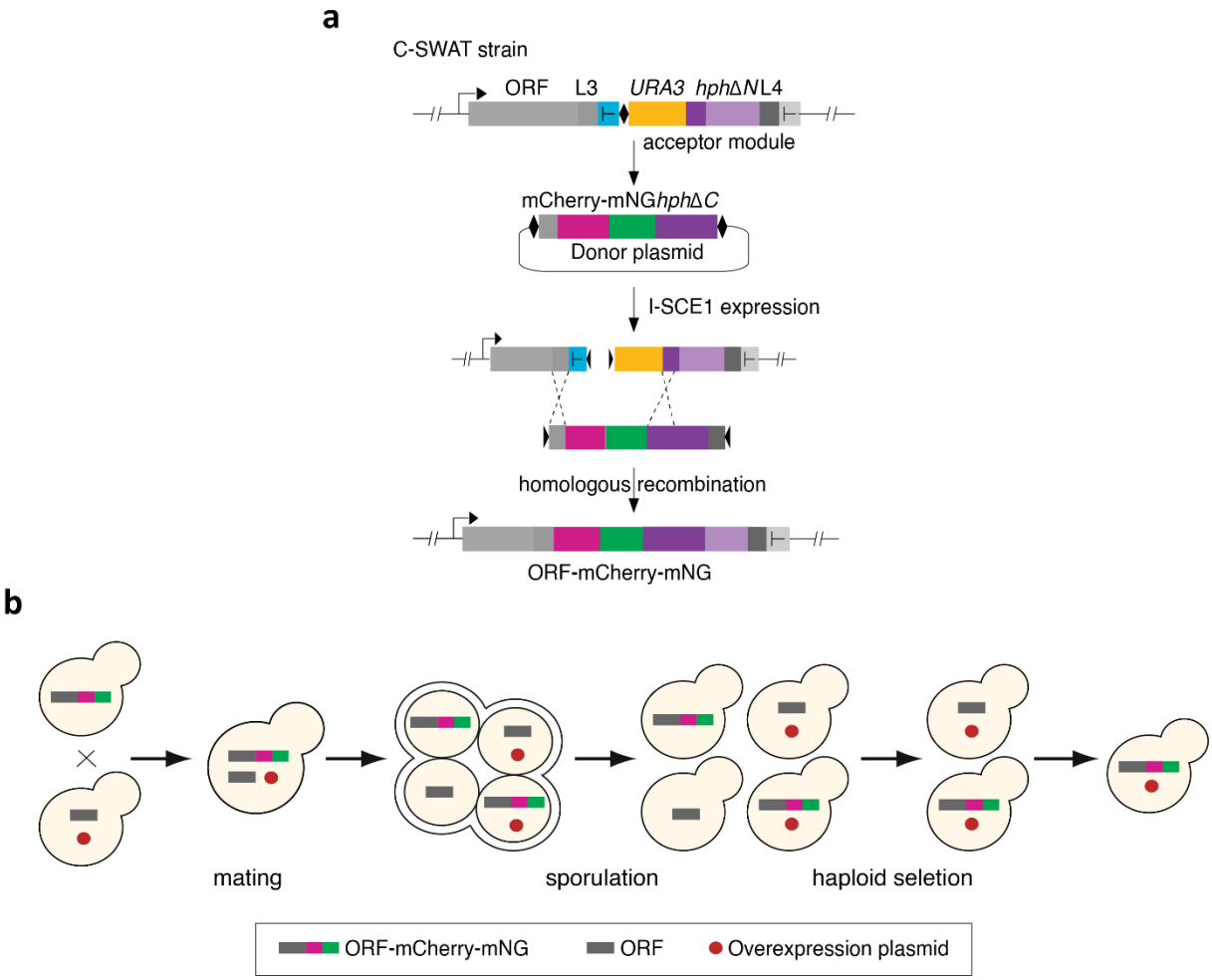


Figure 7. (a) The C-SWAT approach for high-throughput tagging of mCherry-mNG at the C-terminal of each ORF in *Saccharomyces cerevisiae*. The acceptor module consists of homology arms (L3 and L4, for subsequent recombination with the mCherry-mNG tFT tag, a heterologous transcription terminator (T), a restriction site for the I-SceI endonuclease, a selection/counter-selection marker (*URA3*), and a second truncated selection marker (*hph*ΔN). I-SceI induced double-strand breaks (DSBs) at the indicated positions (♦) in the acceptor module and the donor plasmid. DSB repair by homologous recombination led to replacement of the acceptor module by the tag. (b) Cartoon of modeling protein mislocalization through overexpression combined with the tFT assay to follow the fate of each protein.

Next, using synthetic genetic array (SGA) methodology, the tFT library was crossed with two overexpression libraries, respectively. The strain carrying the plasmid with ORF *YGR045C* was used to simulate the native condition. After mating, sporulation and several steps of selection,

the final haploid strains carrying both the tFT-tagged protein and the corresponding overexpression plasmid were used for fluorescence measurements via the high-throughput plate reader (Figure 7b). Fold change of protein levels was calculated by the green fluorescent intensity of each gene in the overexpression condition versus the native condition.

We set a conservative arbitrary threshold of the false discovery rate (FDR) less than 0.1 and fold change of protein levels of at least 20% decrease as significant attenuation. Based on these criteria, we found that protein abundance does not change upon overexpression in the screen with the MoBY2.0 library for the majority of the proteome, but 14.6% of 3894 proteins are significantly attenuated (Figure 8a). Surprisingly, in the screen performed with the mORF library, we found 1864/4000 (46.6%) ORFs show significant attenuation upon overexpression (Figure 8b). Meanwhile, we also observed a large number of strains have a significant decrease in colony size upon overexpression induced by *GAL1* promoter with 1113 significantly attenuated ORFs affected (Figure 8c). Whereas only 186 proteins exhibit significant reduction in colony size in the MoBY2.0 overexpression screen (Figure 8d).

It is likely that a large amount of proteins expressed from *GAL1* promoter saturates protein quality control system. Aberrant accumulation of mislocalized proteins brings harmful effects on cell fitness and physiology, which results in cell growth defects. Based on the results from the mORF library screen, it is hard to determine whether the attenuation of endogenous protein levels is due to protein that are being degraded or some indirect effects in response to cellular environmental changes. Therefore, I decided to use the readout from MoBY2.0 overexpression screen for further studies.

We examined the properties of mislocalized proteins that are subject to quality control. The gene ontology (GO) terms analysis revealed that attenuated proteins were highly enriched with subunits of protein complexes (Figure 9a). To avoid a large number of attenuated proteins coming from one complex biasing the outcome, we excluded the ribosome, one of the most abundant multi-subunit complexes in the cell, from the dataset and still found this enrichment (Figure 9b). This finding is consistent with previous studies that multi-protein complexes are often targets for quality control via recognition of the regions that are normally shielded by interaction partner(s).

Moreover, the sub-cellular localization analysis revealed that the attenuated proteins were enriched in cytosolic and nuclear proteins (Figure 9c). Strikingly, there was a significant depletion of attenuated proteins for mitochondrial proteins and mitochondrial complexes. For example, mitochondrial ribosome subunits, unlike their cytosolic counterparts, did not show down-regulation in the overexpression condition (Figure 9d). This will be further discussed in the last section.

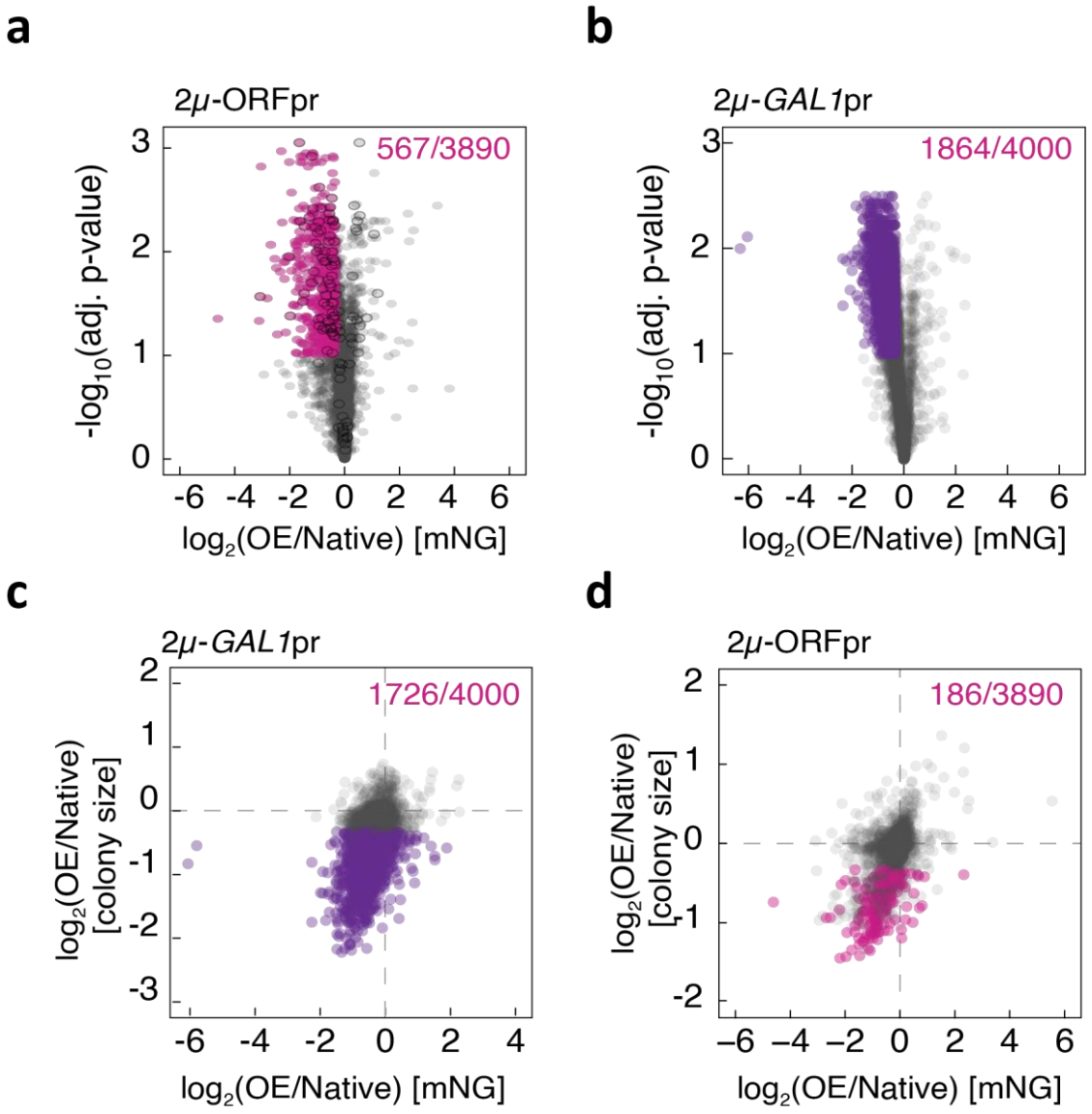


Figure 8. (a) Overexpression screen performed with the MoBY2.0 library. Volcano plot of log2 fold change of mNG intensity values in the overexpression condition versus the native condition. Significantly attenuated proteins (potential substrates) are highlighted in pink. 120 ORFs selected for validation are marked with the thick outline. (b) Overexpression screen performed with the mORF library. Cells were grown onto galactose-containing plates for 24h to induce protein expression. Volcano plot showing the $\log_2$ fold change of mNG intensity values in the overexpression condition versus the native condition. Significantly attenuated proteins are highlighted in purple. (c) Comparison of $\log_2$ fold change in colony size against the $\log_2$ fold change in mNG levels upon overexpression induced by the *GAL1* promoter. Strains with a significant decrease in colony size are highlighted in purple. (d) Comparison of $\log_2$ fold change in colony size against the $\log_2$ fold change in mNG levels in the overexpression screen performed with the MoBY2.0 library. Strains with a significant decrease in colony size are highlighted in pink.

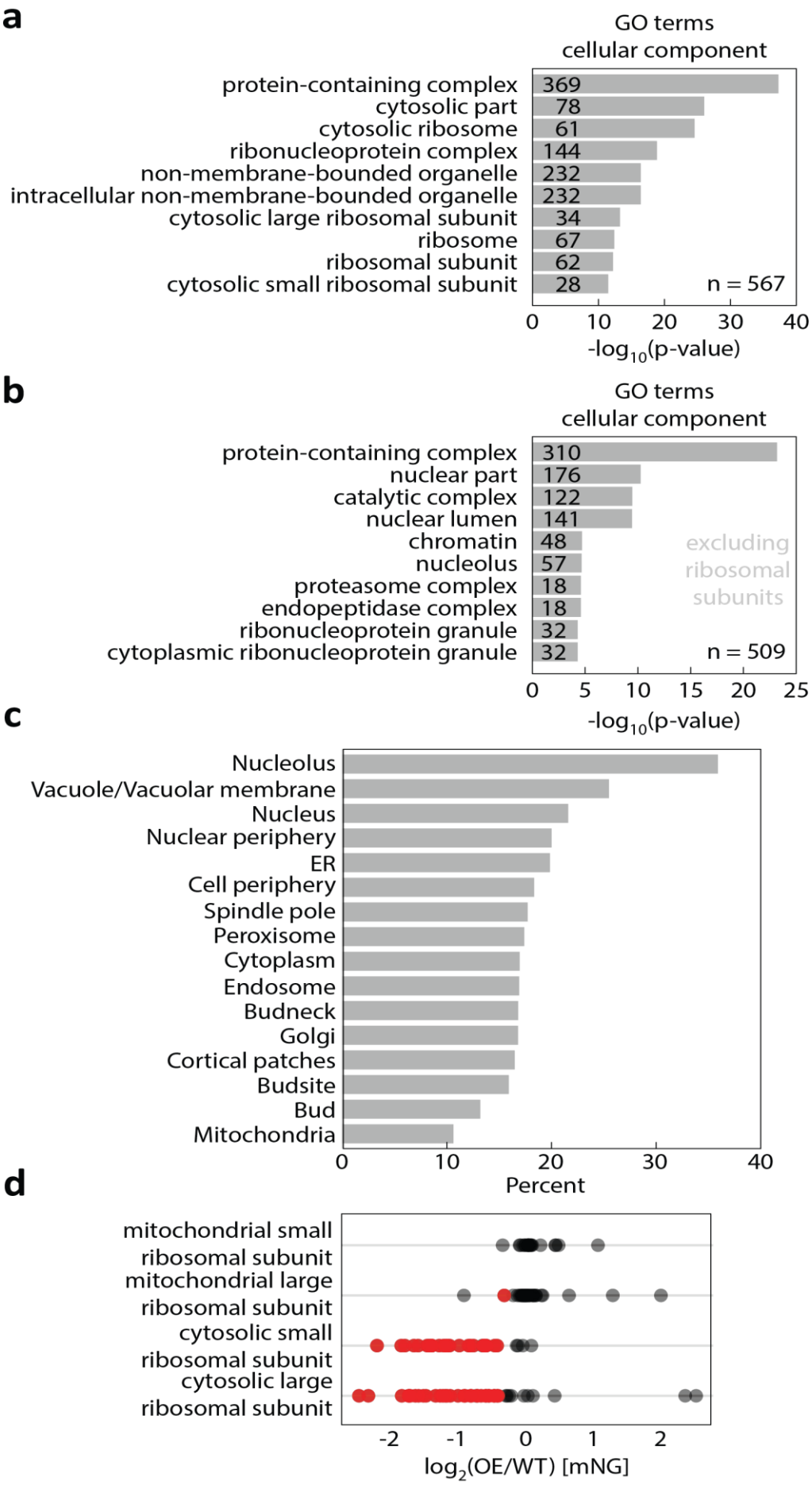

a
GO terms
cellular component
protein-containing complex
cytosolic part
cytosolic ribosome
ribonucleoprotein complex
non-membrane-bounded organelle
intracellular non-membrane-bounded organelle
cytosolic large ribosomal subunit
ribosome
ribosomal subunit
cytosolic small ribosomal subunit
369
78
61
144
232
232
34
67
62
28
n = 567
0
10
20
30
40
-log10(p-value)
b
GO terms
cellular component
protein-containing complex
nuclear part
catalytic complex
nuclear lumen
chromatin
nucleolus
proteasome complex
endopeptidase complex
ribonucleoprotein granule
cytoplasmic ribonucleoprotein granule
310
176
122
141
48
57
18
18
32
32
excluding
ribosomal
subunits
n = 509
0
5
10
15
20
25
-log10(p-value)
c
Nucleolus
Vacuole/Vacuolar membrane
Nucleus
Nuclear periphery
ER
Cell periphery
Spindle pole
Peroxisome
Cytoplasm
Endosome
Budneck
Golgi
Cortical patches
Budsite
Bud
Mitochondria
0
10
20
30
40
Percent
d
mitochondrial small
ribosomal subunit
mitochondrial large
ribosomal subunit
cytosolic small
ribosomal subunit
cytosolic large
ribosomal subunit
-2
-1
0
1
2
log2(OE/WT) [mNG]

Figure 9. (a) and (b) GO term enrichment analysis of attenuated proteins including and excluding ribosomal subunits. (c) Percentage of attenuated ORFs at different subcellular localizations. (d) Comparison of $\log_2$ fold change in mNG intensities upon overexpression between mitochondrial and cytosolic ribosomal subunits. Subunits with at least 20% decrease are marked in red (0.1% FDR).

7.3 VALIDATION OF SCREEN RESULTS

Previously, Dephoure et al. measured protein levels in haploid yeast cells carrying an extra copy of one of the chromosomes by mass spectrometry. In general, the amount of each protein produced by these aneuploid strains increased depending on the number of extra copies of genes found on the extra chromosome. However, they found around 20% of proteins were in lower amounts than expected[112]. The number of proteins showing attenuation was similar to our MoBY2.0 overexpression screen, and both studies found the attenuated proteins were highly enriched with subunits of protein complexes.

Next, we compared the significantly attenuated hits between two datasets but found that there is not a large overlap (n = 129). There were 227 ORFs that are aneuploid-specific observations and 189 ORFs that are specific to the MoBY2 overexpression (Figure 10a). To better understand this discrepancy, we decided to validate our screen results from three aspects: the plasmid identity, growth conditions and the size effect of the large tFT tag respectively.

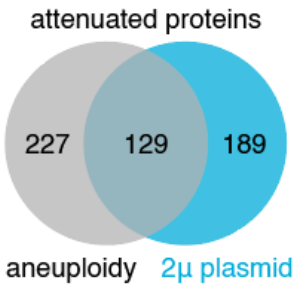


Figure 10. (a) Venn diagram to show the overlap of attenuated proteins detected in the aneuploidy measurements by Dephoure et al and the MoBY2.0 overexpression screen.

7.3.1 VERIFYING THE IDENTITY OF PLASMIDS IN THE MOBY2.0 LIBRARY

As mentioned before, overexpression plasmids are available from the MoBY2.0 library and present in yeast. It contains 4,981 uniquely barcoded ORFs, representing ~90% of all non-dubious ORFs annotated in the Saccharomyces Genome Database (SGD) distributed 71 plates in 96 format[109].

To verify the plasmid identity, we designed a forward ORF-specific primer annealing within the ORF and a generic reverse primer annealing with the plasmid backbone for colony PCR amplification (Figure 11a). A control PCR was set up with the library background strain BY4741 as a source of template DNA. Correct plasmids were identified by the presence of a PCR product of expected size in the sample but not in the control PCR.

We took one yeast strain from each plate to represent the MoBY2.0 library and found the presence of PCR products of expected size in all the tested 71 samples. Thus, we infer that all plasmids in the MoBY2.0 library have correct identity (Figure 11b and Supplementary Figure1).

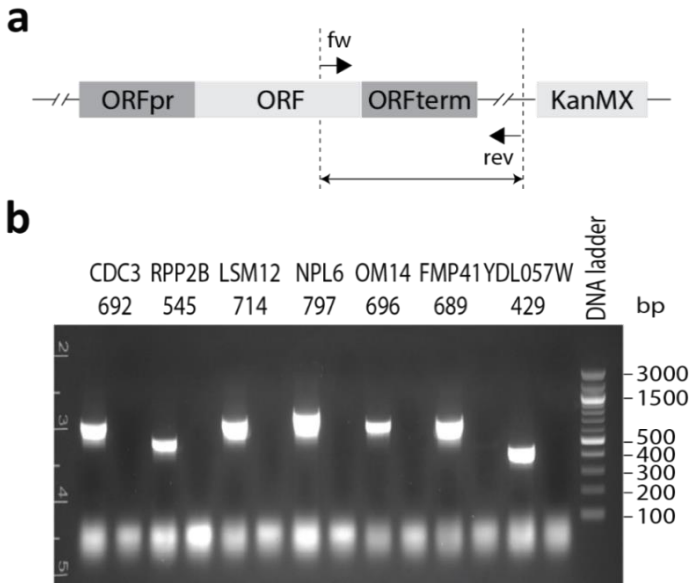


Figure 11. (a) Cartoon of primer-binding sites within the MoBY2.0 overexpression plasmid. (b) One representative image of the agarose gel. The exhibited bands are PCR products of different plasmids with expected sizes. Last lane was loaded with 100bp DNA ladder.

7.3.2 ESTIMATING THE REPRODUCIBILITY OF PROTEIN ABUNDANCE CHANGES

7.3.2.1 SELECTION OF 120 PROTEINS FOR VALIDATION

Next, we selected 120 ORFs to test the reproducibility of the screen in different growth conditions. All 3890 tested proteins were sorted by brightness according to mNG intensities in unperturbed conditions. By filtering the top 50% bright strains, improved signal to background ratio would facilitate the reduction of error rates between different fluorescence detection techniques. Since our broad interest in mislocalized proteins that are subject to quality control, 80 ORFs were randomly chosen from the significantly attenuated group and 20 ORFs from the not affected and significantly increased group, respectively (selected ORFs were marked with the thick outline in Figure 8a).

7.3.2.2 THE SCREEN READOUTS ARE REPRODUCIBLE IN DIFFERENT CULTURE CONDITIONS

Then, the selected 120 tFT strains were crossed to the MoBY2.0 strains to promote protein mislocalization via overexpression. SGA methodology, fluorescence measurements and downstream analysis were performed in the exact same way as the screening. To assess the reproducibility of the screen results, we compared the fold change of protein levels between the screen and the replicate experiment. A high correlation coefficient indicated that our screen data is reproducible on agar plates (Figure 12a) (The readout of each protein is shown in Supplementary Table4).

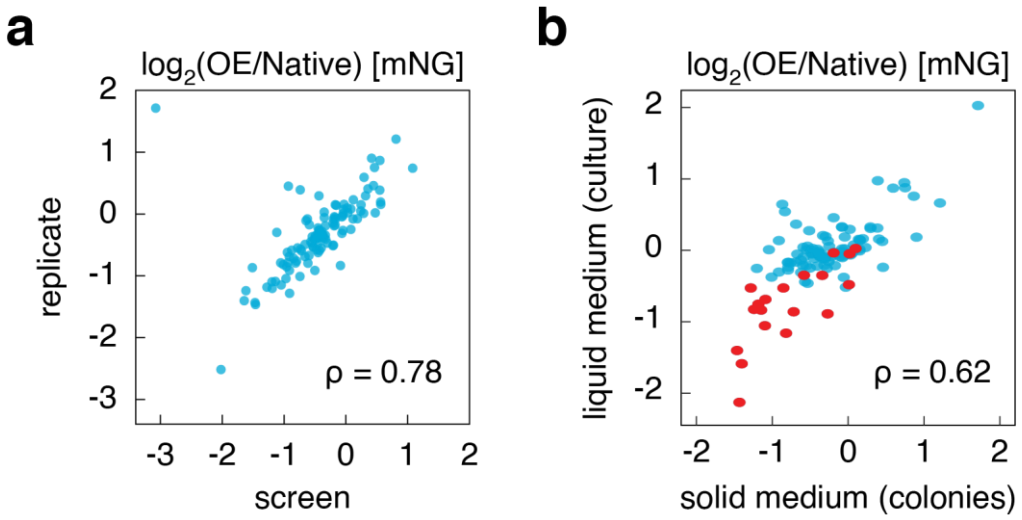


Figure 12. (a) The scatterplot displays the correlation of $\log_2$ fold change in mNG intensities between measurements in the overexpression screen and the replicate experiment on agar plates. Pairwise comparison showed a Spearman correlation coefficient ρ=0.78. (b) Comparison between changes in protein abundance estimated with the replicate experiment on agar plates and liquid measurements by flow cytometry. Pairwise comparison showed a Spearman correlation coefficient ρ=0.62. Dots highlight in red represented the MYC-tagged proteins whose behavior changes upon overexpression were checked by western blot.

Yeast cells not only can grow on solid agar plates but also in liquid culture. Proteins may behave differently in different culture conditions. To test the reproducibility of the overexpression screen in liquid, we cultured yeast cells in liquid medium and then measured the fluorescence intensity at a single cell level via flow cytometry when they reach the exponential phase. Protein abundance was determined by the median of green fluorescence intensities from 50,000 single cell readout. Then $\log_2$ fold change of protein abundance was calculated with the overexpression condition against the native condition. Although more proteins clustered around 0 in the liquid medium than on solid agar plates, we still can observe a linear correlation between two different growth conditions (Figure 12b) (The readout of each protein is shown in Supplementary Table5). Therefore, we concluded that different growth conditions would not cause large-scale changes in mislocalized protein behaviors.

7.3.3 EVALUATING THE POTENTIAL INFLUENCE OF THE TFT TAG ON PROTEIN BEHAVIOR

In contrast to the large tFT tag used in our experiment, Dephoure et al. measured protein levels in aneuploid strains by mass spectrometry without tags[112]. In order to estimate the influence of the tFT tag on protein behavior, we decided to replace the timer of selected 120 ORFs with the short MYC tag and analyze the impact of overexpression on the MYC-tagged protein with immunoblotting. Using the C-SWAT approach, we endogenously tagged selected 120 ORFs with the MYC tag[110] (Figure 13a). Since the MYC-tagged strains will be used to test effects of proteasome inhibition later, they were selected with *pdr5*Δ background.

To check the integration of tagging, we randomly tested three strains from the constructed MYC-tagged library by immunoblotting. All three proteins with expected band sizes were

clearly detected by immunoblotting with MYC antibody (Figure 13b). Therefore, we concluded that the selected 120 proteins have been efficiently tagged with the MYC tag.

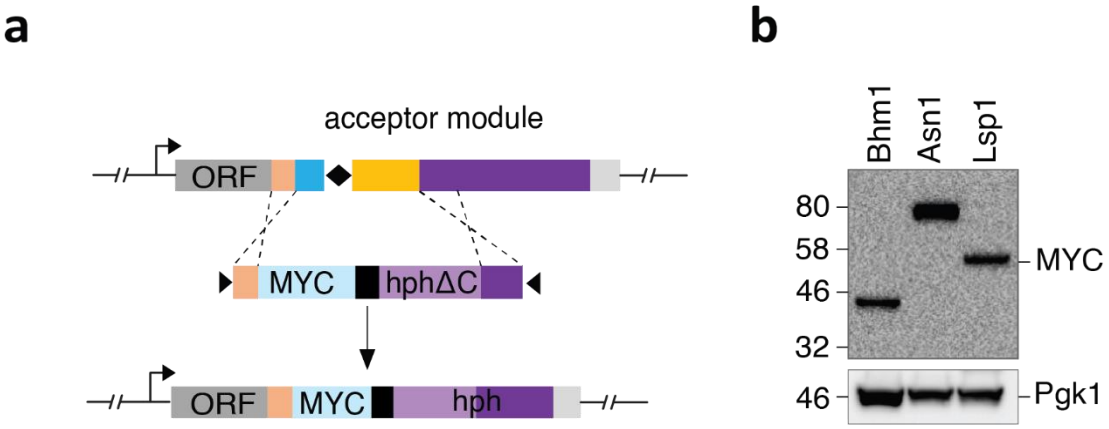


Figure 13. (a) The C-SWAT approach for high-throughput tagging of MYC tag with reconstitution of the hygromycin selection marker (*hph*) at the C-terminal of 120 ORFs. (b) Western Blot to validate the MYC tagging with MYC antibody. Pgk1 was used as a loading control.

Then, the MYC-tagged library was crossed to MoBY2.0 strains to promote protein mislocalization. We checked 20 proteins (highlighted in red in Figure 12b) via western blot and found that 19 ORFs exhibited the same trends as they showed from colony or cytometry measurements with the tFT tag (Figure 13a and b and Supplementary Figure 2). However, only one protein, Rpt2, displayed opposing phenotypes when tagged differently. Specifically, endogenously tFT-tagged Rpt2 exhibited attenuation upon overexpression, whereas, the protein level of MYC-tagged Rpt2 was up-regulated under the same overexpression condition (Figure 13c). It is possible that the tFT tag triggers protein degradation for Rpt2, but in general, we found that the large size of the tFT tag did not significantly impact the behaviors of mislocalized proteins.

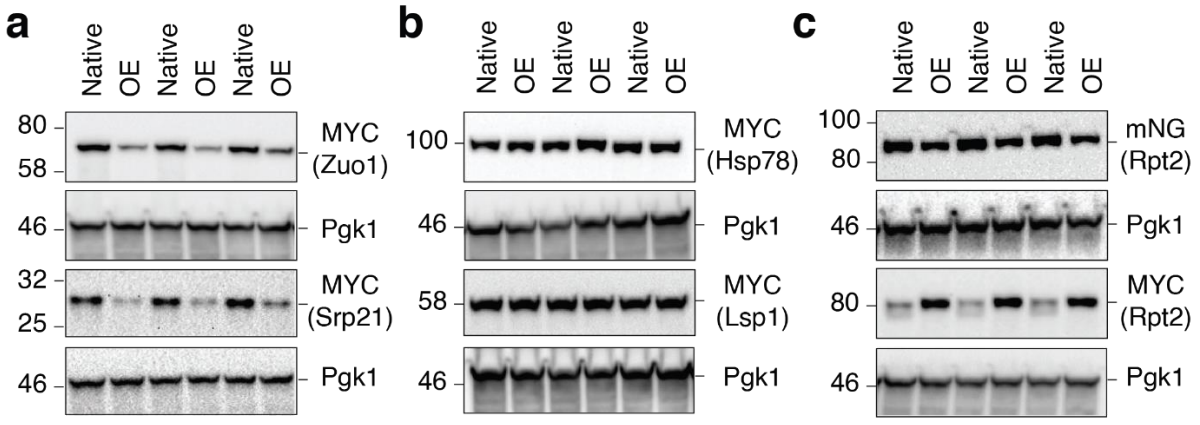


Figure 14. Immunoblotting to detect the endogenous protein level with MYC antibody in the native condition (Native) or cells harboring 2u plasmids (OE). Pgk1was used as loading control. Experiments were done with three biological replicates. (a) Attenuated proteins measured by the plate reader or cytometry exhibited decreased protein levels upon overexpression. (b) Overexpression did not affect protein abundance no matter the size of the tag. (c) tFT-tagged Rpt2 exhibited the opposite protein behavior to the MYC-tagged protein in the overexpression condition.

To sum up, extensive validation experiments proved the reproducibility of the high-throughput screening results. Moreover, cell culture conditions and the large size of the C-terminal tFT tag did not widely affect the protein attenuation phenotypes. Carrying an extra copy of one of the chromosomes in haploid yeast cells and overexpression of individual proteins showed a comparable amount of proteins that are being attenuated, indicating that these genes are more sensitive to alterations in gene copy. Although the overlapping of significantly attenuated hits between two studies is not very high, we can at least conclude that different cell culture conditions and the tag effects are also not major contributors to the discrepancy. It is likely that the significant hits in aneuploid cells may be due to indirect effects from the global response of aneuploidy cells. The MoBY2-specific observations could be due to the higher level of overexpression from the 2µ high copy plasmids as compared to the duplicated chromosomes in aneuploid cells.

7.4 MECHANISMS OF REMOVAL OF MISLOCALIZED PROTEINS

7.4.1 IDENTIFICATION THE SUBSET OF PROTEINS THAT ARE BEING DEGRADED

Next, we would like to investigate the mechanisms of protein attenuation. Decrease in protein abundance can be caused by deceleration of protein synthesis or acceleration of protein degradation. Taking advantage of the tFT readout, whereby the mCherry/mNG ratio reports on protein turnover, we can easily identify which proteins are selectively targeted for degradation. With this notion, we split attenuated proteins into three sub-groups based on their changes in stability upon overexpression. We observed that 23.5% (133/567) of attenuated proteins are destabilized (Figure 15a), indicating that these proteins are likely subject to protein degradation.

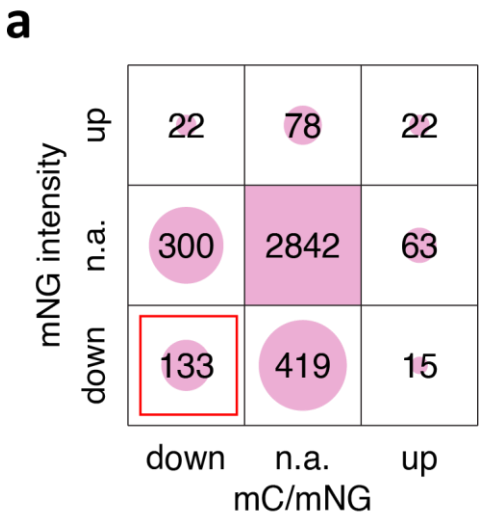


Figure 15. Summary of phenotypic outcomes using mCherry-mNG tFT as a reporter (changes of protein abundance and stability; n.a., not affected).

Before we moved on to the next step, another PhD student in the lab [redacted] examined the performance of different tFTs as reporters of protein degradation kinetics by the ubiquitin-fusion assay[113]. Different tFTs were cloned into a series of constructs for expression of N-terminal ubiquitin (Ubi) fusions. Co-translational cleavage of the ubiquitin moiety exposes a new N-terminus starting with residue X, which determines the stability of the fusion according to the N-end rule[113] (Figure 16a). For all residues, the mCherry-mNG tFT showed the narrowest dynamic range of red-to-green ratio between fast and slow turnover residues than other tFTs (Figure 16b), indicating that very small variations in protein stability cannot be reported by mCherry-mNG tFT.

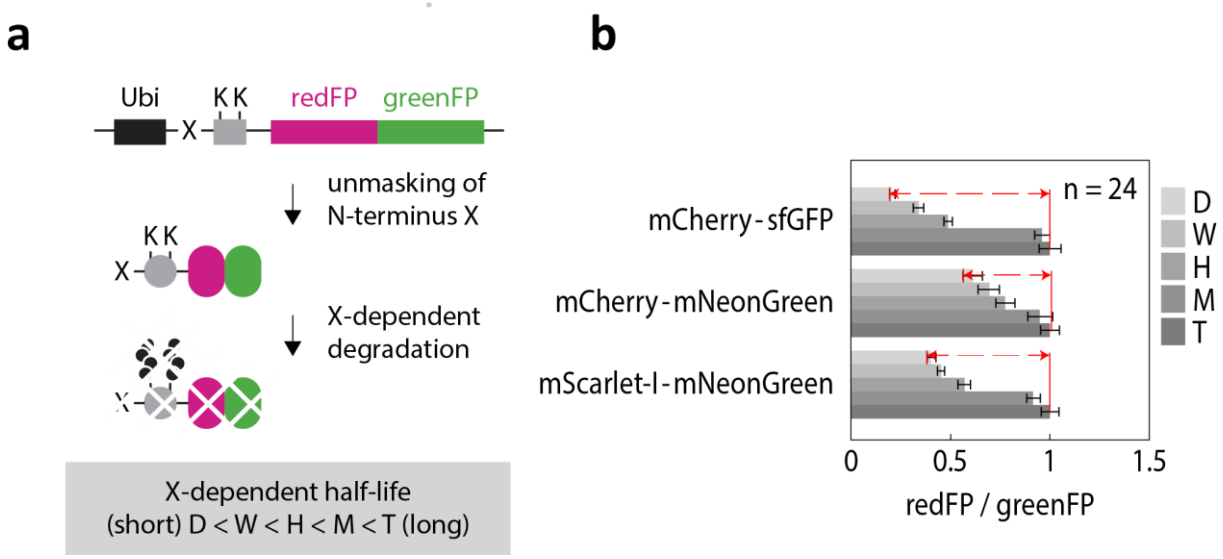


Figure 16. (a) Cartoon of Ubi-X-redFP-greenFP constructs. Cotranslational cleavage of the ubiquitin moiety exposed a new N-terminus starting with residue X. Degradation rate of redFP-greenFP fusions depended on the residue X. (b) Fluorescence measurements of colonies expressing Ubi-X-tFT constructs with the indicated tFTs. Fluorescence intensities were corrected for background autofluorescence and redFP/greenFP ratios were normalized to the stable Ubi-T-tFT construct for each tFT (n = 24 technical replicates per construct, error bars indicate s.d.) Dashed lines marked dynamic range of red to green ratio between fastest and slowest turnover residues.

To verify that the number of attenuated and destabilized proteins is underestimated by the readout of mCherry/mNG, we repeated the MoBY2.0 overexpression screen using mCherry-superfolderGFP (sfGFP) tFT, which displays the highest sensitivity to changes in protein turnover within the ubiquitin-fusion assay. SGA analysis and fluorescence measurements follow the same steps as the MoBY2.0-mCherry-mNG screen.

We first compared the green fluorescence intensities between measurements of two tFTs in the native condition. A strong linear correlation indicates that endogenous protein levels can be well reported by both green fluorescent proteins (Figure 17a). Additionally, we noticed that the low-abundance proteins, which had negative values in sfGFP measurements, were clustered around zero in the mNG channel. This is in line with previously published findings that the molecular brightness of mNG is higher than that of sfGFP[114]. Moreover, upon overexpression of individual proteins, fold change of endogenous protein levels also correlates well between two screenings (Figure 17b).

When we applied the same threshold as before to define the significant attenuation, we observed that only 11.3% of the proteome are significantly attenuated (Figure 17c). In comparison to the mNG readout, a lower number of proteins displayed a reduction in sfGFP intensities upon overexpression. This could be explained by the incomplete degradation of sfGFP by the proteasome[114]. Due to the robust fold of sfGFP, the proteasome could not completely degrade sfGFP or mCherry-sfGFP fusions. Accumulation of processed sfGFP fragments gave a higher readout than the actual protein levels, which led to an

underestimation of the decrease in protein levels. Consequently, fewer proteins were identified as being significantly attenuated than actually were.

Nevertheless, we further analyzed the results of the mCherry-sfGFP screen and categorized them based on changes in protein abundance and stability upon overexpression (Figure 17d). Out of 460 identified attenuated proteins, 200 exhibited a rapid turnover rate. Upon comparing the proportion of destabilized proteins within the attenuated group, we found that a greater number of attenuated proteins displayed destabilization when utilizing the more sensitive tFT to report on protein degradation kinetics (Figure 17e). This result indicates that the mCherry/mNG readout underestimates the actual number of attenuated proteins that are also destabilized upon overexpression.

Based on the above results, we have shown that mCherry-sfGFP tFT is a more sensitive timer than mCherry-mNG tFT to report on protein degradation kinetics. However, incomplete degradation of sfGFP can lead to an underestimation of the attenuation in protein levels. On the other hand, mNG is a brighter green fluorescent protein than sfGFP, which allows for better detection of low-abundance proteins. Therefore, we use results of mCherry-mNG screening to identify mislocalized proteins that are subject to quality control. Additionally, the ratio of mCherry/sfGFP provides a second readout that can be used to identify potential substrates that are likely targeted for protein degradation.

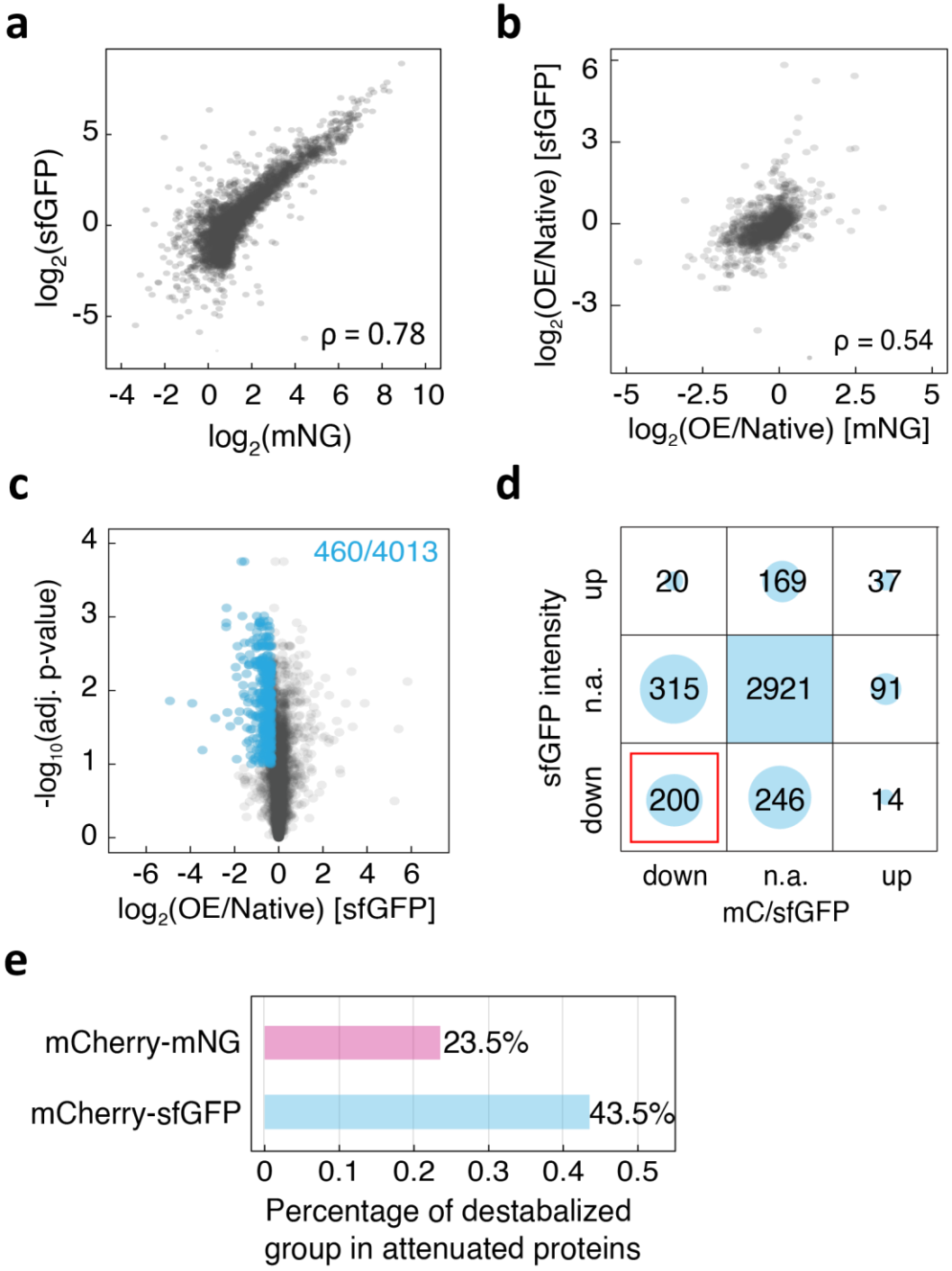


Figure 17. (a) Scatter plot displaying the correlation of measurements of green fluorescence intensities using two different tFTs. Pairwise comparison showed a spearman correlation coefficient ρ=0.78. (b) Correlation of log2 fold change in green fluorescence intensities between two overexpression screens with different tFTs. Pairwise comparison showed a spearman correlation coefficient ρ=0.54. (c) Volcano plot of log2 fold change in sfGFP intensity values in the overexpression condition against the native condition. Significantly attenuated proteins are highlighted in blue. (d) Summary of phenotypic outcomes using mCherry-sfGFP tFT as a reporter (changes of protein abundance and stability; n.a., not affected). (e) Comparing the percentage of destabilized proteins in the attenuated group within mCherry-mNG tFT and mCherry-sfGFP tFT measurements.

Previous studies have demonstrates that ubiquitination at multiple sites is an efficient signal for degradation[115], and an increase in the number of ubiquitination sites can enhance the binding affinity between substrates and proteasomes[116]. To test whether proteins that are attenuated and destabilized are more likely to be degraded by the proteasome, we retrieved ubiquitination data from Swaney's study[117]. This study utilized mass spectrometry to identify the ubiquitination sites on individual proteins. By comparing the number of ubiquitination sites across different categories based on the mCherry/sfGFP readout, we observed that attenuated and destabilized proteins tended to have more ubiquitination sites than proteins that were either not attenuated or attenuated but not destabilized (Figure 18a). This suggested that attenuated proteins that are also destabilized are likely targeted for degradation by the ubiquitin-proteasome system.

To verify this notion, I randomly selected three proteins from attenuated and destabilized subset and checked the effect of proteasome inhibition (with MG132) on the overexpression phenotype. If the attenuation phenotype of these three protein could be reverted by MG132, I infer that proteins within this subset are very likely degraded by the UPS. Our results showed that all three proteins exhibited accumulation effects upon proteasome inhibition. (Figure 18b). I also included one non-destabilized protein, Hmo1, as a negative control. If the protein level of Hmo1 decreases but its turnover rate remains unaffected, it is likely that the decrease is caused by transcriptional regulation. In line with our expectation, proteasome inhibition fails to revert the attenuation phenotype (Figure 18b). These results demonstrated that attenuated and destabilized mislocalized proteins are likely degraded by the ubiquitin-proteasome system. Since not all proteins have been tested with MG132, we cannot exclude the possibility that other degradation pathways, like lysosomal degradation or autophagy, are also involved in detection and removal of mislocalized proteins.

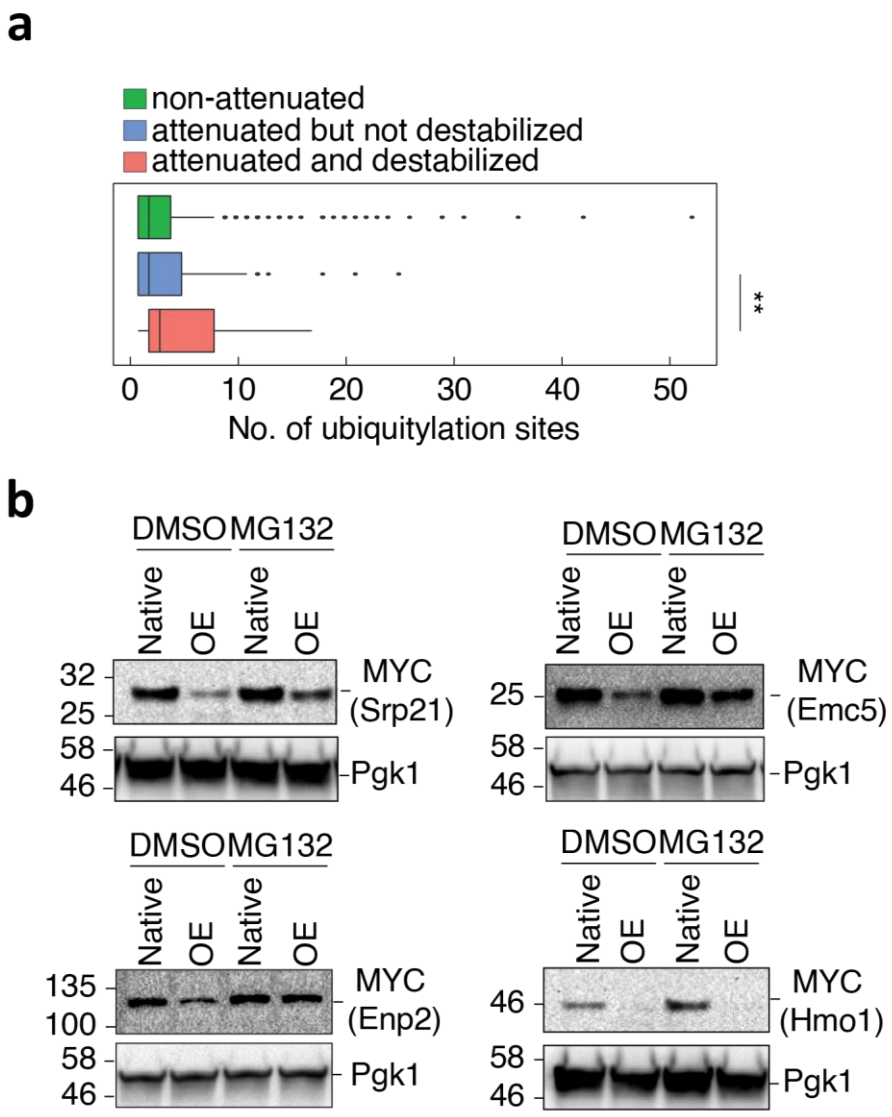


Figure 18. (b) Boxplot displaying number of ubiquitination sites among different protein abundance and turnover categories upon overexpression. (c) Attenuated and destabilized proteins detected were treated with the proteasome inhibitor MG132 for 90 min. Western Blot checked the endogenous protein level in the native condition (Native) or cells harboring 2u plasmids (OE). Pgk1 was used as a loading control. Attenuated but non-destabilized protein Hmo1 was used as a negative control for the proteasome inhibition assay.

7.5 IDENTIFICATION OF QUALITY CONTROL FACTORS OF MISLOCALIZED PROTEINS

Next, we examined which components of the UPS mediate the degradation of mislocalized proteins. To rationalize the efforts, I ranked proteins based on the degree of attenuation according to the readout of mCherry-mNG screening and then selected top 12 attenuated proteins as potential substrates for quality control factors.

First, I validated the attenuation phenotype of these 12 proteins in a low-throughput manner. MoBY2.0 overexpression plasmids were introduced into the corresponding mCherry-mNG tFT strains to facilitate protein mislocalization, and the plasmid carrying the *YGR045C* gene was used to mimic the native condition. Plate reader measurements of colonies showed that all proteins are attenuated upon overexpression (Figure 19b). This reproduced the high throughput results and further confirmed that they are potential substrates of quality control.

Next, we determined the effect of each of UPS mutants on the abundance (mCherry intensity) and turnover (mCherry/mNG ratio) of the tFT-tagged proteins. We crossed 24 strains (12 carrying the corresponding overexpression construct and 12 carrying the control plasmid) with the UPS mutant array (Figure 19a), which encompassed knockout alleles of non-essential UPS components[118] or temperature-sensitive alleles of essential UPS factors[119].

If the UPS component is not the quality control factor for mislocalized proteins, protein abundance or stability should be similar between mutants and WT in the overexpression condition. However, if the UPS component functions as the quality control factor, the mutant is not able to degrade the mislocalized protein, which would revert the overexpression phenotype and exhibit a higher fluorescence intensities and ratios (Figure 19c).

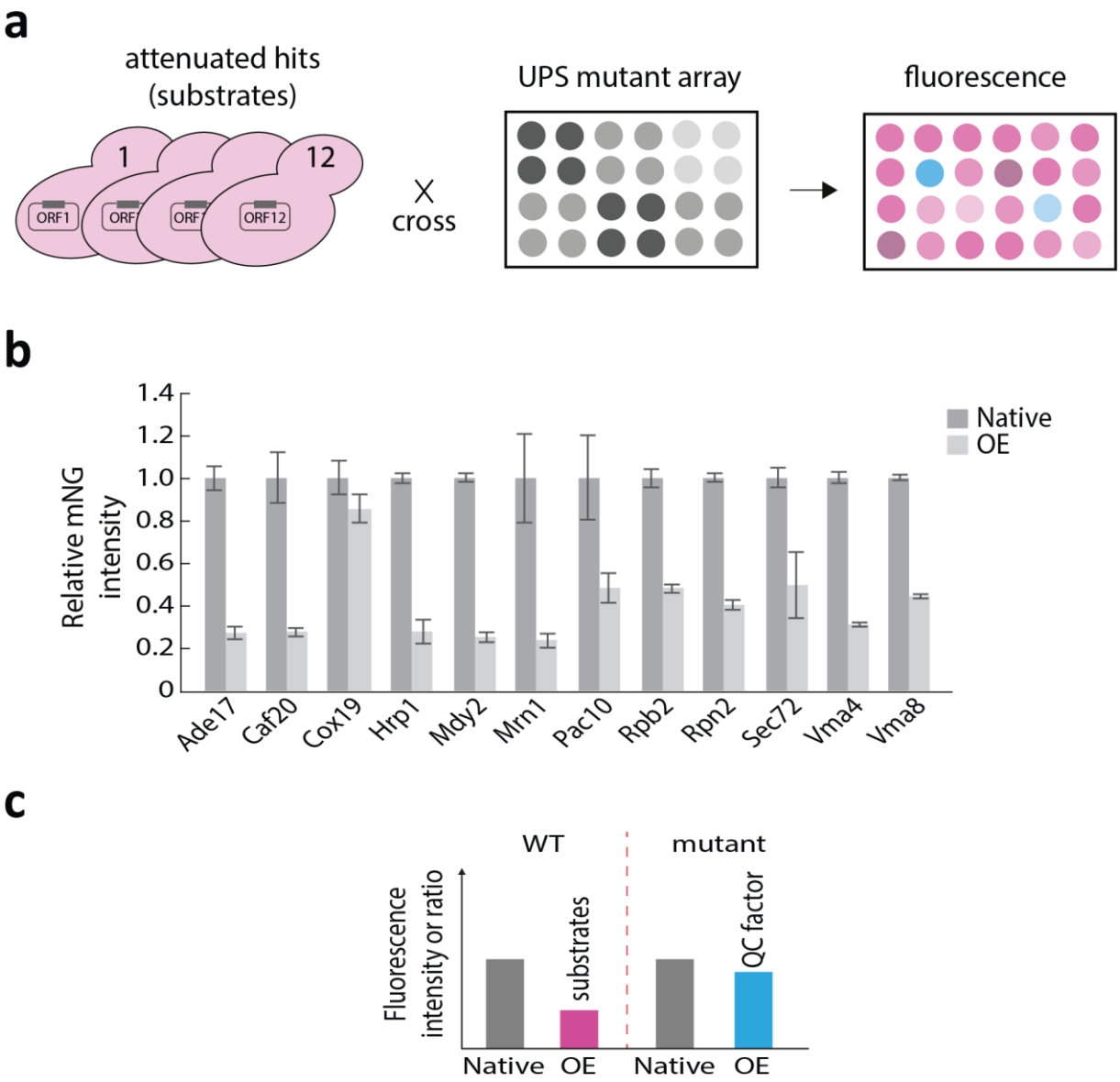


Figure 19. (a) Cartoon of screens to identify quality control factors of mislocalized proteins. (b) Top 12 attenuated substrates were selected for identifying quality control factors. Plate reader measured the fluorescence of cells harboring the corresponding overexpression plasmids (OE) and cells with the control plasmid (Native). Fluorescence intensities were normalized to the native condition for each substrate. (c) The mutant that reverts the overexpression phenotype would be identified as the quality control factor of mislocalized proteins.

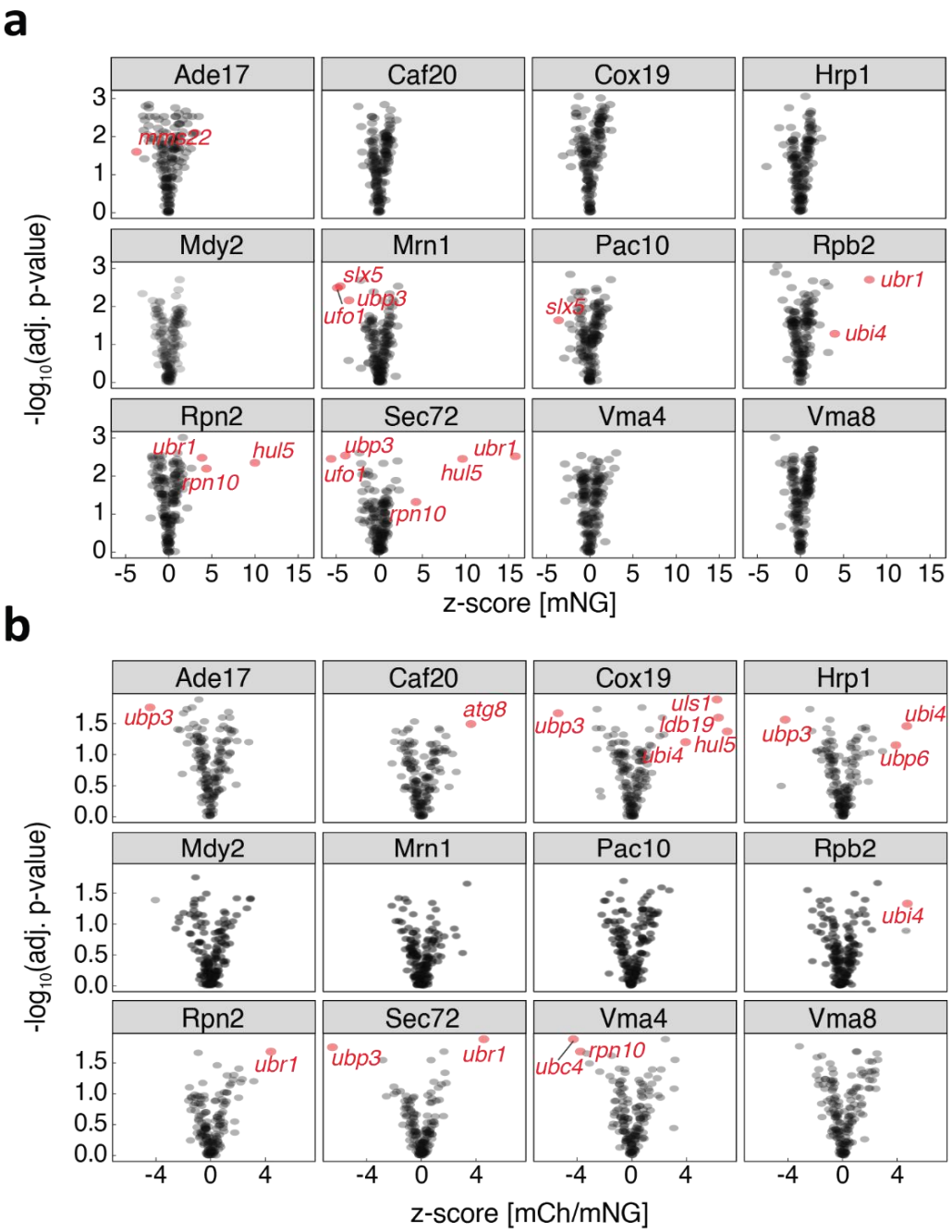


Figure 20. (a) The influence of each mutant of the UPS component on the abundance of its substrate was calculated by z-score of mNG intensities. Mutants with FDR less than 0.1 and z-score >3.5 were highlighted in red. (b) The effect of each mutant of UPS component on the turnover of its substrate was calculated by z-score. Mutants with FDR less than 0.1 and z-score >3.5 were highlighted in red.

The crossing steps follows the SGA procedure and colony fluorescence was measured for two different channels, mCherry and mNG. Changes in protein abundance and stability of each UPS mutant were calculated by z-score in the overexpression condition. We set a conservative threshold of FDR less than 0.1 and z-score of mNG intensities or mcherry/mNG ratios >3.5 as hits.

Ubr1. Based on these criteria, we found that deletion of *UBR1* led to the upregulation of three proteins, Rpb2, Rpn2 and Sec72, respectively (Figure 20a). In addition, Rpn2 and Sec72 were stabilized in the *ubr1*Δ mutant cells (Figure 20b). These results were confirmed in a low-throughput experiment, where three different overexpression plasmids were manually transformed into the corresponding timer strains with *ubr1*Δ backgrounds. All three tested proteins showed a significant increase in protein levels in the *UBR1* deletion strain (Figure 21a left panel). Consistent with the screen results, *ubr1* mutant stabilized Rpn2 and Sec72, but not Rpb2 upon protein overexpression (Figure 21a right panel). Moreover, the screen data also showed that deletion of the deubiquitinase Ubp3 destabilizes several mislocalized proteins, including one of the substrates of Ubr1 (Figure 20b and 21b). To determine the interplay between Ubr1 and Ubp3, we tested the effect of *ubr1*Δ *ubp3*Δ double mutant on Sec72 and found that the protein level was similar to the WT condition (Figure 21c). This indicated that Ubr1 and Ubp3 worked independently. To sum up, our results demonstrated that our screen approach can be used to identify quality control factors of mislocalized proteins.

Rpn10. Rpn10 is a proteasome polyubiquitin receptor. Two sproteins were upregulated in the *rpn10*Δ mutant, suggesting that these proteins were degraded by the ubiquitin-proteasome system. However, 10 out of 12 tested substrates were destabilized upon protein mislocalization based on mCherry/sfGFP readouts, indicating that they are degraded by either the ubiquitin-proteasome system or autophagy. Why other substrates were not affected by any of mutants of degradation components? The proteasome contains many essential subunits, for which temperature-sensitive alleles were used to perturb gene functions. We measured fluorescence intensities of final colonies at 30°C. This temperature was able to partially inhibit growth of all temperature-sensitive mutants. Smaller colony would lead to lower mNG intensity, since a positive correlation between the colony sizes and fluorescence intensities. Therefore, the final readout of stabilization effects may be weakened by the growth defect of proteasomal temperature-sensitive alleles. Moreover, the correlation could also explain why several substrates exhibited low intensity score in *mms22*Δ, *slx5*Δ or *ufo1*Δ cells.

To tackle this problem, we decided to compare relative protein abundance to the native condition, and then calculate z-scores of relative intensity values (Figure 22a). We observed 7 out of 12 proteins were affected by the proteasomal mutants *pre2* and *pre9* (Figure 22b), indicating that the ubiquitin-proteasome system plays a major role in the removal of mislocalized proteins. The absence of *ATG8* stabilized Caf20 suggested that autophagy may also be involved in the degradation of mislocalized proteins.

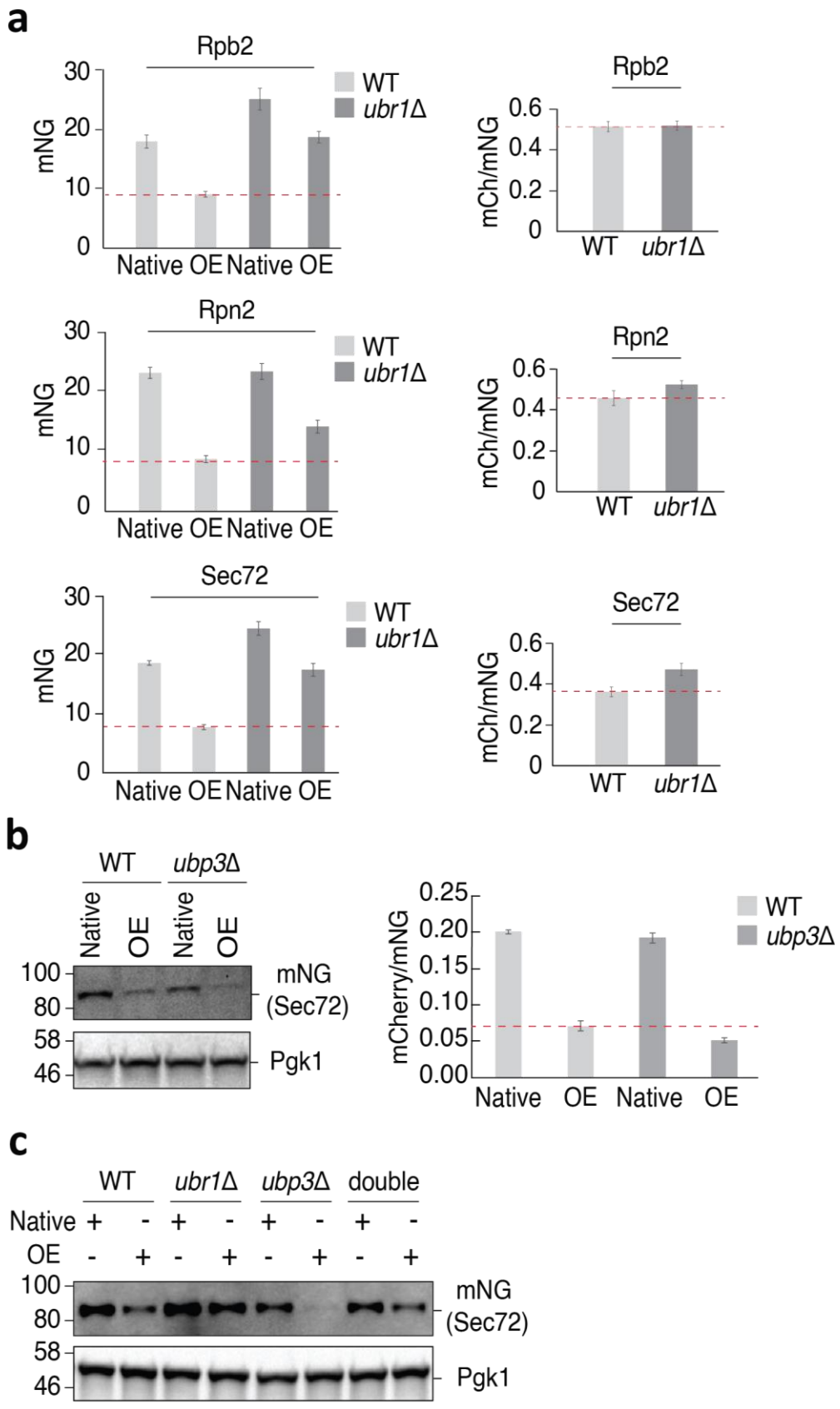


Figure 21. (a) Plate reader measured fluorescence of tFT strains harboring the overexpression plasmid (OE) or the *YGR045C* control plasmid (Native) in the presence and absence of UBR1. Dashed lines marked mNG intensities and ratios measured in overexpression conditions. (b) Immunoblot and flow cytometry detected protein levels of

Sec72 in wild type (WT) and *ubp3*Δ cells. (c) Western blot checked the effect of the *ubr1*Δ mutant, *ubp3*Δ mutant and *ubr1*Δ *ubp3*Δ double mutant on Sec72. Pgk1 was used as a loading control.

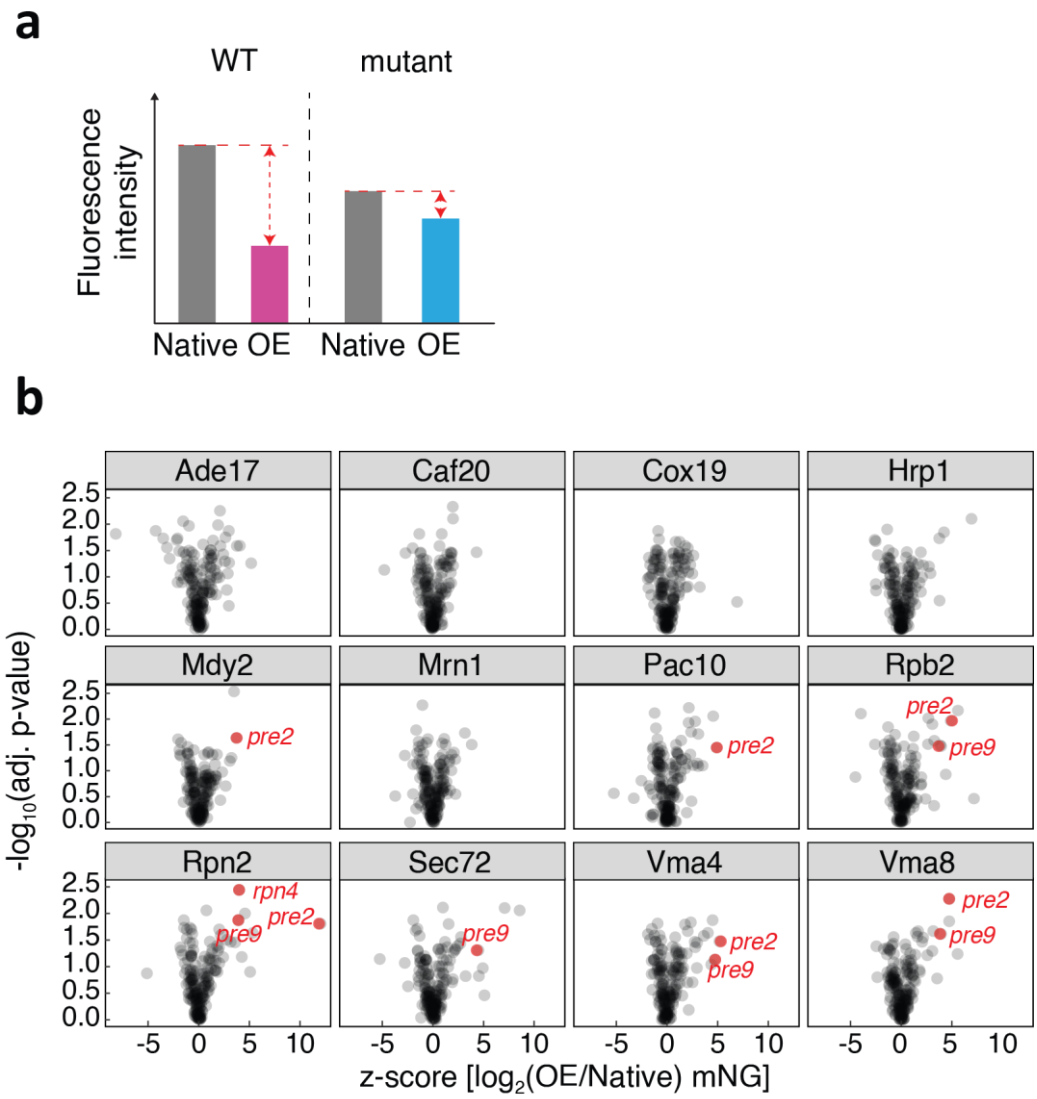


Figure 22. (a) The dashed line marked relative mNG intensity in the overexpression condition to the native condition. (b) Relative abundance on the $\log_2$ scale for each mutant of the UPS components was estimated by z-score. Mutants of proteasomal subunits with FDR less than 0.1 and z-score of relative mNG intensities >3.5 were highlighted in red.

7.5.1 UBR1, SAN1, AND TOM1 ARE CRUCIAL FACTORS FOR THE ELIMINATION OF MISLOCALIZED PROTEINS

The above experiment revealed that our screen approach can be used to identify quality control factors of mislocalized proteins and suggested the importance of cytosolic quality control factor Ubr1. The next aim was to apply this method to all identified destabilized proteins. To rationalize our efforts, we decided to cross all destabilized proteins with mutants of E3s function as quality control factors of misfolded or mislocalized proteins in different subcellular compartments. These E3s include cytoplasmic quality control factors Ubr1 and Tom1[78,94], the endoplasmic reticulum-associated ligases Hrd1, Asi1, and Asi3 involved in the ERAD[100,104,105] and the San1 ubiquitin ligase that mediates destruction of misfolded nuclear

proteins[78]. For each substrate, protein abundance in each mutant were compared to the negative control (*his3*Δ::*natMX*) (Figure 23a). Here, we observed that various proteins accumulated in the absence of *UBR1*, *SAN1* or *TOM1* (Figure 20b and Supplementary Table6), indicating that Ubr1, San1 and Tom1 are crucial factors for the elimination of mislocalized proteins.

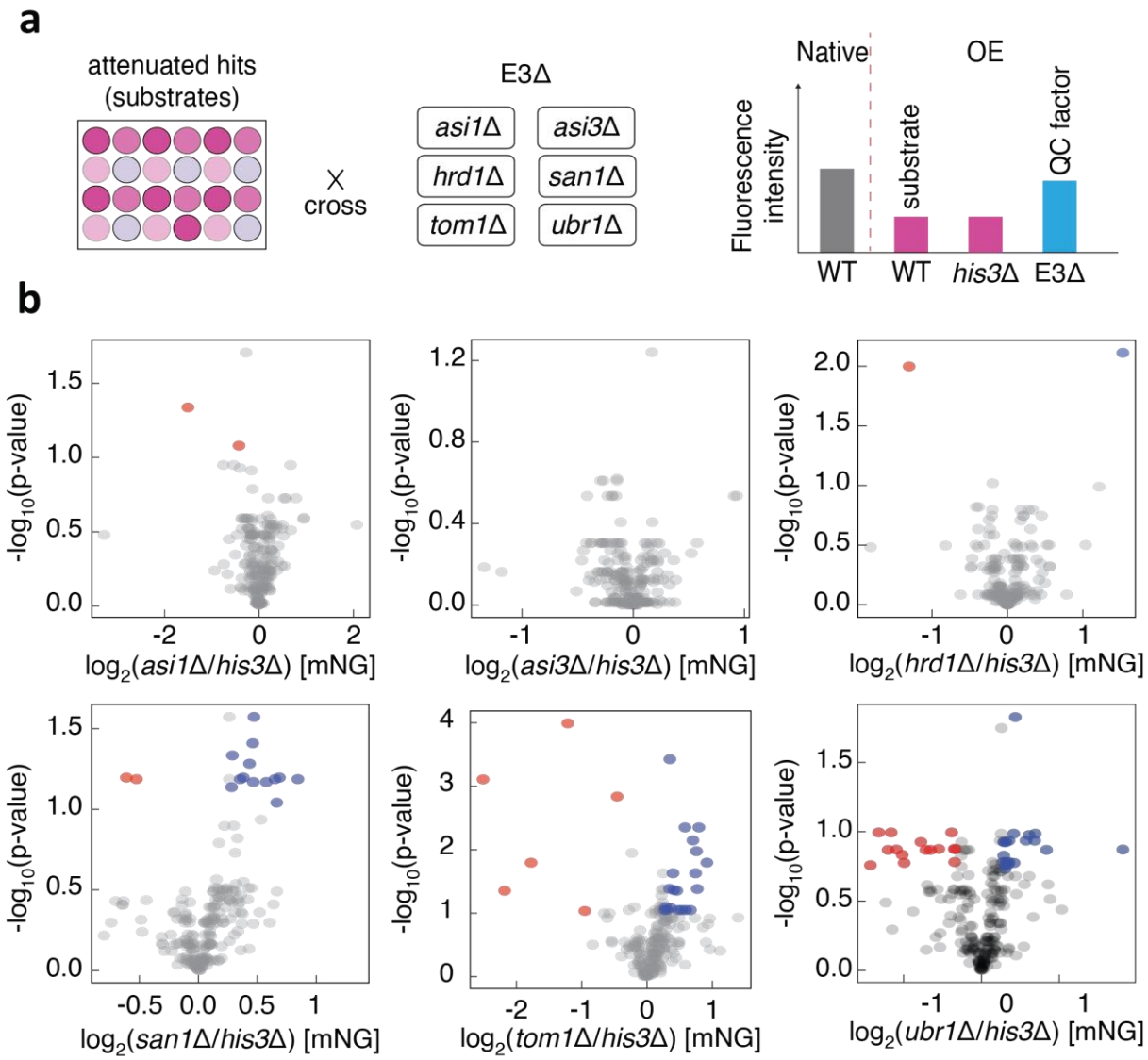


Figure 23. (a) Cartoon of screen approach to identify the indicated six quality control factors of mislocalized proteins. (b) Volcano plots displaying changes in protein abundance in the absence of indicated E3 ligases under the overexpression condition. Protein abundance with $\log_2$ fold change of at least 20% decrease was highlighted in red and with log2 fold change of at least 20% increase was highlighted in blue (0.1% FDR).

Ubr1. The E3 Ubr1 is a cytosolic quality control factor, which ubiquitinates misfolded proteins for degradation. This process depends on the Hsp70 chaperone[85] (Figure 24a). To determine whether the degradation of identified Ubr1 substrates is also dependent on members of the Hsp70 family, we tested protein abundance and turnover in a strain deleted two of the four Hsp70 chaperones (*SSA1*, *SSA2*, *SSA3* and *SSA4)*. We found cells lacking both Ssa1 and Ssa2 exhibited strong stabilization of Ubr1 substrates (Figure 24b), indicating that Ubr1 recognizes

excess proteins via Hsp70 chaperone upon protein misfolding. Therefore, we concluded that protein mislocalization leads to misfolding and subsequent Ubr1-dependent degradation.

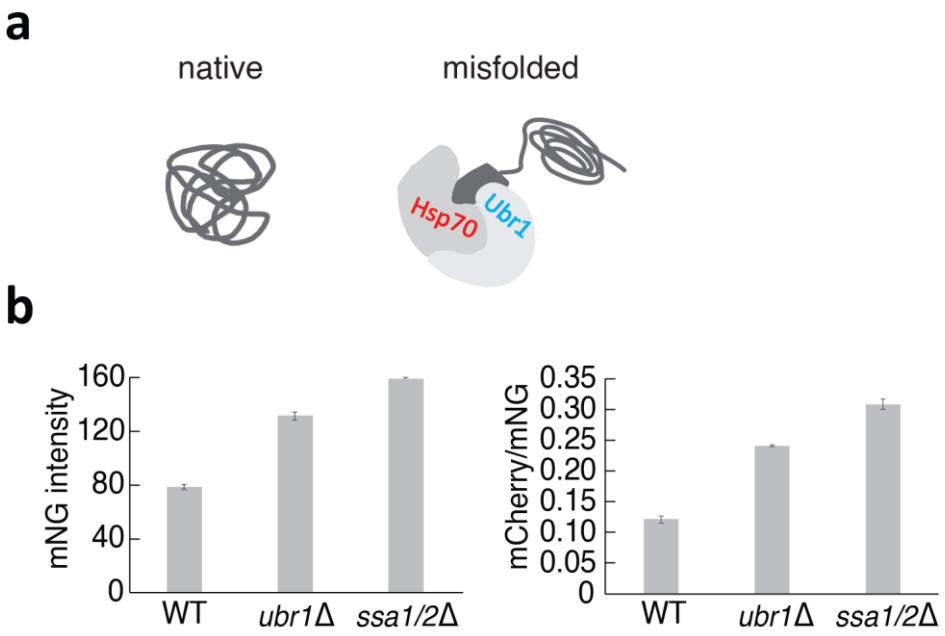


Figure 24. (a) Model of recognition mechanisms of misfolded substrates by Ubr1. (b) Flow cytometry measured the fluorescence of Sec72-tFT strains with the overexpression plasmid but lacking Ubr1 or lacking both Ssa1 and Ssa2.

Tom1. Tom1 is an E3 ubiquitin ligase of the HECT (homologous to E6AP C terminus) family. It recognizes excess ribosomal proteins and mediates them for degradation[94,108]. In our screen, we also observed many ribosomal subunits were affected in the absence of *TOM1* (Figure 25a).

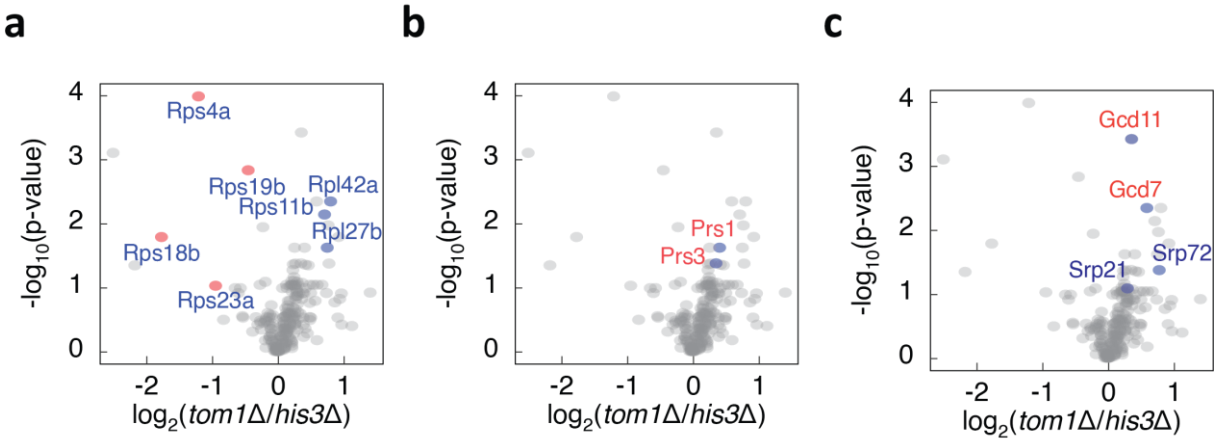


Figure 25. Volcano plot displaying changes in protein abundance in the absence of *TOM1* under the overexpression condition. Subunits belonging to the same complexes highlighted with the same color.

Moreover, Tom1 stabilized other subunits belonging to the same complexes. For example, Prs1 and Prs3 form heterodimer of the 5-phospho-ribosyl-1(alpha)-pyrophosphate synthetase[120]. Both protein levels were upregulated in *tom1*Δ cells (Figure 25b). *SRP21* and *SRP72* encode subunits of the SRP, a ribonucleoprotein complex that mediates the targeting of secretory

proteins to the ER[55]. Additionally, subunits of the translation initiation factor eIF2 Gcd7 and Gcd11 were also identified as potential substrates of Tom1 (Figure 25c). Therefore, our results showed that Tom1 plays an important role in recognizing and degrading unassembled subunits of multi-protein complexes.

7.6 NON-IMPORTED MITOCHONDRIAL RIBOSOMAL PROTEINS ARE NOT SUBSTRATES OF UBR1 AND TOM1

As shown in an earlier section, mitochondrial ribosomal subunits produced in excess do not tend to be degraded. One possible explanation is that overproduced subunits may form aggregates in the cytosol due to the hydrophobicity of mitochondrial targeting signals. Alternatively, these unassembled subunits could be stuck in the mitochondrial matrix because of the absence of UPS in mitochondria. In order to know whether mitochondrial ribosomal subunits can be degraded in the cytosol, I used the dataset from another study where they blocked the mitochondria import with depolarizing agent trifluoromethoxy carbonyl cyanide phenylhydrazone (FCCP). They found non-imported mitochondrial ribosomal proteins mainly localized to the cytoplasm and the nucleus[121]. Importantly, 45% of them showed reduced abundance[121] (Figure 26a), which hints that these proteins are being degraded. To investigate the degradation mechanisms, I selected three proteins that exhibited reduced abundance when import failed and treated the cells with MG132. It turns out that proteasome inhibitors can rescue non-imported mitochondrial ribosomal proteins from degradation, suggesting that the UPS is responsible for the clearance of mislocalized mitochondrial ribosomal proteins upon import failure (Figure 26b). Next, we tested whether Ubr1 or Tom1 mediates the degradation of non-imported mitochondrial ribosomal proteins. Our results showed that the protein levels were not affected in *ubr1*Δ and *tom1*Δ cells (Figure 23c), suggesting that other factors are involved in recognition and degradation of non-imported mitochondrial ribosomal proteins.

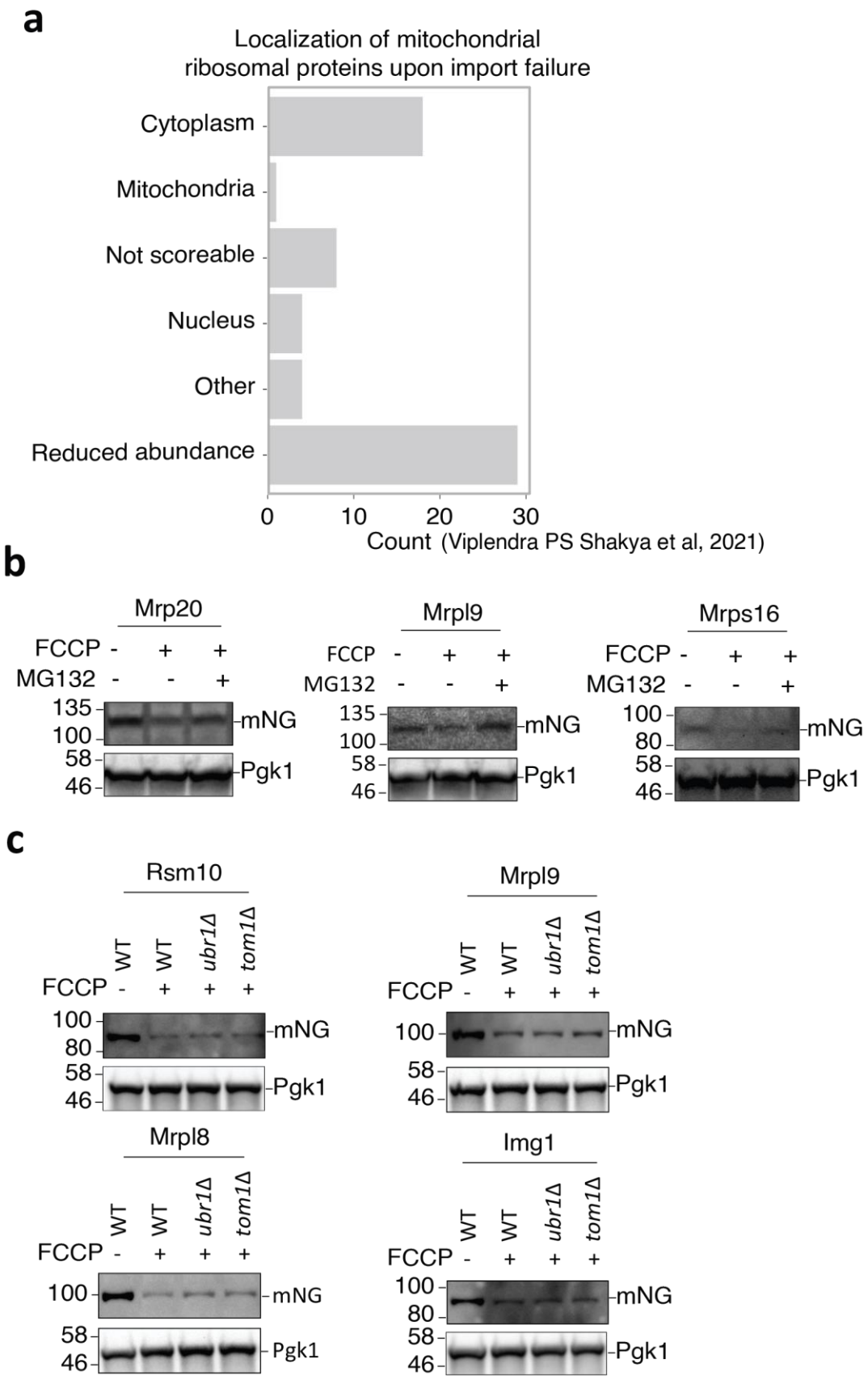


Figure 26. (a) Localization of mitochondrial ribosomal proteins when the import of mitochondria was blocked by FCCP. (b) Western blots checked the level of mCherry-mNG-tagged mitochondrial ribosomal proteins with FCCP or FCCP and MG132 treatments. Pgk1 was used as a loading control. (c) Western blots showing protein levels of mitochondrial ribosomes in WT, *ubr1*Δ and *tom1*Δ strains in the presence of FCCP.

8 DISCUSSION

8.1 STUDYING PROTEIN MISLOCALIZATION THROUGH OVEREXPRESSION

The functional state of a proteome relies not only on correct protein synthesis and folding but also on protein trafficking to the proper compartment and complex assembly with precise stoichiometry[4,41,51]. Proteins that fail to reach their native organelle or assemble into complexes are defined as mislocalized proteins. Protein mislocalization is a constitutive problem and has potentially detrimental impacts on normal cellular functions, loss of which may cause disease. Therefore, eukaryotic cells have evolved quality control systems to handle mislocalized proteins. The major way to handle mislocalized proteins is by degradation, which keeps their abundance as low as possible. This is why it is difficult to identify quality control substrates under non-perturbed conditions. Current studies mainly use artificial ER and mitochondrial proteins lacking the signal sequence, as well as membrane proteins synthesized in rabbit reticulocyte lysate, which force protein mislocalization to the cytosol in order to identify quality control factors. Despite significant discoveries made by applying these methods, the systems performing quality control of mislocalized proteins are unknown for most of the proteome.

To tackle this problem, our work established a platform to promote protein mislocalization through overexpression of individual proteins. We tracked the fate of each protein using a C-terminally tagged tFT. Our analysis revealed that, for the majority of the proteome, protein abundance does not change upon protein overexpression. However, our screen of 3,894 proteins showed that about 14.6% of them are significantly attenuated. Carrying an extra copy of one of the chromosomes in haploid yeast cells leads to a similar amount of proteins attenuated[112]. However, the global response of aneuploidy cells, such as cell cycle delays, metabolic disruptions, genomic instability, alterations to the transcriptome, and proteotoxic stress can also be responsible for protein attenuation[122,123,124,125,126]. Our approach of individual protein overexpression provides a much cleaner readout of what happens to a single protein when it is overexpressed and mislocalized.

One potential issue with our screening approach is that the tFT tag was placed at the C-terminus. Although there is no systematic comparison between N- and C-terminal tagging in terms of their impact on protein behavior, it is generally assumed that C-terminal tagging has fewer artifacts because a substantial fraction of the proteome carries protein localization signal at the very N-terminus[234]. For several protein classes that are known to be affected by C-terminal tagging, such as tail-anchored proteins, we can use the complementary N-terminal tagging library to determine protein behavior upon overexpression[127].

Another potential issue could be point mutations in the overexpression plasmids. In this work, I used colony PCR to verify plasmid identity and observed that all tested 71 plasmids in the MoBY2.0 library have correct identity. However, it is still possible that some plasmids may contain mutations, because mutations can only be verified by sequencing. This may affect the

final readout of the overexpression screen. For example, mutations within the protein interaction domain of the overexpressed protein may prevent it from forming the same protein complex as its endogenous counterpart. As a result, the endogenous counterpart is not being displaced from the protein complex, and mislocalization cannot be induced for any possible attenuation to occur. Therefore, the percentage of attenuated protein may be underestimated.

The overexpression level determined by qPCR analysis was to range between 1-7 folds with an average of 4.8 folds (data not shown). Proteins overproduced within this range do not affect cell fitness and physiology in general. In rare cases, elevated gene expression leads to cell growth defects. This subset of proteins might bring harmful effects on the normal cellular function when overexpressed. However, it could also serve as potential candidates to acutely induce cellular stress responses in the absence of drug treatment. For example, increased protein level of Cox5a results in a smaller colony size in our study. The discovery of the mitochondrial compromised protein import response (mitoCPR) pathway was achieved by overloading the mitochondrial import machinery through overexpression of Cox5a[128]. Therefore, our study provides candidates that may acutely trigger cellular stress, which can be used in discovering the downstream response pathways.

One of the limitation of this study is that the overexpression combined with the tFT assay can only be applied to yeast cells. This is because tagging endogenous proteins in human cells is significantly more time-consuming than in yeast, and achieving proteomic-wide overexpression screen is challenging in human cell lines. However, we could exploit protein abundance and turnover data from our study and aneuploidy yeast and human cells to select protein complexes that are composed of attenuated subunits in all datasets and dissect quality control mechanisms of those selected complexes in human cells.

8.2 ATTENUATED PROTEINS ARE STRONGLY ENRICHED FOR SUBUNITS OF PROTEIN COMPLEXES

The GO terms analysis reveals that attenuated proteins are strongly enriched for subunits of protein complexes. This finding is in line with many other studies that have shown subunits of multi-protein complexes may not exist for long unless they are assembled into their complex[129]. By overexpressing individual proteins and following their fate via the tFT, our analysis has systematically identified subunits that are specifically attenuated when they are produced in the excess of their binding partners.

Complexes can be assembled co-translationally or post-translationally. Recently, it was shown that co-translationally engaged subunits have a higher propensity to misfold or aggregate compared to their fully synthesized partners[51]. Using the selective ribosome profiling method, they analyzed the assembly of a hetero-trimeric complex, the multi-aminoacyl-tRNA synthetase, and found that the assembly initiates by cotranslational interactions of each of its subunits in a network-like manner. In other words, three subunits co-translationally engaged

with each other[51]. In our analysis, all subunits of the multi-aminoacyl-tRNA synthetase exhibits attenuation phenotype (data not shown). Thus, our study provides basic information about complex assembly. We assume that co-translationally engaged subunits are more likely subject to protein quality control.

The complexity of complex assembly increases with the number of subunits within one complex. Assembly order and the involvement of assembly factors further complicate the complex formation process. Our analysis provides a useful dataset regarding which unassembled subunits are specifically subject to quality control. What features distinguish subunits that are attenuated from non-attenuated subunits remain to be determined.

8.3 A SUBSET OF MISLOCALIZED PROTEINS ARE LIKELY DEGRADED BY THE UPS

Our analysis shows that attenuated and destabilized mislocalized proteins are likely degraded by the UPS. Although only four proteins have been tested the accumulation effect with the proteasomal inhibitor (MG132), several lines of evidence have indicated that the UPS plays a major role in removing mislocalized proteins for degradation. First, ribosomal proteins that are produced in stoichiometric excess over either rRNA or other ribosomal subunits are degraded by the UPS[94,108]. However, these excess proteins do not accumulate in the absences of the core autophagy gene *ATG7* or the vacuolar protease *PEP4* in nutrient-rich conditions, indicating that autophagy trafficking to the vacuole is not involved in degradation of overproduced ribosomal subunits[130]. Second, quantitative analysis of aneuploidy in yeast showed that inhibition of the proteasome but not vacuolar degradation could significantly restore the attenuation of protein levels, which demonstrates that the UPS are predominantly responsible for protein degradation of excess proteins on a global scale. Third, various components of the UPS have been recently implicated in quality control systems that recognize and degrade mislocalized proteins[94,95,131,132,133]. General and selective autophagy function to degrade damaged organelles or respond to the need for cellular building blocks in times of nutrient stress. Compared to removal of individual proteins by the UPS, vacuole/autophagy degradation is predominantly respond at cellular organelles or global level. Since individual protein mislocalization promoted by overexpression does not affect cell fitness and physiology in general, I propose that the UPS functions as the main selective degradation system to remove mislocalized proteins. Which substrates are specifically targeted for vacuole/autophagy degradation remains to be determined.

8.4 UBR1 AND TOM1 ARE IMPORTANT FACTORS FOR ELIMINATION OF MISLOCALIZED PROTEINS

Deletion of E3 ligases involved in quality control of misfolded or mislocalized proteins shows that Ubr1 and Tom1 are crucial factors for elimination of mislocalized proteins. This result is consistent with a previous study, which demonstrated that UBR1 acts in parallel with HUWE1 to target unassembled soluble proteins for proteasomal degradation in human cells. Together,

these findings underscore evolutionarily conserved protein quality control factors that target mislocalized proteins for degradation.

Ubr1 can target not only proteins bearing an exposed N-terminal destabilizing amino acids (N-degrons) but also misfolded or unassembled proteins that lack N-degrons[1338478]. Two substrates binding sites of Ubr1 for proteins with destabilizing N-terminal residues are located in the first 700 residues of the Ubr1. Ubr1 substrates whose degradation can be restored by point mutations in N-degron binding pockets are recognized by the Arg/N-degrons pathway. How Ubr1 recognizes misfolded or mislocalized proteins without N-degrons still needs to be determined. It had been shown that the mislocalized protein Δss-CPY* as well as unassembled Fas2, a subunit of the fatty acid synthase complex require Hsp70 chaperones of the Ssa type for degradation[133,78]. It could be that the consensus Hsp70 binding motif (RLLL) forms a substrate recognition module that works directly with Ubr1[134]. Alternatively, it may be that members of the Hsp70 family keep substrates in a soluble form that is recognized by Ubr1.

Moreover, Ubr1 has a third substrate binding site, which targets a transcriptional repressor Cup9 for degradation through an internal degron[84]. The recent discovery of the stress-induced homeostatically regulated protein degradation (SHRED) pathway revealed that the interaction between Ubr1 and cleaved Roq1 could transform substrate specificity of Ubr1, which results in enhanced degradation of misfolded proteins and inhibition of Cup9 degradation[135]. This finding suggests that the recognition of misfolded proteins is independent of the third substrate binding site of Ubr1. Our study identified many physiological substrates of Ubr1, which opens the door for seeking additional recognition sites of Ubr1.

The discovery of Excess Ribosomal protein Quality control (ERISQ) pathway revealed that Tom1 mediates degradation of unassembled ribosomal proteins. In our screen, we observe many ribosomal subunits are affected in the absence of Tom1. If Tom1 degrades unassembled ribosomal proteins, deletion of *TOM1* should accumulate ribosomal subunits. However, our analysis shows that protein abundance of four small ribosome subunits is down-regulated in the *tom1*Δ mutant. Although the yeast genome generally lacks introns, ribosomal genes are an exception[136]. All four ribosomal genes with introns allow gene expression to be controlled after transcriptional steps. It is possible that accumulation of these ribosomal subunits in the *tom1*Δ mutant triggers post-transcriptionally feedback regulation, which results in a decrease in the protein expression level.

Tom1 localizes in the nucleus in yeast. Its homolog in mammals HUWE1 is thought to be primarily in the cytosol. Although HUWE1 was implicated in degradation of the excess ribosomal subunit RPL26 in HEK293 cells, recognition of overproduced ribosomal proteins at distinct cellular compartments between Tom1 and HUWE1 suggest distinct quality control mechanisms in yeast and mammals. In a separate study, the hybrid E2/E3 enzyme UBE2O targets excess ribosomal proteins for degradation in the cytosol[92]. However, there is no obvious homolog of UBE2O in budding yeast. Whether unassembled ribosomal proteins are recognized as abnormal before import into the nucleus and which factors are involved in the clearance of cytosolic excess ribosomal subunits in yeast remain to be determined.

In addition to ribosomal proteins, Tom1 stabilizes many other subunits of multi-protein complexes and targets different subunits within the same complexes for degradation. Previous structural analysis showed that Tom1 ubiquitinated Rpl26a and Rpl4 on residues that are normally shielded in the intact ribosome[94]. We propose that Tom1 recognizes the binding interfaces of multi-protein complexes which are exposed due to overexpression. In mammal, an unbiased proteomic screen for HUWE1 substrates identified 72 candidates. Some substrates were immediately degraded when their respective binding partners are missing, suggesting that the degradation is caused by inefficient assembly.

Stable isotope labeling with amino acids (SILAC)-based quantitative mass spectrometry of HUWE1 and HUWE1 null cells indeed are able to identify HUWE1 substrates. However, the substrate diversity of this giant E3 ligase is very high, ranging from specific proliferation factors to unassembled complex subunits[13795]. The proteomic analysis cannot distinguish unassembled subunits from all 72 identified substrates. In addition, weak or transient association between Tom1 and its substrates might not be detected by mass spectrometry. This may explain why ribosomal proteins failed to be identified as HUWE1 targets in the proteomic screen. Our method is able to specifically identify Tom1 targets that are mislocalized, which provides the possibility to find the consensus sequence or sequence features that are necessary and sufficient for Tom1 recognition and degradation. Moreover, structure analysis revealed HUWE1 adopts a snake-like shape. This ring structure is highly dynamic and flexible, enabling the HECT domain to reach protein targets presented by the various acceptor sites[138]. It is also shown in this study that the HUWE1 uses the IDR1 region to recognized histone H1 and presumably positively charged proteins. By taking advantage of the Alphafold Protein Structure database and structure information of HUWE1, we can further explore the Tom1 binding domain that is specific for unassembled complex subunits.

Moreover, according to the Yeast GFP Localization database, identified Tom1 substrates localize at different sub-cellular compartments, including ER, cytoplasm and nucleus. It is likely that additional substrate adaptors might be needed to shuffle Tom1 substrates to nucleus for degradation.

One of the limitation of this method is that colony sizes affect the fluorescence readouts. Several components of the UPS that degrade mislocalized proteins are critical for cell fitness, such as Not4, Cdc48. Deletion of these factors or culture the temperature-sensitive alleles of essential UPS components at 30°C cause severe growth defect, which weaken the final readouts of protein accumulation effects.

Another potential issue of our study is the redundancy in the UPS system. Single knockouts of non-essential UPS components or temperature-sensitive alleles do not always lead to a loss of function in the UPS, as the system has redundancies that allow it to compensate and maintain its normal activity. Mislocalized proteins might fail to be stabilized by any of these single mutants due to the redundancies within the UPS. This may explain why Mdy2 and Vma8 were stabilized in the proteasomal *pre2*Δ and *pre9*Δ mutants, but no single mutant has been

identified as being responsible for protein degradation. It is also possible that these two mislocalized proteins are recognized and degraded by unknown factors, which could be identified through a genome-wide screening.

9 MATERIAL AND METHODS

Yeast strains used in this study are listed in Supplementary Tables 1. All strains are in S288c background, derivatives of BY4741 or Y8205. Yeast genome manipulations (gene deletion) were performed using PCR targeting and lithium acetate transformation. Experiments were performed at 30°C in synthetic complete (SC) medium with 2% (w/v) glucose as carbon source, unless stated otherwise.

PCR cassettes containing nat-marker were amplified as follows:

PCR mix		PCR program
10x HiFi-buffer dNTPs (10 mM)	5µL	1. 95°C 3 mins
dNTP-Mix	5µL	2. 95°C 20 sec
MgCl2 (50 mM)	1.75µL	3. 72°C 30 sec
S1/S2 primers (10 µM)	2.5 µL / each	4. 72°C 45 sec / 1kb
Template (0.5 µg/µL)	1 µL	5. Go to 2, 30 cycles
HiFi polymerase	1 µL	6. 72°C 5 mins
Betaine (5 M)	5 µL	7. Pause at 4°C
H2O	26.25 µL	

PCR cassettes without nat-marker were amplified as follows:

PCR mix		PCR program
10x HiFi-buffer dNTPs (10 mM)	5µL	1. 95°C 3 mins
dNTP-Mix	5µL	2. 95°C 20 sec
MgCl2 (50 mM)	1.25µL	3. 66°C 30 sec
S1/S2 primers (10 µM)	2.5 µL / each	4. 72°C 45 sec / 1kb
Template (0.5 µg/µL)	1 µL	5. Go to 2, 30 cycles
HiFi polymerase	1 µL	6. 72°C 5 mins
Betaine (5 M)	5 µL	7. Pause at 4°C
H2O	26.25 µL	

The gene deletions were confirmed by colony PCR. Forward primers annealed about 200 bp upstream of the deleted gene of interest. The generic reverse primer annealed inside the selection marker.

Colony PCR reaction was made as follows:

PCR mix		PCR program
10x long incubation buffer	4µL	
dNTP-Mix (10mM)	1.6µL	1. 97°C 7 mins
MgCl2 (50 mM)	1µL	2. 95°C 30 sec
Betaine (5 M)	5 µL	3. 58-61°C 30 sec
ORF-specific forward primer (5 µM)	4 µL	4. 72°C 60 sec / 1kb
Generic reverse primer (100 µM)	0.2 µL	5. Go to 2, 38 cycles
Taq DNA polymerase	1 µL	6. 72°C 5 mins
A dense culture of each strain	1 µL	7. Pause at 4°C
H2O	23.6 µL	

9.1 PLASMID CONSTRUCTION

Plasmids used in this study are listed in Supplementary Table 2. Construction of pLZY001 was done using NEBuilder HiFi DNA Assembly. The the DNA fragment containing 3xMyc tag and two restriction enzyme sites was amplified from the pYM4 plasmid. PCR reaction and restriction enzyme digestion were made as follows:

PCR mix		PCR program
5x HiFi-buffer dNTPs (10 mM)	10µL	1. 98°C 2 mins
dNTP-Mix	5µL	2. 95°C 30 sec
S1/S2 primers (10 µM)	2.5 µL / each	3. 58°C 30 sec
pYM4	100 ng	4. 72°C 30 sec / 1kb
Velocity DNA polymerase	1 µL	5. Go to 2, 30 cycles
H2O	31 µL	6. Pause at 4°C

PCR products and the type II plasmid (pMaM484) were digested by BamHI and SpeI restriction enzymes, followed by ligation and E.coli propagation. The integration of the 3xMyc tag into the type II donor plasmid was confirmed by Sanger sequencing.

Restriction enzyme digestion were made as follows:

Restriction enzyme digestion		Program
BamH1-HF	10µL	
Spei-HF	5µL	PCR product: 37°C 1h
Cutsmart	2.5 µL / each	pYMam484: 37°C 2h
pYM4	100 ng	
PCR product / pYMam484	1 µL	

Yeast plasmid extraction was performed using Zymolyase and Miniprep method. 1.0-1.5×10^7 cells were spun down by centrifugation at 600g for 2 min. The supernatant was then removed and the pellet was suspended in the resuspension buffer containing Zymolyase to digest yeast cell walls. The mixture was incubated at 37^{o}C for 15-60 mins. Subsequent steps were the same as for E.coli plasmid miniprep. Next, 5-10 ul of isolated DNA was used to transform into E.coli cells for propagation. Plasmids were then extracted using the mini-prep kit, followed by Sanger sequencing to confirm DNA sequences.

9.2 MEASUREMENTS OF PROTEIN ABUNDANCE AND TURNOVER UPON OVEREXPRESSION

The annotation file of the MoBY2.0 overexpression library contains the coordinates of each ORF. The new layout of the tFT library and the corresponding untagged non-fluorescence library was generated according to the positions of each strain in the MoBY2.0 library. For re-arraying, the tFT and non-fluorescence libraries were first pinned onto agar plates. Colonies were picked from source agar plates to new targeted plates based on coordinates using a colony picker robot. The tFT and non-fluorescence libraries were re-arrayed into the format in which the ORF positions matched the MoBY2.0 overexpression library.

The re-arrayed tFT library was then assembled with the untagged library in a 1536-array format, with 3 technical replicates of each tFT-tagged strain next to replicates of a non-fluorescence strain. A set of reference strains spanning the full range of protein abundances and stabilities were placed on each plate. Dummy colonies were placed on the outer rows and columns to minimize the influence of nutrient access on colony size and fluorescence readout.

Each strain was then crossed with the overexpression strain carrying the corresponding ORF and the non-overexpression strain carrying the plasmid with ORF *YGR045C*. The crossing was done by sequentially pinning the strains on the appropriate media according to the SGA method as follows:

1. mating of tFT and MoBY2 overexpression library strains on YPD plates
2. selection of diploids on SC(MSG)-Leu + hygromycin plates – done twice
3. sporulation for 7 days on SPO plates at 23 °C

4. selection of haploids, step 1, on SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine plates
5. selection of haploids, step 2, on SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine/hygromycin plates

For the fluorescence measurements, strains from the final step were pinned onto SC-His/Arg/Lys/Leu + Adenine plates. After 24 hours of growth, The colonies were imaged on the PhenoBooth and fluorescence intensities were measured using Infinite M1000 or Infinite M1000 Pro plate readers (Tecan) equipped with stackers for automated plate loading (Tecan) and custom temperature control chambers.

Measurements in mCherry (587/10 nm excitation, 610/10 nm emission, optimal detector gain) and sfGFP (488/10 nm excitation, 510/10 nm emission, optimal detector gain) channels were performed at 400 Hz frequency of the flash lamp, with ten flashes averaged for each measurement. The colony fluorescence measurement data and images were then imported into R for data analysis.

Colony sizes were determined by the R package gitter. Measurements of border colonies were removed and all measurements of empty positions, based on the images, were replaced with NA to be certain which values should be missing. Next, absolute fluorescence intensities were normalized across plates by the median of the reference strains on each plate. Cellular autofluorescence was corrected by subtracting the median of the non-fluorescent control for each tFT strain. Technical replicates were summarized by taking the median. Then, changes of green fluorescence intensity and ratio were calculated by dividing the $\log_2$ transformed values in overexpression conditions versus non-overexpression conditions. P values were computed using a t-test and then adjusted for multiple testing using the method of Benjamini-Hochberg to determine significance.

9.3 CONSTRUCTION OF THE MYC-TAGGED LIBRARY

120 C-SWAT strains were picked from the full C-SWAT library and transferred to new plates using the colony-picking robot. The selected C-SWAT strains were then crossed with the 3Myc donor strains yLZY0048 in a 384-array format on agar. The crossing and tag swapping was done by sequentially pinning the strains on the appropriate media according to the SGA method as follows:

1. mating of C-SWAT strians and donor strains on YPD plates
2. selection of diploids on SC(MSG)-Ura + G418 plates – done twice
3. sporulation for 7 days on SPO plates at 23 °C
4. selection of haploids, step 1, on SC(MSG)-His/Arg/Lys/Ura + canavanine/thialysine plates
5. selection of haploids, step 2, on SC(MSG)-His/Arg/Lys/Ura + canavanine/thialysine/G418 plates

6. selection of haploids, step 3, on SC(MSG)-His/Arg/Lys/Ura + canavanine/thialysine/G418/clonNAT plates
7. induction of I-SceI enzyme for tag swapping on YP + Raf/Gal plates - done twice
8. counter-selection against the acceptor module on SC-His + 5-FOA plates
9. recovery after counter-selection on SC(MSG)-His/Arg/Lys + canavanine/thialysine/ hygromycin/clonNAT plates

The integration of 3MYC tag was checked by immunoblotting.

9.4 IMMUNOBLOTTING

(i) For log phase experiments with ORF-3MYC and ORF-tFT strains, cells were grown to $2x10^7$-$3x10^7$ cells/ml in SC medium or SC medium lacking leucine (to select for plasmids). 1 ml samples were mixed with 150 µl of 1.85 M NaOH.

(ii) For proteasome inhibition experiments with the ORF-3xMYC pdr5Δ strains, log phase cultures were treated with MG132 to 80 µg/ml final concentration or DMSO as control for 90 min, followed by cell harvesting as above.

For (i) and (ii), whole cell protein extracts were prepared by precipitation with 150 µl of 55% (w/v) of trichloroacetic acid, followed by centrifugation to remove the supernatant. The pellet was resuspended in 50µl of HU buffer per $1x10^7$ cells, followed by SDS-PAGE separation and immunoblotting.

Membranes were probed with mouse anti-MYC or rabbit anti-mNG antibodies for myc- or tFT-tagged strains, respectively. Secondary goat anti-mouse immunoglobulin G (IgG)-horseradish peroxidase (HRP) and goat anti-rabbit IgG-HRP antibodies were used for detection. The loading control Pgk1 was detected using mouse anti-Pgk1 antibodies followed by the same anti-mouse secondary antibodies. The ChemiDoc MP imaging system (Bio-Rad) was used to image the membranes after addition of the Pierce ECL Plus Western Blotting Substrate. Image quantification was performed using ImageLab.

9.5 THE UPS SCREEN FOR QUALITY CONTROL FACTORS OF MISLOCALIZED PROTEINS

Timer-tagged strains carrying the corresponding overexpression construct were crossing with the UPS mutant assay. The crossing was done by sequentially pinning the strains on the appropriate media according to the SGA method as follows:

1. mating of tFT strains carring the corresponding overexpression or *YGR045C* control plasmid with the UPS mutant array on YPD plates
2. selection of diploids on SC(MSG)-Leu + hygromycin/clonNAT. plates – done twice
3. sporulation for 7 days on SPO plates at 23 °C
4. selection of haploids, step 1, on SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine plates
5. selection of haploids, step 2, on SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine/hygromycin plates

6. selection of haploids, step 3 SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine/hygromycin/ clonNAT

Raw data values were corrected for spatial effects by local regression, scaled across plates by the median of the reference strains on each plate. Cellular autofluorescence was corrected by subtracting the median of the non-fluorescent control for each tFT strain. Changes in protein abundance (mNG intensity) and stability (mCherry/mNG ratio) were estimated by calculating z-scores.

9.6 E3 LIGASES SCREEN FOR POTENTIAL SUBSTRATES

Timer-tagged strains and the corresponding non-fluorescent strains were crossed with knockout alleles of *UBR1*, *SAN1*, *ASI1*, *ASI3*, *TOM1*, *HRD1* and *his3*Δ::*natNT2* were used as negative control. The crossing was performed by sequential pinning of yeast colonies on appropriate selective media using pinning robots as follows:

1. mating of tFT strains (and the corresponding non-fluorescent strains) and knockout alleles on YPD plates
2. selection of diploids on YPD + clonNAT + hygromycin plates – done twice
3. sporulation for 7 days on SPO plates at 23 °C
4. selection of haploids, step 1, on SC(MSG)-His/Arg/Lys + canavanine/thialysine plates
5. selection of haploids, step 2, on SC(MSG)-His/Arg/Lys + canavanine/thialysine/hygromycin plates
6. selection of haploids, step 3 SC(MSG)-His/Arg/Lys + canavanine/thialysine/hygromycin/clonNAT plates

The knockout alleles of each ligase were assembled with the *his3*Δ::*natNT2* negative control on alternative raws. The new layout of each timer-tagged strains and the corresponding non-fluorescent strains was generated in consistent with the position of each MoBY2.0 overexpression strain in 384 format. Then new assembled tFT library in 384 format were transferred into 1536 format with 3 technical replicates. The reference strain (a radomly chosen stain with strong mNG fluorescence signal) was spaced evenly across each plate at the fourth position. The MoBY2.0 overexpression strains were replicated four times by pinning them from 384 format to a 1536 format. Then the two libraries were crossing with each other by sequential pinning of yeast colonies on appropriate selective media using pinning robots as follows:

1. mating of tFT (and the corresponding non-fluorescent strains) and MoBY2 overexpression library strains on YPD plates
2. selection of diploids on SC(MSG)-Leu + hygromycin/ clonNAT plates – done twice
3. sporulation for 7 days on SPO plates at 23 °C
4. selection of haploids, step 1, on SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine plates

5. selection of haploids, step 2, on SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine/hygromycin plates
6. selection of haploids, step 3 SC(MSG)-His/Arg/Lys/Leu + canavanine/thialysine/hygromycin/clonNAT

Strains from the final step were pinned onto SC-His/Arg/Lys/Leu + Adenine plates. After 24 hours of growth, The colonies were imaged on the PhenoBooth and fluorescence intensities were measured at 30°C as detailed above.

Data from dummy colonies were first removed and then mNG fluorescence intensities were normalized to the median of all reference strains on each plate. Fluorescence intensities were then corrected for spatial effects by normalizing to the median of the local reference strain in the 4x4 quadrant and corrected for cellular autofluorescence by subtracting the median of the non-fluorescent control strain. Once all normalization procedures were completed, we compared the green fluorescence intensities of each E3 knockout with the knockout control for each tFT-tagged strain. p-values were corrected using the Benjamini-Hochberg correction method.

9.7 FLOW CYTOMETRY

High-throughput flow cytometry was performed on a BD LSRFortessa Cell Analyzer (BD Biosciences) equipped with a high-throughput sampler loader. Settings on the high-throughput sampler were as below:

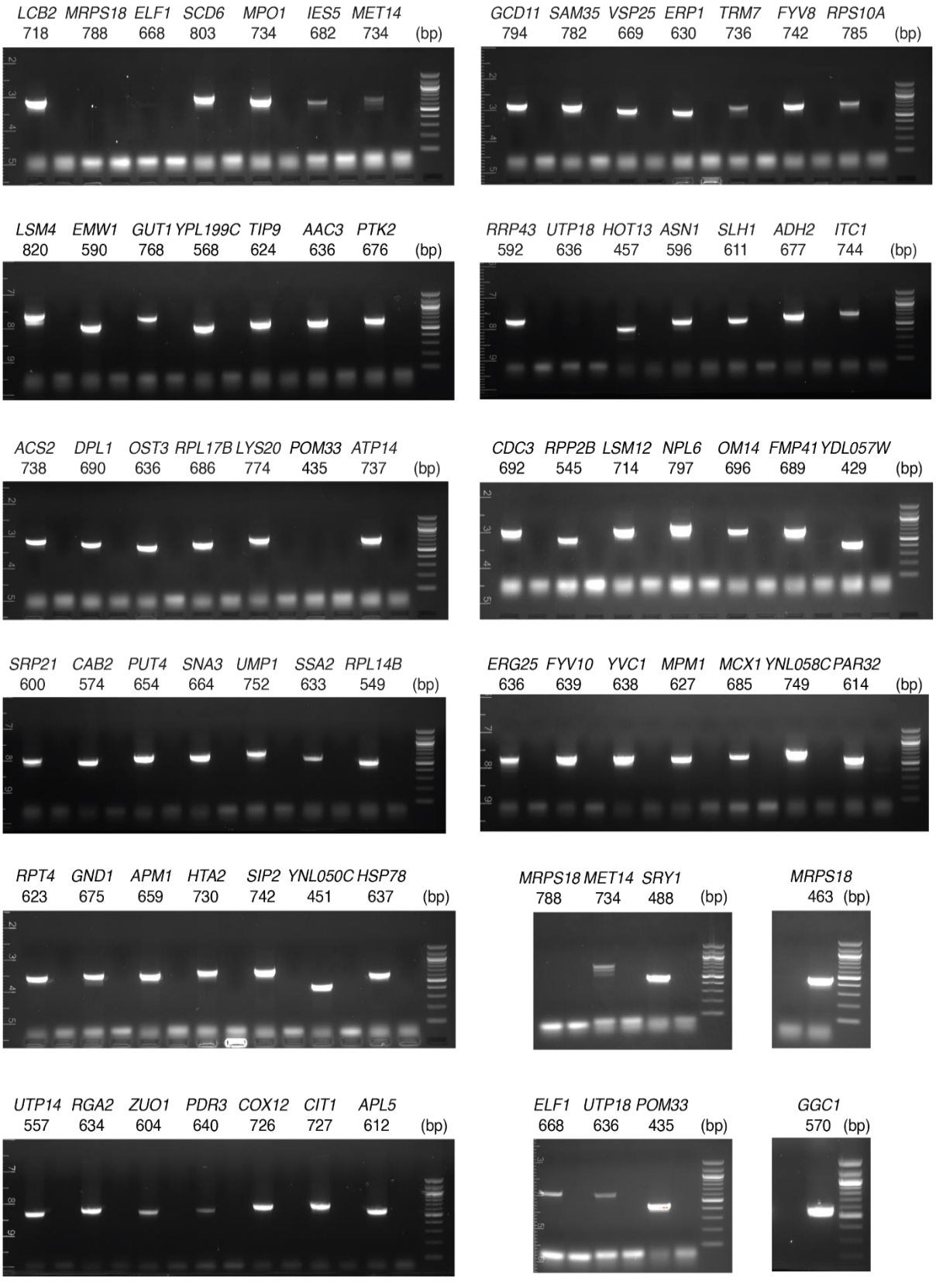
sample volume	15ul	mixing speed	180ul/sec
flow rate	0.5ul/sec	mixing volume	50ul
number of mixes	3	wash volume	400ul

Strains were grown to saturation in 96-well plates, diluted into fresh SC medium lacking leucine (to select for plasmids), and grown for 6 hours to 6-8×10^6 cells/ml. Single-cell fluorescence intensities were measured by 488- and 561-nm lasers for the red and green signals respectively. Cell size and complexity were determined by forward and side scatter (FSC and SSC). At least 50000 cells were analyzed for each strain.

10 APPENDIX I

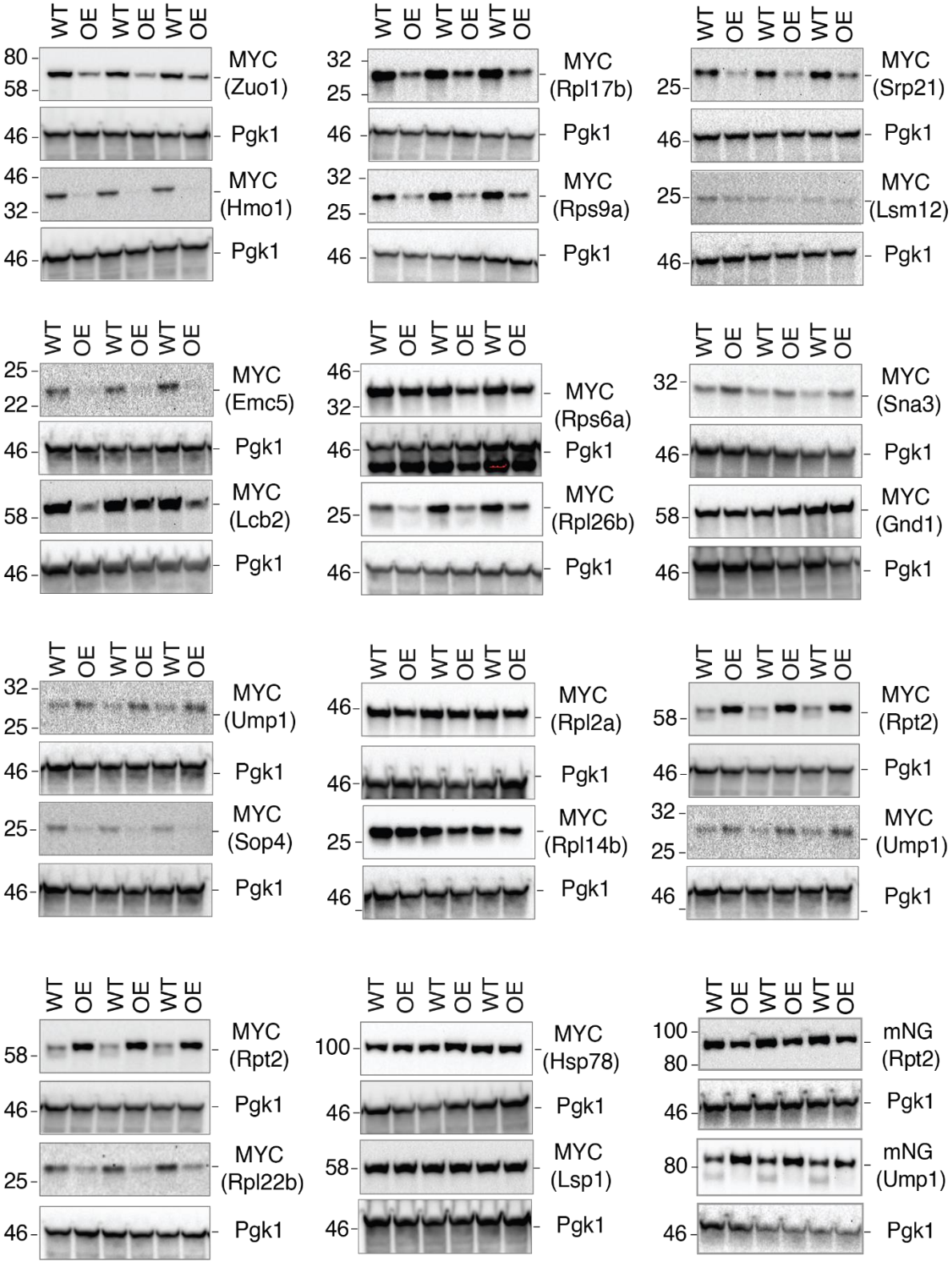
10.1 SUPPLEMENTARY

10.1.1 FIGURE S1. COLONY PCR VERIFICATION OF PLASMIDS IDENTITY



Supplementary Figure 1. Agarose gel electrophoresis (AGE) image of the colony PCR products. The exhibited bands are PCR products of different plasmids with expected sizes. A control PCR was loaded next to each sample PCR. Last lane was loaded with 100bp DNA ladder.

10.1.2 FIGURE S2. IMMUNOBLOTTING TO DETECT ENDOGENOUS PROTEIN LEVELS UPON OVEREXPRESSION



Supplementary Figure 2. Immunoblotting to detect endogenous protein levels with MYC antibody in the native condition (WT) or cells harboring 2u plasmids (OE). Pgk1was used as loading control. Experiments were done with three biological replicates.

10.1.3 TABLE S1. YEAST STRAINS

Strain	Genotype	Source
BY4741	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	(Baker Brachmann et al., 1998)
Y8205	MATalpha his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 can1Δ::STE2pr-SpHIS5 lyp1Δ::STE3pr-LEU2	(Tong and Boone, 2007)
yMaM639	Y8205 leu2Δ0::GAL1pr-NLS-I-SCEI-natNT2	(Meurer et al., 2018)
yLZY0042	MATalpha lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 can1Δ::STE3pr-SpHIS5-ADH1term-GAL1pr-NLS-I-SCEI pdr5Δ::natNT2	this study
yLZY0047	MATalpha lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 can1Δ::STE3pr-SpHIS5-ADH1term-GAL1pr-NLS-I-SCEI PDR5Δ::natNT2 pRS41K-ADH1term-TEFpr-hphΔC	this study
yLZY0048	MATalpha lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 can1Δ::STE3pr-SpHIS5-ADH1term-GAL1pr-NLS-I-SCEI PDR5Δ::natNT2 pRS41K-3xMYC-ADH1term-TEFpr-hphΔC	this study
yLZY0055	BY4741 pMoBY2_YGR045C	this study
yLZY0064	MATα, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lyp1Δ, can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI, Ubi4-mCherry-mNG-hphNT1 pMoBY2_YGR045C	this study
yLZY0065	MATα, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lyp1Δ, can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI, Ubi4-hphNT1 pMoBY2_YGR045C	this study
yLZY0068	MATα, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lyp1Δ, can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI, YJR096W-mCherry-mNG-hphNT1 pMoBY2_YGR045C	this study
yLZY0069	MATα, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lyp1Δ, can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI, YJR096W-hphNT1 pMoBY2_YGR045C	this study
yLZY0152	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRP20-mCherry-mNG-hphNT1	this study
yLZY0153	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPL9-mCherry-mNG-hphNT1	this study
yLZY0159	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPS16-mCherry-mNG-hphNT1	this study
yLZY0162	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRP20-mCherry-mNG-hphNT1 pdr5Δ::natNT2	this study
yLZY0163	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPL9-mCherry-mNG-hphNT1 pdr5Δ::natNT2	this study
yLZY0169	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPS16-mCherry-mNG-hphNT1 pdr5Δ::natNT2	this study
yLZY0184	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ, can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 pMoBY2_YGR045C	this study
yLZY0185	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ, can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 pMoBY2_YLR292C	this study
yLZY0186	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ubp3Δ::natNT2 pMoBY2_YGR045C	this study
yLZY0187	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ubp3Δ::natNT2 pMoBY2_ YLR292C	this study
yLZY0188	BY4741 hrd1Δ::natNT2	this study
yLZY0189	BY4741 san1Δ::natNT2	this study
yLZY0190	BY4741 asi1Δ::natNT2	this study
yLZY0191	BY4741 asi3Δ::natNT2	this study
yLZY0192	BY4741 doa10Δ::natNT2	this study
yLZY0194	BY4741 his3Δ::natNT2	this study
yLZY0195	BY4741 ubr1Δ::natNT2	this study
yLZY0199	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI RSM10-mCherry-mNG-hphNT1 ubr1Δ::natNT2	this study
yLZY0200	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI RSM10-mCherry-mNG-hphNT1 tom1Δ::natNT2	this study
yLZY0201	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPL8-mCherry-mNG-hphNT1 ubr1Δ::natNT2	this study
yLZY0202	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPL8-mCherry-mNG-hphNT1 tom1Δ::natNT2	this study

Strain	Genotype	Source
yLZY0203	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI IMG1-mCherry-mNG-hphNT1 ubr1Δ::natNT2	this study
yLZY0204	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI IMG1-mCherry-mNG-hphNT1 tom1Δ::natNT2	this study
yLZY0207	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPL9-mCherry-mNG-hphNT1 ubr1Δ::natNT2	this study
yLZY0208	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI MRPL9-mCherry-mNG-hphNT1 tom1Δ::natNT2	this study
yLZY0210	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ubr1Δ::CaURA3 ubp3Δ::natNT2 pMoBY2_YGR045C	this study
yLZY0211	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ubr1Δ::CaURA3 ubp3Δ::natNT2 pMoBY2_ YLR292C	this study
yLZY0212	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1	this study
yLZY0213	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ubr1Δ::natNT2	this study
yLZY0220	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 nta1Δ::natNT2 pMoBY2_YGR045C	this study
yLZY0221	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 nta1Δ::natNT2 pMoBY2_ YLR292C	this study
yLZY0222	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ate1Δ::natNT2 pMoBY2_YGR045C	this study
yLZY0223	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ate1Δ::natNT2 pMoBY2_ YLR292C	this study
yLZY0239	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ssa2Δ::natNT2 ssa1Δ::CaURA3 pMoBY2_YGR045C	this study
yLZY0240	MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 lyp1Δ can1Δ::STE3prα -SpHIS5-GAL1pr-I-SCEI SEC72-mCherry-mNG-hphNT1 ssa2Δ::natNT2 ssa1Δ::CaURA3 pMoBY2_ YLR292C	this study

10.1.4 TABLE S2. PLASMIDS

Plasmid	Description	Source
pMoBY2.0	p5476 (2μ LEU2 AmpR)	(Ho, 2011; Magtanong et al., 2011)
pMoBY2.0_YGR045C	p5476 (2μ LEU2 AmpR) YGR045C KanMX	(Ho, 2011; Magtanong et al., 2011)
pMoBY2.0_RPL26B	p5476 (2μ LEU2 AmpR) RPL26B KanMX	(Ho, 2011; Magtanong et al., 2011)
pMoBY2.0_PEX15	p5476 (2μ LEU2 AmpR) PEX15 KanMX	(Ho, 2011; Magtanong et al., 2011)
pMaM484	pRS41K CEN ARS kanMX (type II donor template)	(Meurer et al., 2018)
pFA6a-natNT2	pFA6a-natNT2 (deletion cassette, natNT2 selection marker)	(Janke et al., 2004)
pLZY001	pRS41K-3xMYC-ADH1term-TEFpr-hphΔC	this study
pLZY005	pSD-N1-L1-mCherry-mNeonGreen-L2	this study
pEI080	p413TEF-NLS-SCEI	(Ivanova, 2012)

10.1.5 TABLE S3. LIBRARIES

Name	Description	Source
MoBY2	BY4741 MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 Plasmid: p5476 2μ LEU2 AmpR ORF KanMX	(Ho, 2011; Magtanong et al., 2011)
mORF	Y258 MATa, pep4-3, his4-580, ura3-52, leu2-3,112 Plasmid: BG1805 URA3 GAL1pr-ORF-His6-HA^{epitope}-3C$^{Protease\ site}$-ZZ domain^{ProteinA}-GAL1term	(Gelperin et al., 2015)
C-SWAT	BY4741 ORF-C-SWAT-URA3	(Meurer et al., 2018)
tFT (mCherry-mNG) type II	MATalpha ORF-mCherry-mNG-hphNT1 lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 can1Δ::STE3pr-SpHIS5-ADH1term-GAL1pr-NLS-I-SCEI	(Jia Jun Fung, 2023)
tFT (mCherry-sfGFP) type II	MATalpha ORF-mCherry-sfGFP-hphNT1 lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 can1Δ::STE3pr-SpHIS5-ADH1term-GAL1pr-NLS-I-SCEI	(Jia Jun Fung, 2023)
UPS mutant array	BY4741 ORF::kanMX	(Kong et al., 2021)

10.1.6 TABLE S4. PLATE READER MEASUREMENTS OF COLONIES OF SELECTED 120 STRAINS

ORF	gene	log_2(OE/Native) [mNG].screen	-log_{10}(adj.p-value).screen	log_2(OE/Native) [mNG].replicate	-log_{10}(adj.p-value).screen
YAL033W	POP5	0.233913666	0.046165532	0.049056979	0.293068127
YAR002C-A	ERP1	-0.205563585	0.294006227	-0.087394849	0.155922991
YBR061C	TRM7	-0.339219957	0.01332051	-0.341851189	0.014288696
YBR067C	TIP1	-0.356795807	0.005556809	0.031025262	0.111031101
YBR173C	UMP1	0.279557377	0.076859142	0.010686668	0.907990124
YBR227C	MCX1	0.07342338	0.63292217	-0.248451549	0.111031101
YBR230C	OM14	-0.973682655	0.006550917	-1.043972978	0.002439142
YBR279W	PAF1	-0.738255492	0.01242481	-0.56793368	0.019923073
YCR048W	ARE1	0.108693399	0.483527347	-0.079120147	0.534927304
YDL007W	RPT2	-1.619927606	0.005070202	-1.240134004	0.010952702
YDL106C	PHO2	-0.195591154	0.049889203	-0.191041742	0.012115484
YDL182W	LYS20	0.540117768	0.000887523	0.38586175	0.006195323
YDL185W	VMA1	-0.455003495	0.009987009	-0.623575363	0.003143516
YDL198C	GGC1	-3.076242557	0.027092658	1.712454125	0.039927159
YDR062W	LCB2	-1.144786708	0.003831679	-1.089555265	0.00354253
YDR079C-A	TFB5	-0.177892534	0.059150684	0.149821386	0.366872916
YDR174W	HMO1	-2.022915206	0.011157147	-2.514792956	0.003349542
YDR177W	UBC1	0.555860752	0.051762128	0.155123461	0.041595698
YDR205W	MSC2	-0.439331434	0.095062909	-0.25625495	0.173954164
YDR258C	HSP78	0.014182892	0.713787434	0.098972804	0.45462645
YER021W	RPN3	-0.748061299	0.031130872	-1.009268884	0.003143516
YER025W	GCD11	-1.195905719	0.001200323	-1.202828613	0.001031312
YER069W	ARG5,6	-0.3966367	0.06391712	-0.280281521	0.00656885
YER071C	TDA2	-0.04453057	0.6655439	0.075336133	0.45377352
YER077C	MRX1	-0.1630647	0.1678594	0.141226259	0.173954164
YER112W	LSM4	0.4550607	0.04564908	0.751784675	0.006673836
YER146W	LSM5	-0.8029567	0.02271901	-0.760350321	0.014884667
YER177W	BMH1	-0.6193637	0.03302233	-0.795244618	0.00354253
YFL034C-A	RPL22B	-1.016273	0.009960026	-0.817892251	0.00603948
YFR031C-A	RPL2A	-0.4376979	0.03858356	-0.269847447	0.280740249
YGL010W	MPO1	-0.03721173	0.6354412	0.155595223	0.13313179
YGL133W	ITC1	-0.7856501	0.05660591	-0.547787364	0.391186663
YGL227W	VID30	-0.9201743	0.05176213	-0.685931828	0.222083896
YGR034W	RPL26B	-1.281061	0.004639775	-1.183132922	0.003143516
YGR060W	ERG25	0.5535461	0.004485936	0.194299178	0.043996852
YGR118W	RPS23A	-0.5967583	0.04714807	-0.173820009	0.148759365
YGR145W	ENP2	-0.5745818	0.007965195	-0.449586523	0.023569848
YGR156W	PTI1	-0.5304869	0.02051901	-0.535172222	0.019159769
YGR163W	GTR2	-0.09400677	0.1221705	-0.832988304	0.410307103
YGR189C	CRH1	0.2382177	0.06771484	0.149902095	0.257482417
YGR218W	CRM1	-0.9349632	0.01864873	0.45023901	0.47156054
YGR285C	ZUO1	-1.465766	0.005657875	-1.464141954	0.003143516
YHL001W	RPL14B	-0.4814743	0.05370249	-0.581366066	0.016333427
YHL004W	MRP4	-0.05155527	0.5483537	0.005529475	0.907990124
YHL032C	GUT1	0.8025545	0.04345582	1.210615687	0.207286778
YHR013C	ARD1	-0.4603178	0.00304024	-0.477987112	0.002603127
YHR025W	THR1	0.2865138	0.017173	0.592945317	0.020498227
YHR034C	PIH1	-0.4133874	0.04205282	-0.302947122	0.257482417
YHR047C	AAP1	-0.6497583	0.05540141	-0.107020998	0.570019898
YHR083W	SAM35	-0.4665908	0.017173	-0.329243367	0.043572164
YHR121W	LSM12	-0.9623344	0.04448426	-0.853141186	0.003143516
YHR122W	CIA2	-0.4563463	0.06018237	-0.22470457	0.021396536
YHR154W	RTT107	-1.122213	0.01551992	-0.297179088	0.397012826
YHR183W	GND1	-0.3726814	0.007965195	-0.340633624	0.014245546
YHR196W	UTP9	-0.6280793	0.09009455	-0.079485014	0.632221842
YIL027C	EMC5	-1.050718	0.005814858	-1.146691535	0.001506667

YIL057C	RGI2	-0.3908906	0.04450171	-0.33062497	0.149097363
YIL097W	FYV10	-0.501015	0.0220489	-0.352787712	0.022686668
YIR008C	PRI1	-0.05739452	0.6003255	-0.030509884	0.775494692
YJL002C	OST1	-0.5537015	0.004278381	-0.468186377	0.015102283
YJL004C	SYS1	-0.1926257	0.06782839	-0.044393465	0.630173037
YJL066C	MPM1	-0.9515846	0.1185725	-0.620045069	0.13313179
YJL069C	UTP18	0.0737381	0.4383108	0.080958883	0.344283401
YJL151C	SNA3	0.4672487	0.005070202	0.018697486	0.604325092
YJL177W	RPL17B	-1.647774	0.000887523	-1.402038412	0.00050659
YJL192C	SOP4	-0.9154465	0.002386406	-1.282916993	0.00050659
YJR013W	GPI14	0.1783422	0.1214979	-0.075964684	0.225414735
YJR094W-A	RPL43B	-0.4217051	0.01174584	-0.378561682	0.020502237
YJR102C	VPS25	0.4081906	0.005946676	0.90144628	0.002603127
YKL090W	CUE2	-0.2146822	0.4965834	-0.512831313	0.330262438
YKL122C	SRP21	-0.8499588	0.003831679	-0.717582407	0.003585141
YKL160W	ELF1	-0.2141251	0.03766363	-0.149986179	0.081542467
YKR004C	ECM9	-1.51619	0.03027219	-0.86757055	0.229420357
YKR057W	RPS21A	-0.6020841	0.01873919	-0.670930422	0.002603127
YKR066C	CCP1	-0.4415839	0.06048529	0.29411577	0.632221842
YLR040C	AFB1	-0.180933	0.1337246	-0.17720503	0.07171096
YLR060W	FRS1	-0.6653499	0.02639038	-0.49148874	0.007921708
YLR142W	PUT1	-0.7247519	0.009867253	-0.596086472	0.040947695
YLR153C	ACS2	0.5420177	0.06799709	0.866050856	0.043572164
YLR166C	SEC10	-0.5506528	0.00494629	-0.367857123	0.089126922
YLR295C	ATP14	0.3106263	0.04205282	0.294359104	0.014884667
YLR378C	SEC61	-0.537386	0.007041179	-0.533395027	0.003965869
YLR442C	SIR3	-0.05414389	0.526455	-0.265906987	0.651581627
YML093W	UTP14	0.3499702	0.003590395	0.406841315	0.007804447
YML102W	CAC2	0.02506336	0.8159432	0.021466117	0.611510069
YMR047C	NUP116	-0.3335945	0.02591892	-0.406949254	0.024101565
YMR091C	NPL6	-0.6666333	0.01242481	-0.687998334	0.020498227
YMR303C	ADH2	-0.7437284	0.05370249	0.390729891	0.550471428
YMR314W	PRE5	-0.9426006	0.0051507	-0.785859303	0.001506667
YNL100W	MIC27	-0.5176695	0.003722906	-0.403647802	0.005015958
YNL168C	FMP41	0.1135295	0.6109683	0.230023535	0.077377383
YNL240C	NAR1	-0.07065387	0.3887743	-0.31232562	0.003585141
YNL305C	BXI1	-0.3616945	0.01246395	-0.38705452	0.067446062
YNL313C	EMW1	-0.07109273	0.4850607	-0.044842282	0.585109374
YOL012C	HTZ1	-0.8183379	0.01137485	-0.792628323	0.00603948
YOL021C	DIS3	-0.4822298	0.0136243	-0.395942224	0.005015958
YOR085W	OST3	-0.1837157	0.1184124	-0.059591875	0.405327747
YOR206W	NOC2	-0.9180419	0.009144354	-0.91210613	0.003143516
YOR245C	DGA1	-0.5073407	0.05500477	-0.912628395	0.391186663
YOR259C	RPT4	0.4388929	0.006331455	0.459457104	0.002603127
YOR348C	PUT4	-0.4321389	0.005099595	-0.589358634	0.008155557
YOR361C	PRT1	-1.05907	0.008313373	-0.789145874	0.069475828
YPL004C	LSP1	-0.3381046	0.04127903	-0.191064283	0.04432936
YPL037C	EGD1	-0.4193737	0.00494629	-0.311760879	0.022720652
YPL043W	NOP4	-0.1745719	0.1412515	-0.081174919	0.419067682
YPL081W	RPS9A	-1.473388	0.02514036	-1.432533842	1.34E-05
YPL090C	RPS6A	-1.200135	0.0198996	-1.096594449	0.007998086
YPL112C	PEX25	-0.4051643	0.03107421	-0.549630086	0.045431266
YPL195W	APL5	-0.2075492	0.02271901	-0.502963538	0.222083896
YPL259C	APM1	-0.881291	0.006236033	-0.585858063	0.009016648
YPR034W	ARP7	-0.4657865	0.03113087	-0.390475199	0.114834576
YPR129W	SCD6	1.071366	0.006780823	0.740828485	0.016228286
YPR138C	MEP3	-0.3550248	0.06594843	-0.057607893	0.531710015
YPR145W	ASN1	-0.3059621	0.05370249	-0.470053158	0.021396536

10.1.7 TABLE S5. FLOW CYTOMETRY MEASUREMENTS OF LOG-PHASE CULTURES AND PLATE READER MEASUREMENTS OF COLONIES OF 120 SELECTED STRAINS

ORF	gene	log_2(OE/Native) [mNG].replicate	log_2(OE/Native) [mNG].liquid_I	log_2(OE/Native) [mNG].liquid_II	log_2(OE/Native) [mNG].liquid_III	log_2(OE/Native) [mNG].liquid_median
YAL033W	POP5	0.04905698	-0.03954341	-0.106687	0.0156639	-0.043522186
YAR002C-A	ERP1	-0.0873948	0.016755417	-0.0398744	-0.027123	-0.016747436
YBR061C	TRM7	-0.3418512	-0.07993893	-0.0363342	-0.137887	-0.08472009
YBR067C	TIP1	0.03102526	0	-0.0868841	-0.105929	-0.06427106
YBR173C	UMP1	0.01068667	-0.51493572	-0.4511817	-0.475151	-0.480422916
YBR227C	MCX1	-0.2484515	0.061219581	0.03378172	0.0697371	0.054912786
YBR230C	OM14	-1.043973	0.045931975	-0.1094594	0.09132	0.009264203
YBR279W	PAF1	-0.5679337	-0.41395499	-0.4971746	-0.410401	-0.440510077
YCR048W	ARE1	-0.0791201	0.197351794	0.37014573	0.4026374	0.323378308
YDL007W	RPT2	-1.240134	-0.80335982	-0.7440563	-0.930305	-0.825907161
YDL106C	PHO2	-0.1910417	0.440004898	0.42010849	0.5107379	0.456950434
YDL182W	LYS20	0.38586175	0.225685928	0.33775576	0.3725759	0.312005863
YDL185W	VMA1	-0.6235754	-0.32979931	0.02052547	-0.104105	-0.137792983
YDL198C	GGC1	1.71245413	2.488404527	1.19425753	2.3975924	2.026751473
YDR062W	LCB2	-1.0895553	-0.70394547	-0.6261591	-0.730731	-0.686945283
YDR079C-A	TFB5	0.14982139	0.262334966	-0.1658839	-0.123148	-0.00889905
YDR174W	HMO1	-2.514793	-2.0642944	-2.3432642	-2.299304	-2.235620724
YDR177W	UBC1	0.15512346	0.170675735	0.10775694	0.1541881	0.144206928
YDR205W	MSC2	-0.256255	-0.02105841	-0.1622829	-0.285887	-0.156409288
YDR258C	HSP78	0.0989728	0.07833436	-0.1097089	0.1094538	0.026026426
YER021W	RPN3	-1.0092689	-0.41773505	-0.3329755	-0.373826	-0.374845648
YER025W	GCD11	-1.2028286	-0.26114394	-0.2363614	-0.261355	-0.252953488
YER069W	ARG5,6	-0.2802815	-0.21535838	-0.2603243	-0.282489	-0.252724018
YER071C	TDA2	0.07533613	0.148072227	0.11590265	-0.035796	0.076059696
YER077C	MRX1	0.14122626	0.062110454	0.21442478	0.1859834	0.154172879
YER112W	LSM4	0.75178468	1.10359599	0.81878277	0.7074266	0.876601793
YER146W	LSM5	-0.7603503	-0.09646739	-0.2199969	-0.208695	-0.175052951
YER177W	BMH1	-0.7952446	-0.15417683	-0.1574316	-0.195238	-0.168948887
YFL034C-A	RPL22B	-0.8178923	-1.17971046	-1.1406999	-1.158606	-1.159672131
YFR031C-A	RPL2A	-0.2698474	-0.89164566	-0.9346584	-0.839798	-0.888700718
YGL010W	MPO1	0.15559522	0.239659685	0.02703724	0.1633644	0.143353786
YGL133W	ITC1	-0.5477874	0	0.08534443	-0.0608	0.008181336
YGL227W	VID30	-0.6859318	0.120943993	0.53861168	0.4394217	0.366325776
YGR034W	RPL26B	-1.1831329	-0.76483511	-0.7419771	-0.754781	-0.753864548
YGR060W	ERG25	0.19429918	0.187775805	0.19407668	0.1069536	0.162935352
YGR118W	RPS23A	-0.17382	-0.1402804	-0.2607887	-0.240172	-0.213746976
YGR145W	ENP2	-0.4495865	-0.08655057	-0.123886	-0.132113	-0.114183233
YGR156W	PTI1	-0.5351722	0.07286354	0.05693715	0.0419113	0.05723733
YGR163W	GTR2	-0.8329883	0.252590007	0.98167223	0.39079	0.541684062
YGR189C	CRH1	0.1499021	-0.2074305	0.15387701	0.1014972	0.015981234
YGR218W	CRM1	0.45023901	0.265739009	0.09681074	0.0136847	0.125411472
YGR285C	ZUO1	-1.464142	-1.39270528	-1.3691302	-1.443838	-1.401891321
YHL001W	RPL14B	-0.5813661	-0.39102563	-0.376454	-0.278642	-0.348707117
YHL004W	MRP4	0.00552948	-0.13433318	0	0.022732	-0.03720039
YHL032C	GUT1	1.21061569	0.731549558	0.71041564	0.5469016	0.662955583
YHR013C	ARD1	-0.4779871	-0.00722131	-0.0487787	-0.09893	-0.051643355
YHR025W	THR1	0.59294532	1.254831537	0.42693118	0.9296447	0.870469127
YHR034C	PIH1	-0.3029471	-0.20228736	0	-0.106088	-0.102791811
YHR047C	AAP1	-0.107021	-0.15454664	-0.0136136	-0.100896	-0.089685499
YHR083W	SAM35	-0.3292434	-0.04164913	-0.2534772	-0.05715	-0.117425516
YHR121W	LSM12	-0.8531412	-0.55979375	-0.527189	-0.492214	-0.526398918
YHR122W	CIA2	-0.2247046	-0.03642596	-0.1565431	-0.080338	-0.091102296
YHR154W	RTT107	-0.2971791	0.40818784	0.17457848	0	0.19425544
YHR183W	GND1	-0.3406336	-0.30648942	-0.463499	-0.279742	-0.349909992
YHR196W	UTP9	-0.079485	0.008976665	0.03287868	0.0191278	0.020327723
YIL027C	EMC5	-1.1466915	-0.82787319	-0.8981077	-0.78265	-0.836210373
YIL057C	RGI2	-0.330625	0.145695059	0.19213795	0.2831265	0.206986504

YIL097W	FYV10	-0.3527877	-0.01937255	-0.0543295	-0.051796	-0.041832623
YIR008C	PRI1	-0.0305099	-0.70938325	-0.4917544	-0.346384	-0.515840707
YJL002C	OST1	-0.4681864	-0.03727834	-0.026901	-0.062329	-0.042169507
YJL004C	SYS1	-0.0443935	0.208980579	0.18654502	0.1664683	0.187331287
YJL066C	MPM1	-0.6200451	-0.24040907	-0.30318	-0.382461	-0.308683354
YJL069C	UTP18	0.08095888	-0.02379103	0	0.0282222	0.001477048
YJL151C	SNA3	0.01869749	0.065054254	-0.0772885	-0.133441	-0.048558497
YJL177W	RPL17B	-1.4020384	-1.45899308	-1.6667497	-1.635405	-1.58704926
YJL192C	SOP4	-1.282917	-0.54461345	-0.5217065	-0.514867	-0.527062136
YJR013W	GPI14	-0.0759647	-0.57344364	0.13430109	0.1348419	-0.101433536
YJR094W-A	RPL43B	-0.3785617	0	-0.1053576	-0.12552	-0.076959329
YJR102C	VPS25	0.90144628	0.113997899	0.18653884	0.2521574	0.184231372
YKL090W	CUE2	-0.5128313	0.186199537	0.69647739	-0.056527	0.275383416
YKL122C	SRP21	-0.7175824	-0.8261718	-0.8601785	-0.893085	-0.859811705
YKL160W	ELF1	-0.1499862	0.036748146	-0.1008627	-0.102916	-0.055676758
YKR004C	ECM9	-0.8675706	0.537553571	0.62284524	0.7713552	0.643917999
YKR057W	RPS21A	-0.6709304	-0.2295535	0.03496957	0.034952	-0.053210658
YKR066C	CCP1	0.29411577	0.319763735	0.38511772	0.2766576	0.327179686
YLR040C	AFB1	-0.177205	-0.18016456	0.23473607	-0.139051	-0.028159793
YLR060W	FRS1	-0.4914887	-0.06510685	-0.0272848	0.0976604	0.001756278
YLR142W	PUT1	-0.5960865	-0.03798173	-0.5460025	-0.047901	-0.210628504
YLR153C	ACS2	0.86605086	0.747832043	0.84157738	0.6843351	0.757914839
YLR166C	SEC10	-0.3678571	-0.10662972	-0.1089001	-0.065242	-0.093590685
YLR295C	ATP14	0.2943591	0.316488121	0.28111613	0.3242029	0.307269047
YLR378C	SEC61	-0.533395	-0.14124867	-0.3911248	-0.851814	-0.461395883
YLR442C	SIR3	-0.265907	0.06133838	0	0.3298117	0.130383358
YML093W	UTP14	0.40684132	0.1815625	0.16212881	0.1297944	0.157828579
YML102W	CAC2	0.02146612	0.288457971	0.26870651	0.3243288	0.293831082
YMR047C	NUP116	-0.4069493	-0.2539938	-0.2312	-0.177127	-0.220773581
YMR091C	NPL6	-0.6879983	-0.16727677	-0.0541638	-0.269923	-0.16378795
YMR303C	ADH2	0.39072989	0.429864917	NA	1.5193742	0.974619538
YMR314W	PRE5	-0.7858593	-0.00557636	-0.2818808	-0.217227	-0.168228186
YNL100W	MIC27	-0.4036478	0.04808093	-0.0679663	0.0650543	0.015056302
YNL168C	FMP41	0.23002354	0.016637067	0.08323148	0.0191257	0.039664753
YNL240C	NAR1	-0.3123256	-0.32549903	-0.0655635	-0.233767	-0.20827658
YNL305C	BXI1	-0.3870545	0.119932907	-0.165927	0	-0.015331376
YNL313C	EMW1	-0.0448423	0.095223871	-0.0643001	0.0370298	0.022651205
YOL012C	HTZ1	-0.7926283	-0.23240737	-0.2841727	-0.31007	-0.275550159
YOL021C	DIS3	-0.3959422	0	-0.0632724	-0.057229	-0.040167096
YOR085W	OST3	-0.0595919	-0.10398139	-0.0801926	-0.057007	-0.080393833
YOR206W	NOC2	-0.9121061	-0.29065268	-0.3377881	-0.283719	-0.30405335
YOR245C	DGA1	-0.9126284	0.175188074	0.16995194	0.065095	0.136745013
YOR259C	RPT4	0.4594571	-0.24469907	-0.2301167	-0.238329	-0.237714836
YOR348C	PUT4	-0.5893586	0.094749511	0	0.0334733	0.042740926
YOR361C	PRT1	-0.7891459	-0.27881361	-0.2119844	-0.215994	-0.235597451
YPL004C	LSP1	-0.1910643	-0.03787822	0.00759426	-0.075475	-0.035253072
YPL037C	EGD1	-0.3117609	-0.07625267	-0.1946397	-0.091369	-0.120753738
YPL043W	NOP4	-0.0811749	0.885819257	-0.1159203	0.2174439	0.329114274
YPL081W	RPS9A	-1.4325338	-2.18048392	-2.0613197	-2.141461	-2.127754845
YPL090C	RPS6A	-1.0965944	-1.13901803	-1.0703608	-0.955622	-1.055000433
YPL112C	PEX25	-0.5496301	-0.33594003	-0.1094858	-0.190475	-0.211967018
YPL195W	APL5	-0.5029635	-0.1832517	-0.2187325	-0.340711	-0.247565213
YPL259C	APM1	-0.5858581	-0.3514978	-0.0350571	0.1013426	-0.095070746
YPR034W	ARP7	-0.3904752	0.028489227	-0.0721678	-0.05997	-0.03454953
YPR129W	SCD6	0.74082849	1.063717457	0.91872605	0.857873	0.946772157
YPR138C	MEP3	-0.0576079	-0.41997691	-0.2413142	-0.472125	-0.377805263
YPR145W	ASN1	-0.4700532	0.073893888	0.03851375	-0.031526	0.0269605

10.1.8 TABLE S6. IDENTIFICATION OF POTENTIAL SUBSTRATES OF UBR1, SAN1 AND TOM1

E3 ligase	ORF	gene	subcellular_localization	Complex_subunits
ubr1	YER025W	GCD11	Cytoplasm	Eukaryotic translation initiation factor 2B complex
ubr1	YGR113W	DAM1	None	DASH complex
ubr1	YKR038C	KAE1	Cytoplasm, Nucleus	EKC-KEOPS protein complex
ubr1	YLR446W	YLR446W	None	NA
ubr1	YOR340C	RPA43	Nucleus, nucleolus	RNA polymerase I complex
ubr1	YPR190C	RPC82	Cytoplasm, Nucleus	RNA polymerase III complex
ubr1	YBL002W	HTB2	Nucleus, Anti-nucleolar nucleus	Kap114-Nap1 histone complex
ubr1	YDR006C	SOK1	Nucleus	Putative Sok1-Rtn1-Crp1-Ids2-Pep4 complex
ubr1	YER148W	SPT15	Nucleus	RNA polymerase I transcription factor complex (GO)
ubr1	YHR121W	LSM12	Cytoplasm, Nucleus	NA
ubr1	YLR052W	IES3	Nucleus	INO80 complex
ubr1	YDR404C	RPB7	Nucleus	RNA polymerase II complex
ubr1	YDL007W	RPT2	None	Proteasome regulatory particle
ubr1	YDL191W	RPL35A	Cytoplasm	NA
ubr1	YDR062W	LCB2	ER	Serine C-palmitoyltransferase complex
ubr1	YDR156W	RPA14	Nucleus, nucleolus	RNA polymerase I complex
ubr1	YJL203W	PRP21	Nucleus	snRNP U2
ubr1	YKL119C	VPH2	ER	Vacuolar proton-transporting V-type ATPase, V0 domain
ubr1	YMR284W	YKU70	Cytoplasm, Nucleus	Ku70:Ku80 complex
ubr1	YPR110C	RPC40	Nucleus, nucleolus	RNA polymerase I complex
san1	YAL013W	DEP1	None	Rpd3L complex
san1	YGL172W	NUP49	Nuclear Periphery	Nucleoporin Nsp1 subcomplex
san1	YGR078C	PAC10	Cytoplasm	Prefoldin complex
san1	YGR113W	DAM1	None	DASH complex
san1	YGR156W	PTI1	Nucleus	CPF complex
san1	YLR052W	IES3	Nucleus	INO80 complex
san1	YOL068C	HST1	Cytoplasm, Nucleus, Anti-nucleolar nucleus	Set3C
san1	YDR502C	SAM2	None	NA
san1	YOL013C	HRD1	ER	Hrd1 ubiquitin ligase ERAD-L complex
san1	YOL135C	MED7	Nucleus	Srb-mediator complex
san1	YPR010C	RPA135	Nucleus, nucleolus	RNA polymerase I complex
san1	YDL007W	RPT2	None	Proteasome regulatory particle
san1	YGR005C	TFG2	None	RNA polymerase II complex
hrd1	YJL192C	SOP4	ER	NA
tom1	YBR048W	RPS11B	Cytoplasm	Cytosolic small ribosomal subunit
tom1	YER025W	GCD11	Cytoplasm	Eukaryotic translation initiation factor 2B complex
tom1	YER092W	IES5	Nucleus	INO80 complex
tom1	YGR119C	NUP57	Nuclear Periphery	Nucleoporin Nsp1 subcomplex
tom1	YGR158C	MTR3	Cytoplasm, Nucleus	Exosome (RNase complex)
tom1	YHL011C	PRS3	Cytoplasm	Prs1-Prs3 heterodimer
tom1	YLR291C	GCD7	Cytoplasm, punctate composite	Eukaryotic translation initiation factor 2B complex
tom1	YNL162W	RPL42A	Cytoplasm	NA
tom1	YDR502C	SAM2	Ambiguous	NA
tom1	YGR135W	PRE9	Cytoplasm, Nucleus	Proteasome core complex
tom1	YHR066W	SSF1	Nucleus, nucleolus	NA
tom1	YJL031C	BET4	None	Rab-protein geranylgeranyltransferase complex
tom1	YJL192C	SOP4	ER	NA
tom1	YKL122C	SRP21	ER	Signal recognition particle
tom1	YMR060C	SAM37	Mitochondria	Mitochondrial sorting and assembly machinery complex
tom1	YPL178W	CBC2	Nucleus	Commitment complex (GO)
tom1	YDR471W	RPL27B	Cytoplasm	Cytosolic large ribosomal subunit
tom1	YKL065C	YET1	ER	NA
tom1	YKL181W	PRS1	Cytoplasm	Prs1-Prs3 heterodimer
tom1	YPL210C	SRP72	ER	Signal recognition particle

- Protein localization data is obtained from Yeast GFP Localization database. Subunits of complexes data is obtained from Joris J Benschop *et al.* (2010)[139]. NA: Database does not contain the information for the protein.

10.1.9 TABLE S7. PRIMERS

Primer	Sequence
MoBY2-R	GGCCGCATGTTACTTACCACA
COX12	TGCCCGATATCCCCATCC
RPP2B	AATCTGTCGGTGCTGAAGTCG
GUT1	CTGTCGGTTCTGGCAGTGG
CDC3	GGGGTGATCGAAGTGGACAA
MRPS18	TGACATTTTCAAGCGTTGTGG
UTP14	TACCTGGTTGGGGTGAATGG
UMP1	TCCTTGCCAGACACATTACGC
MET14	TGGTGACAACATTCGTTTTGG
LYS20	CCTATCACCGGGTTCTGTGC
LSM4	TTCAGAAGACAATGCTGAGAGC
ACS2	CACGAAAACGTCTCGGAAGC
RPT4	TTGTTGACCCAGATGGATGG
VPS25	TGGTCGCAAATGACCAAAGA
SCD6	GTAATCGCCCTCCTCAATCC
SNA3	GACCATGACCATCGAATGAGC
ERG25	GCTGAACACCACGATTTGCAT
ATP14	AACTTGAACCGTCTGCGTACC
DPL1	TCCAGGAAAACATTCCAGACC
MPM1	TGACAAGTCGCACTCTCATGG
PTK2	TTGGAAAAACCAACACCAACC
HSP78	TCCACCTGAGTTCATCAATCG
APL5	CCGCTCCCAAAAGGAAAGAT
ELF1	GATTTGTGGGCAGTCGTTCC
UTP18	AAATAGGTGGCTTGCGGTAGG
ASN1	AAGGCATTCGACACCACAGG
FYV8	TGAAGAAATGGCGAAATTGG
MCX1	AAAGGCGCTAAAGAAAGTTGC
OST3	CAGGATCATCTGGGCAACC
FMP41	GCCAAAAAGAAGGGGCTACC
ERP1	CTTTCCTTGCGTCTGATTCG
YNL313C	GCAAGATGTTGGCGACTGG
YGL010W	CCATATTTGTCCCCACAATCC
RPL17B	GAAATGGGACATGCAAAAAGG
RPS10A	TTTCGAAAGGACAGCACAAGG
GCD11	GGTGGTCTGATTGGTGTTGG
SRP21	ATTAGGCCCACGAGGCGTAT
LCB2	ATGCCGGCTCCTGTTTTAGC
PUT4	AGCACCATTTCCGGGTTTCT
TIP1	TTGCCATGGTACACCACAAGA
ZUO1	AAGCTGAAGCTAAGGCCAAGG
APM1	GACGCGGATACACCGACTTT
OM14	TTTTCCCCAGTGCTAAACACG
GND1	CATCGCTTTGATGTGGAGAGG
IES5	TTGAATGGCTGTGGGAAACC
NPL6	CGCAAGGGCCTTAGAAACC
TRM7	TGTGGCAAAGATCTTCAGAGG
SAM35	CTCAGATGCGCAATTCATGG
FYV10	ACGGAAACTGCCACAACAGC
SIP2	TCCAGACGATTTTGGTGACG
LSM12	CAAGCATTTGGAAGTCATTGG

ADH2	GGACATTGTTAGCGCAGTCG
RPL14B	TGGTCCAAAAGCTGGTGTCC
ITC1	GTTGGAGGTGATTGCAGACG
RRP43	TGGAGTTAGACCCCGAATGC
CAB2	GAATCATGGGCCACTCAAGC
RGA2	CGACAAGTTTTTGGATTCACC
HOT13	TGCTTTGAATGCCACCAAGA
YVC1	TCCATGGGATTTGACAGATGG
YPL199C	AAGGGTTGCATTCTCAAAACG
HTA2	CGGTGGTAAAGGTGGTAAAGC
PDR3	CCAGACCAAGGAAATTCAAGC
SLH1	AAACGTTTCGTCAGTGTCTGC
AAC3	TGCCATCAGTAGTGGGTATCG
POM33	CGCCTTTGCTCAATGGAGAG
SSA2	TGAAGCCGAAAAATTCAAGG
YNL058C	CACCATTCAACCCAATTCAGG
YNL050C	GCGTGATACCAAAGCTCGTG
CIT1	GAACGCAGGGAGAGTTGTTCC
YDL057W	CCAATTAGCGCAGCCTCAAA
PAR32	AAAAGGTGGCCAGCAATGG
SRY1	ATGCACTTTCTTGCGGAACG
3MYC-F	ATCGGATCCTCTAGAGGTGAACAA
3MYC-R	CCACTAGTCTAGACTCTAGATGATCCG
S1-HRD1	TCATCTATCAATTGCAATTTGTAAGAGAAGGGGAGAAAGACAAAATAATAATATGcgtacgctgcaggtcgac
S2-HRD1	TCCAGTAGTTTTTTTCTTTAAAAAAAACTATGTATAATATAAAACATGCAATCTAatcgatgaattcgagctcg
S1-SAN1	TTTTTCCCCTTTGTTTTCTCTCATAGTCTTGTAACCTCAGCTTTTGTTCATTATGcgtacgctgcaggtcgac
S2-SAN1	GACTGCCAATAGGACATATTTTCATATTAACATACTTCAGAAGCGGTATTGTTTAatcgatgaattcgagctcg
S1-ASI1	TAATTTTATCTGGGTTTTTTTCTTCTTTTTACAAAGAAACTATGCTAAGAAATATGcgtacgctgcaggtcgac
S2-ASI1	AGCTCCCAAACGAAAAAACCTCTTTTAGATACCATGCAAAAGTTCTTAAACTATTAatcgatgaattcgagctcg
S1-ASI3	CTACAATTTGTCAGGAACAGTCATTACGTAGGGATTTTCAAAAGTTTGACTGATGcgtacgctgcaggtcgac
S2-ASI3	AAAAAAAAAAATTCCTATGATGTCTTAAATACGTATACCTAATAAAATAATTCTAatcgatgaattcgagctcg
S1-UBR1	ACTGAAGTCCCTAATCTTTACAGGTCACACAAATTACATAGAACATTCCAATATGcgtacgctgcaggtcgac
S2-UBR1	AAGTTTTTATATACAAATATGTCAACTATAAAACATAGTAGAGGGCTTGAATCTAatcgatgaattcgagctcg
S1-TOM1	AAAATAGTGCATATTTAGTTTACTTTTTGCCTTTGATTGAAAATATATATTCATGcgtacgctgcaggtcgac
S2-TOM1	GACGTTCTAAAATACTTGGTTACATGGCGCTATAAATTTACACGAAAAATGATCAatcgatgaattcgagctcg
S1-UBP3	GTCTGCTACCATCATCCAGGTACCGCTTTCCTTTCCATCATCATTAAAAAAAATGcgtacgctgcaggtcgac
S2-UBP3	ATTATTGCTATATTATTTTTTATGTATTTTGTCTATAATACCACCCCCCGTCTTAatcgatgaattcgagctcg
S1-HIS3	TTCTTCGAAGAATATACTAAAAAATGAGCAGGCAAGATAAACGAAGGCAAAGATGcgtacgctgcaggtcgac
S2-HIS3	TATATACACATGTATATATATCGTATGCTGCAGCTTTAAATAATCGGTGTCACTAatcgatgaattcgagctcg
S1-SSA1	AAAAGTAATCAAGTATTACAAGAAACAAAAATTCAAGTAAATAACAGATAATATGcgtacgctgcaggtcgac
S2-SSA1	CCCAGATCATTAAAAGACATTTTCGTTATTATCAATTGCCGCACCAATTGGCTTAatcgatgaattcgagctcg
S1-SSA2	AATTGATTAATTCCAACAGATCAAGCAGATTTTATACAGAAATATTTATACAATGcgtacgctgcaggtcgac
S2-SSA2	GAGGAAAGCAAAAGTAAAACTTTTCGGATATTTTACAGGGCGATCGCTAAGCTTAatcgatgaattcgagctcg
SSA1_KO	GCACACGGAACGTTTAGAAGC
SSA2_KO	TCTCTCTTGGGCCCTTATTCG
His&Kan_Tag2	GCTGCGCACGTCAAGACTG
HRD1_KO	CGCAGTAGCCGTTATGTACG
SAN1_KO	GCAAAAAGCCGGGTAAAACA
ASI1_KO	TCTCGCGTATTCCTTTGAGG
ASI3_KO	ACGACGCTCAAAAACGTGCT
UBR1_KO	TGCTTTATTCCTCATTGGGCTTT
TOM1_KO	CGCCGAAGGAATACCCTTCTA
UBP3_KO	TATCGACAGGGGCATTGAGC
HIS3_KO	CGACGCTTTGTCTTCATTCA

10.1.10 TABLE S8. LIST OF GENERAL CONSUMABLES, MACHINES AND SOFTWARE USED

General solutions or chemicals	
Item	**Supplier**
Velocity DNA polymerase	Meridian Bioscience
Taq DNA Polymerase	IMB Protein Production Core Facility
HF-DNA Polymerase	IMB Protein Production Core Facility
Gel Loading Dye, Purple (6X)	New England Biolabs (NEB)
Canavanine	C1625, Sigma-Aldrich
Thialysine	A2636, Sigma-Aldrich
G418	A291-25, Biochrom
Hygromycin B Gold	ant-hg-5, Invivogen
clonNAT	5.0, Werner BioAgents
Adenine	A8626, Sigma-Aldrich
Monosodium glutamic acid (MSG)	G1626, Sigma-Aldrich
Bacto yeast extract	Thermo Fisher Scientific
Bacto peptone	Thermo Fisher Scientific
Bacto agar	Thermo Fisher Scientific
Bacto yeast nitrogen base without amino acids and ammonium sulfate	BD Biosciences
Gentra Puregene Yeast/Bact. Kit B	Qiagen
Plasmid Miniprep Kit	Qiagen
QIAquick gel extraction kit	Qiagen
Color Prestained Protein Standard	New England Biolabs (NEB)
TGX™ FastCast™ Acrylamide Kit	BioRad
Trans-Blot® Turbo™ Mini nitrocellulose transfer packs	BioRad
Pierce ECL Plus Western Blotting Substrate	Thermo Fisher Scientific
MG132	Enzo Life Sciences
5-Fluoroorotic acid	Apollo Scientific
Mouse monoclonal anti-Pgk1	Thermo Fisher Scientific
Mouse monoclonal anti-MYC	IMB Protein Production Core Facility
Rabbit polyclonal anti-mNeonGreen	Cell Signaling Technology

10.1.11 TABLE S9. LIST OF SOFTWARE USED

Software	
Item	**Supplier**
Adobe Illustrator 2020	Adobe
Adobe Photoshop 2020	Adobe
Microsoft Office	Microsoft
R studio 1.2.5033	R studio PBC
ImageJ	JAVA
ImageLab	BioRad
FlowJo	BD Biosciences
BD FACSDiva (Ver. 9.0.1)	BD Biosciences

10.1.12 TABLE S10. LIST OF MACHINES USED

Machines	
Item	**Supplier**
ROTOR	Singer Instruments
PIXL	Singer Instruments
Infinite M1000 plate reader	Tecan
Gel electrophoresis chamber	BioRad
Trans-Blot Turbo transfer system	BioRad
ChemiDoc imaging system	BioRad
BD LSRFortessa SORP	BD Biosciences
Thermoshaker	Eppendorf

10.2 LIST OF ABBREVIATIONS

SV40 - simian virus 40

NLS - nuclear localization signal

NPC - nuclear pore complex

ER - endoplasmic reticulum

TMD - transmembrane domain

CRISPR - clustered regularly-interspaced short palindromic repeats

SRP - signal recognition particle

TA - tail-anchored

GET - guided entry of tail-anchored protein

HSP70 - the 70 kDa heat shock proteins

UBQLNs - ubiquilin family proteins

EMC - ER-membrane protein complex

TOM - translocase of the outer membrane

TIM - translocase of the inner membrane

IM - inner membrane

PAM - presequence translocase-associated motor

MPP - mitochondrial processing peptidase

MIA - mitochondrial intermembrane space import and assembly

IMS - intermembrane space

Puf3 - Pumilio-homology domain family protein 3

NAC - nascent polypeptide associated complex

MTS - mitochondrial targeting sequences

OM - outer membrane

ER-SURF - ER surface-mediated protein targeting

ALS - Amyotrophic lateral sclerosis

UPS - ubiquitin proteasome system

E1 – Ub activating enzyme

E2 – Ub conjugating enzyme

E3 – Ub ligase

HECT - homologous to E6-AP C-terminus

RING - really interesting new gene

DUBs - deubiquitinating enzymes

CPY - carboxypeptidase Y

UBQLNs - proteins of the Ubiquilin family

ERAD - ER-associated degradation

mitoTAD - mitochondrial protein translocation-associated degradation

tFT - tandem fluorescent protein timer

OE - overexpression

mNG – mNeonGreen

MoBY2.0 - Molecular Barcoded Yeast ORF 2.0

2μ - 2micron

AAA - ATPase associated with diverse cellular activities

SWAT - SWAp-Tag

ORF - open reading frame

DSBs - induced double-strand breaks

MORF - moveable open reading frame

GO – Gene Ontology

FDR - false discovery rate

GO - gene ontology

SGD - Saccharomyces Genome Database

SGA - synthetic genetic array

sfGFP – superfolderGFP

mitoCPR - mitochondrial compromised protein import response

FCCP - trifluoromethoxy carbonyl cyanide phenylhydrazone

SILAC - Stable isotope labeling with amino acids

ERISQ - excess ribosomal protein quality control

11 REFERENCES

1. Wickner, W., and Schekman, R. (2005). Protein translocation across biological membranes. Science. *310*, 1452–1456. 10.1126/science.1113752.

2. Lu, J., Wu, T., Zhang, B., Liu, S., Song, W., Qiao, J., and Ruan, H. (2021). Types of nuclear localization signals and mechanisms of protein import into the nucleus. Cell Commun. Signal. *19*, 1–10. 10.1186/s12964-021-00741-y.

3. Hegde, R.S., and Keenan, R.J. (2022). The mechanisms of integral membrane protein biogenesis. Nat. Rev. Mol. Cell Biol. *23*, 107–124. 10.1038/s41580-021-00413-2.

4. Bykov, Y.S., Rapaport, D., Herrmann, J.M., and Schuldiner, M. (2020). Cytosolic events in the biogenesis of mitochondrial proteins. Trends Biochem. Sci. *45*, 650–667. 10.1016/j.tibs.2020.04.001.

5. Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., and Yanagawa, H. (2009). Six classes of nuclear localization signals specific to different binding grooves of importinα. J. Biol. Chem. *284*, 478–485. 10.1074/jbc.M807017200.

6. Kalderon, D., Richardson, W.D., Markham, A.F., and Smith, A.E. (1984). Sequence requirements for nuclear location of SV40 largeT antigen. Nature *311*, 33–38. 10.1038/311033a0.

7. Lange, A., Mills, R.E., Lange, C.J., Stewart, M., Devine, S.E., and Corbett, A.H. (2007). Classical nuclear localization signals: Definition, function, and interaction with importin α. J. Biol. Chem. *282*, 5101–5105. 10.1074/jbc.R600026200.

8. Xu, D., Farmer, A., and Chook, Y.M. (2010). Recognition of nuclear targeting signals by Karyopherin-β proteins. Curr. Opin. Struct. Biol. *20*, 782–790. 10.1016/j.sbi.2010.09.008.

9. Chook, Y.M., and Süel, K.E. (2011). Nuclear import by karyopherin-βs: recognition and inhibition. Biochim. Biophys. Acta - Mol. Cell Res. *1813*, 1593–1606. 10.1016/j.bbamcr.2010.10.014.

10. Rout, M.P., Blobel, G., and Aitchison, J.D. (1997). A distinct nuclear import pathway used by ribosomal proteins. Cell *89*, 715–725. 10.1016/S0092-8674(00)80254-8.

11. Pumroy, R.A., and Cingolani, G. (2015). Diversification of importin-α isoforms in cellular trafficking and disease states. Biochem. J. *466*, 13–28. 10.1042/BJ20141186.

12. Karasev, M.M., Baloban, M., Verkhusha, V. V., and Shcherbakova, D.M. (2022). Nuclear localization signals for optimization of genetically encoded tools in neurons. Front. Cell Dev. Biol. *10*, 1–14. 10.3389/fcell.2022.931237.

13. Hasle, N., Cooke, A., Srivatsan, S., Huang, H., Stephany, J.J., Krieger, Z., Jackson, D., Tang, W., Pendyala, S., Monnat, R.J., et al. (2020). High-throughput, microscope-based sorting to dissect cellular heterogeneity. Mol. Syst. Biol. *16*. 10.15252/msb.20209442.

14. Xin Zhang, S.S. (2015). Fidelity of co-translational protein targeting by the signal recognition particle. Annu Rev Biophys *43*, 381–408. 10.1146/annurev-biophys-051013-022653.

15. Nyathi, Y., Wilkinson, B.M., and Pool, M.R. (2013). Co-translational targeting and translocation of proteins to the endoplasmic reticulum. Biochim. Biophys. Acta - Mol. Cell Res. *1833*, 2392–2402. 10.1016/j.bbamcr.2013.02.021.

16. Voorhees, R.M., and Hegde, R.S. (2016). Structure of the Sec61 channel opened by a signal sequence. Science. *351*, 88–89. 10.1126/science.aad4992.

17. Li, L., Park, E., Ling, J.J., Ingram, J., Ploegh, H., and Rapoport, T.A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. Nature *531*, 395–399. 10.1038/nature17163.

18. Liaci, A.M., Steigenberger, B., Telles de Souza, P.C., Tamara, S., Gröllers-Mulderij, M., Ogrissek, P., Marrink, S.J., Scheltema, R.A., and Förster, F. (2021). Structure of the human signal peptidase complex reveals the determinants for signal peptide cleavage. Mol. Cell *81*, 3934-3948.e11. 10.1016/j.molcel.2021.07.031.

19. Guna, A., and Hegde, R.S. (2018). Transmembrane domain recognition during membrane protein biogenesis and quality control. Curr. Biol. *28*, R498–R511. 10.1016/j.cub.2018.02.004.

20. Hegde, R.S., and Keenan, R.J. (2011). Tail-anchored membrane protein insertion into the endoplasmic reticulum. Nat. Rev. Mol. Cell Biol. *12*, 787–798. 10.1038/nrm3226.

21. Denic, V., Dötsch, V., and Sinning, I. (2013). Endoplasmic reticulum targeting and insertion of tail-anchored membrane proteins by the GET pathway. Cold Spring Harb. Perspect. Biol. *5*, 1–11. 10.1101/cshperspect.a013334.

22. Abell, B.M., Rabu, C., Leznicki, P., Young, J.C., and High, S. (2007). Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones. J. Cell Sci. *120*, 1743–1751. 10.1242/jcs.002410.

23. Guna, A., Volkmar, N., Christianson, J.C., and Hegde, R.S. (2018). The ER membrane protein complex is a transmembrane domain insertase. Science. *359*, 470–473. 10.1126/science.aao3099.

24. Morgenstern, M., Stiller, S.B., Lübbert, P., Peikert, C.D., Dannenmaier, S., Drepper, F., Weill, U., Höß, P., Feuerstein, R., Gebert, M., et al. (2017). Definition of a high-confidence mitochondrial proteome at quantitative scale. Cell Rep. *19*, 2836–2852. 10.1016/j.celrep.2017.06.014.

25. den Brave, F., Engelke, J., and Becker, T. (2021). Quality control of protein import into mitochondria. Biochem. J. *478*, 3125–3143. 10.1042/BCJ20190584.

26. Hill, K., Model, K., Ryan, M.T., Dietmeier, K., Martint, F., Wagner, R., and Pfanner, N. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature *395*, 516–521. 10.1038/26780.

27. Ahting, U., Thieffry, M., Engelhardt, H., Hegerl, R., Neupert, W., and Nussberger, S. (2001). Tom40, the pore-forming component of the protein-conducting TOM channel in the outer membrane of mitochondria. J. Cell Biol. *153*, 1151–1160. 10.1083/jcb.153.6.1151.

28. Neupert, W. (1997). Protein import into mitochondria. Annu. Rev. Biochem *66*, 863–917. 10.1016/B978-0-12-378630-2.00203-6.

29. Endo, T., and Kohda, D. (2002). Functions of outer membrane receptors in mitochondrial protein import. Biochim. Biophys. Acta - Mol. Cell Res. *1592*, 3–14. 10.1016/S0167-4889(02)00259-8.

30. Wiedemann, N., Frazier, A.E., and Pfanner, N. (2004). The protein import machinery of mitochondria. J. Biol. Chem. *279*, 14473–14476. 10.1074/jbc.R400003200.

31. Malhotra, K., Sathappa, M., Landin, J.S., Johnson, A.E., and Alder, N.N. (2013). Structural changes in the mitochondrial Tim23 channel are coupled to the proton-motive force. Nat. Struct. Mol. Biol. *20*, 965–972. 10.1038/nsmb.2613.

32. Weinhäupl, K., Lindau, C., Hessel, A., Wang, Y., Schütze, C., Jores, T., Melchionda, L., Schönfisch, B., Kalbacher, H., Bersch, B., et al. (2018). Structural basis of membrane protein chaperoning through the mitochondrial intermembrane space. Cell *175*, 1365-1379.e25. 10.1016/j.cell.2018.10.039.

33. Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuán Szklarz, L.K., Schulze-Specking, A., Truscott, K.N., Guiard, B., Meisinger, C., and Pfanner, N. (2004). Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. EMBO J. *23*, 3735–3746. 10.1038/sj.emboj.7600389.

34. Christopher C. Williams, Calvin H. Jan, J.S.W. (2014). Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling. Science. *346*, 748–751. 10.1126/science.1257522.

35. Saint-Georges, Y., Garcia, M., Delaveau, T., Jourdren, L., Le Crom, S., Lemoine, S., Tanty, V., Devaux, F., and Jacq, C. (2008). Yeast mitochondrial biogenesis: A role for the PUF RNA-binding protein Puf3p in mRNA localization. PLoS One *3*. 10.1371/journal.pone.0002293.

36. Gerber, A.P., Herschlag, D., and Brown, P.O. (2004). Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. PLoS Biol. *2*. 10.1371/journal.pbio.0020079.

37. García-Rodríguez, L.J., Gay, A.C., and Pon, L.A. (2007). Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast. J. Cell Biol. *176*, 197–207. 10.1083/jcb.200606054.

38. Marvin Wickens, David S. Bernstein, J.K. and R.P. (2002). A PUF family portrait: 3'UTR regulation as a way of life. TRENDS Genet. *18*, 150–157. 10.1016/s0168-9525(01)02616-6.

39. George, R., Beddoe, T., Landl, K., and Lithgow, T. (1998). The yeast nascent polypeptide-associated complex initiates protein targeting to mitochondria in vivo. Proc. Natl. Acad. Sci. U. S. A. *95*, 2296–2301. 10.1073/pnas.95.5.2296.

40. George, R., Walsh, P., Beddoe, T., and Lithgow, T. (2002). The nascent polypeptide-associated complex (NAC) promotes interaction of ribosomes with the mitochondrial surface in vivo. FEBS Lett. *516*, 213–216. 10.1016/S0014-5793(02)02528-0.

41. Lesnik, C., Cohen, Y., Atir-Lande, A., Schuldiner, M., and Arava, Y. (2014). OM14 is a mitochondrial receptor for cytosolic ribosomes that supports co-translational import into mitochondria. Nat. Commun. *5*. 10.1038/ncomms6711.

42. M, G., MA, H., T, F., and E, D. (2015). The principle of antagonism ensures protein targeting specificity at the endoplasmic reticulum. Science. *348*, 201–207.

43. Young, J.C., Hoogenraad, N.J., and Hartl, F.U. (2003). Molecular chaperones Hsp90 and Hsp70 deliver preproteins to the mitochondrial import receptor Tom70. Cell *112*, 41–50. 10.1016/s0092-8674(02)01250-3.

44. Rosenzweig, R., Nillegoda, N.B., Mayer, M.P., and Bukau, B. (2019). The Hsp70 chaperone network. Nat. Rev. Mol. Cell Biol. *20*, 665–680. 10.1038/s41580-019-0133-3.

45. Saitoh, T., Igura, M., Obita, T., Ose, T., Kojima, R., Maenaka, K., Endo, T., and Kohda, D. (2007). Tom20 recognizes mitochondrial presequences through dynamic equilibrium among multiple bound states. EMBO J. *26*, 4777–4787. 10.1038/sj.emboj.7601888.

46. Brix, J., Dietmeier, K., and Pfanner, N. (1997). Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70. J. Biol. Chem. *272*, 20730–20735. 10.1074/jbc.272.33.20730.

47. Katja G Hansen, Naama Aviram, Janina Laborenz, Chen Bibi, Maren Meyer, Anne Spang, Maya Schuldiner, J.M.H. (2018). An ER surface retrieval pathwaysafeguards the import of mitochondrialmembrane proteins in yeast. Science. *1122*, 1118–1122. 10.1126/science.aar8174.

48. Krogan, N.J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., Tikuisis, A.P., et al. (2006). Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature *440*, 637–643. 10.1038/nature04670.

49. Zinani, O.Q.H., Keseroğlu, K., and Özbudak, E.M. (2022). Regulatory mechanisms ensuring coordinated expression of functionally related genes. Trends Genet. *38*, 73–81. 10.1016/j.tig.2021.07.008.

50. McShane, E., Sin, C., Zauber, H., Wells, J.N., Donnelly, N., Wang, X., Hou, J., Chen, W., Storchova, Z., Marsh, J.A., et al. (2016). Kinetic analysis of protein stability reveals age-dependent degradation. Cell *167*, 803-815.e21. 10.1016/j.cell.2016.09.015.

51. Shiber, A., Döring, K., Friedrich, U., Klann, K., Merker, D., Zedan, M., Tippmann, F., Kramer, G., and Bukau, B. (2018). Cotranslational assembly of protein complexes in eukaryotes revealed by ribosome profiling. Nature *561*, 268–272. 10.1038/s41586-018-0462-y.

52. Brunet, S., Sardon, T., Zimmerman, T., Wittmann, T., Pepperkok, R., Karsenti, E., and Vernos, I. (2005). The efficiency of protein compartmentalization into the secretory pathway. Mol Biol Cell *16*, 279–291. 10.1091/mbc.E04-06-0508.

53. Itakura, E., Zavodszky, E., Shao, S., Wohlever, M.L., Keenan, R.J., and Hegde, R.S. (2016). Ubiquilins chaperone and riage Mitochondrial Membrane Proteins for Degradation. Mol. Cell *63*, 21–33. 10.1016/j.molcel.2016.05.020.

54. Kalderon, D., Roberts, B.L., Richardson, W.D., and Smith, A.E. (1984). A short amino acid sequence able to specify nuclear location. Cell *39*, 499–509. 10.1016/0092-8674(84)90457-4.

55. Costa, E.A., Subramanian, K., Nunnari, J., and Weissman, J.S. (2018). Defining the

physiological role of SRP in protein-targeting efficiency and specificity. Science. *359*, 689–692. 10.1126/science.aar3607.

56. Kang, S.W., Rane, N.S., Kim, S.J., Garrison, J.L., Taunton, J., and Hegde, R.S. (2006). Substrate-specific translocational attenuation during ER stress defines a pre-emptive quality control pathway. Cell *127*, 999–1013. 10.1016/j.cell.2006.10.032.

57. Wright, G., Terada, K., Yan, M., Sergeev, I., and Mori, M. (2001). Oxidative stress inhibits the mitochondrial import of preproteins and leads to their degradation. Exp. Cell Res. *263*, 107–117. 10.1006/excr.2000.5096.

58. Springer, M., Weissman, J.S., and Kirschner, M.W. (2010). A general lack of compensation for gene dosage in yeast. Mol. Syst. Biol. *6*, 1–8. 10.1038/msb.2010.19.

59. Pidasheva, S., Canaff, L., Simonds, W.F., Marx, S.J., and Hendy, G.N. (2005). Impaired cotranslational processing of the calcium-sensing receptor due to signal peptide missense mutations in familial hypocalciuric hypercalcemia. Hum. Mol. Genet. *14*, 1679–1690. 10.1093/hmg/ddi176.

60. Rajpar, M.H., Koch, M.J., Davies, R.M., Mellody, K.T., Kielty, C.M., and Dixon, M.J. (2002). Mutation of the signal peptide region of the bicistronic gene DSPP affects translocation to the endoplasmic reticulum and results in defective dentine biomineralization. Hum. Mol. Genet. *11*, 2559–2565. 10.1093/hmg/11.21.2559.

61. McLane, L.M., and Corbett, A.H. (2009). Nuclear localization signals and human disease. IUBMB Life *61*, 697–706. 10.1002/iub.194.

62. Sabherwal, N., Schneider, K.U., Blaschke, R.J., Marchini, A., and Rappold, G. (2004). Impairment of SHOX nuclear localization as a cause for Léri-Weill syndrome. J. Cell Sci. *117*, 3041–3048. 10.1242/jcs.01152.

63. Mizutani, A., Matsuzaki, A., Momoi, M.Y., Fujita, E., Tanabe, Y., and Momoi, T. (2007). Intracellular distribution of a speech/language disorder associated FOXP2 mutant. Biochem. Biophys. Res. Commun. *353*, 869–874. 10.1016/j.bbrc.2006.12.130.

64. Djordjevic, S., Zhang, X., Bartlam, M., Ye, S., Rao, Z., and Danpure, C.J. (2010). Structural implications of a G170R mutation of alanine:glyoxylate aminotransferase that is associated with peroxisome-to-mitochondrion mistargeting. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. *66*, 233–236. 10.1107/S1744309109054645.

65. Clarke, G.M., and Higgins, T.N. (2000). Laboratory investigation of hemoglobinopathies and thalassemias: Review and update. Clin. Chem. *46*, 1284–1290. 10.1093/clinchem/46.8.1284.

66. Feng, L., Gell, D.A., Zhou, S., Gu, L., Kong, Y., Li, J., Hu, M., Yan, N., Lee, C., Rich, A.M., et al. (2004). Molecular mechanism of AHSP-mediated stabilization of α-hemoglobin. Cell *119*, 629–640. 10.1016/j.cell.2004.11.025.

67. Bloom, G.S. (2014). Amyloid-β and tau: The trigger and bullet in Alzheimer disease pathogenesis. JAMA Neurol. *71*, 505–508. 10.1001/jamaneurol.2013.5847.

68. Winton, M.J., Igaz, L.M., Wong, M.M., Kwong, L.K., Trojanowski, J.Q., and Lee, V.M.Y. (2008). Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. J. Biol.

Chem. *283*, 13302–13309. 10.1074/jbc.M800342200.

69. Purice, M.D., and Taylor, J.P. (2018). Linking hnRNP function to ALS and FTD pathology. Front. Neurosci. *12*, 1–12. 10.3389/fnins.2018.00326.

70. Dormann, D., Rodde, R., Edbauer, D., Bentmann, E., Fischer, I., Hruscha, A., Than, M.E., MacKenzie, I.R.A., Capell, A., Schmid, B., et al. (2010). ALS-associated fused in sarcoma (FUS) mutations disrupt transportin-mediated nuclear import. EMBO J. *29*, 2841–2857. 10.1038/emboj.2010.143.

71. Finley, D., Ulrich, H.D., Sommer, T., and Kaiser, P. (2012). The ubiquitin-proteasome system of Saccharomyces cerevisiae. Genetics *192*, 319–360. 10.1534/genetics.112.140467.

72. Kleiger, G., and Mayor, T. (2014). Perilous journey: A tour of the ubiquitin-proteasome system. Trends Cell Biol. *24*, 352–359. 10.1016/j.tcb.2013.12.003.

73. Kong, K.Y.E., Fischer, B., Meurer, M., Kats, I., Li, Z., Rühle, F., Barry, J.D., Kirrmaier, D., Chevyreva, V., San Luis, B.J., et al. (2021). Timer-based proteomic profiling of the ubiquitin-proteasome system reveals a substrate receptor of the GID ubiquitin ligase. Mol. Cell *81*, 2460-2476.e11. 10.1016/j.molcel.2021.04.018.

74. Huibregtse, J.M., Scheffner, M., Beaudenon, S., and Howley, P.M. (1995). A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. Proc. Natl. Acad. Sci. U. S. A. *92*, 2563–2567. 10.1073/pnas.92.7.2563.

75. Deshaies, R.J., and Joazeiro, C.A.P. (2009). RING domain E3 ubiquitin ligases. Annu. Rev. Biochem. *78*, 399–434. 10.1146/annurev.biochem.78.101807.093809.

76. Peng, J., Schwartz, D., Elias, J.E., Thoreen, C.C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S.P. (2003). A proteomics approach to understanding protein ubiquitination. Nat. Biotechnol. *21*, 921–926. 10.1038/nbt849.

77. Mevissen, T.E.T., and Komander, D. (2017). Mechanisms of deubiquitinase specificity and regulation. Annu. Rev. Biochem. *86*, 159–192. 10.1146/annurev-biochem-061516-044916.

78. Heck, J.W., Cheung, S.K., and Hampton, R.Y. (2010). Cytoplasmic protein quality control degradation mediated by parallel actions of the E3 ubiquitin ligases Ubr1 and San1. Proc. Natl. Acad. Sci. U. S. A. *107*, 1106–1111. 10.1073/pnas.0910591107.

79. Kaneko, T., and Murata, S. (2012). Ubiquitin Family Modifiers and the Proteasome 10.1007/978-1-61779-474-2_30.

80. Farzin Khosrow-Khavar, Fang, N.N., Ng, A.H.M., Winget, J.M., Comyn, S.A., and Mayor, T. (2012). The yeast ubr1 ubiquitin ligase participates in a prominent pathway that targets cytosolic thermosensitive mutants for degradation. G3 Genes, Genomes, Genet. *2*, 619–628. 10.1534/g3.111.001933.

81. Pan, M., Zheng, Q., Wang, T., Liang, L., Mao, J., Zuo, C., Ding, R., Ai, H., Xie, Y., Si, D., et al. (2021). Structural insights into Ubr1-mediated N-degron polyubiquitination. Nature *600*, 334–338. 10.1038/s41586-021-04097-8.

82. Bartel, B., Wunning, I., and Varshavsky, A. (1990). The recognition component of the

N-end rule pathway. EMBO J. *9*, 3179–3189. 10.1002/j.1460-2075.1990.tb07516.x.

83. Bachmair, A., and Varshavsky, A. (1989). The degradation signal in a short-lived protein. Cell *56*, 1019–1032. 10.1016/0092-8674(89)90635-1.

84. Xia, Z., Webster, A., Du, F., Piatkov, K., Ghislain, M., and Varshavsky, A. (2008). Substrate-binding sites of UBR1, the ubiquitin ligase of the N-end rule pathway. J. Biol. Chem. *283*, 24011–24028. 10.1074/jbc.M802583200.

85. Singh, A., Vashistha, N., Heck, J., Tang, X., Wipf, P., Brodsky, J.L., and Hampton, R.Y. (2020). Direct involvement of Hsp70 ATP hydrolysis in Ubr1-dependent quality control. Mol. Biol. Cell *31*, 2669–2686. 10.1091/mbc.E20-08-0541.

86. Hessa, T., Sharma, A., Mariappan, M., Eshleman, H.D., Gutierrez, E., and Hegde, R.S. (2011). Protein targeting and degradation are coupled for elimination of mislocalized proteins. Nature *475*, 394–399. 10.1038/nature10181.

87. Rodrigo-Brenni, M.C., Gutierrez, E., and Hegde, R.S. (2014). Cytosolic quality control of mislocalized proteins requires RNF126 recruitment to bag6. Mol. Cell *55*, 227–237. 10.1016/j.molcel.2014.05.025.

88. Mariappan, M., Li, X., Stefanovic, S., Sharma, A., Mateja, A., Keenan, R.J., and Hegde, R.S. (2010). A ribosome-associating factor chaperones tail-anchored membrane proteins. Nature *466*, 1120–1124. 10.1038/nature09296.

89. Itakura, E., Zavodszky, E., Shao, S., Wohlever, M.L., Keenan, R.J., and Hegde, R.S. (2016). Ubiquilins chaperone and triage mitochondrial membrane proteins for degradation. Mol. Cell *63*, 21–33. 10.1016/j.molcel.2016.05.020.

90. Nguyen, K.T., Mun, S.H., Lee, C.S., and Hwang, C.S. (2018). Control of protein degradation by N-terminal acetylation and the N-end rule pathway. Exp. Mol. Med. *50*. 10.1038/s12276-018-0097-y.

91. 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. 10.1016/j.molcel.2013.03.018.

92. Yanagitani, K., Juszkiewicz, S., and Hegde, R.S. (2017). UBE2O is a quality control factor for orphans of multiprotein complexes. Science. *357*, 472–475. 10.1126/science.aan0178.

93. Yip, M.C.J., Sedor, S.F., and Shao, S. (2022). Mechanism of client selection by the protein quality-control factor UBE2O. Nat. Struct. Mol. Biol. *29*, 774–780. 10.1038/s41594-022-00807-6.

94. 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*, 1–28. 10.7554/elife.19105.

95. Xu, Y., Anderson, D.E., and Ye, Y. (2016). The HECT domain ubiquitin ligase HUWE1 targets unassembled soluble proteins for degradation. Nat. Publ. Gr., 1–16. 10.1038/celldisc.2016.40.

96. Zavodszky, E., Peak-Chew, S.Y., Juszkiewicz, S., Narvaez, A.J., and Hegde, R.S. (2021). Identification of a quality-control factor that monitors failures during proteasome assembly. Science. *373*, 998–1004. 10.1126/science.abc6500.

97. Murata, S., Yashiroda, H., and Tanaka, K. (2009). Molecular mechanisms of proteasome assembly. Nat. Rev. Mol. Cell Biol. *10*, 104–115. 10.1038/nrm2630.

98. Krshnan, L., van de Weijer, M.L., and Carvalho, P. (2022). Endoplasmic reticulum–associated protein degradation. Cold Spring Harb. Perspect. Biol. *14*. 10.1101/cshperspect.a041247.

99. Bays, N.W., Gardner, R.G., Seelig, L.P., Joazeiro, C.A., and Hampton, R.Y. (2001). Hrd1p/Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. Nat. Cell Biol. *3*, 24–29. 10.1038/35050524.

100. Carvalho, P., Goder, V., and Rapoport, T.A. (2006). Distinct Ubiquitin-Ligase Complexes Define Convergent Pathways for the Degradation of ER Proteins. Cell *126*, 361–373. 10.1016/j.cell.2006.05.043.

101. Matsumoto, S., Nakatsukasa, K., Kakuta, C., Tamura, Y., Esaki, M., and Endo, T. (2019). Msp1 clears mistargeted proteins by facilitating their transfer from mitochondria to the ER. Mol. Cell *76*, 191-205.e10. 10.1016/j.molcel.2019.07.006.

102. Dederer, V., Khmelinskii, A., Huhn, A.G., Okreglak, V., Knop, M., and Lemberg, M.K. (2019). Cooperation of mitochondrial and er factors in quality control of tail-anchored proteins. Elife *8*, 1–23. 10.7554/eLife.45506.

103. Habeck, G., Ebner, F.A., Shimada-Kreft, H., and Kreft, S.G. (2015). The yeast ERAD-C ubiquitin ligase Doa10 recognizes an intramembrane degron. J. Cell Biol. *209*, 261–273. 10.1083/jcb.201408088.

104. Khmelinskii, A., Blaszczak, E., Pantazopoulou, M., Fischer, B., Omnus, D.J., Dez, G. Le, Brossard, A., Gunnarsson, A., Barry, J.D., Meurer, M., et al. (2014). Protein quality control at the inner nuclear membrane. Nature *516*, 410–413. 10.1038/nature14096.

105. Foresti, O., Rodriguez-Vaello, V., Funaya, C., and Carvalho, P. (2014). Quality control of inner nuclear membrane proteins by the Asi complex. Science. *346*, 751–755. 10.1126/science.1255638.

106. Natarajan, N., Foresti, O., Wendrich, K., Stein, A., and Carvalho, P. (2020). Quality Control of Protein Complex Assembly by a Transmembrane Recognition Factor. Mol. Cell *77*, 108-119.e9. 10.1016/j.molcel.2019.10.003.

107. Okreglak, V., and Walter, P. (2014). The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. Proc. Natl. Acad. Sci. U. S. A. *111*, 8019–8024. 10.1073/pnas.1405755111.

108. Sung, M.-K., Hess, S., Deshaies, R.J., Sweredoski, M.J., and Reitsma, J.M. (2016). Ribosomal proteins produced in excess are degraded by the ubiquitin-proteasome system. Mol. Biol. Cell *27*, 2642–2652. 10.1091/mbc.e16-05-0290.

109. Ho, C.H. (2011). Molecular barcoded plasmid yeast ORF library: linking bioactive compounds to their cellular targets and mapping dosage suppressor networks. PhD Thesis. . http://hdl.handle.net/1807/29745.

110. Meurer, M., Duan, Y., Sass, E., Kats, I., Herbst, K., Buchmuller, B.C., Dederer, V., Huber, F., Kirrmaier, D., Štefl, M., et al. (2018). Genome-wide C-SWAT library for high-throughput yeast genome tagging. Nat. Methods *15*, 598–600. 10.1038/s41592-018-0045-8.

111. 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 Dev. *19*, 2816–2826. 10.1101/gad.1362105.

112. 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*, 1–27. 10.7554/eLife.03023.

113. Pérez-Torrado Editor, R. (2022). The Unfolded Protein Response Methods and Protocols.

114. Khmelinskii, A., Meurer, M., Ho, C.T., Besenbeck, B., Füller, J., Lemberg, M.K., Bukau, B., Mogk, A., and Knop, M. (2016). Incomplete proteasomal degradation of green fluorescent proteins in the context of tandem fluorescent protein timers. Mol. Biol. Cell *27*, 360–370. 10.1091/mbc.E15-07-0525.

115. Lu, Y., Lee, B.H., King, R.W., Finley, D., and Kirschner, M.W. (2015). Substrate degradation by the proteasome: A single-molecule kinetic analysis. Science. *348*, 183–184. 10.1126/science.1250834.

116. Liu, F., and Walters, K.J. (2010). Multitasking with ubiquitin through multivalent interactions. Trends Biochem. Sci. *35*, 352–360. 10.1016/j.tibs.2010.01.002.

117. Swaney, D.L., Beltrao, P., Starita, L., Guo, A., Rush, J., Fields, S., Krogan, N.J., and Villén, J. (2013). Global analysis of phosphorylation and ubiquitylation cross-talk in protein degradation. Nat. Methods *10*, 676–682. 10.1038/nmeth.2519.

118. Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., et al. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science. *285*, 901–906. 10.1126/science.285.5429.901.

119. Li, Z., Vizeacoumar, F.J., Bahr, S., Li, J., Warringer, J., Vizeacoumar, F.S., Min, R., Vandersluis, B., Bellay, J., Devit, M., et al. (2011). Systematic exploration of essential yeast gene function with temperature-sensitive mutants. Nat. Biotechnol. *29*, 361–367. 10.1038/nbt.1832.

120. Ozturk, D.H., Dorfman, R.H., Scapin, G., Sacchettini, J.C., and Grubmeyer, C. (1995). Locations and functional roles of conserved lysine residues in Salmonella typhimurium Orotate Phosphoribosyltransferase. Biochemistry *34*, 10755–10763. 10.1021/bi00034a007.

121. Shakya, V.P.S., Barbeau, W.A., Xiao, T., Knutson, C.S., Schuler, M.H., and Hughes, A.L. (2021). A nuclear-based quality control pathway for non-imported mitochondrial proteins. Elife *10*, 1–21. 10.7554/eLife.61230.

122. 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. 10.1126/science.1142210.

123. Williams, B.R., Prabhu, V.R., Hunter, K.E., Glazier, C.M., Whittaker, C. a, Housman, D.E., and Amon, A. (2008). Aneuploidy affects proliferation and spontaneous immortalization in mammalian cells. October *322*, 703–710. 10.1126/science.1160058.

124. Pavelka, N., Rancati, G., Zhu, J., Bradford, W.D., Saraf, A., Florens, L., Sanderson, B.W., Hattem, G.L., and Li, R. (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. Nature *468*, 321–325. 10.1038/nature09529.

125. Humpton, T.J., Brito, I.L., Hiraoka, Y., Niwa, O., and Amon, A. (2011). Aneuploidy drives genomic instability in Yeast. Science. *333*, 1026–1030. 10.1126/science.1206412.

126. Oromendia, A.B., Dodgson, S.E., and Amon, A. (2012). Aneuploidy causes proteotoxic stress in yeast. Genes Dev. *26*, 2696–2708. 10.1101/gad.207407.112.

127. Yofe, I., Weill, U., Chuartzman, S., Zalckvar, E., Schütze, C., Wiedemann, N., Meurer, M., Knop, M., Khmelinski, A., and Schuldiner, M. (2016). One library to make them all: Streamlining yeast library creation by a SWAp-Tag (SWAT) strategy. Nat Methods *13*, 371–378. 10.1038/nmeth.3795.

128. Weidberg, H., and Amon, A. (2018). MitoCPR—A surveillance pathway that protects mitochondria in response to protein import stress. Science. *360*. 10.1126/science.aan4146.

129. Juszkiewicz, S., and Hegde, R.S. (2018). Quality control of orphaned proteins. Mol. Cell *71*, 443–457. 10.1016/j.molcel.2018.07.001.

130. An, H., and Harper, J.W. (2020). Ribosome abundance control via the ubiquitin–proteasome system and autophagy. J. Mol. Biol. *432*, 170–184. 10.1016/j.jmb.2019.06.001.

131. Kowalski, L., Bragoszewski, P., Khmelinskii, A., Glow, E., Knop, M., and Chacinska, A. (2018). Determinants of the cytosolic turnover of mitochondrial intermembrane space proteins. BMC Biol. *16*, 1–22. 10.1186/s12915-018-0536-1.

132. Ishikawa, K., Ishihara, A., and Moriya, H. (2020). Exploring the complexity of protein-level dosage compensation that fine-tunes stoichiometry of multiprotein complexes. PLoS Genet. *16*, 1–26. 10.1371/journal.pgen.1009091.

133. 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. 10.1074/jbc.M114.596064.

134. Abildgaard, A.B., Voutsinos, V., Petersen, S.D., Larsen, F.B., Kampmeyer, C., Johansson, K.E., Stein, A., Ravid, T., Andréasson, C., Jensen, M.K., et al. (2023). HSP70-binding motifs function as protein quality control degrons. Cell. Mol. Life Sci. *80*, 1–17. 10.1007/s00018-022-04679-3.

135. Szoradi, T., Schaeff, K., Garcia-Rivera, E.M., Itzhak, D.N., Schmidt, R.M., Bircham, P.W.,

Leiss, K., Diaz-Miyar, J., Chen, V.K., Muzzey, D., et al. (2018). SHRED is a regulatory cascade that reprograms Ubr1 substrate specificity for enhanced protein quality control during stress. Mol. Cell *70*, 1025-1037.e5. 10.1016/j.molcel.2018.04.027.

136. Ares M., J., Grate, L., and Pauling, M.H. (1999). A handful of intron-containing genes produces the lion's share of yeast mRNA. RNA *5*, 1138–1139. 10.1017/S1355838299991379.

137. Adhikary, S., Marinoni, F., Hock, A., Hulleman, E., Popov, N., Beier, R., Bernard, S., Quarto, M., Capra, M., Goettig, S., et al. (2005). The ubiquitin ligase HectH9 regulates transcriptional activation by Myc and is essential for tumor cell proliferation. Cell *123*, 409–421. 10.1016/j.cell.2005.08.016.

138. Grabarczyk, D.B., Petrova, O.A., Deszcz, L., Kurzbauer, R., Murphy, P., Ahel, J., Vogel, A., Gogova, R., Faas, V., Kordic, D., et al. (2021). HUWE1 employs a giant substrate-binding ring to feed and regulate its HECT E3 domain. Nat. Chem. Biol. *17*. 10.1038/s41589-021-00831-5.

139. Benschop, J.J., Brabers, N., van Leenen, D., Bakker, L. V., van Deutekom, H.W.M., van Berkum, N.L., Apweiler, E., Lijnzaad, P., Holstege, F.C.P., and Kemmeren, P. (2010). A consensus of core protein complex compositions for Saccharomyces cerevisiae. Mol. Cell *38*, 916–928. 10.1016/j.molcel.2010.06.002.

12 APPENDIX II

12.1 ACKNOWLEDGEMENTS

12.2 CURRICULUM VITAE