Proteomic Analysis of the Cdc48/Ubx Network Identifies a Role for Ubx2 in the Regulation of Lipid Biosynthesis

Thesis by

Natalie Kolawa

In Partial Fulfillment of the Requirement

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2013

(Defended May 10, 2013)

Acknowledgements

There are many people to thank for help getting to this point. First and foremost, I would like to thank Dr. Raymond Deshaies. I am tremendously grateful to have had the opportunity to work in his lab . Not only a brilliant scientist, Ray has been wonderful mentor, who with his unwavering support and encouragement, has guided me throughout this journey. I have learned so much from him.

I would also like to acknowledge the members of my thesis committee: Dr. David Chan, Dr. Shu-ou Shan, and Dr. Paul Sternberg. I would like to thank them for kindly accepting to review this work and for their help, suggestions, and time throughout the years.

I am also very indebted to the all members of the Deshaies Laboratory. I would like to thank them for creating such an enjoyable daily work environment and community, and for their continued support and advice. To my fellow Deshaies Lab graduate students, Tara Gomez, KJ Chang, Nathan Pierce, and Ruzbeh Mosadeghi, thank you for helping me survive the trails and tribulations of graduate school. What an amazing group of people to go through it with. Additionally, other people who deserve mention include: Rob Oania for all his patience and technical support; Rati Verma, for her limitless expertise and advice and for pushing my project forward, even on days when I thought it had nowhere left to go; Willem denBesten for all his help and insightful discussions on science and life in general.

Without Sonja Hess, Mike Sweredoski, Bobby Graham, and Geoff Smith at the Proteome Exploration Laboratory at Caltech, none of this work would have been

possible. Working with them has been one of the highlights of my time at Caltech. I am grateful for their mentorship and willingness to try anything.

Lastly, I would like to thank my incredible family and friends. To my mother, David, and Adam, your love and constant encouragement is what got me here today. And I dedicate this thesis to my father, who was consistently my biggest fan and supporter. He would have been more thrilled than anyone to witness this achievement.

Abstract

Cdc48/p97 is an essential, highly abundant hexameric member of the AAA (ATPase associated with various cellular activities) family. It has been linked to a variety of processes throughout the cell but it is best known for its role in the ubiquitin proteasome pathway. In this system it is believed that Cdc48 behaves as a segregase, transducing the chemical energy of ATP hydrolysis into mechanical force to separate ubiquitin-conjugated proteins from their tightly-bound partners.

Current models posit that Cdc48 is linked to its substrates through a variety of adaptor proteins, including a family of seven proteins (13 in humans) that contain a Cdc48-binding UBX domain. As such, due to the complexity of the network of adaptor proteins for which it serves as the hub, Cdc48/p97 has the potential to exert a profound influence on the ubiquitin proteasome pathway. However, the number of known substrates of Cdc48/p97 remains relatively small, and smaller still is the number of substrates that have been linked to a specific UBX domain protein. As such, the goal of this dissertation research has been to discover new substrates and better understand the functions of the Cdc48 network. With this objective in mind, we established a proteomic screen to assemble a catalog of candidate substrate/targets of the Ubx adaptor system.

Here we describe the implementation and optimization of a cutting-edge quantitative mass spectrometry method to measure relative changes in the *Saccharomyces cerevisiae* proteome. Utilizing this technology, and in order to better understand the breadth of function of Cdc48 and its adaptors, we then performed a global screen to identify accumulating ubiquitin conjugates in *cdc48-3* and *ubxΔ*

mutants. In this screen different *ubx* mutants exhibited reproducible patterns of conjugate accumulation that differed greatly from each other, pointing to various unexpected functional specializations of the individual Ubx proteins.

As validation of our mass spectrometry findings, we then examined in detail the endoplasmic-reticulum bound transcription factor Spt23, which we identified as a putative Ubx2 substrate. In these studies *ubx2Δ* cells were deficient in processing of Spt23 to its active p90 form, and in localizing p90 to the nucleus. Additionally, consistent with reduced processing of Spt23, *ubx2Δ* cells demonstrated a defect in expression of their target gene *OLE1*, a fatty acid desaturase. Overall, this work demonstrates the power of proteomics as a tool to identify new targets of various pathways and reveals Ubx2 as a key regulator lipid membrane biosynthesis.

Table of Contents

List of Tables and Figures

Chapter 1:

Introduction to the Ubiquitin Proteasome System and the Cdc48/p97 ATPase

The Ubiquitin Proteasome System

The ubiquitin proteasome system (UPS) is a critical pathway found in eukaryotes that regulates various processes throughout the cell. By targeting different proteins for degradation to the proteasome, a large ATP-dependent protease complex, the UPS maintains cellular homeostasis and controls many cellular responses. The system gives directionality to pathways via the quick degradation of regulatory elements and degrades damaged and misfolded proteins in the cell. The UPS participates in a large array of events such as transcriptional regulation, signal transduction, and the cell cycle[1]. Impaired UPS activity has been found to play a role in neurodegenerative diseases, such as Parkinson's[2, 3] and Alzheimer's[4], and multiple cancers. Additionally, proteasomal activity has been implicated in immune response and development[1]. Furthermore, due to this system's intricate role in the cell, it has been the focus of many therapeutic investigations[1].

In terms of process, ubiquitin, a small conserved 76 amino acid protein, is covalently attached to substrates via an enzymatic cascade involving three enzymes known as E1s, E2s, and E3s. In the first step of the reaction, the ubiquitin-activating enzyme E1 activates free ubiquitin by forming a thioster bond. The activated ubiquitin, in another similar reaction, is then transferred to a cysteine residue of the E2 ubiquitin-conjugating enzyme. In a final step, ubiquitin is attached to a specific substrate via an isopeptide bond at the C-terminal glycine of ubiquitin and a lysine residue on the substrate protein. This final reaction is mediated by the E3 ubiquitin ligase enzyme. The E1 enzyme is encoded by a single gene, Uba1, in all

eukaryotes[5]. However there are a large variety of E2 and E3 enzymes. These different E2s and E3s act on specific substrates allowing for the targeting of proteins to the proteasome in a very regulated and specific manner[6-8] (Figure 1.1).

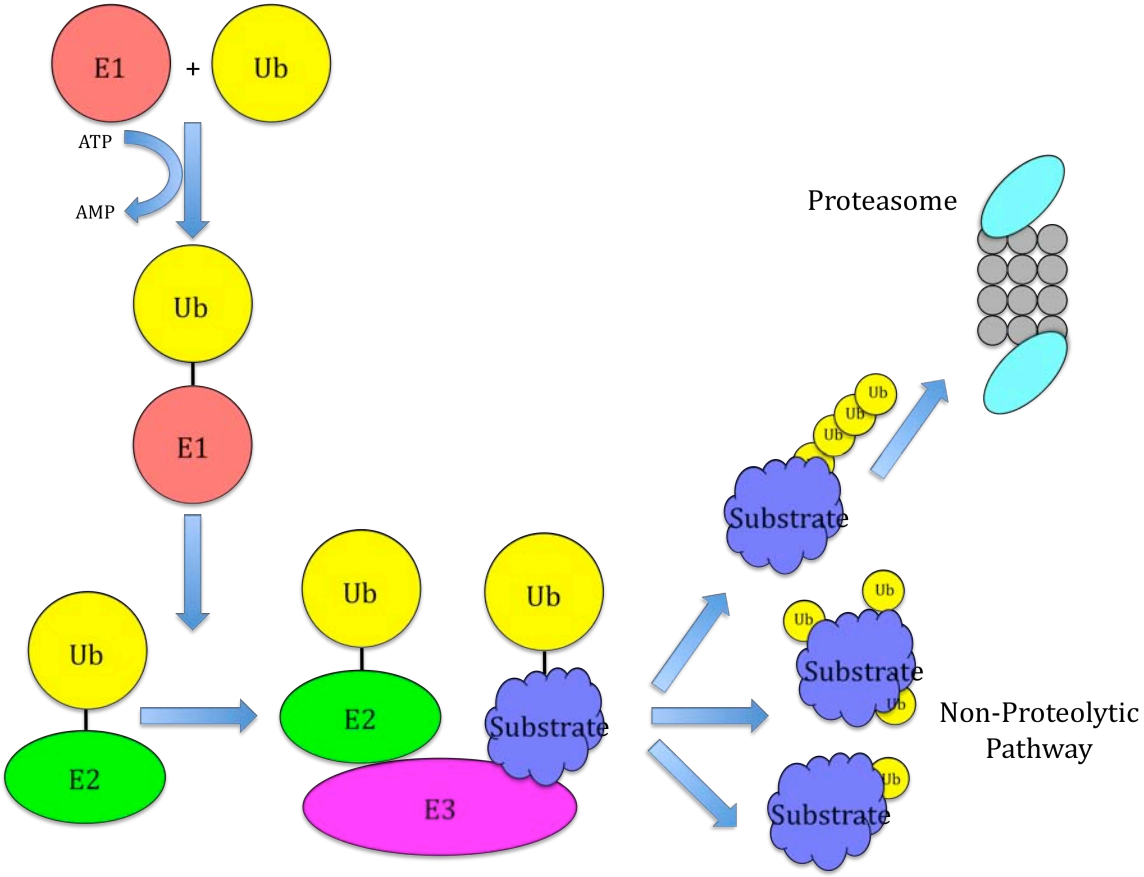


Figure 1.1

Modification of a protein by ubiquitin occurs via a three-enzyme cascade. Ubiquitin is activated by covalent attachment to an E1 enzyme, the ubiquitin-activating enzyme, which couples ATP hydrolysis to the formation of a thioester bond between itself and the ubiquitin molecule. Following activation, ubiquitin is transferred to an E2, an ubiquitin-conjugating enzyme. Lastly the E3 enzyme, an ubiquitin ligase, facilitates the transfer of ubiquitin from the E2 to the target protein. The target protein can either be left monoubiquitinated, or the cycle can be repeated with the protein modified with ubiquitin at different lysine residues known as multi-ubiquitination. Additionally the protein

can be poly-ubiquitinated, with an ubiquitin chain building at the same lysine residue. Mono-ubiquitination and multi-ubiquitination are generally involved in protein regulation in a proteolysis-independent manner. Most often poly-ubiquitinated proteins are subsequently degraded by the 26S proteasome.

This cascade continues until a substrate has built a polyubiquitin chain. Polyubiquitin chains on substrates are formed via the sequential addition of ubiquitin molecules forming isopeptide linkages to one another[9]. The ubiquitin molecule has multiple different lysines allowing for the formation of different combinations of ubiquitin chain linkages. Different chain linkages play a variety of roles in the cell. For instance, Lys-48 polyubiquitin linkages target substrates to the proteasome for degradation[7]. A minimal tetra-ubiquitin chain with Lys-48 linkages is believed to be required for proteasomal recognition. Other ubiquitin chain linkages such as Lys-63 and Lys-29 have been found to be involved in non-proteasomal roles such as DNA-damage repair, endocytosis, and intracellular trafficking[10]. The function of other ubiquitin linkages such as Lys-11, Lys-27, and Lys-33 is not well understood. Additionally, proteins can be left mono-ubiquitinated or multi-ubiquitinated with ubiquitin modifications attached to different lysines on the protein. These types of modifications are also involved non-proteolytic events.

The 26S Proteasome

The 26S proteasome regulates most intracellular degradation. Proteins are delivered to the proteasome and proteolyzed into small peptides that can be

recycled to make new proteins. The proteasome is a large 2.5 mDa ATP-dependent proteolytic complex, consisting of 33 unique protein subunits. The complex is composed of two main modules, a central catalytic core known as the 20S and the 19S regulatory cap subunit. The inside of the 20S catalytic core is hollow, providing an enclosed cavity where proteins are degraded and the openings at the ends of the core allow for target proteins to enter. Each end of the catalytic core associates with a 19S cap that has multiple ubiquitin binding sites to interact with substrates. The 19S recognizes and unfolds polyubiquitinated substrates, allowing them to pass into the narrow pore of the catalytic core.

In terms of makeup, the 19S consists of two subcomplexes, a base that interacts with the 20S, and a lid that recognizes ubiquitinated substrates. The 19S lid is composed of 11 non-ATPase subunits while the 19S base is composed of two non-ATPase subunits that act as scaffolding units and a heterohexameric ring of ATPase subunits that require energy for the unfolding of substrates[11-13].
The catalytic 20S is made of a cylindrical core. It contains two inner hetero-heptameric and two outer heteroheptameric rings[14]. Subunits of these rings have various proteolytic activities (chymotrypsin, trypsin, and caspase-like) with different cleavage preferences, in order to degrade substrates into small peptides[8](Figure 1.2).

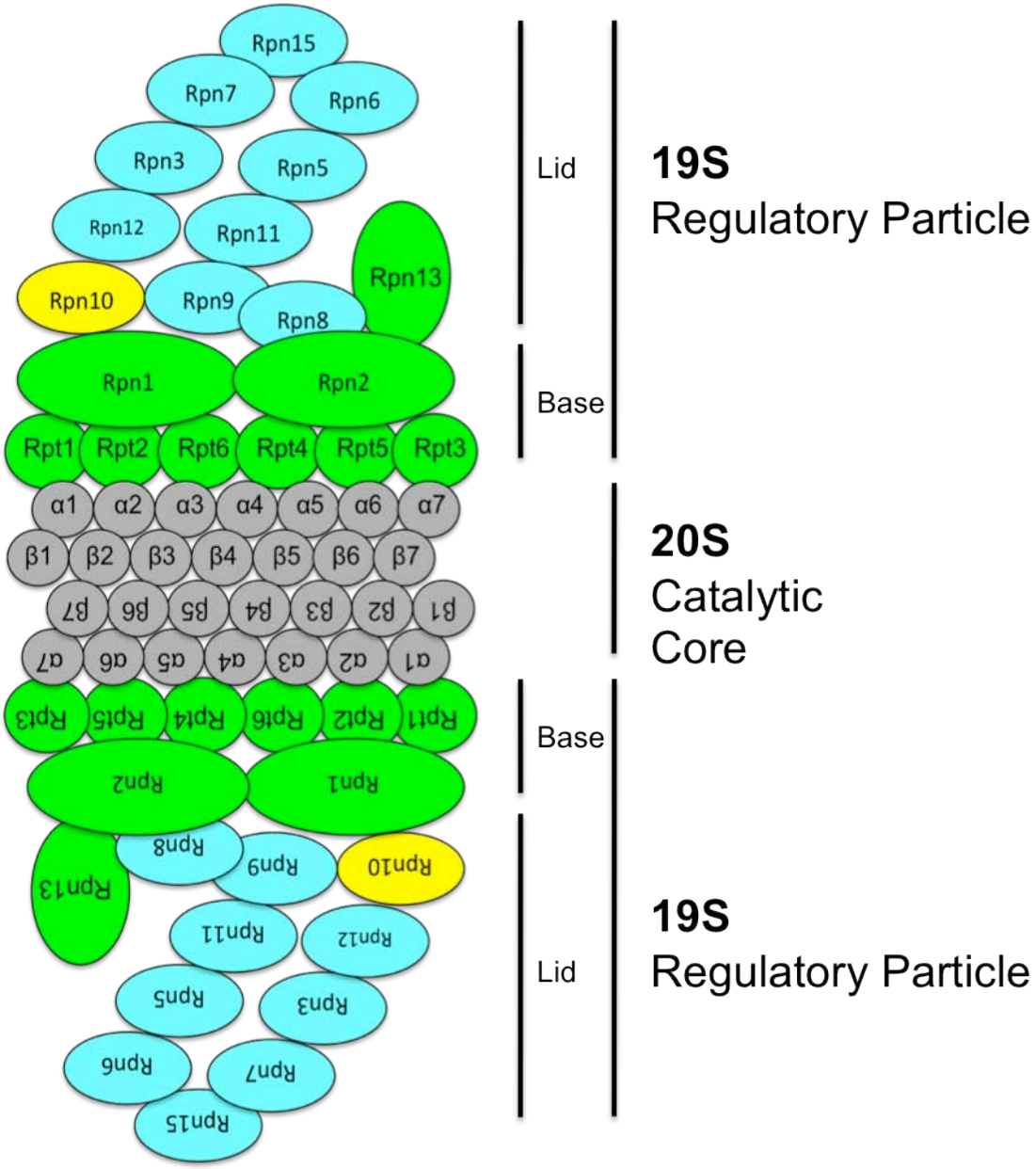


Figure 1.2

The 26S eukaryotic proteasome. A schematic representation of the proteasome which consists of the catalytic 20S core and the 19S regulatory particle. The 19S can be further divided into the base (shown in green) and lid (shown in blue) subcomplexes. Rpn10 is shown in yellow since it is known to be located at the base-lid interface.

In order to undergo selective degradation, receptor proteins recognize ubiquitinated substrates and shuttle them to the proteasome. Rpn10 was the first discovered proteasomal receptor, and was found to be necessary for the efficient delivery of ubiquitinated substrates at the proteasome. Rpn10 is an intrinsic ubiquitin receptor of the proteasome and is believed to form a hinge between the 19S and 20S domains. It contains a UIM domain (ubiquitin-interacting domain) that recognizes polyubiquitinated chains[15-18]. Interestingly, a Rpn10 deletion strain did not display a phenotype indicative of a large-scale deficiency in cellular protein degradation suggesting that there are multiple proteins involved in recognition and delivery of substrate to the proteasome [19]. Not surprisingly, several other receptors, including Rad23, Dsk2, and Ddi1, known as UBA-UBL proteins, were later discovered[13, 20-22]. These proteins loosely associate with the proteasome, rapidly binding and unbinding the complex, delivering ubiquitinated substrates. The proteins recognize and interact with ubiquitin chains on substrates via their UBA domain (ubiquitin-associated domain) and dock to the proteasome via their UBL domain (ubiquitin-like domain)[23, 24]. However, a triple Rpn10, Rad23, and Dsk2 mutant is not lethal, suggesting that there are yet more proteasome receptors playing a role in proteasomal delivery[22].

Additionally, it has been suggested that Cdc48, a large ATP-dependent molecular chaperone, and its adaptors play a role in delivery of substrates. Cdc48 is believed to work in conjunction with the receptors Rad23/Dsk2 in aiding target proteins to the proteasome. A working model is that an ubiquitinated substrate is recruited to Cdc48 by its adaptors Ufd1 and Npl4. After recruitment the substrate is

ubiquitinated further by the chain elongation factor, Ufd2, an adaptor of Cdc48. Ufd2 then recruits Rad23/Dsk2 to the substrate-bound Cdc48 complex and the substrate is targeted for degradation by the proteasome[15, 25]. However, the exact role for Cdc48 at the proteasome is a subject of much debate. It is unclear whether Cdc48's function is to shuttle substrate to the proteasome like Rad23/Dsk2/Ddi1, or if it is required at the proteasome for remodeling rather than delivery.

The Cdc48/p97 ATPase

The previously mentioned Cdc48 (p97/VCP in mammals) is an essential, highly abundant, ATP-driven machine in the cell. Discovered over twenty-five years ago in a screen for conditional mutants causing cell-cycle arrest in yeast, the complex has been shown to play a central role in ubiquitin-dependent degradation. It is a member of the AAA ATPases (ATPases associated with various cellular activities) and has a large array of functions throughout the cell (Figure 1.3). The protein is best known for its critical role in ERAD (endoplasmic reticulum associated protein degradation by the ubiquitin proteasome system), but is involved in a myriad of processes including cell-cycle progression, homotypic membrane fusion, chromatin remodeling, autophagy, and transcriptional and metabolic regulation[26-29]. It is believed that Cdc48 acts as a "segregase" in these processes, using the energy from ATP hydrolysis to extract substrate ubiquitinated proteins from protein complexes or membranes. Cdc48 has been the subject of much attention over the last few years due to its causal link to amyotrophic lateral sclerosis (ALS) and inclusion body myopathy, Paget's disease of the bone, and frontotemporal dementia

(IBMPFD), as well as its implied role in a variety of diseases including cancer[30-34]. As a result, p97 has been the subject of multiple drug development efforts[35, 36].

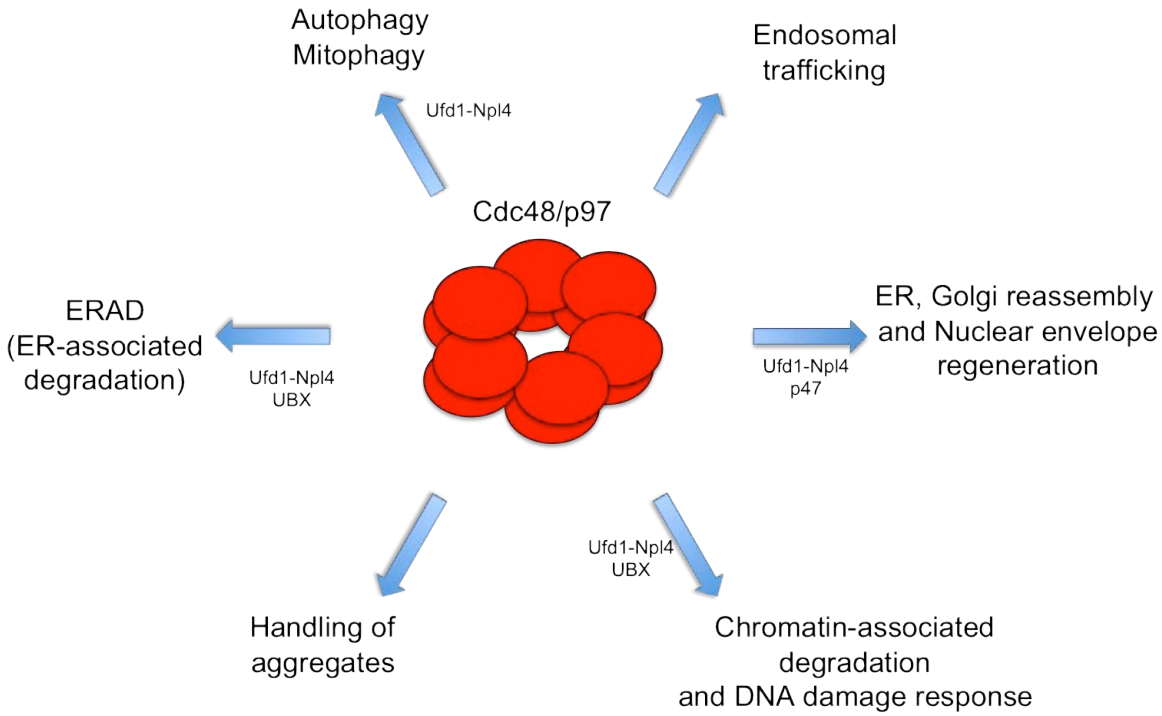


Figure 1.3

Multiple functions of Cdc48/p97. N-terminal cofactors involved in each function are indicated.

In terms of structure, Cdc48 forms a homohexameric complex, with the six subunits arranged in a barrel shape with a central tunnel. Each subunit consists of two ATPase domains, D1 and D2, connected by a short linker (the two domains form stacked rings), a N-terminal domain, and a C-terminal tail. Cdc48 has in total 12 different ATPase sites, with the D1 domain showing weak ATPase activity under physiological conditions, leaving a majority of the activity at the D2 domain[26, 37].

During ATP hydrolysis, Cdc48 undergoes conformational changes which result in the opening and closing of the D1 and D2 rings[38]. It is believed that these conformational changes provide the force for Cdc48 to disassemble protein complexes and extract substrates from intracellular structures. However, the detailed mechanism of substrate handling by Cdc48 is still not well understood. Additionally, even though Cdc48 has been largely deemed a "segregase," others believe the complex primarily functions as "unfoldase", unfolding substrates and aiding in their degradation by the proteasome[39]. The detailed mechanism for both of these suggested functions is the subject of much debate within the field.

Cdc48 Cofactors and Binding Partners

Cdc48 is associated with a multitude of cofactors and binding partners. It is believed that these cofactors help facilitate and provide specificity to the numerous processes Cdc48 is involved in, mediating a stable interaction between the Cdc48 complex and ubiquitinated proteins. With the exception of a few proteins, most cofactors interact with Cdc48 via the N-terminal domain. Perhaps the best characterized adaptors of Cdc48 are Ufd1 and Npl4, which form a heterodimer and are involved in ubiquitin dependent degradation. These proteins are both essential in yeast. Ufd1 and Npl4 have been shown to bind Cdc48 alone or in conjunction with other cofactors. [40-43].

Another class of cofactors is the UBX 'Ubiquitin regulatory X' domain containing family of proteins. Although the largest family of Cdc48 cofactors, it is still very poorly characterized with the cellular function of many of the Ubx proteins

largely unknown. To date, there have been thirteen UBX proteins identified in mammals, seven of which are found in yeast. The carboxy-terminal UBX domain (which is also found in Npl4) is similar in structure to ubiquitin and serves as a binding module to Cdc48. Among the UBX proteins, there is a subfamily of proteins containing a UBA domain. These amino-terminal domains are involved in binding ubiquitinated substrates directly[41, 42](Figure 1.4).

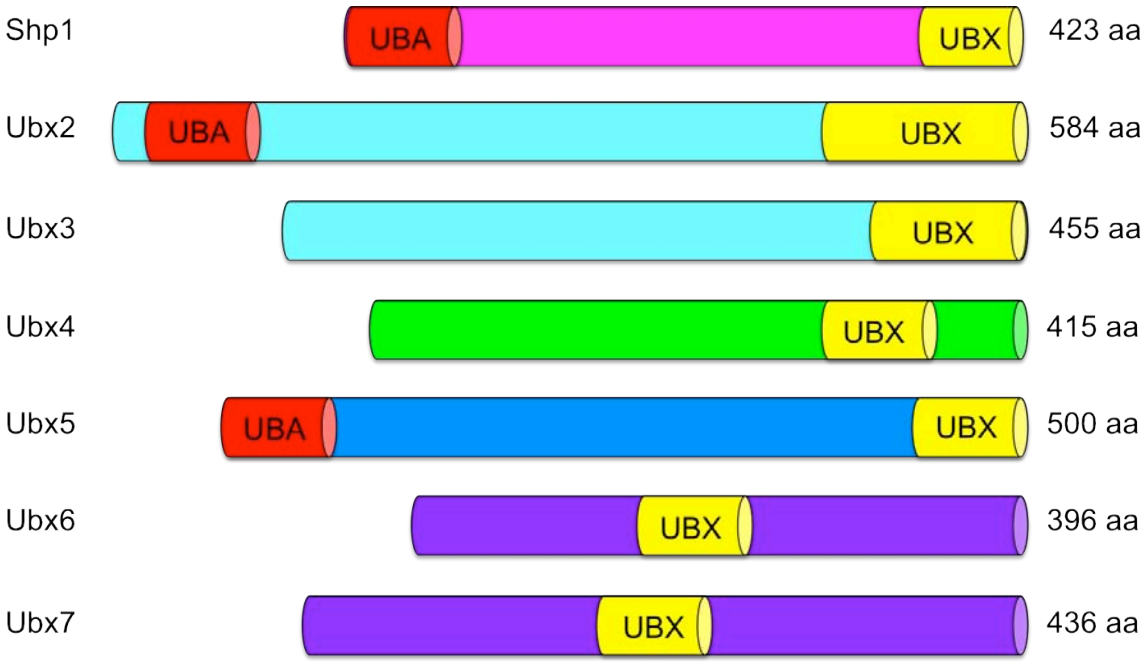


Figure 1.4

UBX proteins found in *Saccharomyces cerevisiae*. UBX (yellow) and UBA (red) domains of proteins are indicated. Proteins demonstrating substantial homology outside of these domains are shown in similar colors.

As mentioned, very little is known about UBX proteins, the processes they are involved in, and their potential substrates. Yeast Shp1, also known as Ubx1, is an

UBA/UBX domain-containing adaptor with a mammalian ortholog, p47. Shp1 seemingly plays an important role in cells, as a deletion strain is hardly viable. Shp1 and p47 have been linked to homotypic membrane fusion and autophagy[26, 28, 29]. The Buchberger lab recently demonstrated Cdc48/Shp1 complex involvement in the promotion of cell cycle progression by the positive regulation the protein phosphatase Glc7 in yeast. It is theorized that the complex moderates Glc7 by regulating its interaction with the activator protein Glc8[44]. Lastly, Shp1 has been shown to interact with Cdc48 exclusively of Ufd1 and Npl4[39, 45].

Contrastingly, the UBX protein Ubx2 has been demonstrated to work in complex with Ufd1 and Npl4. Ubx2 (also known as Sel1) and its mammalian ortholog UBXD8 are ER membrane UBA-UBX proteins. In both systems, the proteins have been linked to the regulation of lipid droplet homeostasis and are involved in ERAD[46-49].

Also localized at the ER and throughout the cytoplasm is the adaptor Ubx4. Ubx4, containing only a UBX domain, has been linked to ERAD, the degradation of the metabolic enzyme FBPase, and the degradation of the largest subunit of RNA Polymerase II, Rpb1 after UV treatment along with Ubx5[50-52]. Ubx5, another UBA/UBX protein, is localized to the nucleus on chromatin and throughout the cytoplasm. It has a mammalian ortholog, UBXD7 that has been found to link the interaction of p97 and its substrate HIF1α, as well as play a role in linking in p97 to neddylated cullin ring ligases via its unique UIM domain[53, 54]. The other UBX containing cofactors are Ubx3, Ubx6, and Ubx7. Very little is known about these

proteins. Ubx6 and Ubx7 share large sequence homology and a triple mutant *ubx4Δ ubx6Δ ubx7Δ* has a known meiotic defect[55].

Npl4, Ufd1, and the UBX proteins are all known as substrate-recruiting factors, which are believed to facilitate the decision between major cellular pathways and to improve binding of Cdc48 to substrate. There is another class of cofactors, known as substrate-processing cofactors. These cofactors have additional enzymatic activity and are involved in substrate regulation such as deubiquitination or chain elongation. Some of these proteins include Otu1, Ufd3, and Ufd2[25, 42]. The large variety of proteins associated with Cdc48 allows the protein to assemble a large array of cofactor complexes with distinct functions and interaction networks. Although many cofactors of Cdc48 have been discovered, their number is still growing.

Cdc48 and Proteasome-Dependent Degradation

As mentioned above, the best understood role of Cdc48 is its role in ERAD and proteasome-dependent processes. In ERAD, Cdc48, in complex with Npl4 and Ufd1, are recruited to the ER membrane via Ubx2, an integral membrane protein. During this process, the complex binds ubiquitinated ERAD substrates, which are misfolded or incompletely assembled. The target proteins are retrotranslated and segregated from the ER membrane or their binding partners and are then delivered to the proteasome for degradation. Ubx4 has also been shown to play a role in ERAD, however its role in the process is not well understood[47].

Recently Cdc48 was linked to protein quality control/mitochondrial stress response at the mitochondrial outer membrane. Upon exposure of cells to rapamycin or hydrogen peroxide, Cdc48 is recruited to the mitochondrial membrane via a new adaptor, Vms1, where it binds ubiquitinated proteins and transports them to the cytosol for degradation. Npl4 was found to be involved in this process, forming a heterodimer with Vms1[56, 57]. The Cdc48-Npl4-Vms1 complex activates degradation of Fzo1, a yeast mitofusin protein, which is degraded by the proteasome via the SCF-^{Mdm30} complex. Human orthologs of Fzo1, MFN1 and MFN2, are also degraded by the UPS and the human E3 ligase Parkin in a p97 dependent manner[58]. Detailed mechanism of Cdc48/p97 at the mitochondria is not yet understood, but it is feasible that the Cdc48-dependent mitochondrial retrotranslocation system works similarly to the one in ERAD. It is has also been speculated that Cdc48-mediated degradation may have a regulatory role in mitochondrial membrane fusion (mitophagy), preventing the fusion of defective and healthy mitochondria, and facilitating the degradation of damaged mitochondria via autophagy[59].

Additionally, Cdc48 is involved in cytoplasmic and nuclear protein degradation. Cytoplasmic degradation via Cdc48 was first demonstrated with synthetic substrates of the UFD pathway. Following this discovery, native cytoplasmic proteins, fructose-1,6-bisphosphatase (FBPase) and phosphoenolpyruvate carboxykinase (PEPCK) were found to undergo Cdc48-dependent degradation during the gluconeogenic switch, the catabolite inactivation

upon glucose signaling. The UBX protein Ubx4 was later linked to the degradation of FBPase in this pathway[26, 51, 60].

In the nucleus, Cdc48 was found to play a role in the UV-induced proteasomal degradation of chromatin-bound Rpb1 (RNA polymerase II) at DNA lesions. Once RNA polymerase has stalled and undergoes ubiquitination by the ligase Cul3, the UBX cofactor Ubx5 recruits Cdc48 to remove Rpb1 from the chromatin bound complex[52]. A similar chromatin-associated function was found for p97 prior to this finding with the extraction of Aurora B kinase from mitotic chromosomes. These studies linked Cdc48/p97 to cell cycle regulation, as the function of Aurora B kinase is to regulate the attachment of mitotic spindle to the centromere. It was shown that during mitosis, p97-Ufd1-Npl4, by removing the kinase, inhibits its activity and allows for chromatin decondensation and nuclear envelope formation[61, 62]. In yeast, Cdc48 has been linked to the degradation of Far1, a G1-cyclin-dependent kinase inhibitor, thereby promoting the cell-cycle[63].

Cdc48 and Proteasome-Independent Processes

Although Cdc48 was first believed to be involved exclusively in proteasome-dependent processes, the complex was recently linked to the other major degradative system in the cell, the lysosome. Cdc48 and Ubx1/Shp1 are involved in macroautophagy, specifically in autophagosome biogenesis. The first implication of a role for Cdc48/p97 in autophagy was from studies on the pathogenesis of Paget's disease of the bone and frontotemporal dementia (IBMPFD), a disease with missense mutations in p97. A phenotype seen in these patient's tissues is vacuoles

rimmed with LC3 (Atg8 in yeast), implying aberrant autophagosome formation. LC3/Atg8 are proteins involved in the maturation of vacuoles and are known targets of the Cdc48-Shp1/p97-p47 complex. This phenotype was shown again in transgenic mice overexpressing disease mutants of p97[64-67]. However, even though Cdc48 has been linked to this process, the exact molecular function is unknown.

In terms of other proteasome-independent processes, Cdc48 has been linked to endolysosomal trafficking, membrane fusion, and reformation of the Golgi complex following mitotic disassembly[42, 68, 69].

Lastly, Cdc48/p97 has been linked to transcriptional regulation. It was reported that p97 interacts with IkBα, the inhibitory factor of the NF-kB pathway. A possible regulatory role for p97 in this pathway may be control of IkBα stability[29]. Cdc48 has been linked to a similar processing pathway in yeast, the OLE pathway, which regulates the expression of genes involved in fatty acid metabolism. In this pathway the proteasome is involved in processing but the substrate only undergoes partial degradation[70].

The OLE Pathway

The fluidity of yeast membranes is largely controlled by Ole1, a Δ9-fatty acid desaturase enzyme that converts saturated fatty acids into unsaturated fatty acids by introducing double bonds into their carbon chains. Known as the OLE pathway, this process is essential for cell viability and is regulated by the UPS and Cdc48-

Npl4-Ufd1 complex[71]. In this process, ER membrane-bound transcription factors are processed and transported to the nucleus to promote OLE1 transcription.

The first step of the OLE pathway involves the ubiquitin-proteasome dependent processing of Spt23 and Mga2. Spt23 and Mga2 are homologous transcription factors that are essential for activation of *OLE1* expression. Deletion of either gene has little effect, however, the double deletion causes fatty acid auxotrophy and is lethal but can be rescued by providing an external source of unsaturated fatty acids. Spt23 and Mga2 exist in two forms: an inactive ER membrane-bound 120kDa precursor protein and a processed soluble 90kDa protein that is localized to the nucleus and cytoplasm. The inactive precursor forms of Spt23 and Mga2 are bound to the ER as homodimers via a single 20 amino acid hydrophobic domain at their C-termini. Upon fatty acid depletion, Spt23 and Mga2 are ubiquitinated by the HECT E3 ligase Rsp5 leading to the processing of the p120 by the proteasome. During processing it is believed that the protein forms a hairpin loop with the proteasome beginning processing internally, degrading through the C-terminus, and leaving the N-terminal active form intact[72-77].

After the active p90 molecule is produced, the protein remains tethered to its partner p120 molecule at the ER membrane. In a second step, the transcriptionally active p90 fragment is released from the membrane by the Cdc48-Npl4-Ufd1 complex. The p90 protein is then translocated to the nucleus to activate *OLE1* transcription to create more unsaturated fatty acids in the cell. Once the active protein is translocated it is thought to be very unstable[75, 78](Figure 1.5).

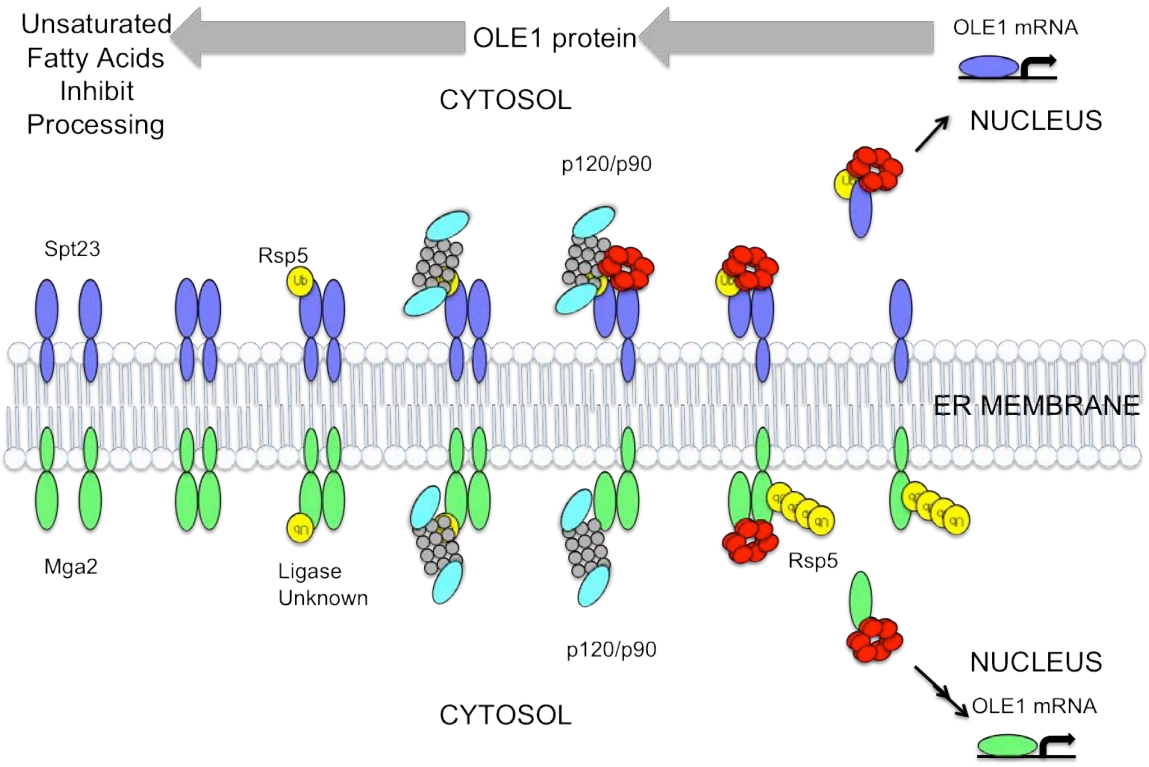


Figure 1.5

The OLE Pathway. Spt23 (blue) and Mga2 (green) are activated upon a depletion of unsaturated fatty acids in the cell. Spt23 is monoubiquitinated by the Rsp5 ubiquitin ligase which is required for processing by the proteasome (gray and blue barrel). Post-processing, the Cdc48-Ufd1-Npl4 complex (red) liberates the active p90 form of Spt23 from its p120 binding partner and p90 is translocated to the nucleus for transcription. In the case of Mga2, Rsp5 is dispensable for initiation of processing and the E3 ligase at this step is unknown. Post–processing by the proteasome the active p90 form is released from its p120 binding partner by the Cdc48-Ufd1-Npl4 complex. However, for this step polyubiquitination of the p120 binding partner by Rsp5 is required for release. Mga2 is then translocated to the nucleus for transcription. Production of Ole1 leads to increased unsaturated fatty acid levels, which thereby inhibit processing of Spt23.

In addition, Cdc48 and its binding partners have been linked to Spt23 p90 stability. Ufd2, an E4 ligase that binds to Cdc48, extends the ubiquitin chain on the

transcription factor enabling the targeting of the protein to the proteasome. In *ufd2* deletion strains, Spt23 p90 is stabilized, causing excessive expression of Ole1 and sensitivity of the strain to unsaturated fatty acids. In addition, once translated, the Ole1 protein is a highly unstable protein that is degraded by ERAD. Ole1 stability is regulated by unsaturated fatty acid levels in the cell, thereby creating a tightly regulated negative feedback loop[25, 79].

Although much of the OLE pathway has been characterized, there is still much to be understood. The actual mechanism of fatty acid regulation of Spt23/Mga2 activation is still unclear. As mentioned above, the pathway is based on feedback regulation[79]. Production of Ole1p leads to the production of unsaturated fatty acids which in turn has been shown to inhibit Spt23 processing. It has been demonstrated that dimerization of Spt23 is necessary for ubiquitination. This has led to the belief that processing is regulated at either the level of dimerization or upon ubiquitination by Rsp5[71]. The UFA sensor, whether it is Rsp5 or upstream of Rsp5, is still largely unknown.

Additionally, although homologous, Spt23 and Mga2 have demonstrated some differences in processing. At the level of ubiquitination, Mga2 uses an Rsp5-independent mode of processing and the E3 ligase is unknown. Post-processing, Rsp5 dependent ubiquitination appears to be required for mobilization of the processed Mga2[76, 80]. In the case of Spt23, processing requires Rsp5 dependent ubiquitination. Furthermore, the level of ubiquitination post-processing of the two transcription factors is different. The active Spt23 p90 form is monoubiquitinated[75] but in the Mga2 pathway, the inactive p120 binding partner

of p90 is polyubiquitinated prior to mobilization[76]. Lastly, addition of unsaturated fatty acids represses the processing of Spt23. In contrast, Mga2 processing is unaffected by the addition of unsaturated fatty acids. However, OLE1 transcription is still repressed indicating that fatty acid mediated OLE1 repression seems to act post-processing for Mga2 and prior to processing of Spt23[81]. All this evidence indicates that the processing pathways are much more complex and Spt23 and Mga2 may play slightly different roles[82].

Finally, the role of the Cdc48-Ufd1-Npl4 complex in the OLE pathway is still unclear. As mentioned above, a role for Cdc48-Ufd1-Npl4 in the release of p90 from the ER membrane after proteasomal processing is known. The dismantling of the active p90 from its p120 binding partner was demonstrated in vitro with the addition of Cdc48 to cell extracts. However, another group suggested that Cdc48, Ufd1, and Npl4 may also be involved in proteasome-dependent processing[83]. In temperature sensitive mutant strains of the mentioned proteins Spt23 and Mga2 were shown to be less efficiently processed as compared to wild-type strains. However, the role of the complex in this part of the pathway is not yet understood. Some possible functions include extraction of the C-terminus of one of the transcription factor monomers to aid degradation by the proteasome. However, if the transcription factors undergo internal processing a non-“extraction” role for the complex is possible. Perhaps the complex recruits proteasome to the substrate and aids in processing. With all these questions and unknowns, it is evident that much more work needs to be done to better understand the intricacies of the OLE pathway[75, 83, 84].

Shotgun Mass Spectrometry

As can be deduced from the summary above, the UPS/Cdc48 system plays an integral role in the cell. We have only begun to understand its breadth of influence and have yet to discover many new substrates. However, the complexity and the intricacies of this system make it extremely challenging to study in a global manner. As such, much of the work has been on the details of specific pathways and discovery of new targets has proven difficult.

In the last decade there have been major advances in shotgun mass spectrometry techniques and data analysis software that have enabled quantitative analysis of changes in protein level on a proteome-wide scale. These new technologies have enabled one to look at complex samples and analyze and quantify a large number of proteins under varying conditions, revolutionizing the field of proteomics[85-87].

At the heart of whole proteome sequencing is utilizing the proper pre-fractionation technique to expand dynamic range. Often, high abundance peptides prevent the sequencing of low abundance ones, which limits sequencing depth. Historically, the GeLC [88-91]and MudPIT[92-94] approaches have been the most common pre-fractionation techniques. These approaches separate proteins by molecular weight (GeLC) and peptides by cationic exchange and reverse phase fractionation (MudPIT) and decrease analyte complexity. However, no group has been able to sequence at the *Saccharomyces cerevisiae* proteome depth using these approaches. In response to this continued issue of scale, the Mann Laboratory

developed a few new methodologies and performed the most comprehensive identification and quantification of the budding yeast proteome to date[95]. In these experiments, an Agilent OFF-GEL, an apparatus that performs isoelectric focusing of proteins and peptides in solution separating them into different wells on a pH gradient, was used for fractionation[96-98]. Additionally, the Mann Group developed a strong anionic exchange (SAX) stage tip method as another fractionation means for deep sequencing[99]. They also developed a comprehensive software package, MaxQuant, for very accurate identification and quantitation of protein ratios[100]. Adaptation and implementation the above system to carry out accurate and reproducible quantitative mass spectrometry on a proteome-wide scale would be a powerful tool to discover new substrates and function of the UPS/Cdc48 system.

Research Objectives

With the current tools available discovering new roles and substrates of Cdc48 and its adaptors has proven difficult. In this dissertation I describe the implementation and optimization of a quantitative mass spectrometry method to monitor changes in *Saccharomyces cerevisiae* protein levels on a proteome-wide scale (Chapter 2). After establishing this technology at Caltech, I developed a screen to survey changes in the ubiquitin proteome in *cdc48* and *ubx* mutants in the hope to find new targets of the Cdc48 network (Chapter 3). Different mutants displayed reproducible patterns of ubiquitin conjugate accumulation that differed greatly from each other supporting the idea of functional and pathway specializations of

individual Ubx proteins. The results of this screen suggested a wide variety of new functions and potential new substrates of various Ubx proteins of which previously very little was known.

To further validate my mass spectrometry findings, I examined in detail the ER-membrane bound transcription factor Spt23 of the OLE pathway, which was identified as a putative Ubx2 substrate. I verified Spt23 as an Ubx2 substrate, demonstrating a role for Ubx2 in Spt23 processing and stability, and as a regulator of lipid biosynthesis (Chapter 3). In addition, this work illustrates the value of proteomics to identify new targets for ubiquitin receptor pathways and the dataset is a rich resource for future investigations into the role of the Cdc48 network.

Chapter 2:

Implementation and Optimization of a Mass Spectrometry-Based Method to Perform Proteome Quantification of *Saccharomyces cerevisiae*

Introduction

Proteomics is a rapidly evolving science created to analyze large amounts of proteins, even an entire proteome, in a single experiment. Mass spectrometry has become the most popular tool in the field and is used to identify and characterize proteins in complex mixtures. In contrast to gene expression studies in genomics, proteomics directly analyzes the level of gene products present and can analyze various other protein details such a post-translational modifications, protein interactions, and subcellular distributions[101, 102].

Proteomics first began with two-dimensional electrophoresis (2-DE). In this method, protein staining patterns across samples were compared to identify "unregulated" and "downregulated" proteins. This method was roughly quantitative, as one could have a vague idea of relative amounts of protein between the different samples. However, a large caveat to this method was limited dynamic range and, eventually, over the next few years, 2-DE gels were replaced with the use of mass spectrometry-based proteomics[102-104].

The current sample analysis workflow in mass spectrometry emerged from peptide mass fingerprinting. This technique utilized proteolytic digestion of simple protein mixtures with site-specific proteases and subsequent mass spectrometric analysis of the resulting peptides. These peptides were analyzed against a theoretical digest of a protein database and the proteins with the closest theoretical digest were matched[105-108]. However, this method was largely unusable for highly complicated mixtures. In the next few years, with the development of collision induced dissociation (CID), electrospray ionization (ESI), and the

introduction of a new class of mass analyzers (quadropole-ion trap mass spectrometers), the proteomics field dramatically changed[109-111]. These changes allowed for new advances in complicated polypeptide mixture analysis and, eventually, led to deep proteome sequencing[95].

With the ability to rapidly identify large numbers of proteins with mass spectrometry, the focus in the field then shifted to the reliable quantitation of the proteins detected. Many different techniques and strategies for absolute or relative quantitation of protein mixtures emerged. One of the most popular methods, known as stable isotope labeling of amino acids in cell culture (SILAC), has proven to be one of the most powerful approaches. SILAC consists of growing two (or three) cell populations in media containing a heavy or light form of an essential amino acid. After several cell doublings, the complete cellular proteome is labeled with the incorporated amino acid and every peptide pair is separated by the mass difference introduced by the labeled amino acid. This method then employs mass spectrometry and sophisticated software packages to accurately identify and measure relative protein levels.[112]

As far as the field has come in the last two decades, due to the complicated and sophisticated nature of the instruments and techniques, few laboratories in the world have been able to achieve this high level of mass spectrometry-based quantitative proteomic sequencing. In this chapter we describe the implementation of a high-throughput SILAC-based quantitative mass spectrometry setup at Caltech. Although our laboratory did own the most advanced mass spectrometer, an LTQ-Orbitrap (an instrument with the combination of an ion trap with a high mass

accuracy analyzer), little was known about optimal sample preparation, fractionation techniques, chromatography and instruments settings and setup, and data analysis. Caltech had an experimental arrangement that had not evolved with the rapidly changing field. As such, we attempted to revamp our entire experimental method and emulate the setup of the Mann Lab, a laboratory at the forefront of quantitative mass-spectrometry based proteomics. This Chapter gives, in detail, a description of the changes we made to our system in order to achieve our goal of deep sequencing and quantitation at the level of the *Saccharomyces cerevisiae* proteome. The objective was to then utilize this technology to quantify changes in the proteome upon various perturbations or mutations in yeast cells, specifically in the UPS.

Methods

Yeast Cell Culture

The wild-type *Saccharomyces cerevisiae* strain RJD 4614 (S288C background) was used throughout all optimization experiments. Cells were grown to log-phase (OD_{600} 0.7-1.0) in yeast extract peptone dextrose (YPD) liquid medium. Cells were harvested by centrifugation for 5 minutes at 4,000 x g, washed with cold TBS, and flash frozen. For SILAC experiments, a cell strain (S288C, arg4::HYG, lys2Δ0) that is an auxotroph for lysine and arginine was used. Cells were grown in Complete Synthetic Media (CSM) with 2% dextrose containing 20 mg/L lysine and arginine or in "heavy" medium with 20 mg/L $^{13}C_6{}^{15}N_2$-lysine and $^{13}C_6$-arginine (Cambridge Isotope Laboratories). Cells were grown for 10 generations to log-phase (OD_{600} 1.0-

2.0) and equal amounts of the normal and "heavy" SILAC-labeled cells (as determined by OD_{600}) were mixed 1:1, harvested, washed, and flash frozen.

Lysate preparation via Milling Device

Lysis of frozen cells was performed by mechanical disruption in a milling device (CryoMill MM400, Retsch). Cells were lysed during 3 cycles of 3-minute milling at 20 Hz with continuous cooling.

Lysate and Sample Preparation via Urea Lysis and In-solution digestion

Cells were resuspended in 8M Urea Lysis Buffer (8M Urea, 300mM NaCl, 50 mM sodium phosphate, 10 mM Tris-HCl pH 8.0, 0.2% Triton) and lysed three times by bead beating in a FastPrep (MPBio) at top speed. The lysate was clarified by spinning at 13,600 rpm for 15 minutes at room temperature. Lysate was reduced with 3mM (tris(2-carboxyethyl)phosphine) (TCEP) (Sigma) by incubation for 20 minutes at room temperature. Next, reduced cysteines were alkylated by the addition of 10mM chloroacetamide (Sigma) and incubated at room temperature in the dark for 15 minutes. Proteins were proteolytically digested at room temperature with the addition of Lys-C (Wako) at a ratio of 1:100 (protease: protein) for four hours at room temperature. Samples were then diluted to 2M Urea by the addition of 100mM Tris-HCl pH 8.5 and adjusted to 1mM $CaCl_2$ (Sigma). Sequencing grade trypsin (Promega) was added at the ratio of 1:200 (protease: protein) and the lysate was incubated overnight at room temperature in the dark.

The next day, samples were acidified with 1% TFA and desalted via a C_{18} peptide MacroTrap (Microm Bioresources) on the Alliance-HT (Waters Corporation).

Filter Aided Sample Preparation (FASP)

Cells were milled with the milling device and powder was resuspended in SDT-lysis buffer (4% SDS, 100mM Tris-HCl pH 7.6, 100 mM DTT). Cells were boiled for 5 minutes and sonicated to reduce the viscosity of the sample. Samples were clarified by centrifugation at 16,000 x g for 5 minutes. The FASP II protocol on a 30k filter was then closely followed as previously described[113]. Digested peptides were collected, desalted, lyophilized, and resuspended prior to use.

Isoelectric Focusing (IEF) of peptides

100-200 μg of peptides were separated according to isoelectric point using the 3100 OFFGEL Fractionator (Agilent). The fractionation was set up according to the manual protocol with the High Res Kit, pH 3-10. 24 cm Immobiline DryStrip, pH 3-10 (GE Healthcare) and IPG Buffer, pH 3-10 (GE Healthcare) were used. Gel strips were rehydrated for 20 minutes with 50 μl of rehydration buffer. 150 μl of rehydrated peptide solution was then added to each well of the OFFGEL device. The wells were covered to prevent evaporation of liquid and peptides were focused for 50 kVh at a maximum current of 50 uA, maximum voltage of 8000 V, and maximum power of 200mW. Run time was ~36-40 hours. Each peptide fraction was acidified by the addition of 3% ACN, 1% TFA, and 0.5% acetic acid and then desalted with C18 Stage Tips.

SDS-Lysis for GeLC

Frozen cells were boiled for three minutes at 100°C to prevent protease activity. Cells were then lysed in SDS-PAGE buffer (2% SDS, 50mM Tris pH 6.8, 10% Glycerol, Bromophenol blue, 0.1% 2-Mercaptoehtanol) by bead-beading on a FastPrep (MPBio) for 60 seconds at top speed. The cells were boiled for 3 minutes and centrifuged in a microcentrifuge at top speed for 5 minutes. Lysate was then collected and used for SDS-PAGE.

GeLC

Lysate (~50-100 μg) was separated by 1D SDS-PAGE, using 4-12% NuPage Novex Bis-Tris gels (Invitrogen) and NuPage MES SDS running buffer. The gel was run at 100 milliAmps for 1 hour. The gel was stained using the Colloidal Blue Staining Kit (Invitrogen) and cut into 12 slices. The peptides were then prepped via in-gel digestion. The gel bands, which were cut into 1 mm^3 cubes, were washed with 100 mM ammonium bicarbonate for 5 minutes and then washed with a 1:1 100mM ammonium: ACN for five minutes in order to destain. This was repeated until the gel pieces were no longer blue. The proteins were then reduced in-gel with 10mM DTT in 100mM ammonium bicarbonate by incubating at 50°C for 30 min. Next, the gel pieces were incubated in 50 μl of 55mM chloroacetamide made in 100mM ammonium bicarbonate for 20 minutes at room temperature for alkylation. The samples were then washed and dehydrated with ACN, rehydrated with a 6ng/ul trypsin (Promega) solution, and incubated overnight at 37°C. The next day peptides

were extracted by various washes with 0.1% formic acid, ACN, and water. The peptides were then lyophilized and resuspended prior to MS analysis.

Strong Anionic Exchange Fractionation

Desalted peptides from whole cell lysate (~50-100 μg) were fractionated on in-house built strong anion exchange (SAX) pipettes which were assembled based on the StageTip protocol [99] by stacking 6 layers of an Empore Anion Exchange disk (3M Purifications) into a 200 μl pipette tip. Equilibration and elution of fractions was done according to protocol with Britton and Robinson Buffer at various pH. Peptides were loaded at pH 11 and the flow-through was captured on Stage Tip [114] containing 3 layers of C_{18} membrane. Fractions were subsequently eluted from SAX with buffer solutions of pH 8, 6, 5, 4, 3, respectively. The eluted peptide fractions were captured on C_{18} StageTips, desalted, and eluted with 80% acetonitrile, 0.5% acetic acid, lyophilized, and resuspended prior to MS analysis.

LC-MS/MS

After optimization, mass spectrometry experiments were performed on an EASY-nLC (Thermo Scientific) connected to a hybrid LTQ-Orbitrap Classic with a nanoelectrospray ion source (Thermo Scientific) in a setup and settings[115] very similar to those previously described[116]. Binding and separation of the peptides took place on an in-house packed 15 cm silica analytical column (75 μm inner diameter) packed with reversed phase ReproSil-Pur $C1_8$AQ 3 μm resin (Dr Maisch GmbH, Ammerbuch-Entringen, Germany). Samples were run for varying amounts of

time on a 5% to 25% acetonitrile in 0.2% formic acid gradient at a flow rate of 350 nL per minute. The mass spectrometer was programmed to acquire data in a data-dependent mode, automatically switching between full-scan MS and tandem MS acquisition. Survey full scan MS spectra (from m/z 300 to 1,7000) were acquired in the Orbitrap after the accumulation of 500,000 ions, with a resolution of 60,000 at 400 m/z. The ten most intense ions were sequentially isolated, and after the accumulation of 5,000 ions, fragmented in the linear ion trap by CID (collisional energy 35% and isolation width 2 Da). Precursor ion charge state screening was enabled and singly charged and unassigned charge states were rejected. The dynamic exclusion list was set for a 90s maximum retention time, a relative mass window of 10 ppm, and early expiration was enabled.

MudPIT

Peptides were pressure-loaded onto an in-house packed triphasic microcapillary column as described previously[117]. A fused silica capillary with a 75 μm inner diameter and a 5 μm tip was packed with 6.5 cm of C_{18} AQUA reverse-phase material, 3.5 cm of strong cationic exchange material, and another 2.5 cm of C_{18} AQUA. The sample-loaded column was placed inline between an Agilent HPLC and the LTQ-Orbitrap. Samples were separated on a six-step chromatography program totaling over 12 hours.

Data Analysis

All raw data files were analyzed by MaxQuant (v 1.0.13.13) [100] and searched against the Saccharomyces Genome Database. The search parameters included tryptic digestion with a maximum of two missed cleavages. A fixed carboxyamidomethyl modification and the variable modifications of oxidation of methionine and protein N-terminus acetylation were all included in the search. A 1% False Discovery Rate (FDR) threshold for both peptide and proteins was used. At least two peptides were required for protein identification and at least two different scanning events were required for protein quantitation.

Results

As mentioned, our objective was to optimize conditions of sample preparation and sequencing in order to achieve deep proteome sequencing. In order to accomplish this goal, we overhauled the entire experimental method from protein extraction, to sequencing, to analysis (Figure 2.1). It was crucial to investigate any changes and enhancements that could be made. Our hope was to achieve deep yeast proteome sequencing at a level of ~3,600 proteins which had just been demonstrated by the Mann group at Max-Planck[95].

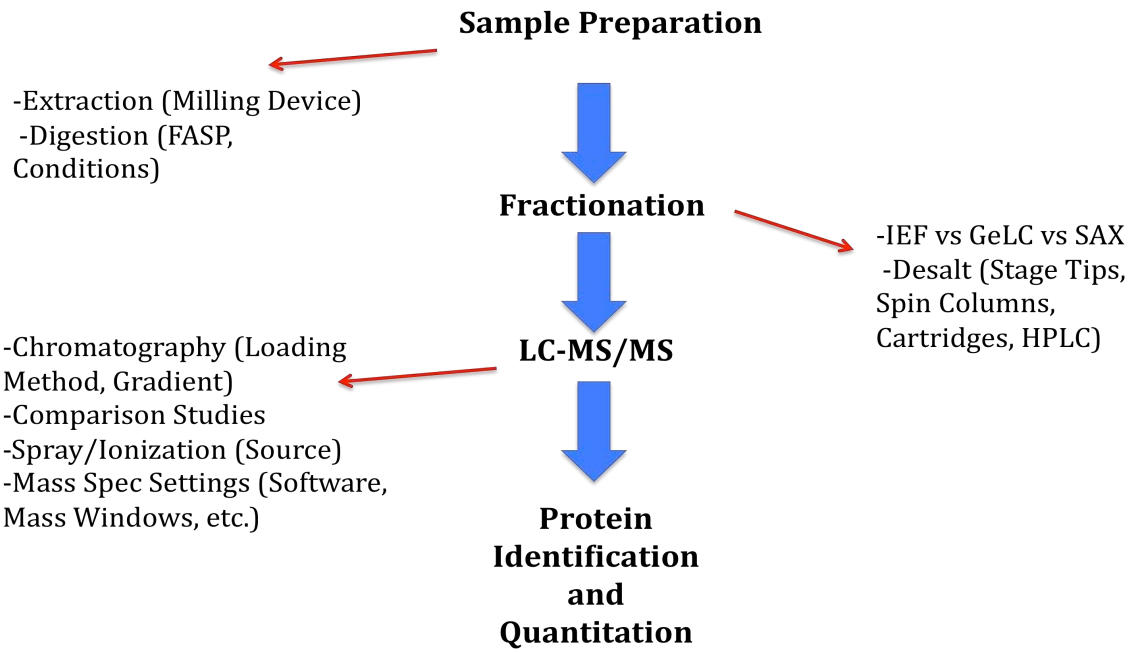


Figure 2.1

Experimental workflow. To optimize the experimental method, conditions and parameters in all stages of sample preparation were tested as shown above.

Sample Preparation

To begin we looked at sample preparation, first focusing on protein extraction. The two methods compared were standard lysis with bead beading on a FastPrep and extraction using a CryoMill milling device. The FastPrep lysis method uses a harsh and rapid method of extraction. Cells were resuspended in an 8M Urea Buffer and then disrupted by "bead beating." The addition of urea and detergent to the buffer allowed for complete denaturation and extraction of proteins. The CryoMill method of lysis is similar to cryogenic grinding. The sample underwent lysis with ball milling under constant cooling with liquid nitrogen. The process resulted in efficient grinding and extraction. The powder was then resuspended in an 8M Urea buffer.

Unexpectedly, we found a significant difference in sequencing depth when comparing the two methods of preparation. During 160 minute sequencing gradients the samples prepared by the CryoMill consistently resulted in more protein identifications. The CryoMill samples in four of five experimental runs resulted in 30-40% higher protein identifications than the samples prepared by the FastPrep (Table 2.1). This was further demonstrated when the samples where fractionated by isoelectric focusing. Again, the CryoMill samples delivered much more sequencing depth across all fractions, indicating that the CryoMill method of lysis leads to more efficient and thorough extraction of different proteins.

Yeast Lysate	**Fast Prep (Protein Identifications)**	**Milling Device (Protein Identifications)**
Sample Set 1	344	518
Sample Set 2	338	452
Sample Set 3	452	447
Sample Set 4	321	422
Sample Set 5	299	436

Table 2.1

Sample lysis with a Milling Device leads to deeper sequencing. In a comparison of sample preparation with a Fast Prep versus a Milling Device, yeast lysate samples prepared with the milling device led to more proteins identified across various experimental replicates.

Post-lysis we focused on various digestion conditions. We investigated different proteases for digestion, digestion temperature, and explored the use of filter sample-aided preparation (FASP) in yeast. In terms of digestion, a LysC only digest or LysC/Trypsin digest were considered. LysC is a serine protease that hydrolyzes proteins specifically at the carboxyl terminus of lysine. LysC is often used in conjunction with trypsin, another serine protease that specifically cleaves at the carboxy termini of lysine and arginine. We examined the sequencing depth of samples prepared by digestion with LysC alone versus samples digested with LysC and Trypsin. Digestion with LysC resulted in fewer peptides and led to less crowded MS spectra. However, in our hands, we found the LysC digested samples to have many highly charged, large spectra suggesting inefficient digestion. Additionally, we found the use of LysC/Trypsin combination led to a 20% increase in proteins identified in yeast lysate samples.

After deciding on a dual protease digestion, we investigated the effects of temperature on digestion efficiency. Our digests were originally prepared according to MudPit protocols at 37°C[117]. However after comparing room temperature and 37°C digestion we discovered room temperature digests led to a higher digestion efficiency, with a lower amount of peptide, spectrum, and sequence missed cleavages (Figure 2.2).

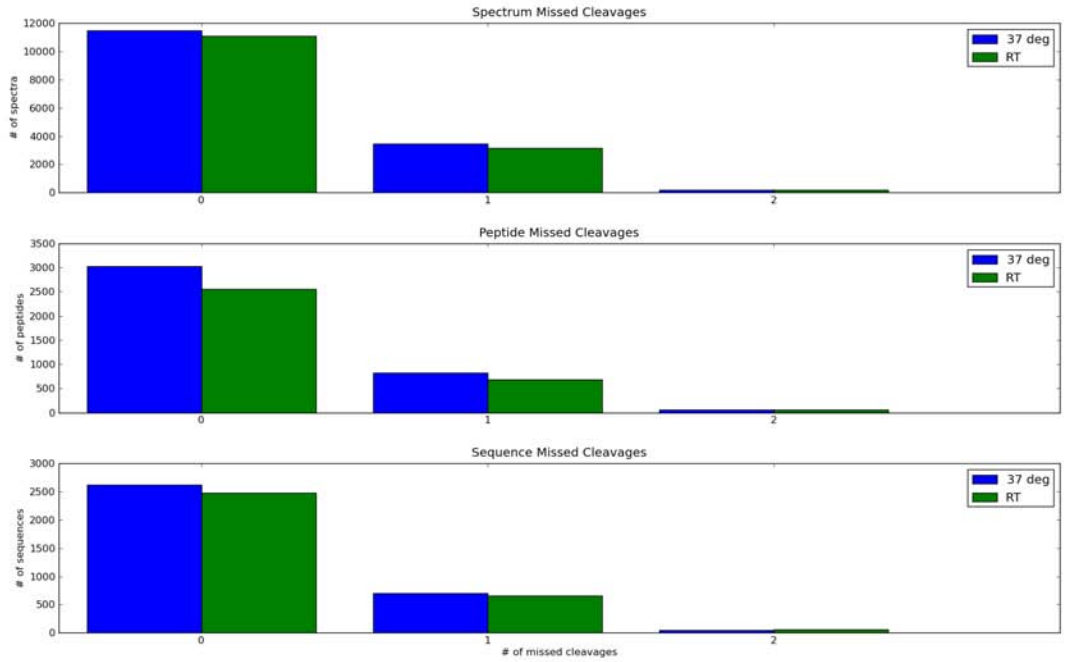


Figure 2.2

Urea digestions are more efficient at room temperature. The amount of spectrum, peptide, and sample missed cleavages after digestion of yeast lysate samples at room temperature versus 37°C demonstrated an advantage to digesting at room temperature.

These results may be due to carbamylation, an effect of performing the digestions in 8M Urea. Urea (a chaotrope), when in solution, is in equilibrium with ammonium cyanate. Carbamylation occurs when a form of cyanate, isocyanic acid, reacts with the amino groups of proteins. Cyanate additionally reacts with the side chains of lysine and arginine residues, leading to peptide missed cleavages. For sequencing purposes, carbamylation and missed cleavages add another dimension of modifications and search parameters, making sequencing results complicated and, at times, difficult to match. A urea solution, if left at room temperature, will always

degrade to isocyanic acid. The reason we believe we witnessed less effective digestion at higher temperatures was due to more carbamylation. The degradation of urea is accelerated at higher temperatures. Indeed, when we adjusted and began doing digests at room temperature, this led to a ~10% increase in protein identifications across all experiments.

Lastly, we implemented a FASP (Filter-Aided Proteome Preparation) methodology in yeast. This method was originally developed for mammalian cells and lead to an almost 40% increase in protein identifications in lysate samples[113]. At the heart of the method is complete solubilization of the proteome in sodium dodecyl sulfate (SDS) which is exchanged by urea on a standard filtration device. Peptides, eluted after in-solution digestion on the filter, are extremely pure, which leads to deep proteome coverage. A key difference between yeast and mammalian cells is their cell walls. Yeast cells are much more durable and difficult to lyse, making the recommended 4% SDS lysis by boiling difficult. We attempted to adapt this methodology by milling cells and resuspending the powder in 4% SDS buffer. After doing FASP, we noticed very clean spectra but saw a minimal improvement in the number of protein identifications in yeast lysate.

Fractionation

Although we focused on optimization of sample processing and sequencing, we simultaneously investigated fractionation techniques to increase proteome sequencing depth. We compared four different fractionation methods including GeLC, Multi-Dimensional Protein Identification Technology (MudPIT), Strong

Anionic Exchange (SAX), and Isoelectric Focusing (IEF). Protein lysate samples are hugely complex with a vast array of protein modifications and a large dynamic range. Fractionation is used to distribute samples into multiple simpler fractions, thereby decreasing the sample dynamic range and allowing for deeper sequencing.

The GeLC-MS/MS fractionation strategy of in-gel digestion of proteins isolated by gel electrophoresis has been used in the mass spectrometry proteomics field for many years[90]. In this method, proteins are separated by molecular weight using SDS-PAGE gel electrophoresis. The gel is then cut into multiple fragments, and as mentioned above, the proteins undergo in-gel digestion. The peptides are then extracted and each fraction is most often run on a 60-90 minute gradient. A large caveat to this method, although one can use a large amount of starting material there is a large loss of sample during processing with inefficient and uneven elution of peptides.

We next investigated the MudPIT strategy. In this fractionation technique, a sample, or in our case, a yeast cell lysate, is digested in urea and loaded directly onto a LC column where it undergoes various stages of fractionation on the column prior to sequencing. The peptides undergo three phases of fractionation on the column: reverse-phase, followed by strong cationic exchange (SCX), followed by another reverse-phase fractionation via a step gradient of acetonitrile and ammonium acetate buffers for 12 hours. MudPIT is commonly used as an alternative to gel-based protein separation[118, 119]. An advantage of this method is that peptide samples undergo fractionation on column and are directly eluted in-line to an ion trap mass spectrometer minimizing sample loss.

A fractionation method known as strong anionic exchange (SAX) was also tested. Strong anionic exchange (SAX) has been used in the field for many years with offline fractionation prior to running on a mass spectrometer. We implemented a protocol utilizing in-house built SAX pipette tips resulting in six different fractions that were analyzed for 160 minutes each[99].

Lastly, we implemented and tested an OFF-GEL Isoelectric Focusing method, where peptides are separated based on their isoelectric points. This is achieved by carrier ampholytes in an immobilized pH gradient gel strip. Samples were collected in 24 liquid fractions, which were then analyzed for 160 minutes each. Analysis of all fractions was around 72 hours.

After implementation of all the fractionation techniques, not surprisingly, the most sequencing depth was achieved with OFF-GEL fractionation. For the purpose of yeast proteome identification, sequencing of 24 IEF fractions with the large amount of sequencing time was used in order to achieve a depth of ~3,500 proteins. However, the SAX fractionation served as a convenient alternative, with sequencing for 24 hours leading to around 2,300 protein identifications. The GeLC and MudPIT strategies, although shorter sequencing methods, were not reproducible enough and did not give enough sequencing depth to be considered for large-scale sequencing experiments.

LC-MS/MS

A large effort was undertaken into the optimization of chromatography methods and mass spectrometric settings. All chromatography parameters were

studied, including loading methods, gradients, elution times, and ion spray tips. As for the mass spectrometry settings, many changes were made to have the most optimized and efficient sequencing of complex mixtures.

To begin, we purchased a brand new LC system. This system, the EASY-nLC from Proxeon, is a high-pressure chromatography system capable of nanoflow elution. This system allows for extremely sharp chromatography and is extremely sensitive and reproducible. Even though this is a very sophisticated and cutting edge system there were still many parameters to optimize on the EASY-nLC.

After implementing the new system, we investigated whether to install a pre-column filter to the standard chromatography setup. When using high-pressure chromatography systems, peptide mixtures are often passed through an online pre-column filter. This is done to protect the column from dust, salts, or particles that pass from the HPLC pump onto the column. This prevents the system from clogging which can lead to a shut down of the system and a loss of the experimental run. In addition this keeps the mass spectrometer clean. However, use of a pre-column can lead to sample loss and at times may be unnecessary if ample precaution is taken during sample preparation. We compared the use of pre-column to direct loading (no pre-column) in identical yeast lysate samples and gradients (Table 2.2).

Loading Method	Protein Identifications
UPLC (Pre-Column)	521, 542
Direct Load	643, 712

Table 2.2

Comparison of LC loading methods. An evaluation of the number of proteins identified using the pre-column versus the direct loading methods indicates a substantial increase in proteome coverage with the use of the direct load method.

From these experiments, we deduced that use of the direct load method led to a consistent 25-30% increase in protein identifications in yeast lysate. This was further confirmed by intensities of experimental runs. When loading the same amount of sample, the direct load experiments consistently showed up to a magnitude higher signal than that of the run with a pre-column.

We then tested several other parameters in the chromatography system, upstream of ion injection. We investigated the use of several capillary column widths and lengths and tested various C_{18} bead sizes within the column. Smaller beads provide more surface area and optimal separation, improving chromatographic resolution. However the optimal size of beads is dependent on the pressure limits of the chromatography system. Additionally, the larger the diameter of the capillary column the lower the electron spray ionization sensitivity due to sample dilution. We settled on what had previously been the lab standard of 15 cm

silica analytical columns with a 75 μm inner diameter packed with reversed phase C_{18}AQ 3μm resin.

As a continuation of these questions, we also investigated the effects of packing the electron spray ionization tip with C_{18}AQ resin. Prior to packing the tip, our setup included a 15cm column with a connector fitting, linking the column to a thin glass capillary tip. This was done in order to easily switch out tips upon problems with electron spray ionization (ESI) without having to replace an entire 15 cm column. Our comparison indicated the packed tips led to slightly longer retention time compared to empty glass tips, indicating that packed tips may have an effect on ESI. However, these results were not conclusive and we settled on unpacked tips due to the convenience in maintenance of the system. We also tested whether gold-plated or steel tips for ESI would be superior to the glass capillary tips. We saw no difference in retention time or spray quality between all three types of tips.

Additionally, the composition of buffers and a gradient method to have the maximum amount of sequencing time per experiment were optimized. We ran yeast lysate samples with various acentonitrile elution gradients and examined which composition led to the shortest peptide retention times with continuous sequencing throughout the entire gradient. On our system we settled on a 5%-25% acetonitrile gradient. This gradient led to continuous elution of peptides throughout the entire experimental run (Figure 2.3).

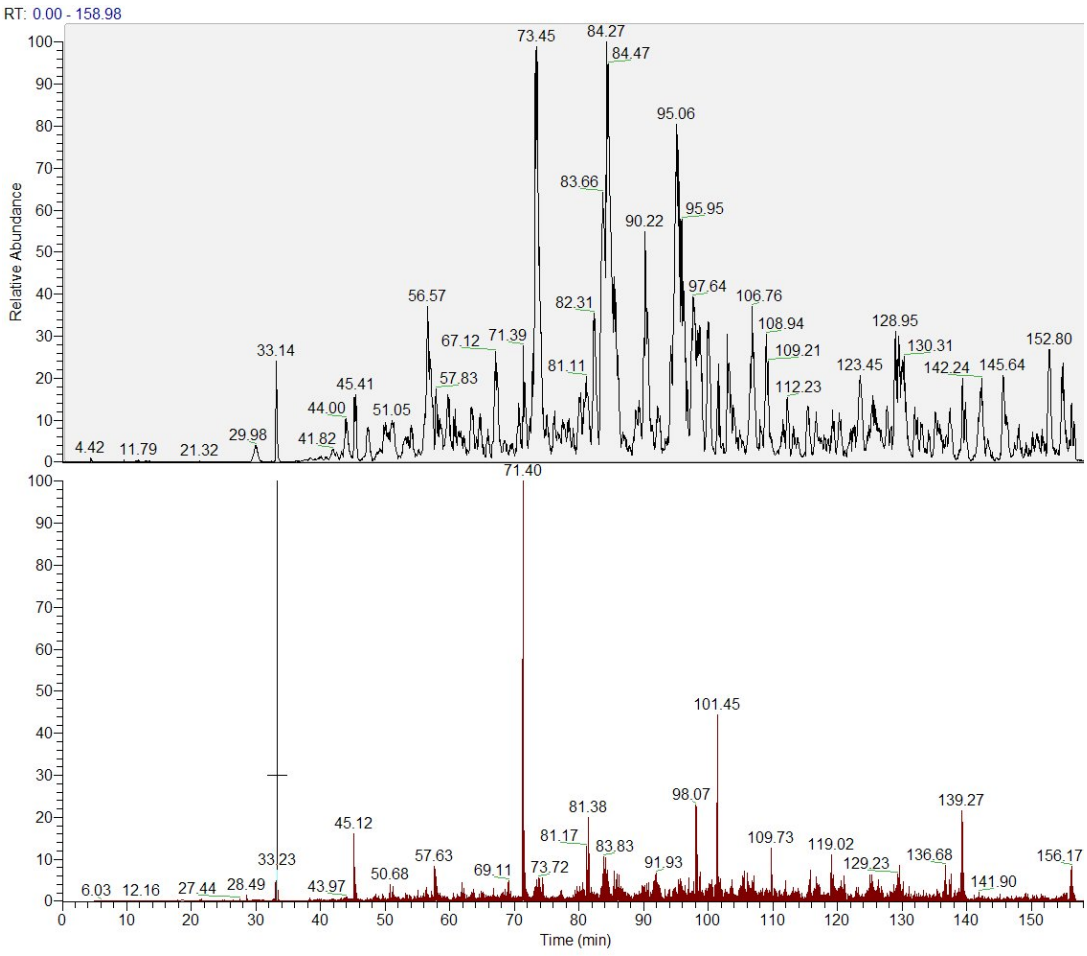


Figure 2.3

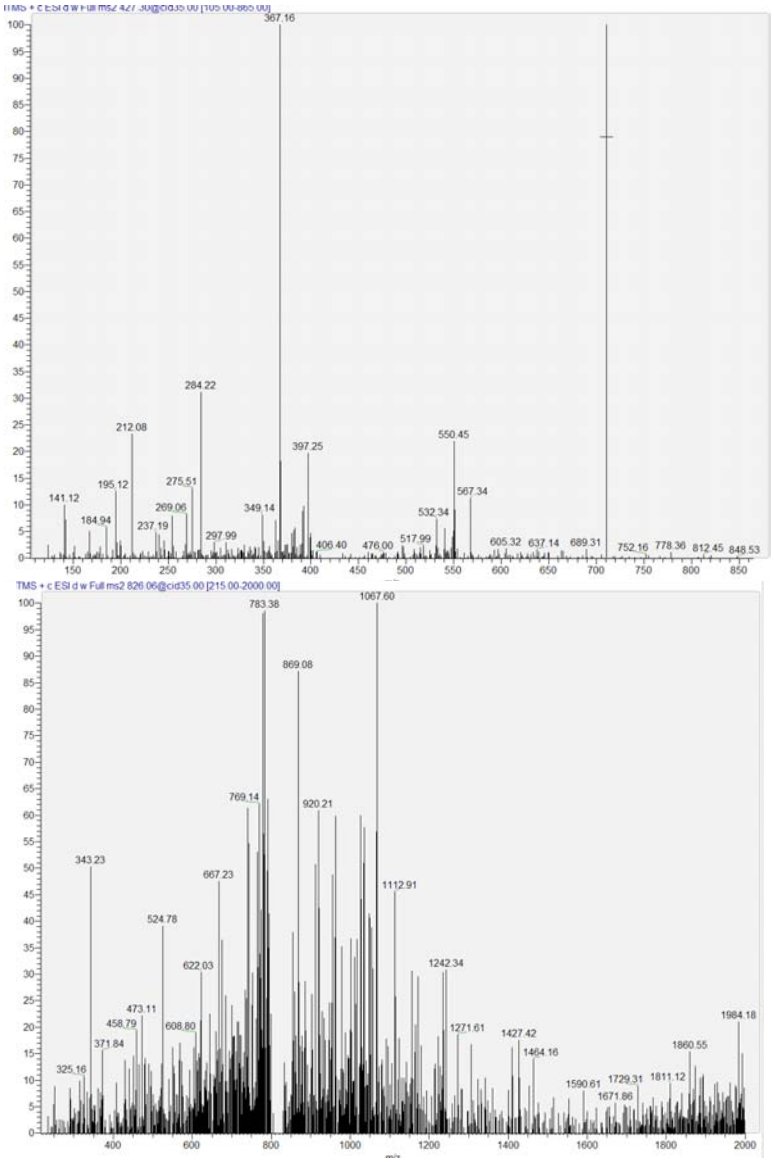
An optimized chromatography gradient. A chromatogram of peptide elution (top) and MS/MS events (bottom) triggered throughout the length of an entire experimental run demonstrated the quality of the newly implemented elution gradient.

Downstream of chromatography, significant efforts were made to optimize sequencing, with a complete overhaul of instrumental methods. Various parameters were tested, including lockmass calibration, the size of mass windows, sequencing range, and charge state recognition and sequencing. Our end goal was to have extremely efficient sequencing, with the highest number of sequencing events of different peptides.

To begin the overhaul of methods, we first altered our method of acquisition. In our original experimental methods, acquisition was done by utilizing long high resolution initial MS scans prior to selecting the most intense ion for fragmentation and sequencing. We tested a new approach of acquiring data, known as the data-dependent mode, where the mass spectrometer automatically switches between full-scan MS and tandem MS acquisition. Upon the accumulation of a certain threshold of ions, the instrument does an initial high-resolution scan in the Orbitrap. Following the scan, the ten most intense ions are isolated and fragmented in the linear ion trap. This mode leads to a larger number of sequencing events, however, there is still a fine balance between large number of scans and scan quality. In order to achieve clear spectra, a higher threshold of ions is preferred, but in order to accumulate ions one must sacrifice time for accumulation. The ion threshold was monitored to see which threshold gave the best spectra with minimal sacrifice of sequencing events. We discovered that the data-dependent configuration led to a four-fold increase of MS2 events and a large increase in protein identifications. Additionally we made other minor changes in sequencing settings. We allowed for charge state screening, permitting the instrument to monitor charge states of peptides, rejecting singly and assigned charge peptides. This mode decreased the likelihood of unassigned sequences.

Additionally, we optimized our precursor mass window for sequencing. Upon the selection of an ion for fragmentation, the instrument accumulates ions within a certain mass range to allow for minimal mass deviation. It has been shown in complex mixtures that up to 95% of peptides are within 1.5ppm of the correct value

and almost none deviate by more than 2ppm[120]. Our initial settings allowed for 3ppm, a conservative window to include all correct peptide hits. However, upon testing a 2ppm window we found an almost 20% increase in identifications and much cleaner MS2 spectra (Figure 2.4).



Yeast Lysate	2ppm Window	3ppm Window
Sample 1	421 Identifications	357 Identifications

Sample 2	372 Identifications	291 Identifications

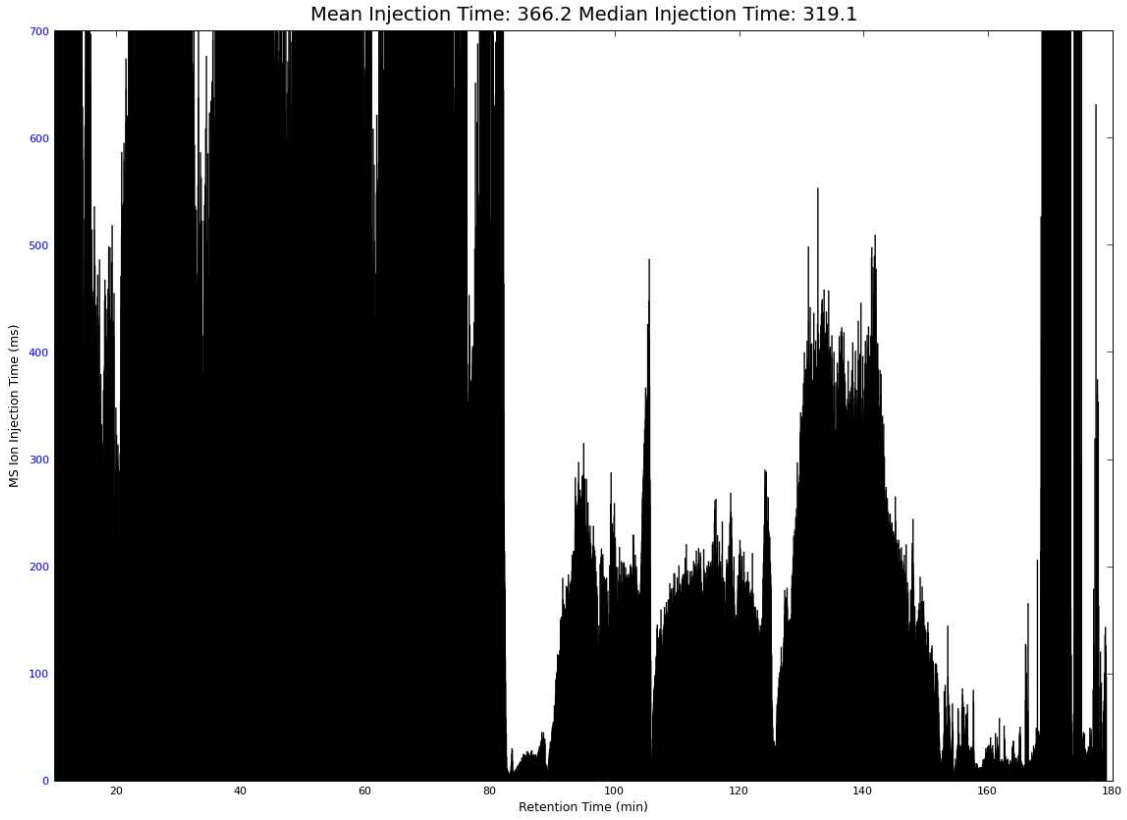
Figure 2.4

A smaller mass window leads to deeper sequencing of yeast lysate. MS/MS spectra of peptides using a 2ppm (top) and 3ppm (bottom) mass window are shown. Evaluation of the spectra and a comparison of the number of proteins identified in yeast lysate samples (two experimental replicates) using the different mass windows indicated an increase in proteome covering using the smaller mass window.

Furthermore, we investigated the usage of lock mass calibration. Mass accuracy is a very important in proteomic experiments and is key to peptide identification. To compensate for drifts in instrument calibration during an experimental run, a 'lock mass,' an internal mass standard for on-going instrument calibrations during the run, is used. The lock mass most often added to LTQ-Orbitrap spectra is found in laboratory air. We found that this recalibration had little impact on our overall protein identifications.

Lastly, after making various changes to the instrument settings and chromatography, we investigated the ion injection time of our experimental runs. As mentioned earlier in this chapter, in the analytical cycle of the data-dependent mode, ions are first injected into the Orbitrap and once a threshold of ions has filled the trap the most abundant parent ions are isolated and sequenced. To achieve the maximal number of sequencing during an experimental run, it is crucial to go through as many analytical cycles as possible. Ion injection is usually the rate-determining step in this process for if there is a delay in ion injection MS2 events will not be triggered. These delays may be due to inefficient spray or

chromatography and even improper settings. Prior to our overhaul of the experimental setup our mean injection time was 366 ms with a median of 319 ms. When studying the distribution of injection times throughout an experimental run, it was clear that ions were not being efficiently delivered to the trap and sequencing was suboptimal (Figure 2.5).



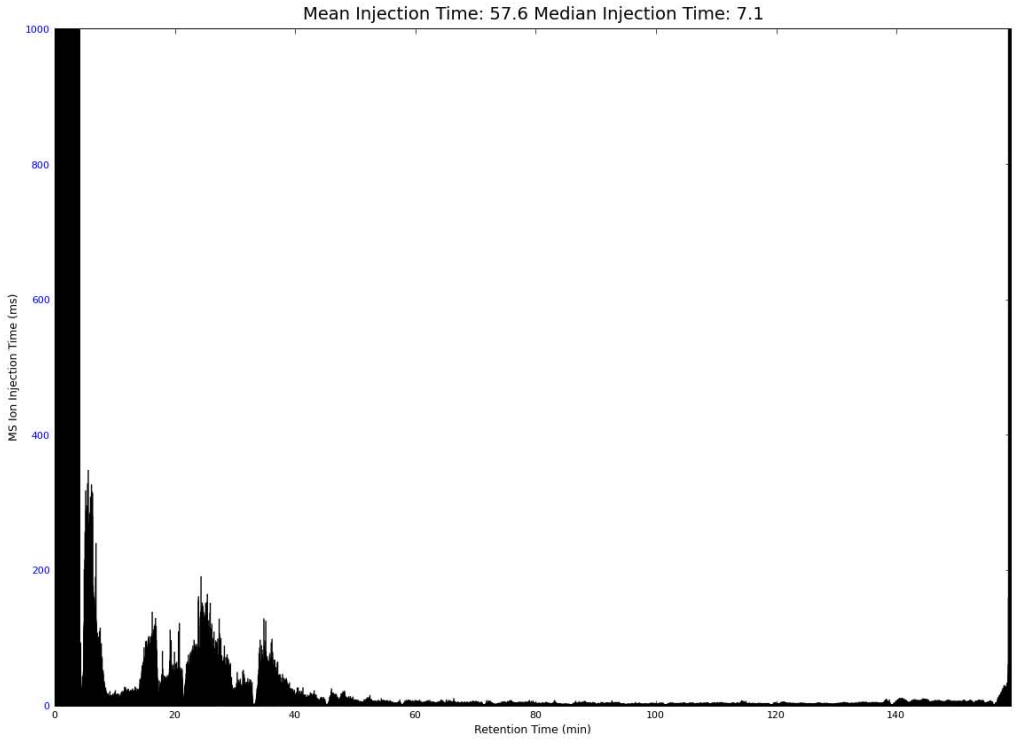


Figure 2.5

Improvement in the distribution of injection times in experimental runs. Comparison of the efficiency of ion delivery prior to (top) and after (bottom) the optimization of the experimental method.

Subsequent to our optimization there was a significant change in ion injection times. With our new setup, the mean injection time was 57.6 ms with a median of 7.1 ms. These results indicated that ion delivery and sequencing ere extremely efficient.

It is also important to mention that we implemented the MaxQuant quantitative proteomics software package into our pipeline. This was, and currently still is, the most advanced tool for analysis of large mass spectrometric data sets. We consistently used this software to track our progress in the development of the experimental method.

Essentially, all possible parameters and changes were investigated to overhaul the experimental pipeline. As mentioned early on, we decided that although time-consuming, the IEF fractionation method would provide the best sequencing depth. Additionally, we settled on the SILAC labeling technique, as in the field this seemed to be the most reliable and simple method for relative quantitation of samples. After implementation of various experimental changes, we monitored our progress, using MaxQuant, as we overhauled the experimental pipeline (Figure 2.6)

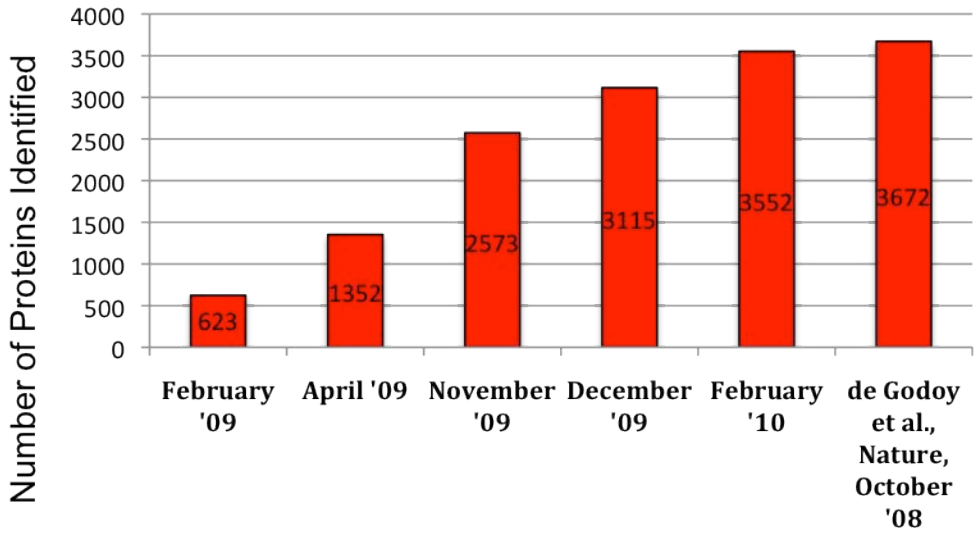


Figure 2.6

Successful implementation of a strategy for in-depth sequencing of the yeast proteome. A graph of the number of proteins identified in 24-fraction IEF experiments upon the implementation of changes to the experimental pipeline. Experiments were compared to a yeast proteome sequencing data set from Godoy et al. The number of proteins identified in the February '10 experiment indicated we had successfully optimized our experimental method. Lysates are SILAC labeled.

Over the course of a year-long effort we saw a significant increase in protein identifications and were finally able to reach our goal of yeast proteome sequencing at a similar level to the deepest sequencing that had been previously published[95]. In the end our method consisted of the following: milling and mixing of SILAC labeled “heavy” and “light” cells, followed by tryptic digestion and IEF fractionation of peptides into 24 samples, followed by the analysis of each sample on a three-hour gradient.

Discussion

Here we have demonstrated the successful implementation of a mass spectrometry-based method to perform accurate quantitative analysis of the yeast proteome. We have examined and optimized every step in the experimental process and have performed a complete overhaul of our previous methodologies. We have optimized our instrument and chromatography system to achieve a maximal sequencing speed while still sequencing with effective accuracy, sensitivity, and dynamic range. In terms of sample preparation our method now employs an isoelectric focusing-based fractionation technique in addition to SILAC, a method of labeling cells to produce 1:1 pairs of true peptide signals for relative quantification.

The applications and potential of this technology are tremendous. Using this technique, with minimal sample, one can identify several thousand proteins in the yeast proteome to look at the impacts of different treatments or mutations on the global proteome. We suspect that the sequencing depth we have achieved is sufficient and represents most of the expressed proteome, as all 6,000 yeast

proteins may not be expressed at once. However, although a promising and exciting experimental method this setup does have its limitations.

The two largest limitations of the experimental setup are cost and the long duration of sequencing time of about 72 hours. However, if maximal sequencing depth is not necessary this setup can be modified and used in conjunction with other fractionation techniques described earlier in the chapter. For instance, we applied this technology and utilized a SAX fractionation technique to sequence changes in the ubiquitin proteome in *cdc48* and *ubx* mutants, in order to look for potential new substrates in the Cdc48 network. These experiments, which only required 24 hours of sequencing time, are described in detail in the next chapter (Chapter 3). In addition, our experimental setup was modified and utilized in various other publications[121-123]. Essentially, with an optimized chromatography and mass spectrometric system and a knowledge and understanding of the various fractionation techniques, we now have the capability to customize methods according to experimental sequencing requirements. As mentioned, the possibilities are endless.

Despite these achievements and capabilities, quantitative proteomics by mass spectrometry is still a rapidly evolving field. Although one can do deep sequencing of the yeast proteome, our capabilities are still not as sensitive or do not have as much coverage as necessary for larger proteomes, such as human, which express up to three times more genes than yeast. Additionally, in yeast to have a more complete understanding of the proteome higher peptide sequencing coverage per protein is necessary. Sequencing of this caliber cannot be achieved with the

current technology. However, attaining this goal seems to be the trajectory of the field. New instruments, such as the LTQ Orbitrap Velos, are pushing the limits of sequencing depth and new hybrid approaches of targeted and shotgun proteomics are making larger proteome sequencing much more achievable in the not-so-distant future.

Chapter 3:

Perturbations to the Ubiquitin Conjugate Proteome in *Δubx* Mutants Identify Ubx2 as a Regulator of Membrane Lipid Composition*

*This chapter, currently under review for publication, was written by Natalie Kolawa, Michael J. Sweredoski, Robert L.J. Graham, Robert Oania, Sonja Hess, and Raymond J. Deshaies.

Abstract

Yeast Cdc48 (p97/VCP in human cells) is a hexameric AAA ATPase that is thought to transduce the chemical energy of ATP hydrolysis into mechanical force that can be used to segregate ubiquitin-conjugated proteins from tightly-bound partners. Current models posit that the mechanochemical activity of Cdc48 is linked to its substrates through adaptor proteins, including a family of seven proteins (13 in human) that contain a Cdc48-binding UBX domain. However, few substrates for specific UBX proteins are known, and hence the generality of this hypothesis remains largely untested. Here we report comprehensive identification of ubiquitin conjugates that accumulate in *cdc48* and *ubx* mutants using quantitative mass spectrometry. Different *ubx* mutants exhibit reproducible patterns of conjugate accumulation that differ greatly from each other and that point to unexpected functional specializations of individual Ubx proteins. To validate our mass spectrometry findings, we examined in detail the endoplasmic reticulum-bound transcription factor Spt23, which we identified as a putative Ubx2 substrate. Mutant *ubx2Δ* cells are deficient in both cleaving the ubiquitinated 120 kD precursor of Spt23 to form active p90 and in localizing p90 to the nucleus. Consistent with reduced production of active p90, *ubx2Δ* cells are deficient in expression of its target gene *OLE1*, a fatty acid desaturase. Our findings illustrate the utility of proteomics to identify ligands for specific ubiquitin receptor pathways and uncover Ubx2 as a key player in the regulation of membrane lipid biosynthesis.

Introduction

Cdc48/p97 is an essential, highly abundant member of the AAA (ATPase associated with various cellular activities) protein family. Cdc48 has been linked to numerous functions throughout the cell but is best known for its critical role in ERAD (endoplasmic reticulum associated protein degradation), which occurs via the ubiquitin proteasome system (UPS). It is also involved in cell-cycle progression, homotypic membrane fusion, chromatin remodeling, autophagy, and transcriptional and metabolic regulation[26-29]. Cdc48 has been the subject of much attention over the last few years due to its causal links to amyotrophic lateral sclerosis (ALS) and inclusion body myopathy, Paget's disease of the bone, and frontotemporal dementia (IBMPFD) as well as its implied role in a variety of diseases including cancer[30-34]. As a result, p97 has been the target of multiple drug development efforts[35, 36].

Cdc48/p97 interacts with a large number of putative substrate adaptors and co-factors, including a family of proteins (seven in yeast, thirteen in human cells) that contain an UBX domain[41, 42]. The UBX domain binds to the N-terminal region of p97 and proteins bearing this domain have been suggested to serve as interchangeable adaptors that target Cdc48/p97 to specific substrates. Although the functions and mechanism of action of Cdc48/p97 remain poorly understood, it is generally presumed that it uses ATP hydrolysis to fuel the extraction of ubiquitinated proteins from multi-subunit complexes or membranes as a prelude to their degradation by the proteasome. Cdc48/p97 may also remodel protein:protein and protein:nucleic acid complexes in a manner that is not coupled to either prior ubiquitination[124] or subsequent degradation by the proteasome[44, 125].

Based on the abundance of Cdc48/p97 and the complexity of the network of adaptor proteins for which it serves as the hub, Cdc48/p97 has the potential to exert a profound influence on the UPS. However, the number of known substrates of Cdc48/p97 remains relatively small and smaller still is the number of substrates that have been linked to a specific UBX domain protein. Indeed, only a handful of specific cases are known and in half of them proteasomal degradation is not the final outcome[44, 51-53, 125-128]. To understand why there is such a profusion of UBX domain proteins in cells it will be important to determine what these proteins do to their substrates, which will require knowing what their substrates are. Therefore, a major goal of this work is to enable future investigations into the function and regulation of UBX proteins by assembling a catalog of candidate substrates/targets.

A key paradigm that guides our current understanding of Cdc48 function emerged from studies on the processing of the transcription factors Spt23 and Mga2 [73, 75, 83]. These seminal studies revealed the mechanism by which cells control their ratio of saturated to unsaturated fatty acids to maintain an appropriate lipid composition in cellular membranes. Central to this regulation are the transcription factors Spt23 and Mga2 which are released from the endoplasmic reticulum (ER) membrane so that they can translocate to the nucleus and activate expression of *OLE1* which encodes the stearoyl-Δ9 desaturase that governs the conversion of saturated fatty acids (SFAs) to unsaturated fatty acids (UFAs)[71, 72]. Spt23 and Mga2 are initially produced as 120 kD precursors (p120) embedded in the ER by C-terminal signal/anchor domains, such that the bulk of each protein projects into the cytosol. The p120 form of each protein is ubiquitinated by Rsp5 and cleaved by the proteasome, which yields a p90 form that

lacks the transmembrane signal/anchor [73, 76, 77, 80, 83]. However, p90 remains tethered to the ER through dimerization to an uncleaved p120, until it is disengaged from its partner by the 'segregase' activity of Cdc48, acting in concert with Ufd1–Npl4[75, 78, 79]. The released p90 can then travel to the nucleus where it activates expression of *OLE1*[75]. The p90 species are quite unstable and Cdc48 also promotes their degradation[129]. In addition to acting as a transcription factor, Mga2 also influences the stability of *OLE1* mRNA[130], although the degree to which this function is carried out by p90 or p120 is not understood.

Methods

Yeast Strains and Growth Conditions

Strains, plasmids, and primers used in this study are described in Supplemental Tables 3.1-3.3. All yeast strains are derivates of either the wild-type strain RJD 4614 (S288C background from the Open Biosystems Yeast Knockout Library) or the wild-type strain RJD 360 (W303 background). Standard genetic techniques were used. Unless otherwise stated strains were grown at 30°C and cultured on YPD.

SILAC Labeling of Cells

For SILAC experiments strains auxotrophic for lysine and arginine were used. Cells expressing ^{His8}ubiquitin were grown in complete synthetic medium with 2% dextrose containing 20 mg/L lysine and arginine or in "heavy" medium with 20 mg/L $^{13}C_6{}^{15}N_2$-lysine and $^{13}C_6$-arginine (Cambridge Isotope Laboratories). All mutant strains with the exception of *cdc48-3* and *ubx2Δ* expressed ^{His8}ubiquitin and endogenous ubiquitin in a 1:1 ratio. *Cdc48-3* and *ubx2Δ cells* expressed lower levels of ^{His8}ubiquitin, which we corrected for during quantification. Cells were grown for 10 generations to log-phase (OD_{600} 1.0-2.0) and equal amounts of the "light" and "heavy" -labeled cells (as determined by OD_{600}) were mixed, harvested, washed, and flash frozen.

Purification of Ub Conjugates

Purifications were done similarly to the protocol described previously [131]. 200 OD_{600} units of cells were lysed in urea buffer (8M urea, 300mM NaCl, 100mM Na_2PO_4, 10mM Tris-HCl pH 8.0, 0.2% Triton X-100, 20mM imidazole, and 5mM N-ethylmaleimide (NEM)) by vortexing with glass beads. Ubiquitinated proteins were purified by the addition of nickel-NTA bead slurry (50µl beads/5-10mg of lysate) to clarified lysate (13,200 rpm for 15 minutes in an Eppendorf Centrifuge 5417R) and mixed at room temperature for 90 minutes. Beads were washed three times with 20 bed volumes of buffer. A two-step tryptic digestion was performed directly on proteins bound to the beads [121].

Fractionation and LC-MS/MS

The tryptic peptides were desalted on a C18 macrotrap (Michrom Bioresources) and concentrated *in vacuo*. Dried samples were resuspended and subjected to StageTip-based strong anionic exchange (SAX) as previously described [99]. Samples were eluted, concentrated, and then acidified prior to mass spectrometric analysis. Mass spectrometry experiments were performed on an EASY-nLC (Thermo Scientific) connected to a hybrid LTQ-Orbitrap Classic with a nanoelectrospray ion source (Thermo Scientific) in a setup and settings [115] very similar to those previously described [116]. Binding and separation of the peptides took place on a 15cm silica analytical column (75 µm ID) packed in-house with reversed phase ReproSil-Pur C_{18AQ} 3 µm resin (Dr Maisch GmbH, Ammerbuch-Entringen, Germany). Samples were run for 240 minutes on a 5% to 25% acetonitrile gradient in 0.2% formic acid at a flow rate of 350 nL per minute. The

mass spectrometer was programmed to acquire data in a data-dependent mode, automatically switching between full-scan MS and tandem MS acquisition. Survey full scan MS spectra (from m/z 300 to 1,700) were acquired in the Orbitrap after the accumulation of 500,000 ions, with a resolution of 60,000 at 400 m/z. The ten most intense ions were sequentially isolated and, after the accumulation of 5,000 ions, fragmented in the linear ion trap by CID (collisional energy 35% and isolation width 2 Da). Precursor ion charge state screening was enabled and singly charged and unassigned charge states were rejected. The dynamic exclusion list was set for a 90s maximum retention time, a relative mass window of 10 ppm, and early expiration was enabled.

Mass Spec Analysis and Quantification:

Thermo raw data files were analyzed by MaxQuant (v 1.3.0.5)[100] and were searched against the SGD yeast database (5911 sequences) and an in-house contaminant database (259 sequences) including human keratins and proteases. All default options were used except as follows: match between runs was enabled (2 min. maximum), GlyGly on Lys (+114.042927), and multiplicity of 2 with heavy labels Arg6 (+6.020129) and Lys8 (+8.014199). Tryptic digest was specified with up to two missed cleavages. Peptide, protein, and site false discovery rates were fixed at 1% using the target-decoy approach with a reversed database[100]. Further data processing was performed to calculate ratios and standard errors of the ratios for each mutant using in-house scripts described previously [132]. Briefly, hierarchical models were constructed for each mutant where the overall ratio for each protein is

the geometric mean of the biological replicates and the biological replicate ratio is the median of all of the peptide ratios in the replicate. The standard error of the overall protein ratio is calculated by estimating the global peptide ratio standard error using pooled variance (calculated separately for peptide ratios based on requantified isotopic patterns) and using a bootstrapping procedure to resample at each level in the hierarchical model.

All quantified proteins were used to construct Figures 3.1B and Supplemental Figures 3.2A,B. Proteins for which the probability that they changed more than 10% in the WT:WT control experiment was >99% or were not significantly enriched (p-value ≥ 0.05) in the ^{His8}Ub expressed cells in the untagged/tagged control experiment were filtered out due to the high likelihood that they would yield false positive identifications in the mutant:wild-type comparisons. Analyses relating to comparisons with SCUD and the response of the ubiquitination enzymes in the *cdc48-3* mutant used all proteins that were not filtered out. All proteins that were not filtered out and were quantified in all mutants were used in the remaining bioinformatic analyses. In Figures 3.1C and D, thresholds for significance were greater than 10% change from 1:1 (p-value < 0.01) (i.e., 99% likely to have changed more than 10%). Yeast GO-Slim annotations for Figures 3.2A and B were downloaded from the SGD. Proteins considered significantly enriched had a more than 10% change from 1:1 (p-value < 0.05). Annotations for Supplemental Figure 3.3A and Supplemental Figure 3.4A-E were downloaded from DAVID [133, 134]. Displayed terms were deemed significantly different at a p-value < 0.05 (e.g., the 25th percentile of the ratios of the proteins annotated with the term was

significantly greater than random) after correcting for multiple hypotheses testing using the Benjamini and Hochberg method[135].

Confirmation of Ub Conjugate Accumulation for Spt23

Ubiquitin conjugates were purified from 100 OD_{600} units of cells expressing tagged ubiquitin as described in the purification of Ub conjugates section. However, after washing beads an equal volume of 2X SDS-PAGE buffer was added prior to boiling the beads for 5 minutes to elute conjugates. Boiled aliquots were resolved by SDS-PAGE and immunoblotted for endogenously tagged $^{myc}Spt23^{V5}$.

Confirmation of Ub Conjugate Accumulation for Mga2

Cells (100 OD_{600} units) were disrupted using a FastPrep-24 in lysis buffer containing 50mM Tris-HCl (pH 7.5), 150mM NaCl, 1mM EDTA, 10mM NEM, 0.5mM AEBSF and Protease Inhibitor Cocktail Tablet (Roche). Lysates were clarified by centrifugation (14,200 rpm for 20 minutes on an Eppendorf Centrifuge 5417R) and bound to 30µL of TUBE2 agarose beads (Boston Biochem)/5-10 mg lysate for 2-3 hours with rotation at 4°C. Beads were washed 3 times with lysis buffer and an equal volume of 2X SDS-PAGE buffer was added prior to boiling the beads for 5 minutes to elute conjugates. Boiled aliquots were resolved by SDS-PAGE and immunoblotted.

Quantitative Reverse Transcription PCR

RNA was isolated from 3 OD_{600} units of cells using Trizol (Invitrogen) according to the manufacturer's recommendations. cDNA was prepared using the Superscript first strand synthesis kit (Invitrogen) and quantitative PCR was performed using the SYBR GreenER super mix (Invitrogen) with primers described in Supplemental Table 5.

Immunoblot Analysis

Cell pellets were first boiled for 3 minutes and then after the addition of SDS-PAGE buffer supplemented with 5 mM NEM and glass beads, lysed by vortexing in a Fast Prep-24 (MP) for 1 minute at a setting of 6.5 and boiled again for 3 minutes. Boiled lysates were centrifuged at 16,000 x g for 5 minutes. Aliquots were resolved by SDS-PAGE, transferred to nitrocellulose, and stained with Ponceau S to determine equivalent loading of protein extracts. The nitrocellulose filters were immunoblotted with the desired antibody and developed by ECL or SuperSignal (Invitrogen). For detecting $^{myc}Spt23^{V5}$, $^{myc}Spt23^{HA}$, and $^{myc}Mga2$, anti-myc (Covance), anti-V5 (Invitrogen), and anti-HA (16B12, Roche) antibodies were used at a dilution of 1:3000. Other immunoblots were performed with antibodies specific for Ubiquitin (Stressgen Biotechnologies) dilution 1:1000, Dpm1 (Invitrogen) dilution 1:3000, tubulin (Santa Cruz) dilution 1:30000, TAP tag (Thermo Scientific) dilution 1:3000, Rsp5 (gift from Linda Hicke) dilution 1:3000, Cdc48 (gift from Thomas Sommer and Ernst Jarosch) dilution 1:1000, and Ufd1 (custom polyclonal antibody from Covance) dilution 1:10000. For Ubx2 domain analysis experiments mutants were derived from strains gifted by Chao-Wen Wang.

Growth Assays

For plating assays cells were grown in YPD or SRaffinose-URA and diluted to OD_{600} of 0.3 in water. Serial five-fold dilutions were prepared in water and spotted onto YPD +/- 0.2% oleic acid or minimal medium plates supplemented with various additives as described in the text. Plates were incubated at 30°C or 37°C for 2-4 days.

Turnover of Spt23

RJD 6138, 6139, and 6167 cells were grown in YPD to an OD_{600} ~1.0, at which point 100μg/ml cycloheximide was added to initiate a chase. Samples were taken at the timepoints indicated. To monitor Spt23 following the switch from medium containing oleic acid to medium lacking oleic acid, cells were grown in YPD+0.2% oleic acid, washed in water, and resuspended in YPD supplemented with cycloheximide. Proteins were extracted as described in the Immunoblot Analysis section with boiling SDS-PAGE buffer supplemented with 5mM NEM. Lysates were resolved by SDS-PAGE and immunoblotted.

Immunoprecipitation of Ubx2

RJD 3166, 6172, 6173, and 4614 were grown in YP plus 2% galactose (to induce expression of ^{myc}Spt23^{HA} from the *GAL* promoter) to an OD_{600} ~1.0, harvested, and flash-frozen in liquid nitrogen. Cells were lysed by grinding in liquid nitrogen. Frozen cell powder was resuspended in lysis buffer containing 50mM Tris

(pH 7.5), 100mM NaCl, 1mM EDTA, 5mM $Mg(OAc)_2$, 0.2% Triton X-100, 5mM NEM, and Protease Inhibitor Cocktail Tablet (Roche), clarified by centrifugation (14,200 rpm for 20 minutes on an Eppendorf Centrifuge 5417R) and bound to rabbit IgG conjugated to superparamagnetic Dynabeads (Invitrogen) for three hours with mixing at 4°C. Beads were washed 3 times with lysis buffer and bound proteins were eluted by boiling in SDS-PAGE buffer and resolved by SDS-PAGE prior to immunoblotting.

Immunofluoresence

RJD 6172 and 6182 cells were grown in YP plus 2% galactose (to induce expression of ^{myc}Spt23^{HA} from the *GAL* promoter) to an OD_{600} ~0.5-1.0. Cell cultures were adjusted to 4% formaldehyde and incubated for 15 minutes before cells were collected, washed, and spheroplasted by treatment with 50 units of lyticase (Sigma) for 10 minutes. Spheroplasts were washed and placed on ConA-coated glass slides. Cells were fixed with methanol and acetone washes. Cells were treated with blocking buffer (2% BSA, 0.1% Tween20 in PBS) and incubated with a primary antibody overnight at a dilution of 1:1000. FITC-conjugated secondary antibody was added and cells were incubated for another hour followed by DNA staining with DAPI. Mounting media was placed on cells and images were taken with a Zeiss confocal microscope with a 100x objective.

Results

Identification and Quantification of Changes in the Ubiquitin Proteome

To better understand the breadth of function of Cdc48 and its adaptors that contain an UBX domain, we sought to identify ubiquitin conjugates whose levels increased in yeast cells deficient in these components. The rationale underlying this approach is based on the observation that ubiquitinated Rpb1 substrate accumulates in *ubx5Δ* and *cdc48-3* cells[52]. Wild-type and mutant yeast cells expressing His8-tagged ubiquitin were labeled with 'heavy' and 'light' lysine and arginine, respectively, and immediately prior to harvesting and lysis the cultures were mixed. Ubiquitin conjugates were purified on Ni^{2+}-NTA magnetic beads under denaturing conditions (8M urea) and digested with trypsin. The resulting peptides were subjected to strong anionic exchange (SAX) fractionation prior to sequencing in a mass spectrometer (Figure 3.1A). We performed this analysis in triplicate with *cdc48-3* and in duplicate with six of the seven *ubxΔ* mutant strains (*ubx2Δ*-*ubx7Δ*; *shp1Δ*/*ubx1Δ* was not analyzed because it grew so poorly).

To evaluate the reproducibility of our methodology we compared biological replicates (Supplemental Figures 3.1A,B), which revealed that our datasets were highly correlated and demonstrated low coefficients of variation. To improve further the quality of our proteomic dataset, we performed two separate control experiments, the results of which were used to computationally filter our data prior to further analysis. In the first experiment both the 'light' and 'heavy'-labeled cultures were wild-type. This analysis yielded a narrow distribution of heavy:light ratios centered on 2^0, confirming the significance of the broader ratio distributions seen in the wild-type:mutant comparisons (Figure 3.1B, Supplemental Figure 3.2A). Proteins that exhibited ratios that deviated from 2^0 by >10% (p-value <0.01) in this control experiment were excluded from our overall dataset. In the second control, we compared untagged, light-labeled cells and ^{His8}Ub-

expressing heavy-labeled cells. Proteins that were significantly enriched in the ^{His8}Ub-expressing cells (p-value<0.05) were deemed to be derived from valid ubiquitin conjugates. After computational filtration, our dataset contained 1,916 ratios for ubiquitinated proteins identified and quantified in at least one mutant experiment and 1,733 ubiquitinated proteins across all experiments (all proteins identified are listed in Supplementary Table 1). We additionally identified 67 di-glycine peptide signatures, which are indicative of ubiquitination sites; 25 of the 67 were previously identified (Supplementary Table 2). The proteins in this study covered over 50% of the previously identified ubiquitinated proteins in the *Saccharomyces cerevisiae* ubiquitin database (SCUD, http://scud.kaist.ac.kr/) and included 1,226 proteins not previously listed (Supplemental Figure 3.2B). These metrics suggest that we have sampled a substantial portion of the ubiquitinated proteome.

UBX Proteins Target Different Subsets of Ubiquitin Conjugates

From the initial screen it was evident that *cdc48-3* and *ubxΔ* mutations had wide-ranging and variable effects on the ubiquitin proteome. Comparison of the ratio distributions of the different experiments underscores this conclusion (Figure 3.1B). Consistent with Cdc48 being the hub of its eponymous network, *cdc48-3* had the greatest effect on the ubiquitinated proteome. 26% of ubiquitin conjugates were detected at higher levels in *cdc48-3* compared to wild-type cells whereas 17% were decreased in amount. We do not know the reason for the latter, but it could be due to reduced expression of the protein or increased competition for free ubiquitin because total conjugated ubiquitin increases substantially in *cdc48-3* cells[52]. Among the *ubxΔ* mutants *ubx5Δ* and *ubx6Δ* caused the strongest perturbations; by contrast *ubx3Δ* had the least effect, followed by *ubx4Δ*. To evaluate the impact of the *cdc48-3* and *ubxΔ* mutations with more granularity, we

generated a heatmap that captures the behavior of the 1,733 conjugates quantified in all experiments in our dataset (Figure 3.1C). Three important points are evident in this graphic. First, *ubx6Δ* and *ubx7Δ* displayed remarkable overlap with over 85% of the ubiquitin conjugates accumulating in *ubx7Δ* also accumulating in *ubx6Δ*. Very little is known about Ubx6 and Ubx7, but this result is consistent with their close homology and colocalization at the perinuclear membrane[55]. Second, *cdc48-3* mutants accumulated a set of ubiquitin conjugates (128 proteins) that did not accumulate in any single *ubxΔ* mutant. This suggests that many substrates can engage Cdc48 without an Ubx protein, and may explain why a relatively low fraction of Cdc48/p97 substrates have been shown to be dependent on an UBX domain co-factor. Third, with the exception of *ubx6Δ* and *ubx7Δ*, the different mutations had markedly different effects on the ubiquitin conjugate proteome. This conclusion was also supported by an independent analysis (Figure 3.1D), which indicated that for each *ubxΔ* mutant except *ubx6Δ* and *ubx7Δ*, >40% of the accumulating conjugates were elevated only in that mutant and up to ~90% were elevated in that mutant and at most one other mutant. This supports the hypothesis that individual Ubx proteins target distinct sets of ubiquitin conjugates, which is consistent with their proposed role as substrate adaptors for the Cdc48 engine. Remarkably, of the 1,733 proteins in our dataset, 62% were elevated significantly in at least one mutant (62% was highly significant in comparison to the null hypothesis where peptide ratios are randomly assigned to proteins (p-value < 1e-100)), suggesting that the Cdc48 network has a profound impact on the UPS that is greater than has been previously appreciated.

Ubiquitin Conjugates that are Elevated in Different *ubxΔ* Mutants Map to Diverse Cellular Components and Cellular Processes

To understand the subcellular localization and functions of the ubiquitinated species identified in our analyses, we evaluated Yeast GO-Slim cell components and GO-Slim biological process annotations (Figures 3.2A and 3.2B). Many different trends were apparent from this analysis. Ubiquitin conjugates that were elevated in *cdc48-3* cells (referred to hereafter as '*cdc48-3* conjugates') encompassed numerous known targets including Mga2 and Spt23[75, 76], the HMG-CoA reductases Hmg1 and Hmg2[136], and the Rpb1 subunit of RNA Polymerase II[52]. Consistent with the role of Cdc48 in ERAD, *cdc48-3* conjugates were significantly enriched for proteins of the endomembrane system including the ERP and ERV family of proteins involved in ER to Golgi transport, members of the GPI family of proteins, subunits of the Golgi mannosyltransferase complex, and all seven proteins of the PMT (Protein o-mannosyltransferase) family. The *cdc48-3* conjugates were also enriched for glycosylated proteins, enzymes of lipid and carbohydrate metabolism, and proteins involved in nuclear pore complex, nuclear organization, stress response, and the ubiquitin-proteasome system (Supplemental Figure 3.3A). Further investigation of the latter category revealed that *cdc48-3* cells contained elevated levels of ubiquitin conjugated species of multiple E3 ubiquitin ligases and F-Box proteins (Supplemental Figure 3.3B). This is consistent with recent identification of Cdc48 as an SCF^{Met30} disassembly factor[137] and suggests that Cdc48 may be more widely involved in extraction and/or degradation of ubiquitinated F-box proteins from SCF complexes.

Analysis of the ubiquitin conjugates that were elevated in *ubx2Δ* relative to wild-type cells (i.e. '*ubx2Δ* conjugates') indicated there was a significant enrichment of vacuolar and metabolic transport proteins, including secretory proteins (Figure 3.2A-C). Many of these proteins are synthesized in the ER where Ubx2 is localized. Notably, elevated levels of

ubiquitinated forms of a number of ER proteins were detected in both *ubx2Δ* and *cdc48-3* cells, including members of the ergosterol biosynthesis pathway and the *OLE1* transcription factors Spt23 and Mga2, indicating a possible role for Ubx2–Cdc48 in regulating lipid metabolism.

Deletion of *UBX3* had little impact on the ubiquitin conjugate proteome with no proteins changing consistently in a significant manner. The pool of *ubx4Δ* conjugates was likewise small, but was enriched for ER proteins involved in lipid metabolism including Spt23, Mga2, Hmg1, and Hmg2 (Figure 3.2A,B, Supplemental Figure 3.4A,B).

Ubiquitin conjugates that were elevated in *ubx5Δ* cells were enriched for vacuolar and plasma membrane proteins, including proteins involved in ion and amino acid transport such as the TPO family of proteins involved in polyamine transport (Figure 3.2A,B, Supplemental Figure 3.4C). Ubx5 contains an UIM domain that links it to Rub1-conjugated cullin–RING ubiquitin ligases (CRLs)[54, 138], raising the possibility that a CRL regulates trafficking of membrane proteins in yeast

Lastly, ubiquitin conjugates that were elevated in *ubx6Δ* and *ubx7Δ* were highly enriched for nucleolar and ribosomal proteins, including proteins involved in ribosome assembly (Figure 3.2A,B, Supplemental Figure 3.4D,E). In agreement with a nuclear and nucleolar function, Ubx6 and Ubx7 localize to the nuclear periphery and nucleus[55].

Ubiquitinated Spt23 and Mga2 Accumulate in *ubx2Δ*

The increase in ubiquitinated Spt23 and Mga2 species in *cdc48-3*, *ubx2Δ* and *ubx4Δ* cells (Figure 3.3A) piqued our interest because although it was already known that Cdc48–Ufd1–Npl4 promotes release of Spt23 and Mga2 from the ER membrane and ubiquitinated forms of these transcription factors accumulate in *cdc48*, *npl4*, and *ufd1* cells[75, 76, 78,

83], Ubx proteins had not previously been implicated in this pathway. To evaluate the ubiquitination status of Spt23 in *ubx2Δ* and *ubx4Δ* mutants, we first generated a yeast strain in which endogenous *SPT23* was modified to express proteins bearing an N-terminal myc tag and C-terminal V5 tag (^{myc}Spt23^{V5}). We then purified ubiquitin conjugates from ^{myc}*SPT23*^{V5} strains expressing His8-tagged ubiquitin with Ni^{2+}-NTA magnetic beads. Immunoblotting for tagged Spt23 revealed robust accumulation of high molecular weight conjugates in both *ubx2Δ* and *ubx4Δ* as compared to wild-type (Figure 3.3B). For reasons we do not understand, overexpression of ^{His8}Ub had opposing effects on the accumulation of ubiquitin conjugates on ^{myc}Spt23^{V5} in *ubx2Δ* and *ubx4Δ*. In the absence of ^{His8}Ub, we consistently observed stronger accumulation of ubiquitinated ^{myc}Spt23^{V5} in *ubx2Δ* (e.g. Figure 3.5A, B). To evaluate the ubiquitination status of Mga2 in *ubx2Δ* and *ubx4Δ* mutants, we used Tandem Ubiquitin Binding Entity (TUBE2) beads to purify endogenous ubiquitin conjugates from strains transformed with a plasmid that expressed Mga2 with a myc tag appended to its N-terminus (^{myc}Mga2). Immunoblotting for ^{myc}Mga2 revealed an accumulation of high molecular weight conjugates in *ubx2Δ* cells compared to wild-type (Supplemental Figure 3.5). By contrast, few if any high MW ubiquitin conjugates of Mga2 were detected in *ubx4Δ* cells.

Ubx2 is a Component of the *OLE1* Pathway

In rapidly growing cells that are producing lipids for membrane assembly, the transmembrane p120 forms of Spt23 and Mga2 are ubiquitinated by the ubiquitin ligase Rsp5 and cleaved by the proteasome[73]. Subsequently, Cdc48–Ufd1–Npl4 promotes release of cleaved p90 forms of Spt23 and Mga2 from the ER membrane, allowing these proteins to translocate to the nucleus where they activate transcription of *OLE1*, which

encodes the essential enzyme that generates UFAs[71]. The accumulation of ubiquitin-conjugated Spt23 and Mga2 in *ubx2Δ* raised the question of whether Ubx2 also functions in this pathway. To address this question we first tested for genetic interactions between *rsp5-1* and *ubxΔ* mutants. Similar to the previously reported synthetic lethality of *cdc48* and *ufd1* with *rsp5*[75], *ubx2Δ rsp5-1* cells failed to grow on YPD at 30°C but were rescued by addition of the UFA oleic acid (Figure 3.4A). By contrast, *ubx4Δ rsp5-1* did not exhibit synthetic lethality. To test more directly the hypothesis that Ubx2 promotes expression of *OLE1*, we measured *OLE1* mRNA levels in mutants grown in YPD. During normal log-phase growth in YPD, *ubx2Δ* cells showed a five-fold reduction in *OLE1* mRNA (Figure 3.4B). Moreover, as was previously reported for *rsp5* mutants[73], both *ubx2Δ* and *rsp5Δ* cells failed to accumulate *OLE1* mRNA upon being shifted from a medium containing oleic acid to a medium lacking UFAs (Figure 3.4C). By contrast, *ubx4Δ* was not defective in *OLE1* regulation and did not exacerbate the effect of *ubx2Δ* (Figures 3.4B,C).

The results above suggested that Ubx2 is involved in activation of Spt23 and Mga2. As further confirmation we examined the effect of Spt23 and Mga2 overexpression in mutant cells (Figure 3.4D). Although UFAs are required for survival, the excess production of oleic acid that occurs upon hyperactivation of *OLE1* expression is toxic[139]. Whereas overexpression of Spt23 or Mga2 from a galactose-inducible promoter inhibited growth of wild-type cells, *ubx2Δ* cells, like *rsp5-1* cells[140], were more tolerant (Figure 3.4D) even though they equally overexpressed these *OLE1* activators (data not shown).

Ubx2 Regulates Processing of Spt23

We next wanted to determine if Ubx2 is involved in processing of the *OLE1* activators from their p120 precursor forms to their p90 active forms. For these

experiments we focused on Spt23 because we found it difficult to consistently detect tagged Mga2 expressed from its endogenous locus. We examined steady state levels of ^{myc}Spt23^{V5} in YPD-grown cells deficient in various components of the Cdc48 and proteasome pathways (Figure 3.5A). As anticipated from prior results[75], the *cim3-1* and *rsp5-1* mutants exhibited a severe defect in p90 accumulation (Figure 3.5A, B; *cim3-1* is a temperature-sensitive allele of the gene that encodes proteasome subunit Rpt6). Notably, p90 formation was also significantly impaired in *cdc48-3* and *ufd1-2* strains. In *ubx2Δ* (Figure 3.5A, B), but not in other *ubxΔ* mutants (Figure 3.5B), ubiquitinated high MW forms of ^{myc}Spt23^{V5} accumulated and the p90:p120 ratio was significantly reduced. However, the degree to which processing to p90 was defective varied in different experiments. In the prevailing model, Cdc48–Ufd1–Npl4 acts after the proteasome to promote release of p90 from the ER membrane[39, 75, 78]. However, it has also been reported that Cdc48–Ufd1–Npl4 is required for processing of p120 to p90[83]. Our data are consistent with those of Hitchcock et al. and suggest that Cdc48–Ufd1 and Ubx2 contribute to efficient formation of p90.

After establishing a role for Ubx2 in Spt23 processing, we next investigated the contribution of its UBA and UBX domains to this reaction. We immunoblotted for ^{myc}Spt23^{V5} in *ubx2Δ* strains in which full-length and various domain mutants of Ubx2 were integrated at the *LEU2* locus under control of the *UBX2* promoter[46] (Figure 3.5C). This experiment confirmed that the defect in Spt23 processing in *ubx2Δ* cells was due to lack of Ubx2, and suggested that the UBX domain played a more important role in processing than the UBA domain. RT-PCR analysis of these same strains grown in oleic acid and switched to

oleic acid-free medium revealed a large decrease in *OLE1* mRNA induction in the UBAΔ UBXΔ mutant but not the single mutants (data not shown).

Ubx2 is orthologous to human UBXD8 which may directly sense UFAs[141, 142]. We therefore tested whether Ubx2 was required to sense UFAs in yeast. To address this question we grew wild-type and mutant cells with and without oleic acid and evaluated processing of ^{myc}Spt23^{V5}. We reasoned if Ubx2 is required to sense UFAs that the residual level of p90 formed in *ubx2Δ* would be insensitive to their presence. Processing of ^{myc}Spt23^{V5} in wild type cells was strongly repressed by addition of oleic acid (Figure 3.5D). Likewise, the residual formation of p90 observed in *cim3-1*, *cdc48-3*, and *ubx2Δ* cells remained sensitive to inhibition by oleic acid. Consistent with this, our analysis of *OLE1* mRNA revealed that although its levels were severely reduced in *ubx2Δ* and *ubx2Δ ubx4Δ*, they were diminished further upon addition of oleic acid (Figure 3.4C). Taken together, these data indicate that there must be at least one UFA sensor that remained in *ubx2Δ* cells. Careful inspection of the modification state of ^{myc}Spt23^{V5} in *cdc48-3* and *ubx2Δ* cells grown with or without oleic acid suggested that sensing of UFAs occurred at or before Rsp5-dependent ubiquitination (Figure 3.5D).

To investigate in greater depth the role of Ubx2 in Spt23 processing, we performed cycloheximide chase experiments (Figure 3.5E). The p90 form of ^{myc}Spt23^{V5} was rapidly degraded in wild-type but was modestly stabilized in *ubx2Δ* and strongly stabilized in *cdc48-3* cells. In contrast, p120 was more stable than p90 in all conditions whereas the high MW ubiquitin conjugates that accumulated in *ubx2Δ* and *cdc48-3* were rapidly metabolized. Given that: (i) p90 was unstable in wild type, (ii) more than half of ^{myc}Spt23^{V5} detected at zero-time in wild-type was processed to p90, and (iii) conversion of p120 to p90 was very

slow and inefficient, we suggest that in wild-type cells p120 became conjugated with ubiquitin and rapidly processed to p90 either during or very shortly after completion of synthesis, such that newly-synthesized molecules that escaped modification were largely refractory to subsequent processing.

Ubx2 Associates with Spt23 and affects its Subcellular Localization

To test whether the effects of *ubx2Δ* on Spt23 were likely to be direct, TAP-tagged Ubx2 was immunoprecipitated from cells and immunoblotted for associated proteins (Figure 3.6A). This experiment revealed a direct interaction between Ubx2 and unmodified p120 as well as higher MW ubiquitinated ^{myc}Spt23^{HA}. Additionally, as expected[129], Cdc48 and Ufd1 were detected at levels above background in the Ubx2^{TAP} immunoprecipitate, as well as Rsp5, which had not previously been shown to interact with Ubx2.

To address the effect of *ubx2Δ* on subcellular localization of Spt23, we did immunofluoresence on cells that expressed a pulse of ^{myc}Spt23 from the *GAL1* promoter (attempts to detect ^{myc}Spt23^{V5} expressed from the natural locus were unsuccessful). Consistent with our observation that Ubx2 was required for *OLE1* expression, ^{myc}Spt23 demonstrated a decrease in localization to the nucleus in most *ubx2Δ* cells (Figure 3.6B). Additionally, in ~50% of *ubx2Δ* cells, Spt23 was observed to accumulate in cytosolic punctae (Figure 3.6C). Although we do not understand their exact molecular nature, the existence of these punctae in *ubx2Δ* together with poor nuclear accumulation point to a role for Ubx2 in proper localization of Spt23.

Discussion

The Cdc48 network plays a broad role in shaping the ubiquitin conjugate proteome.

To address globally the impact of Cdc48 and its Ubx adaptors on the UPS, we evaluated changes to the ubiquitin conjugate proteome in yeast cells deleted for individual *UBX* genes or carrying a temperature-sensitive *cdc48-3* allele. Our findings support a broad role for the Cdc48 network in the UPS. Overall, we identified 1,916 putative ubiquitin conjugates and 1,733 were quantified in all experiments. Of the 1,733, over 62% were found at statistically significant elevated levels relative to wild-type in at least one of the mutant strains. Mutation of different *UBX* genes led to markedly distinct effects on the ubiquitin conjugate proteome. For example, our findings suggested novel and unexpected roles for Ubx2, Ubx5, and Ubx6/7 in regulation of vacuolar, plasma membrane, and nucleolar/ribosomal proteins, respectively.

Our approach assumes that ubiquitin conjugates that are elevated in a particular mutant relative to wild-type represent intermediates that accumulate upon attenuation of Cdc48 activity or deletion of the adaptor that normally links them to Cdc48. However it should be borne in mind that increased levels of a conjugate may arise from increased abundance of the protein. Of interest, numerous ubiquitin conjugates were selectively depleted in specific *ubxΔ* strains. This was not investigated further, but during the course of this work it was reported that the human UBX protein SAKS1 protects ERAD substrates from deubiquitination[143] and overproduced UBXD7 binds to and increases the fraction of Nedd8-conjugated Cul2 in human cells[54, 138]. Hence, the large number of conjugate depletion events that we observed may represent species

whose ubiquitin-conjugated forms are stabilized upon binding of Ubx proteins. Our extensive dataset provide a rich resource for future investigations into this and other questions related to the functions and regulation of the Cdc48 network.

Ubx2 and Cdc48 promote proteasome-dependent processing of Spt23

Prior work from the Jentsch and Haines laboratories emphasized a role for Cdc48–Ufd1–Npl4 in release of the processed forms of Spt23[75] and Mga2[78] from the ER membrane. By contrast, data from the Silver laboratory emphasized an upstream role for Cdc48–Ufd1–Npl4 in processing of both proteins[83]. Our data support a role for Cdc48–Ufd1 and Ubx2 in both steps. We observe diminished production of p90 in mutant cells. However, the defect in p90 formation in *ubx2Δ* is variable and in general is less striking than the defect in *OLE1* induction. We propose that Cdc48–Ufd1–Npl4–Ubx2 promotes both the processing of Spt23 p120 to p90 and localization of p90 to the nucleus. It remains unclear why different laboratories have reported discordant results on p90 formation in Cdc48 pathway mutants. There are a number of variables that differ in the various studies including the epitope tags and conditional alleles used and the expression level of p120. Notably, a role for Cdc48–Ufd1 and Ubx2 in proteasome-dependent cleavage of p120 resonates with data on other Cdc48–proteasome substrates, which consistently place Cdc48/p97 function upstream of (or contemporaneous with) the proteasome. Regardless, sequential roles for Ubx2 in the formation and activation of p90 help explain why *ubx2Δ* exhibits a severe defect in *OLE1* regulation.

Consistent with the prior data on *rsp5* mutants[75], *ubx2Δ* cells are severely deficient in induction of *OLE1* expression and double mutants deficient in Ubx2 and

Rsp5 activity fail to grow on normal medium at 30°C but are rescued by the addition of oleic acid, pointing to a crucial role for Ubx2 in maintaining sufficient Ole1 activity to meet the cellular requirement for UFAs. An important and interesting question that remains unresolved is, how are levels of fatty acids sensed so that the cell maintains an appropriate ratio of SFAs:UFAs? Addition of exogenous UFAs suppresses processing of p120 to p90 by the Rsp5–Ubx2–Cdc48 pathway, which reduces expression of the desaturase enzyme encoded by *OLE1*. It was suggested in prior work that the human ortholog of Ubx2, UBXD8, directly senses UFAs[49, 141, 142]. However, addition of UFAs suppresses accumulation of ubiquitinated Spt23 and further reduces the residual expression of *OLE1* in *ubx2Δ*, indicating that these cells retain the ability to sense UFAs. Thus, if Ubx2 is a fatty acid sensor, it cannot be the only sensor in the *OLE1* pathway. Very recently it was reported that Ire1 serves as a sensor for membrane lipid saturation in mouse cells[144]. Our data suggest that either Rsp5 itself or a step upstream of Rsp5-dependent ubiquitination must be capable of sensing UFAs.

Acknowledgements

We thank S. Jentsch, L. Hicke, T. Sommer, E. Jarosch, and C.-W. Wang for gifts of strains, plasmids, antibodies, and W. den Besten for help with strain construction and valuable discussions. We also thank R. Ernst for communicating results prior to publication and M. Mann for hosting N.K. for a two-week visit. This work was supported by the Howard Hughes Medical Institute, of which R.J.D. is an Investigator.

Figure Legends

Figure 3.1

Sequencing and quantification of the ubiquitin proteome reveals proteins regulated by Cdc48 and its UBX domain adaptors.

(**A**) Experimental workflow for identification and quantification of Ub conjugates. Proteins were quantified using the SILAC method. Light (WT) and heavy (mutant) isotope-labeled cells constitutively expressing ^{His8}Ub were grown, collected in log-phase, and mixed. The cells were lysed in denaturing urea buffer and ubiquitinated proteins were purified with nickel-NTA magnetic beads. Proteins were digested into peptides, which were fractionated via strong anionic exchange and analyzed by mass spectrometry. (**B**) Smoothed histogram depicting changes in the ubiquitinated proteome in mutants. (**C**) Heatmap of relative changes to the ubiquitin proteome in various mutants. Only proteins that exhibited a significant change from a 1:1 ratio (mutant: wild-type) in at least one experiment are shown. Abundance changes were deemed significant if the mutant:wild-type ratio deviated from 1:1 with a p-value less than .05 corrected for multiple hypothesis testing by the Benjamini and Hochberg method[135]. (**D**) Graphical representation of how frequently an ubiquitin conjugate found to accumulate in the query mutant also accumulated in other mutants.

Figure 3.2

Changes to the ubiquitin conjugate proteome in *ubxΔ* mutants suggest functional specialization of Ubx proteins.

(**A**) Overview of how ubiquitin conjugates that accumulate in *cdc48-3* and *ubxΔ* mutants are distributed according to GO-Slim cell components and (**B**) GO-Slim biological processes. All of the proteins analyzed in (A) and (B) demonstrated enrichment greater than 10% (with ≥95% confidence) in one or more mutants. (**C**) Box and whisker plot of annotation terms demonstrating a significant enrichment in the Ub-Proteome for a majority of their identified members in the *ubx2Δ* mutant strain. Significance was calculated based on the 25th percentile of the distribution of protein ratios of protein annotated with the term in comparison to all *ubx2Δ* ratios. Enriched GO Cellular Component (CC), GO Biological Process (BP), GO Molecular Function (MF), and Uniprot Keyword (KW) terms are shown.

Figure 3.3

Ubiquitin-conjugated forms of Spt23 and Mga2 accumulate in *ubx2Δ* cells.

(**A**) Table of Spt23 and Mga2 SILAC ratios in *ubx2Δ*:WT and *ubx4Δ*:WT comparisons. The reported ratios reflect the amalgamation of many independent peptide ratio measurements (Spt23 *ubx2Δ*: 7, *ubx4Δ*: 24; Mga2 *ubx2Δ*: 21, *ubx4Δ*: 37). (**B**) Validation of Spt23 ubiquitin conjugate accumulation. Ubiquitin conjugates in WT, *ubx2Δ*, and *ubx4Δ* cells expressing ^{myc}Spt23^{V5} from the natural locus and ^{His8}Ub from a plasmid were purified with nickel-NTA magnetic beads. The input extract and bound fractions were immunoblotted for the myc epitope, Ub, and tubulin (input loading control) as indicated.

Figure 3.4

Ubx2 promotes inducible expression of *OLE1*.

(**A**) Synthetic lethality of *rsp5-1* and *ubx2Δ* mutations. Five-fold serial dilutions of WT, *rsp5-1*, *ubx2Δ*, *ubx4Δ, rsp5-1 ubx2Δ,* and *rsp5-1 ubx4Δ* cells were plated on YPD and YPD+0.2% oleic acid and incubated at 30°C or 37°C for two days. (**B,C**) Expression of *OLE1* mRNA in *ubx2Δ* cells grown in YPD (B) or grown first in YPD +0.2% oleic acid and shifted to YPD for two hours (C). RNA from cells grown to log-phase in YPD medium was used for qPCR to assess *OLE1* mRNA levels. Values were normalized to *ACT1*. Error bars denote SD; n=2. (**D**) Toxicity due to Spt23 and Mga2 overexpression is suppressed by *ubx2Δ.* WT, *ubx2Δ*, and *ubx4Δ* strains transformed with a 2 μm plasmid that contains either *GAL-SPT23* or *GAL-MGA2* were plated onto YPD and synthetic dextrose- or galactose- containing media and grown for 4 days at 30°C.

Figure 3.5

Ubx2 regulates processing and stability of the transcription factor Spt23.

(**A**) Steady state level of Spt23 in *OLE1* pathway mutants. The indicated mutant strains expressing ^{myc}Spt23^{V5} from the endogenous locus were grown in YPD at 30°C unless indicated otherwise, and cell lysates were fractionated by SDS-PAGE and immunoblotted with anti-myc, anti-V5, and anti-tubulin antibodies. (**B**) Steady state levels of Spt23 in *ubxΔ* mutants. The indicated strains were analyzed as in (A). (**C**) Influence of Ubx2's UBA and UBX domains on the steady-state level of Spt23. Same as (A) except that strains in lanes 2-6 were *ubx2Δ* bearing a copy of full-length (FL) *UBX2* or a version lacking the indicated domain integrated at *LEU2*. (**D**) Effect of UFAs on modification and processing of Spt23 in various mutants. Same as (A) except that strains were grown in normal medium or medium supplemented with

0.2% oleic acid as indicated. (**E**) Role of Ubx2 and Cdc48 in p90 degradation. Wild-type, *ubx2Δ,* and *cdc48-3* cells expressing ^{myc}Spt23^{V5} from the endogenous locus were grown in YPD at 30°C (left panels) or at 25°C followed by a shift to 37°C for 2 hours (right panels) prior to addition of cycloheximide to initiate a chase. At the indicated times cells were harvested and processed for SDS-PAGE followed by immunoblotting with anti-myc, or anti-tubulin antibodies.

Figure 3.6

Ubx2 interacts with the Spt23 and influences its subcellular localization.

(**A**) Ubx2 directly interacts with Spt23 and Rsp5. RJD 6173 cells expressing Ubx2^{TAP} from the natural locus were grown at 30°C in YP plus 2% galactose to induce expression of Spt23. Native cell lysates were subjected to anti-TAP immunoprecipitation (IP) and the immunoprecipitates and input lysates were immunoblotted (IB) for ^{myc}Spt23^{HA}, Cdc48, Ufd1, and Rsp5. Input extracts were also blotted for tubulin. (**B**) Immunofluorescence localization of Spt23 in WT and *ubx2Δ* cells. WT (RJD 6172) and *ubx2Δ* (RJD 6182) cells expressing Spt23 from the *GAL* promoter were grown at 30°C for 6 hours in YP plus 2% galactose. All forms of Spt23 containing an intact N-terminus were detected by the anti-myc antibody and the nucleus was marked by staining DNA with DAPI. (**C**) Quantification of percent of cell population from (B) that contain Spt23 punctae. Seven fields of ~300 cells were counted per strain. Error bars denote SD.

Supplemental Figure 3.1

Reproducibility of biological replicates.

(**A**) Pairwise comparisons of three independent *cdc48-3*:WT analyses, demonstrating low coefficients of variation between biological replicates. (**B**) Same as (A), except two independent analyses for each of the *ubxΔ* mutants are plotted. In these replicates the SILAC labels were switched, hence the negative correlation.

Supplemental Figure 3.2

Curation of Ub conjugates and overlap with previous screens listed in SCUD.

(**A**) 2-D plot of protein ratios from the ^{His8}Ub tagged vs untagged control experiment (x-axis) and the WT ^{His8}Ub vs WT ^{His8}Ub control (y-axis). Proteins that were recovered in significantly greater amounts (p-value < 0.05) from the ^{His8}Ub expressing cells in the tagged vs untagged control experiment and whose tagged vs tagged ratios did not differ by significantly greater than 10% (p-value ≥ 0.01) were deemed candidate Ub conjugates. (**B**) Overlap of Ub conjugates in the SCUD database with those identified in this publication.

Supplemental Figure 3.3

Changes in the Ub proteome in *cdc48-3* mutants.

(**A**) Box and whisker plot of annotation terms demonstrating a significant enrichment in the Ub proteome for a majority of their identified members in the *cdc48-3* mutant relative to WT. Significance is defined as the probability that a random subset of *cdc48-3* proteins ratios would have a 25th percentile at least as high as the 25th percentile of the ratios of proteins annotated with the term. Enriched GO Cellular Component (CC) and GO Biological Process (BP) terms are shown. (**B**) Fold changes in enzymes of the Ubiqutin Proteasome System. Error bars indicate the standard error of the ratios and asterisks indicate significant deviation

from 1:1 (p-value < 0.05 after correction by the Bejamini and Hochberg method). Colors indicate magnitude of change.

Supplemental Figure 3.4

Changes in the Ub proteome in *ubxΔ* mutants.

(**A-E**) Box and whisker plots of annotation terms demonstrating a significant enrichment in the Ub proteome for a majority of their identified members in the corresponding *ubx* mutant. Significance is defined as the probability that a random subset of *ubxΔ*:WT protein ratios would have a 25th percentile at least as high as the 25th percentile of the ratios of proteins annotated with the term. Enriched GO Cellular Component (CC), GO Biological Process (BP), GO Molecular Function (MF), KEGG Pathway (KFFF), Uniprot Sequence feature (SF), Uniprot Keyword (KW), and Protein Superfamily (PIR) terms are shown.

Supplemental Figure 3.5

Orthogonal validation of Mga2 ubiquitin conjugate accumulation. WT, *ubx2Δ*, and *ubx4Δ* cells expressing ^{myc}Mga2 from a YEPlac181 plasmid were grown in YPD at 30°C and native lysates were fractionated on a TUBE2 resin. The input extract and bound fractions were immunoblotted for V5 and myc epitopes, Ub, and Dpm1 (input loading control) as indicated.

Figure 3.1 Sequencing and quantification of the ubiquitin proteome reveals proteins regulated by Cdc48 and its UBX domain adaptors.

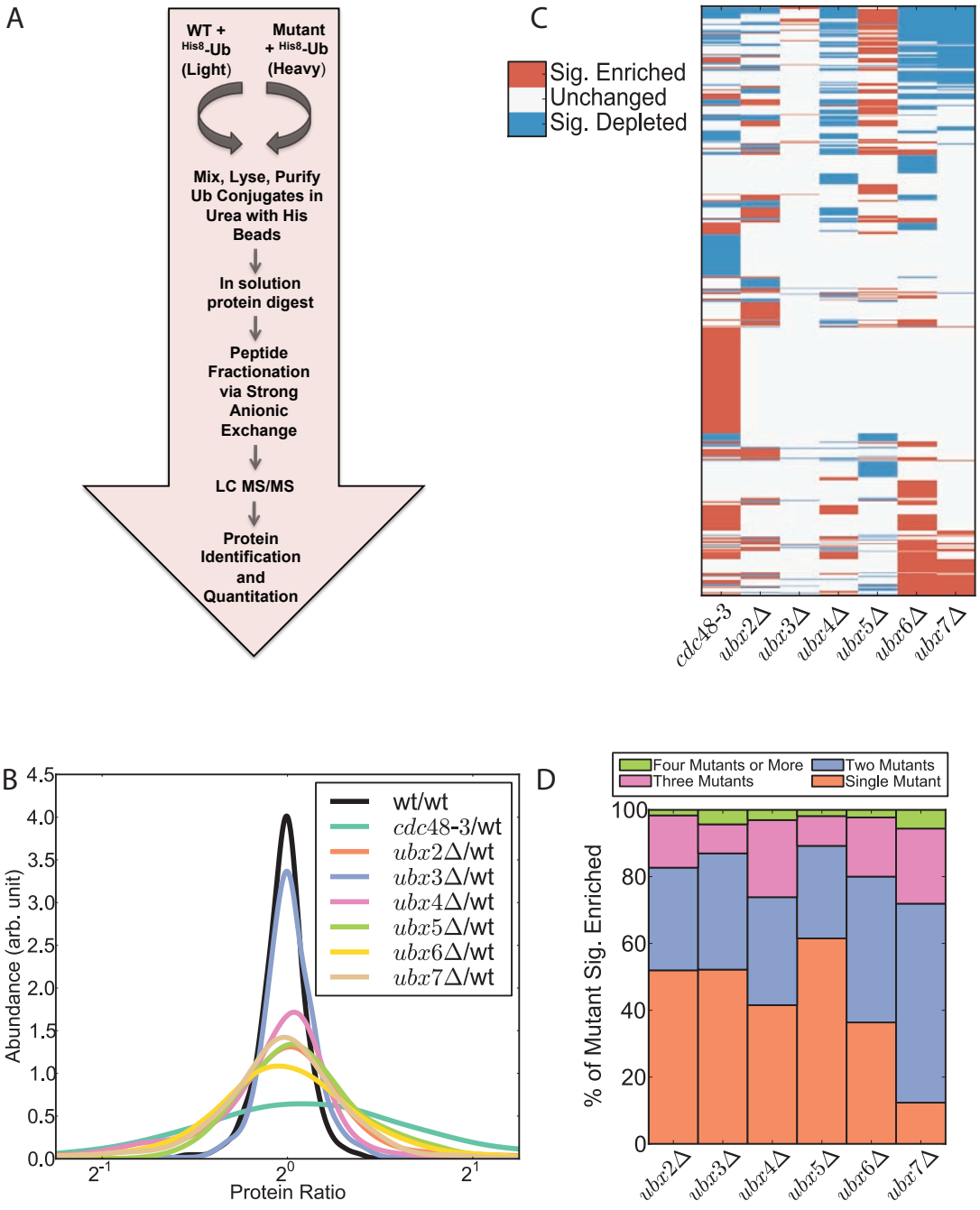


Figure 3.2 Changes to the ubiquitin conjugate proteome in *ubxΔ* mutants suggest functional specialization of Ubx proteins.

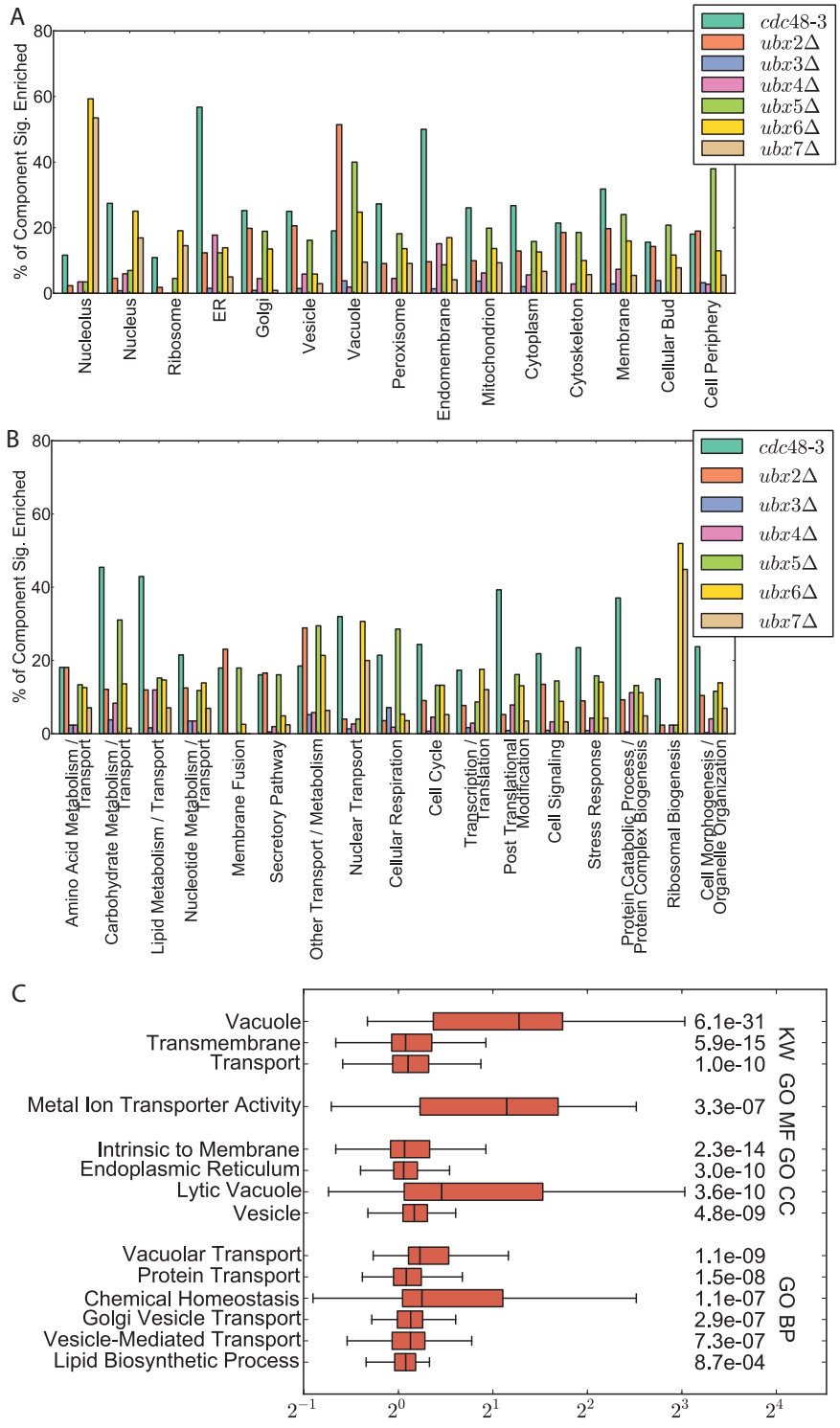


Figure 3.3 Ubiquitin-conjugated forms of Spt23 and Mga2 accumulate in *ubx2Δ* cells.

A

Protein	*ubx2Δ* SILAC Ratio	*ubx4Δ* SILAC Ratio
MGA2	4.66	3.74
SPT23	2.96	3.02

B

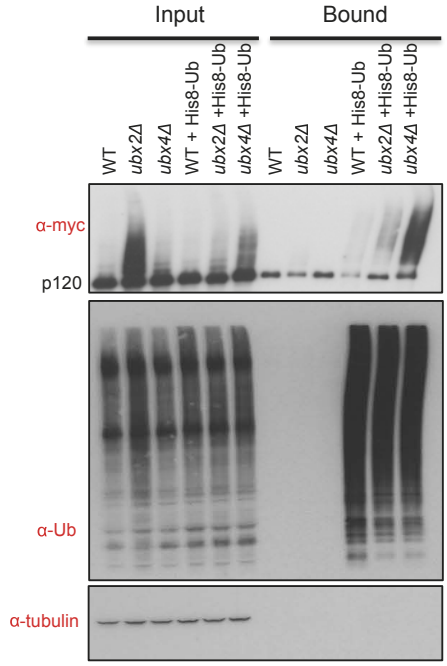


Figure 3.4 Ubx2 promotes inducible expression of *OLE1*.

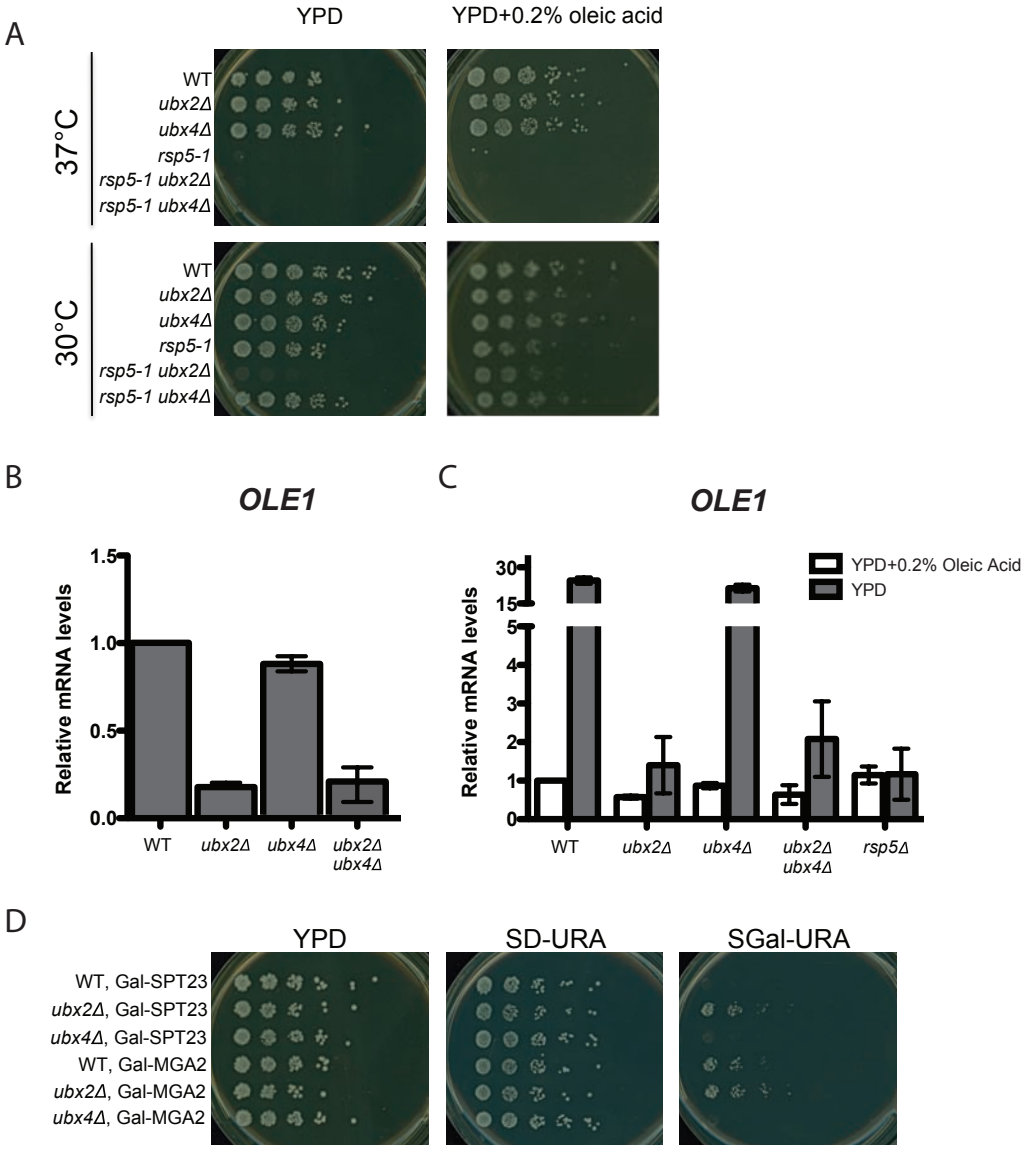


Figure 3.5 Ubx2 regulates processing and stability of the transcription factor Spt23.

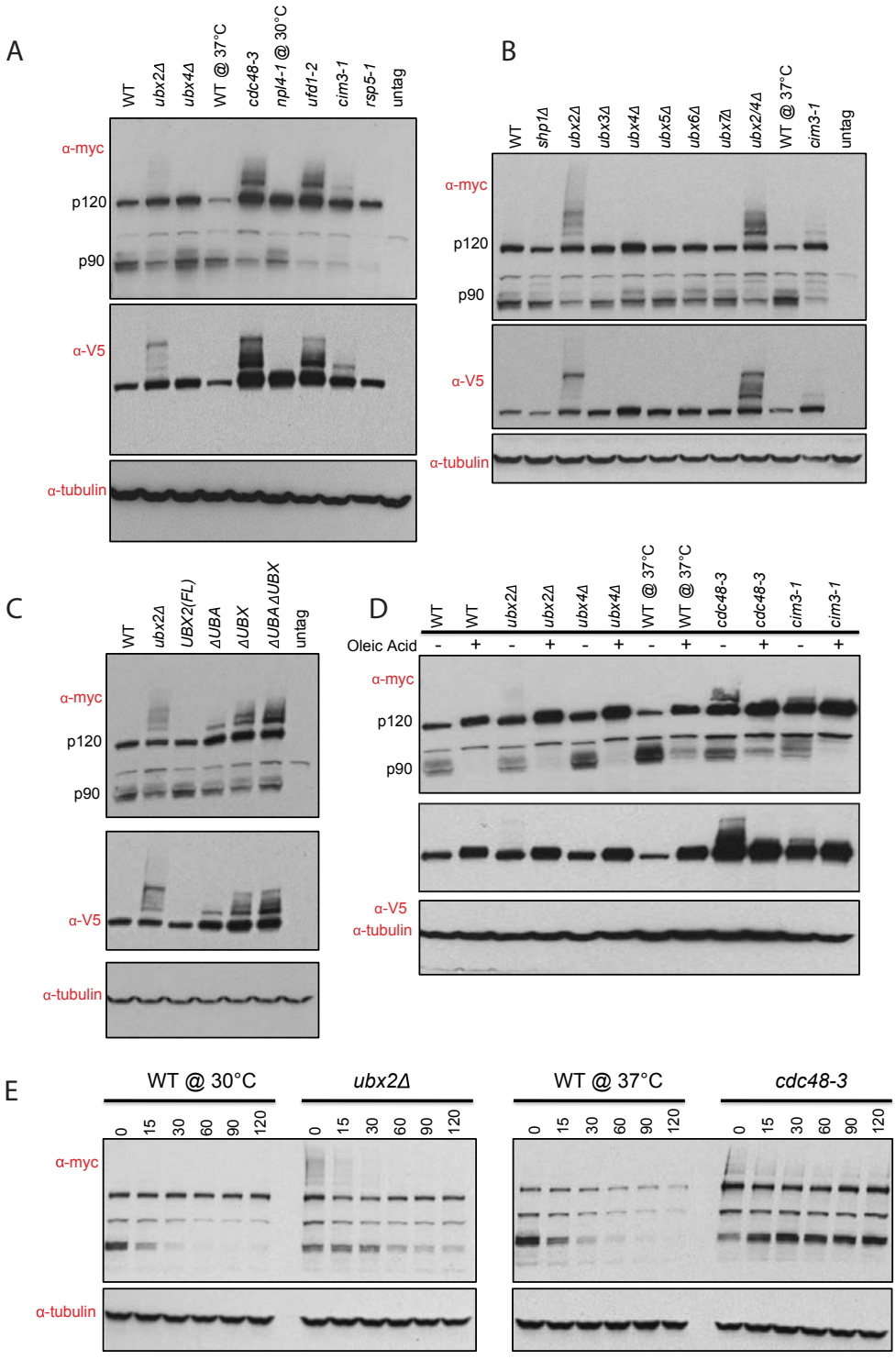
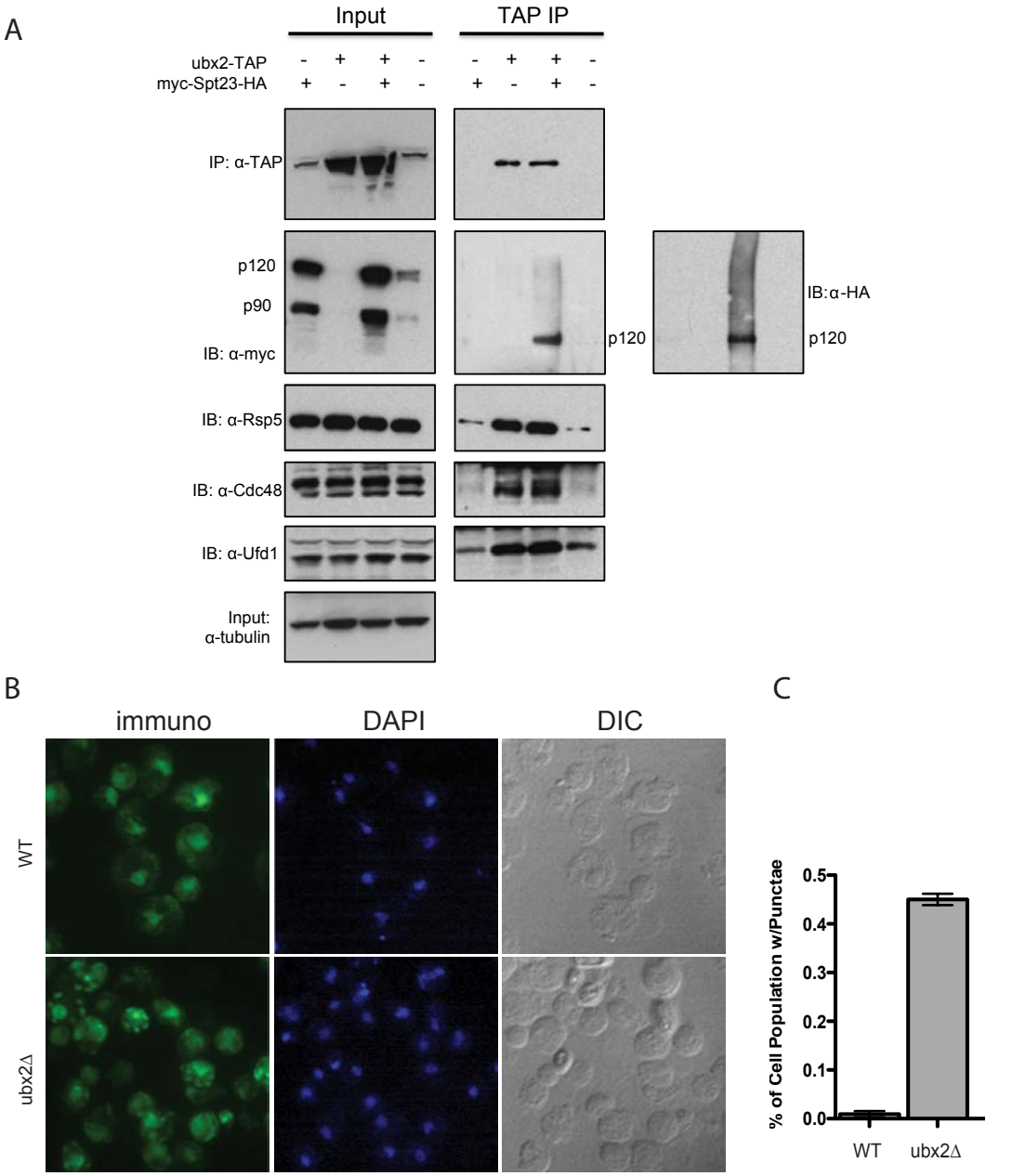
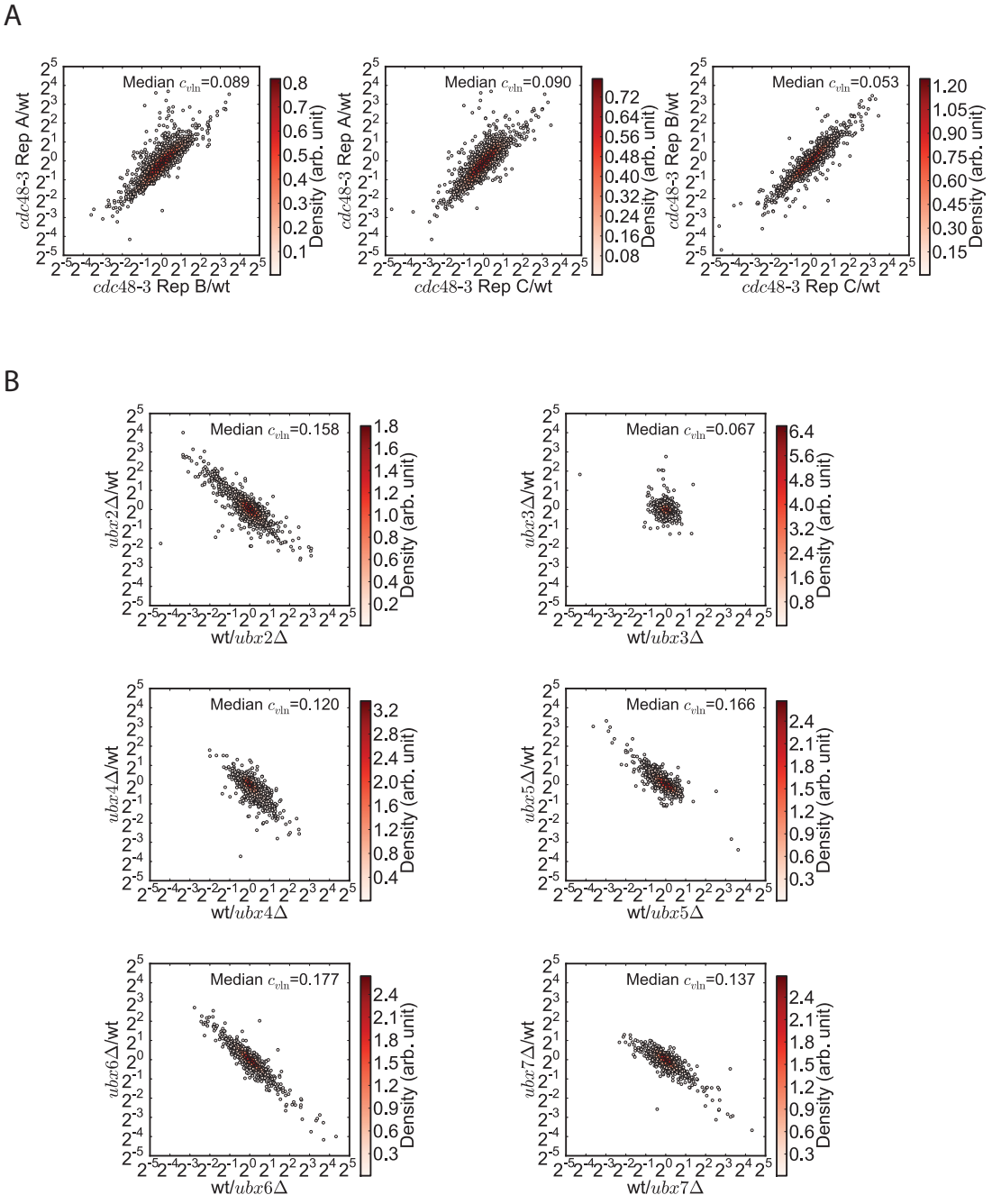


Figure 3.6 Ubx2 interacts with the Spt23 and influences its subcellular localization.

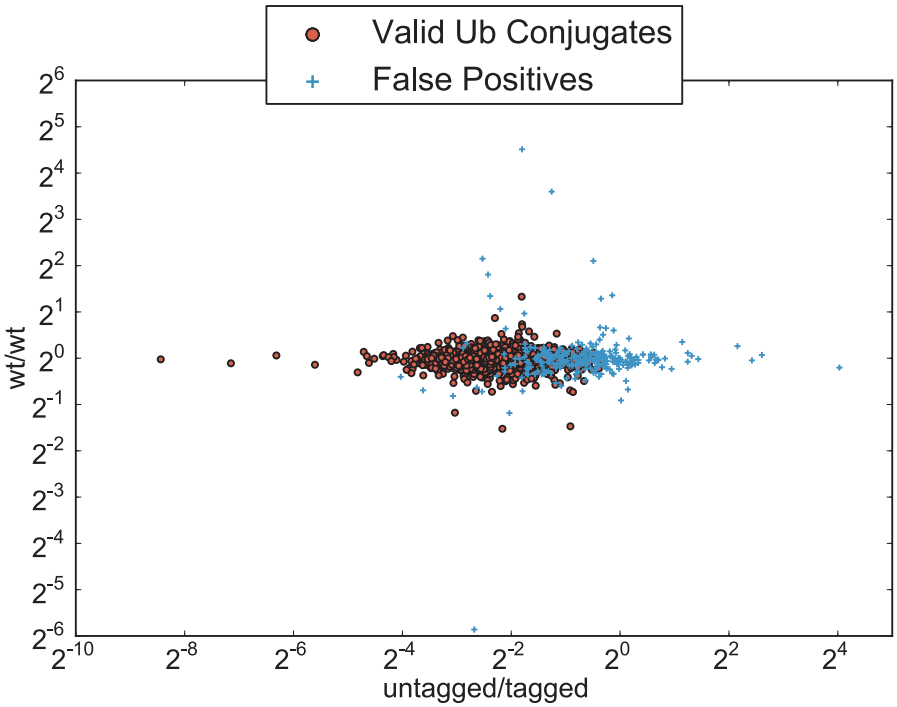


Supplemental Figure 3.1 Reproducibility of biological replicates.

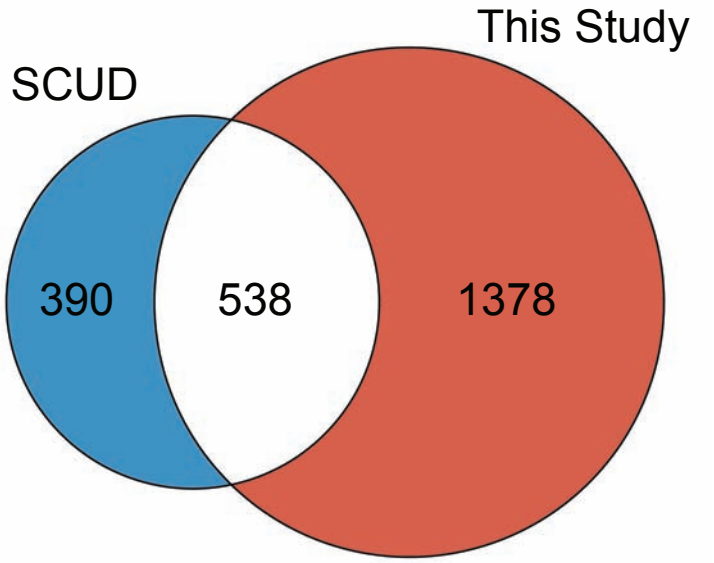


Supplemental Figure 3.2 Curation of Ub conjugates and overlap with previous screens listed in SCUD.

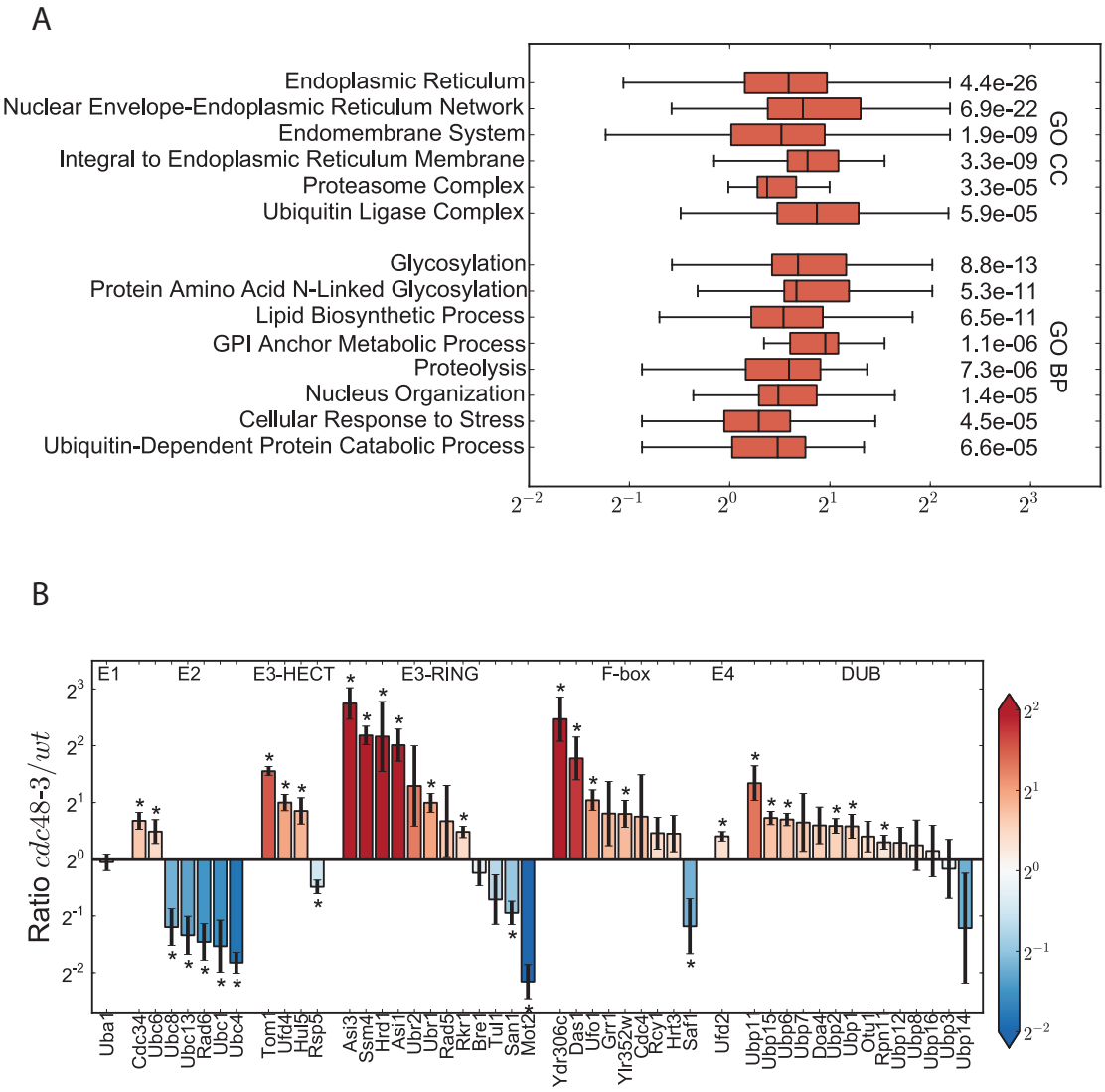
A



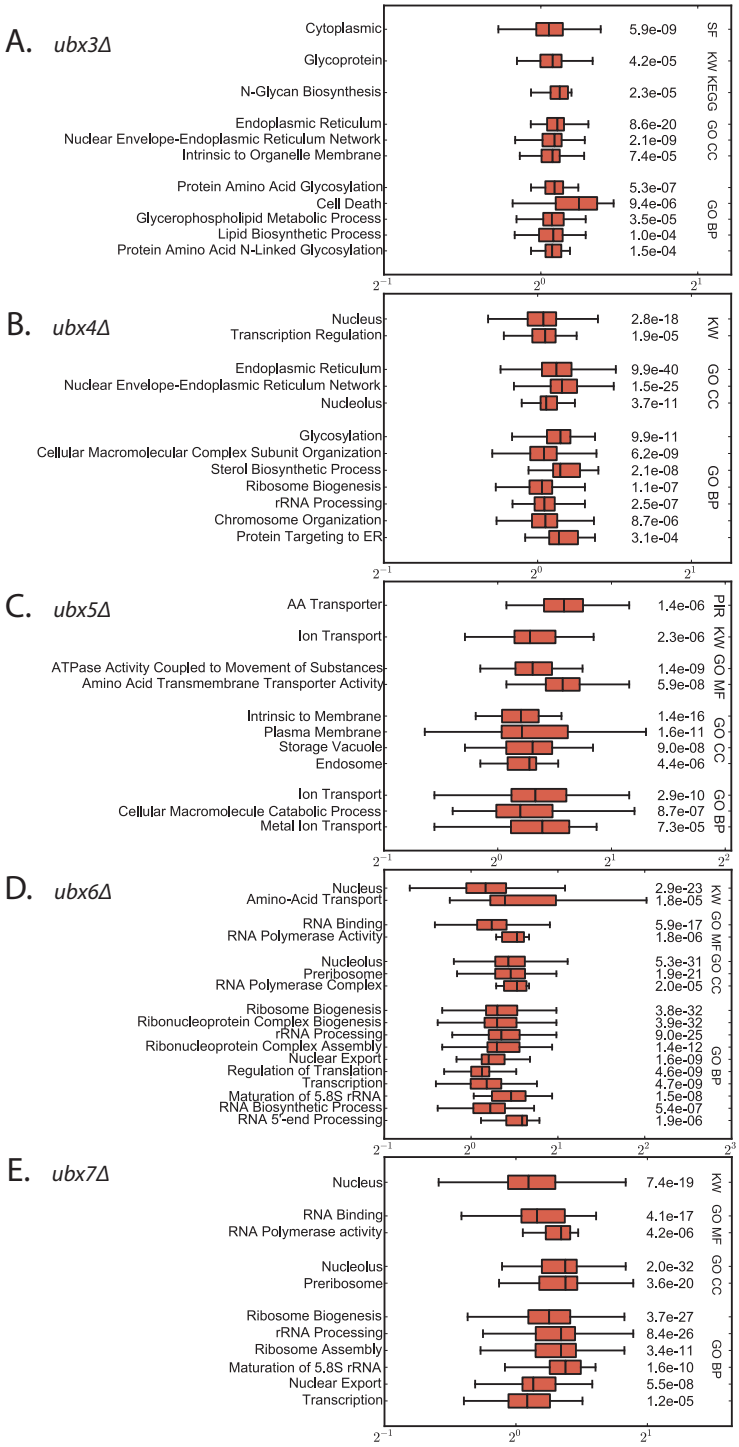
B



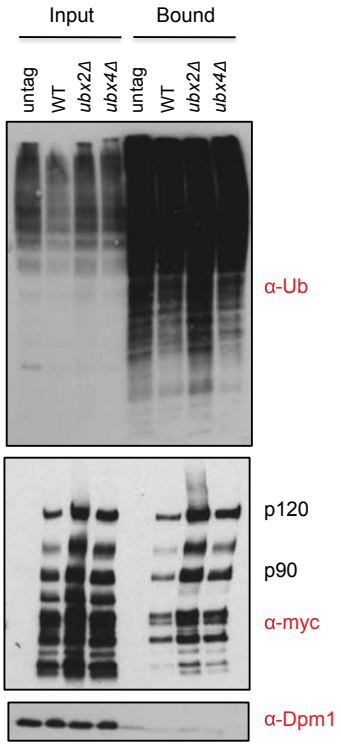
Supplemental Figure 3.3 Changes in the Ub proteome in *cdc48-3* mutants.



Supplemental Figure 3.4 Changes in the Ub proteome in *ubxΔ* mutants.



Supplemental Figure 3.5 Orthogonal validation of Mga2 ubiquitin conjugate accumulation.



Supplemental Table 3.1

Strains used in this study.

Strain	Genotype	Source
RJD 4614	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, MATa*	Open Biosystems
RJD 6137	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0 ubx2::KANMX MATa*	Open Biosystems
RJD 6138	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS] MATa*	This study
RJD 6139	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx2::KANMX MATa*	This study
RJD 6140	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx1::KANMX MATa*	This study
RJD 6141	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx3::KANMX MATa*	This study
RJD 6142	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx4::KANMX MATa*	This study
RJD 6143	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx5::KANMX MATa*	This study
RJD 6144	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx6::KANMX MATa*	This study
RJD 6145	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx7::KANMX MATa*	This study
RJD 6146	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx2::KANMX , ubx4::NATMX, MATa*	This study
RJD 6147	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx2::KANMX, ubx4::KANMX MATa*	This study
RJD 4781	*CAN1, lys2::HISMX, arg4::KANMX, leu2-3,-112 his3-11,-15, trp1-1, ura3-1, ade2-1, [pRS316-His8-Ubiqutin], MATa*	D. Chan
RJD 4977	*cdc48-3, CAN1, lys2::HISMX, arg4::KANMX, leu2-*	This study

	3,-112 his3-11,-15, trp1-1, ura3-1, ade2-1, [pRS316-His8-Ubiqutin], MATa	
RJD 6148	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, [pRS316-His8-Ubiqutin], MATa*	This study
RJD 6149	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, ubx2::KANMX, [pRS316-His8-Ubiqutin], MATa*	This study
RJD 6150	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, ubx3::KANMX, [pRS316-His8-Ubiqutin],MATa*	This study
RJD 6151	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, ubx4::KANMX, [pRS316-His8-Ubiqutin], MATa*	This study
RJD 6152	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, ubx5::KANMX, [pRS316-His8-Ubiqutin], MATa*	This study
RJD 6153	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, ubx6::KANMX, [pRS316-His8-Ubiqutin], MATa*	This study
RJD 6154	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, ubx7::KANMX, [pRS316-His8-Ubiqutin],MATa*	This study
RJD 6155	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pMGA2-MYC-Mga2-LEU], MATa*	This study
RJD 6156	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pMGA2-MYC-Mga2-LEU], ubx2::KANMX, MATa*	This study
RJD 6157	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pMGA2-MYC-Mga2-LEU], ubx4::KANMX, MATa*	This study
RJD 6158	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, rsp5-1, MATa*	This study
RJD 6159	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx4::KANMX, rsp5-1, MATa*	This study
RJD 4472	*rsp5-1, his4-914ΔR5, lys2-128Δ, ura3-52, GAL2+, MATalpha*	J. Huibregtse
RJD 6160	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, rsp5::KANMX, MATa*	N. Shcherbik
RJD 6161	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pGAL-FLAG-SPT23-HA-URA3], MATa*	This Study
RJD 6162	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pGAL-FLAG-SPT23-HA-URA3], ubx2::KANMX, MATa*	This Study
RJD 6163	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pGAL-FLAG-SPT23-HA-URA3], ubx4::KAN, MATa*	This Study
RJD 6164	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pGAL-FLAG-MGA2-HA-URA3], MATa*	This Study
RJD 6165	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pGAL-*	This study

	FLAG-MGA2-HA-URA3], ubx2::KAN, MATa	
RJD 6166	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, [pGAL-FLAG-MGA2-HA-URA3], ubx4::KAN, MATa*	This study
RJD 6167	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], cdc48-3, MATa*	This study
RJD 6168	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], npl4-1, MATa*	This study
RJD 6169	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ufd1-2, MATa*	This study
RJD 6170	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], cim3-1, MATa*	This study
RJD 6171	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], rsp5-1, MATa*	This study
RJD 3166	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, sel1::SEL1-TAP-HIS3MX, MATa*	Open Biosystems
RJD 6172	*his3Δ1, leu2Δ0, met15Δ0, ura3-1, lys2Δ0, ura3::pGAL-MYC-SPT23-HA, MATa*	This study
RJD 6173	*his3Δ1, leu2Δ0, met15Δ0, ura3-1, lys2Δ0, ura3::pGAL-MYC-SPT23-HA, sel1::SEL1-TAP-HIS3MX, MATa*	This study
RJD 6174	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, leu::pUBX2-UBX2(FL), MATa*	C. Wang
RJD 6175	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, SPT23::MYC-SPT23-V5[HIS], ubx2::KANMX, leu::pUBX2-UBX2(FL), MATa*	This Study
RJD 6176	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, leu::pUBX2-UBX2(ΔUBA), MATa*	C. Wang
RJD 6177	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, SPT23:MYC-SPT23-V5[HIS], leu::pUBX2-UBX2(ΔUBA), MATa*	This Study
RJD 6178	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, leu::pUBX2-UBX2(ΔUBX), MATa*	C. Wang
RJD 6179	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, SPT23:MYC-SPT23-V5[HIS], leu::pUBX2-UBX2(ΔUBX), MATa*	This Study
RJD 6180	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, leu::pUBX2-UBX2(ΔUBAΔUBX), MATa*	C. Wang
RJD 6181	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, ubx2::KANMX, SPT23:MYC-SPT23-V5[HIS], leu::pUBX2-UBX2(ΔUBAΔUBX), MATa*	This Study
RJD 6182	*his3Δ1, leu2Δ0, met15Δ0, ura3-1, lys2Δ0, pGAL-MYC-SPT23-HA::URA, ubx2::KANMX, MATa*	This Study
RJD 6183	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0, arg4::HYG, MATa*	This Study

RJD 5246	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0 ubx4::KANMX, MATa*	Open Biosystems
RJD 5772	*his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, lys2Δ0 ubx2::KANMX, ubx4::KANMX, MATa*	This Study

Supplemental Table 3.2

Plasmids used in this study.

Plasmid	Description	Source
RDB 2818	*MYC-MGA2 in YEPlac181*	S. Jentsch
RDB 2819	*GAL-MYC-SPT23-HA in YIPlac211*	S. Jentsch
RDB 1937	*GAL-FLAG-SPT23-HA in YEPlac181*	D. Haines
RDB 1938	*GAL-FLAG-MGA2-HA in YEPlac181*	D. Haines
RDB 1851	*HIS8-Ubiquitin in pRS316*	Our Laboratory

Supplemental Table 3.3

Primers used for quantitative PCR in this study.

Gene	Forward Primer	Reverse Primer
ACT1	5’ -CTGCCGGTATTGACCAAACT-3’	5’-CGGTGATTTCCTTTTGCATT-3’
OLE1	5’-TAATGGGCTCCAAGGAAATG-3’	5’-CATGGTTGTTCGGAGATGTG-3’

Chapter 4:

Findings, Implications, and Future Directions

Summary

The goal of this dissertation was to implement and optimize a quantitative mass spectrometry-based method to monitor changes in *Saccharomyces cerevisiae* protein levels on a proteome-wide scale. Secondly, this technique was then used to survey changes in the ubiquitin proteome in *cdc48* and *ubx* mutants in an effort to better understand their functions in the cell and discover potential new substrates.

One of the major contributions of this work lies in the implementation of the aforementioned deep-sequencing method. This cutting-edge technology has a myriad of applications and leaves Caltech capable of quantitative sequencing of the highest caliber.

Secondly, our proteomic screen provided both confirmation of and new insight into the roles of Cdc48 and its Ubx adaptors in the cell. Our dataset contained 1,196 ratios for ubiquitinated proteins identified in at least one mutant experiment and 1,733 ubiquitinated proteins identified across all experiments. Our study covered over 50% of the previously identified ubiquitinated proteins in the SCUD database and included 1,226 proteins not previously listed. These values suggest that we sampled a substantial portion of the yeast ubiquitinated proteome and in more depth than previously ever done. Notably, different *ubx* mutants exhibited reproducible patterns that differed greatly from each other. This supported the long-standing hypothesis that individual Ubx proteins target distinct sets of ubiquitin conjugates and have functional specializations. This is consistent with their roles as substrate adaptors for the Cdc48 engine.

Lastly, our findings also demonstrated the general value of proteomics and its utility to identify ligands for specific ubiquitin receptor pathways. Our list of changing ubiquitinated proteins among *cdc48* and *ubx* mutants provided a large catalog of potential substrates/targets and suggested new functions of the Ubx adaptors. Some of the key findings for Cdc48 and the Ubx adaptors are listed below.

Cdc48: Mutation of Cdc48 had the largest effect on the ubiquitinated proteome. The *cdc48-3* dataset encompassed many known targets of Cdc48 including the transcription factors Mga2 and Spt23, HMG-CoA reductases Hmg2, and the Rpb1 subunit of the RNA Polymerase II[52, 73, 76, 136]. Additionally, the *cdc48-3* conjugates were highly enriched for proteins of the endomembrane system and, very interestingly, multiple E3 ubiquitin ligases and F-box proteins.

Ubx2: Deletion of *UBX2* also had a large effect on the ubiquitinated proteome. These strains showed a significant enrichment of ubiquitinated vacuolar and transport proteins, and proteins involved in fatty acid metabolism, including Spt23 and Mga2.

Ubx5: Ubiquitin conjugates that were elevated in *ubx5Δ* cells showed a strong enrichment for vacuolar and plasma membrane proteins.

Ubx6 and Ubx7: Close homologs, *ubx6Δ* and *ubx7Δ,* displayed a significant overlap with over 85% of ubiquitin conjugates accumulating *ubx7Δ* also accumulating in *ubx6Δ.* Consistent with Ubx6 and Ubx7 localizing to the nucleus and nuclear

periphery, the accumulating conjugates displayed enrichment for nucleolar and ribosomal proteins.

To validate some of these mass spectrometric findings, we studied in detail the transcription factor Spt23, which was identified, as mentioned above, as a putative Ubx2 substrate. Cdc48 had already been linked to the promotion of release of Spt23 and Mga2 from the ER membrane[75, 78]. However, Ubx proteins had not been previously linked to this pathway. Our work demonstrated Ubx2 as a regulator of processing and stability of Spt23. This work identified Ubx2 as key player in the OLE pathway and as such, a regulator of cellular lipid biosynthesis.

Future Directions

The proteomic screen not only led to the discovery of Ubx2 as a key regulator in lipid biosynthesis but also provided a rich resource for future investigations related to the functions and regulation of the Cdc48 network. Several key findings have potential for future research projects.

1) Cdc48 may act globally in the extraction and/or degradation of ubiquitinated F-box proteins from SCF complexes. Our data revealed that *cdc48-3* cells contained elevated levels of ubiquitin-conjugated forms of multiple E3 ubiquitin ligases and F-box proteins. This idea is consistent with a recent finding of Cdc48 as an SCF^{Met30} disassembly factor[137].

2) Ubx5 as a regulator of plasma membrane proteins. Ubx5 contains a UIM domain that links it to Rub1-conjugated cullin ring-ligases[54, 138]. This raises the possibility that a CRL regulates trafficking of membrane proteins in yeast.
3) Ubx6 and Ubx7 as regulators of ribosomal biogenesis and nuclear import and export of proteins/ribosomes. Very little, aside from their localization, is known about these proteins.
4) Lastly, our biological studies of Spt23 leave several unanswered questions. Although we confirmed Ubx2 as a player in the OLE pathway and as a regulator of Spt23 processing, the mechanism of Ubx2 involvement is not clearly understood. Ubx2 may act in the recruitment of Cdc48 to the complex (perhaps as a sensor), or alternatively, may work as a processivity factor with Cdc48 and the proteasome in activation of Spt23. Additionally, another main question that remains unresolved is, how are fatty acid levels detected in the cell in order to maintain the appropriate ratios of SFAs:UFAs? Our studies indicate that in *ubx2Δ* backgrounds, cells retain the ability to sense UFAs. Thus, if Ubx2 is a fatty acid sensor it cannot be the only sensor in the OLE pathway. Our work suggests that either Rsp5 itself, or a step upstream of Rsp5-dependent ubiquitination is capable of sensing fatty acid levels.

Bibliography

1. Hershko, A. and A. Ciechanover, *The ubiquitin system.* Annu Rev Biochem, 1998. **67**: p. 425-79.
2. Shimura, H., et al., *Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase.* Nat Genet, 2000. **25**(3): p. 302-5.
3. McNaught, K.S., et al., *Failure of the ubiquitin-proteasome system in Parkinson's disease.* Nat Rev Neurosci, 2001. **2**(8): p. 589-94.
4. Oddo, S., *The ubiquitin-proteasome system in Alzheimer's disease.* J Cell Mol Med, 2008. **12**(2): p. 363-73.
5. McGrath, J.P., S. Jentsch, and A. Varshavsky, *UBA 1: an essential yeast gene encoding ubiquitin-activating enzyme.* EMBO J, 1991. **10**(1): p. 227-36.
6. Deshaies, R.J., *Make it or break it: the role of ubiquitin-dependent proteolysis in cellular regulation.* Trends Cell Biol, 1995. **5**(11): p. 428-34.
7. Thrower, J.S., et al., *Recognition of the polyubiquitin proteolytic signal.* EMBO J, 2000. **19**(1): p. 94-102.
8. Finley, D., et al., *The ubiquitin-proteasome system of Saccharomyces cerevisiae.* Genetics, 2012. **192**(2): p. 319-60.
9. Pierce, N.W., et al., *Detection of sequential polyubiquitylation on a millisecond timescale.* Nature, 2009. **462**(7273): p. 615-9.
10. Ikeda, F. and I. Dikic, *Atypical ubiquitin chains: new molecular signals. 'Protein Modifications: Beyond the Usual Suspects' review series.* EMBO Rep, 2008. **9**(6): p. 536-42.
11. Baumeister, W., et al., *The proteasome: paradigm of a self-compartmentalizing protease.* Cell, 1998. **92**(3): p. 367-80.
12. Pickart, C.M. and R.E. Cohen, *Proteasomes and their kin: proteases in the machine age.* Nat Rev Mol Cell Biol, 2004. **5**(3): p. 177-87.
13. Finley, D., *Recognition and processing of ubiquitin-protein conjugates by the proteasome.* Annu Rev Biochem, 2009. **78**: p. 477-513.
14. Groll, M., et al., *Structure of 20S proteasome from yeast at 2.4 A resolution.* Nature, 1997. **386**(6624): p. 463-71.
15. Elsasser, S., et al., *Rad23 and Rpn10 serve as alternative ubiquitin receptors for the proteasome.* J Biol Chem, 2004. **279**(26): p. 26817-22.
16. Deveraux, Q., et al., *Inhibition of ubiquitin-mediated proteolysis by the Arabidopsis 26 S protease subunit S5a.* J Biol Chem, 1995. **270**(50): p. 29660-3.
17. Verma, R., et al., *Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system.* Cell, 2004. **118**(1): p. 99-110.
18. Mayor, T., et al., *Quantitative profiling of ubiquitylated proteins reveals proteasome substrates and the substrate repertoire influenced by the Rpn10 receptor pathway.* Mol Cell Proteomics, 2007. **6**(11): p. 1885-95.
19. Saeki, Y., et al., *Ubiquitin-like proteins and Rpn10 play cooperative roles in ubiquitin-dependent proteolysis.* Biochem Biophys Res Commun, 2002. **293**(3): p. 986-92.

20. Chen, L. and K. Madura, *Rad23 promotes the targeting of proteolytic substrates to the proteasome.* Mol Cell Biol, 2002. **22**(13): p. 4902-13.
21. Rao, H. and A. Sastry, *Recognition of specific ubiquitin conjugates is important for the proteolytic functions of the ubiquitin-associated domain proteins Dsk2 and Rad23.* J Biol Chem, 2002. **277**(14): p. 11691-5.
22. Elsasser, S. and D. Finley, *Delivery of ubiquitinated substrates to protein-unfolding machines.* Nat Cell Biol, 2005. **7**(8): p. 742-9.
23. Chen, L., et al., *Ubiquitin-associated (UBA) domains in Rad23 bind ubiquitin and promote inhibition of multi-ubiquitin chain assembly.* EMBO Rep, 2001. **2**(10): p. 933-8.
24. Wilkinson, C.R., et al., *Proteins containing the UBA domain are able to bind to multi-ubiquitin chains.* Nat Cell Biol, 2001. **3**(10): p. 939-43.
25. Richly, H., et al., *A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting.* Cell, 2005. **120**(1): p. 73-84.
26. Stolz, A., et al., *Cdc48: a power machine in protein degradation.* Trends Biochem Sci, 2011. **36**(10): p. 515-23.
27. Meusser, B., et al., *ERAD: the long road to destruction.* Nat Cell Biol, 2005. **7**(8): p. 766-72.
28. Meyer, H., M. Bug, and S. Bremer, *Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system.* Nat Cell Biol, 2012. **14**(2): p. 117-23.
29. Ye, Y., *Diverse functions with a common regulator: ubiquitin takes command of an AAA ATPase.* J Struct Biol, 2006. **156**(1): p. 29-40.
30. Johnson, J.O., et al., *Exome sequencing reveals VCP mutations as a cause of familial ALS.* Neuron, 2010. **68**(5): p. 857-64.
31. Neumann, M., et al., *TDP-43-positive white matter pathology in frontotemporal lobar degeneration with ubiquitin-positive inclusions.* J Neuropathol Exp Neurol, 2007. **66**(3): p. 177-83.
32. Chapman, E., A.N. Fry, and M. Kang, *The complexities of p97 function in health and disease.* Mol Biosyst, 2011. **7**(3): p. 700-10.
33. Haines, D.S., *p97-containing complexes in proliferation control and cancer: emerging culprits or guilt by association?* Genes Cancer, 2010. **1**(7): p. 753-763.
34. Watts, G.D., et al., *Novel VCP mutations in inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia.* Clin Genet, 2007. **72**(5): p. 420-6.
35. Chou, T.F. and R.J. Deshaies, *Development of p97 AAA ATPase inhibitors.* Autophagy, 2011. **7**(9): p. 1091-2.
36. Bursavich, M.G., et al., *2-Anilino-4-aryl-1,3-thiazole inhibitors of valosin-containing protein (VCP or p97).* Bioorg Med Chem Lett, 2010. **20**(5): p. 1677-9.
37. Davies, J.M., A.T. Brunger, and W.I. Weis, *Improved structures of full-length p97, an AAA ATPase: implications for mechanisms of nucleotide-dependent conformational change.* Structure, 2008. **16**(5): p. 715-26.
38. Pye, V.E., et al., *Going through the motions: the ATPase cycle of p97.* J Struct Biol, 2006. **156**(1): p. 12-28.

39. Jentsch, S. and S. Rumpf, *Cdc48 (p97): a "molecular gearbox" in the ubiquitin pathway?* Trends Biochem Sci, 2007. **32**(1): p. 6-11.
40. Buchberger, A., B. Bukau, and T. Sommer, *Protein quality control in the cytosol and the endoplasmic reticulum: brothers in arms.* Mol Cell, 2010. **40**(2): p. 238-52.
41. Kloppsteck, P., et al., *Regulation of p97 in the ubiquitin-proteasome system by the UBX protein-family.* Biochim Biophys Acta, 2012. **1823**(1): p. 125-9.
42. Schuberth, C. and A. Buchberger, *UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97.* Cell Mol Life Sci, 2008. **65**(15): p. 2360-71.
43. Buchberger, A., *Control of ubiquitin conjugation by cdc48 and its cofactors.* Subcell Biochem, 2010. **54**: p. 17-30.
44. Bohm, S. and A. Buchberger, *The budding yeast Cdc48(Shp1) complex promotes cell cycle progression by positive regulation of protein phosphatase 1 (Glc7).* PLoS One, 2013. **8**(2): p. e56486.
45. Meyer, H.H., Y. Wang, and G. Warren, *Direct binding of ubiquitin conjugates by the mammalian p97 adaptor complexes, p47 and Ufd1-Npl4.* EMBO J, 2002. **21**(21): p. 5645-52.
46. Wang, C.W. and S.C. Lee, *The ubiquitin-like (UBX)-domain-containing protein Ubx2/Ubxd8 regulates lipid droplet homeostasis.* J Cell Sci, 2012. **125**(Pt 12): p. 2930-9.
47. Wolf, D.H. and A. Stolz, *The Cdc48 machine in endoplasmic reticulum associated protein degradation.* Biochim Biophys Acta, 2012. **1823**(1): p. 117-24.
48. Suzuki, M., et al., *Derlin-1 and UBXD8 are engaged in dislocation and degradation of lipidated ApoB-100 at lipid droplets.* Mol Biol Cell, 2012. **23**(5): p. 800-10.
49. Olzmann, J.A., C.M. Richter, and R.R. Kopito, *Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover.* Proc Natl Acad Sci U S A, 2013. **110**(4): p. 1345-50.
50. Alberts, S.M., et al., *Ubx4 modulates cdc48 activity and influences degradation of misfolded proteins of the endoplasmic reticulum.* J Biol Chem, 2009. **284**(24): p. 16082-9.
51. Barbin, L., et al., *The Cdc48-Ufd1-Npl4 complex is central in ubiquitin-proteasome triggered catabolite degradation of fructose-1,6-bisphosphatase.* Biochem Biophys Res Commun, 2010. **394**(2): p. 335-41.
52. Verma, R., et al., *Cdc48/p97 mediates UV-dependent turnover of RNA Pol II.* Mol Cell, 2011. **41**(1): p. 82-92.
53. Alexandru, G., et al., *UBXD7 binds multiple ubiquitin ligases and implicates p97 in HIF1alpha turnover.* Cell, 2008. **134**(5): p. 804-16.
54. den Besten, W., et al., *NEDD8 links cullin-RING ubiquitin ligase function to the p97 pathway.* Nat Struct Mol Biol, 2012. **19**(5): p. 511-6, S1.
55. Decottignies, A., A. Evain, and M. Ghislain, *Binding of Cdc48p to a ubiquitin-related UBX domain from novel yeast proteins involved in intracellular proteolysis and sporulation.* Yeast, 2004. **21**(2): p. 127-39.

56. Tran, J.R., L.R. Tomsic, and J.L. Brodsky, *A Cdc48p-associated factor modulates endoplasmic reticulum-associated degradation, cell stress, and ubiquitinated protein homeostasis.* J Biol Chem, 2011. **286**(7): p. 5744-55.
57. Heo, J.M., et al., *A stress-responsive system for mitochondrial protein degradation.* Mol Cell, 2010. **40**(3): p. 465-80.
58. Xu, S., et al., *The AAA-ATPase p97 is essential for outer mitochondrial membrane protein turnover.* Mol Biol Cell, 2011. **22**(3): p. 291-300.
59. Tanaka, A., et al., *Proteasome and p97 mediate mitophagy and degradation of mitofusins induced by Parkin.* J Cell Biol, 2010. **191**(7): p. 1367-80.
60. Ghislain, M., et al., *Cdc48p interacts with Ufd3p, a WD repeat protein required for ubiquitin-mediated proteolysis in Saccharomyces cerevisiae.* EMBO J, 1996. **15**(18): p. 4884-99.
61. Ramadan, K., et al., *Cdc48/p97 promotes reformation of the nucleus by extracting the kinase Aurora B from chromatin.* Nature, 2007. **450**(7173): p. 1258-62.
62. Dobrynin, G., et al., *Cdc48/p97-Ufd1-Npl4 antagonizes Aurora B during chromosome segregation in HeLa cells.* J Cell Sci, 2011. **124**(Pt 9): p. 1571-80.
63. Fu, X., et al., *Cdc48p is required for the cell cycle commitment point at Start via degradation of the G1-CDK inhibitor Far1p.* J Cell Biol, 2003. **163**(1): p. 21-6.
64. Kiel, J.A., *Autophagy in unicellular eukaryotes.* Philos Trans R Soc Lond B Biol Sci, 2010. **365**(1541): p. 819-30.
65. Tresse, E., et al., *VCP/p97 is essential for maturation of ubiquitin-containing autophagosomes and this function is impaired by mutations that cause IBMPFD.* Autophagy, 2010. **6**(2): p. 217-27.
66. Ju, J.S. and C.C. Weihl, *p97/VCP at the intersection of the autophagy and the ubiquitin proteasome system.* Autophagy, 2010. **6**(2): p. 283-5.
67. Krick, R., et al., *Cdc48/p97 and Shp1/p47 regulate autophagosome biogenesis in concert with ubiquitin-like Atg8.* J Cell Biol, 2010. **190**(6): p. 965-73.
68. Kondo, H., et al., *p47 is a cofactor for p97-mediated membrane fusion.* Nature, 1997. **388**(6637): p. 75-8.
69. Uchiyama, K., et al., *The localization and phosphorylation of p47 are important for Golgi disassembly-assembly during the cell cycle.* J Cell Biol, 2003. **161**(6): p. 1067-79.
70. Bays, N.W. and R.Y. Hampton, *Cdc48-Ufd1-Npl4: stuck in the middle with Ub.* Curr Biol, 2002. **12**(10): p. R366-71.
71. Martin, C.E., C.S. Oh, and Y. Jiang, *Regulation of long chain unsaturated fatty acid synthesis in yeast.* Biochim Biophys Acta, 2007. **1771**(3): p. 271-85.
72. Zhang, S., Y. Skalsky, and D.J. Garfinkel, *MGA2 or SPT23 is required for transcription of the delta9 fatty acid desaturase gene, OLE1, and nuclear membrane integrity in Saccharomyces cerevisiae.* Genetics, 1999. **151**(2): p. 473-83.
73. Hoppe, T., et al., *Activation of a membrane-bound transcription factor by regulated ubiquitin/proteasome-dependent processing.* Cell, 2000. **102**(5): p. 577-86.

74. Rape, M. and S. Jentsch, *Productive RUPture: activation of transcription factors by proteasomal processing.* Biochim Biophys Acta, 2004. **1695**(1-3): p. 209-13.
75. Rape, M., et al., *Mobilization of processed, membrane-tethered SPT23 transcription factor by CDC48(UFD1/NPL4), a ubiquitin-selective chaperone.* Cell, 2001. **107**(5): p. 667-77.
76. Shcherbik, N., et al., *Rsp5p is required for ER bound Mga2p120 polyubiquitination and release of the processed/tethered transactivator Mga2p90.* Curr Biol, 2003. **13**(14): p. 1227-33.
77. Piwko, W. and S. Jentsch, *Proteasome-mediated protein processing by bidirectional degradation initiated from an internal site.* Nat Struct Mol Biol, 2006. **13**(8): p. 691-7.
78. Shcherbik, N. and D.S. Haines, *Cdc48p(Npl4p/Ufd1p) binds and segregates membrane-anchored/tethered complexes via a polyubiquitin signal present on the anchors.* Mol Cell, 2007. **25**(3): p. 385-97.
79. Braun, S., et al., *Role of the ubiquitin-selective CDC48(UFD1/NPL4)chaperone (segregase) in ERAD of OLE1 and other substrates.* EMBO J, 2002. **21**(4): p. 615-21.
80. Shcherbik, N., et al., *A single PXY motif located within the carboxyl terminus of Spt23p and Mga2p mediates a physical and functional interaction with ubiquitin ligase Rsp5p.* J Biol Chem, 2004. **279**(51): p. 53892-8.
81. Jiang, Y., et al., *Mga2p processing by hypoxia and unsaturated fatty acids in Saccharomyces cerevisiae: impact on LORE-dependent gene expression.* Eukaryot Cell, 2002. **1**(3): p. 481-90.
82. Chellappa, R., et al., *The membrane proteins, Spt23p and Mga2p, play distinct roles in the activation of Saccharomyces cerevisiae OLE1 gene expression. Fatty acid-mediated regulation of Mga2p activity is independent of its proteolytic processing into a soluble transcription activator.* J Biol Chem, 2001. **276**(47): p. 43548-56.
83. Hitchcock, A.L., et al., *The conserved npl4 protein complex mediates proteasome-dependent membrane-bound transcription factor activation.* Mol Biol Cell, 2001. **12**(10): p. 3226-41.
84. Shcherbik, N. and D.S. Haines, *Ub on the move.* J Cell Biochem, 2004. **93**(1): p. 11-9.
85. Cravatt, B.F., G.M. Simon, and J.R. Yates, 3rd, *The biological impact of mass-spectrometry-based proteomics.* Nature, 2007. **450**(7172): p. 991-1000.
86. Malmstrom, J., H. Lee, and R. Aebersold, *Advances in proteomic workflows for systems biology.* Curr Opin Biotechnol, 2007. **18**(4): p. 378-84.
87. Peng, J. and S.P. Gygi, *Proteomics: the move to mixtures.* J Mass Spectrom, 2001. **36**(10): p. 1083-91.
88. Havlis, J. and A. Shevchenko, *Absolute quantification of proteins in solutions and in polyacrylamide gels by mass spectrometry.* Anal Chem, 2004. **76**(11): p. 3029-36.
89. Havlis, J., et al., *Fast-response proteomics by accelerated in-gel digestion of proteins.* Anal Chem, 2003. **75**(6): p. 1300-6.

90. Shevchenko, A., et al., *In-gel digestion for mass spectrometric characterization of proteins and proteomes.* Nat Protoc, 2006. **1**(6): p. 2856-60.
91. de Godoy, L.M., et al., *Status of complete proteome analysis by mass spectrometry: SILAC labeled yeast as a model system.* Genome Biol, 2006. **7**(6): p. R50.
92. Delahunty, C.M. and J.R. Yates, 3rd, *MudPIT: multidimensional protein identification technology.* Biotechniques, 2007. **43**(5): p. 563, 565, 567 passim.
93. Wolters, D.A., M.P. Washburn, and J.R. Yates, 3rd, *An automated multidimensional protein identification technology for shotgun proteomics.* Anal Chem, 2001. **73**(23): p. 5683-90.
94. Chen, E.I., et al., *Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT).* Mol Cell Proteomics, 2006. **5**(1): p. 53-6.
95. de Godoy, L.M., et al., *Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast.* Nature, 2008. **455**(7217): p. 1251-4.
96. Horth, P., et al., *Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis.* Mol Cell Proteomics, 2006. **5**(10): p. 1968-74.
97. Chenau, J., et al., *Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling.* Proteome Sci, 2008. **6**: p. 9.
98. Hubner, N.C., S. Ren, and M. Mann, *Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis.* Proteomics, 2008. **8**(23-24): p. 4862-72.
99. Wisniewski, J.R., A. Zougman, and M. Mann, *Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome.* J Proteome Res, 2009. **8**(12): p. 5674-8.
100. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nat Biotechnol, 2008. **26**(12): p. 1367-72.
101. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics.* Nature, 2003. **422**(6928): p. 198-207.
102. Ong, S.E. and M. Mann, *Mass spectrometry-based proteomics turns quantitative.* Nat Chem Biol, 2005. **1**(5): p. 252-62.
103. Wilkins, M.R., et al., *Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it.* Biotechnol Genet Eng Rev, 1996. **13**: p. 19-50.
104. Anderson, N.L. and N.G. Anderson, *Proteome and proteomics: new technologies, new concepts, and new words.* Electrophoresis, 1998. **19**(11): p. 1853-61.
105. Mann, M., P. Hojrup, and P. Roepstorff, *Use of mass spectrometric molecular weight information to identify proteins in sequence databases.* Biol Mass Spectrom, 1993. **22**(6): p. 338-45.

106. Henzel, W.J., et al., *Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases.* Proc Natl Acad Sci U S A, 1993. **90**(11): p. 5011-5.
107. James, P., et al., *Protein identification by mass profile fingerprinting.* Biochem Biophys Res Commun, 1993. **195**(1): p. 58-64.
108. Pappin, D.J., P. Hojrup, and A.J. Bleasby, *Rapid identification of proteins by peptide-mass fingerprinting.* Curr Biol, 1993. **3**(6): p. 327-32.
109. Jonscher, K.R. and J.R. Yates, 3rd, *The quadrupole ion trap mass spectrometer--a small solution to a big challenge.* Anal Biochem, 1997. **244**(1): p. 1-15.
110. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
111. Hunt, D.F., et al., *Protein sequencing by tandem mass spectrometry.* Proc Natl Acad Sci U S A, 1986. **83**(17): p. 6233-7.
112. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Mol Cell Proteomics, 2002. **1**(5): p. 376-86.
113. Wisniewski, J.R., et al., *Universal sample preparation method for proteome analysis.* Nat Methods, 2009. **6**(5): p. 359-62.
114. Rappsilber, J., Y. Ishihama, and M. Mann, *Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics.* Anal Chem, 2003. **75**(3): p. 663-70.
115. Lee, J.E., et al., *The steady-state repertoire of human SCF ubiquitin ligase complexes does not require ongoing Nedd8 conjugation.* Mol Cell Proteomics, 2011. **10**(5): p. M110 006460.
116. Kalli, A. and S. Hess, *Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer.* Proteomics, 2012. **12**(1): p. 21-31.
117. Graumann, J., et al., *Applicability of tandem affinity purification MudPIT to pathway proteomics in yeast.* Mol Cell Proteomics, 2004. **3**(3): p. 226-37.
118. McCormack, A.L., et al., *Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level.* Anal Chem, 1997. **69**(4): p. 767-76.
119. Link, A.J., et al., *Direct analysis of protein complexes using mass spectrometry.* Nat Biotechnol, 1999. **17**(7): p. 676-82.
120. Olsen, J.V., S.E. Ong, and M. Mann, *Trypsin cleaves exclusively C-terminal to arginine and lysine residues.* Mol Cell Proteomics, 2004. **3**(6): p. 608-14.
121. Gomez, T.A., et al., *Identification of a functional docking site in the Rpn1 LRR domain for the UBA-UBL domain protein Ddi1.* BMC Biol, 2011. **9**: p. 33.
122. Chan, N.C., et al., *Broad activation of the ubiquitin-proteasome system by Parkin is critical for mitophagy.* Hum Mol Genet, 2011. **20**(9): p. 1726-37.
123. Verma, R., et al., *Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome.* Elife, 2013. **2**: p. e00308.
124. Wilcox, A.J. and J.D. Laney, *A ubiquitin-selective AAA-ATPase mediates transcriptional switching by remodelling a repressor-promoter DNA complex.* Nat Cell Biol, 2009. **11**(12): p. 1481-6.

125. Kirchner, P., M. Bug, and H. Meyer, *Ubiquitination of the N-terminal Region of Caveolin-1 Regulates Endosomal Sorting by the VCP/p97 AAA-ATPase.* J Biol Chem, 2013. **288**(10): p. 7363-72.
126. Zhang, L., et al., *Fas-associated factor 1 is a scaffold protein that promotes beta-transducin repeat-containing protein (beta-TrCP)-mediated beta-catenin ubiquitination and degradation.* J Biol Chem, 2012. **287**(36): p. 30701-10.
127. Lee, J.J., et al., *Complex of Fas-associated Factor 1 (FAF1) with Valosin-containing Protein (VCP)-Npl4-Ufd1 and Polyubiquitinated Proteins Promotes Endoplasmic Reticulum-associated Degradation (ERAD).* J Biol Chem, 2013. **288**(10): p. 6998-7011.
128. Lee, J.J., et al., *Ubiquitin-associated (UBA) domain in human Fas associated factor 1 inhibits tumor formation by promoting Hsp70 degradation.* PLoS One, 2012. **7**(8): p. e40361.
129. Rumpf, S. and S. Jentsch, *Functional division of substrate processing cofactors of the ubiquitin-selective Cdc48 chaperone.* Mol Cell, 2006. **21**(2): p. 261-9.
130. Kandasamy, P., et al., *Regulation of unsaturated fatty acid biosynthesis in Saccharomyces: the endoplasmic reticulum membrane protein, Mga2p, a transcription activator of the OLE1 gene, regulates the stability of the OLE1 mRNA through exosome-mediated mechanisms.* J Biol Chem, 2004. **279**(35): p. 36586-92.
131. Mayor, T. and R.J. Deshaies, *Two-step affinity purification of multiubiquitylated proteins from Saccharomyces cerevisiae.* Methods Enzymol, 2005. **399**: p. 385-92.
132. Pierce, N.W., et al., *Cand1 Promotes Assembly of New SCF Complexes through Dynamic Exchange of F Box Proteins.* Cell, 2013. **153**(1): p. 206-15.
133. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
134. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists.* Nucleic Acids Res, 2009. **37**(1): p. 1-13.
135. Benjamini, Y., et al., *Controlling the false discovery rate in behavior genetics research.* Behav Brain Res, 2001. **125**(1-2): p. 279-84.
136. DeBose-Boyd, R.A., *Feedback regulation of cholesterol synthesis: sterol-accelerated ubiquitination and degradation of HMG CoA reductase.* Cell Res, 2008. **18**(6): p. 609-21.
137. Yen, J.L., et al., *Signal-induced disassembly of the SCF ubiquitin ligase complex by Cdc48/p97.* Mol Cell, 2012. **48**(2): p. 288-97.
138. Bandau, S., et al., *UBXN7 docks on neddylated cullin complexes using its UIM motif and causes HIF1alpha accumulation.* BMC Biol, 2012. **10**: p. 36.
139. Stukey, J.E., V.M. McDonough, and C.E. Martin, *Isolation and characterization of OLE1, a gene affecting fatty acid desaturation from Saccharomyces cerevisiae.* J Biol Chem, 1989. **264**(28): p. 16537-44.
140. Kee, Y., N. Lyon, and J.M. Huibregtse, *The Rsp5 ubiquitin ligase is coupled to and antagonized by the Ubp2 deubiquitinating enzyme.* EMBO J, 2005. **24**(13): p. 2414-24.

141. Lee, J.N., et al., *Unsaturated fatty acids inhibit proteasomal degradation of Insig-1 at a postubiquitination step.* J Biol Chem, 2008. **283**(48): p. 33772-83.
142. Lee, J.N., et al., *Identification of Ubxd8 protein as a sensor for unsaturated fatty acids and regulator of triglyceride synthesis.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21424-9.
143. LaLonde, D.P. and A. Bretscher, *The UBX protein SAKS1 negatively regulates endoplasmic reticulum-associated degradation and p97-dependent degradation.* J Biol Chem, 2011. **286**(6): p. 4892-901.
144. Volmer, R., K. van der Ploeg, and D. Ron, *Membrane lipid saturation activates endoplasmic reticulum unfolded protein response transducers through their transmembrane domains.* Proc Natl Acad Sci U S A, 2013. **110**(12): p. 4628-33.