Regulation of Protein Turnover in Fission Yeast

Von der Fakultät Energie-, Verfahrens- und Biotechnik der Universität Stuttgart
zur Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

vorgelegt von

Michael Schmidt
aus Augsburg

Hauptberichter: Prof. Dr. Dieter H. Wolf
Mitberichter: PD Dr. Wolfgang Hilt
Tag der mündlichen Prüfung: 04.03.2008

Institut für Biochemie der Universität Stuttgart

2008

TABLE OF CONTENTS

ABBREVIATIONS

1D	one-dimensional
2D	two-dimensional
2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
ASC	adjusted spectrum counts
ATP	adenosine triphosphate
ATP-RS	ATP regeneration system
BTB	Bric-a-brac, Tramtrack, Broad-complex
CRL	cullin-RING ligase
CSN	COP9 signalosome
DTT	1,4-dithio-treitol
DUB	deubiquitylating enzyme
EMM	Edinburgh minimal medium
ESI	electro spray ionisation
FBP	F-box protein
GO	gene ontology
HEAT	huntington-elongation-A subunit-TOR
HECT	homologous to the E6-AP carboxyl-terminus
HPV	human papillomavirus
IAM	iodoacetamide
IB	immunoblot
IEF	immobilized pH gradient
IEF	isoelectric focussing
IP	immun precipitation
IPTG	isopropyl-b-1-thio-D-galactopyranoside
LB	Luria broth
LC	liquid chromatograph
MS	mass spectrometry
MW	molecular weight
NEM	N-ethylmaleimide
OD	optical density
ORF	open reading frame
PCR	polymerase chain reaction
pI	isoelectric point
PMSF	phenylmethylsulfonylfluoride
PVDF	PolyVinyliDene Fluoride
qRT/PCR	quantitative real-time polymerase chain reaction
RING	really interesting new gene
RT	room temperature
RT/PCR	reverse transcriptase polymerase chain reaction
RX	reaction buffer
SAX	strong anion excahenge

SCF	Skp1-Cul1-F-Box
SCX	strong cation exchange
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPASS	Solid Phase Assay for Systematic profiling of ubiquitylation Substrates
TCA	trichloroacetic acid
TCEP	Tris(2-carboxyethyl)phosphine
Ub	ubiquitin
UPS	ubiquitin proteasome system
WB	Western blot
YES	yeast extract and supplements

ABSTRACT

Proteins are molecular machines that perform the majority of cellular functions that characterize living organisms. Protein levels within cells are the result of a delicate balance between protein synthesis and protein degradation. Eukaryotic cells are able to shift this balance to one or the other side to regulate the steady state level of each protein within a cell. If a protein or a functional group of proteins is needed to fulfill a certain duty, cells have evolved a number of pathways to increase the levels of these proteins rapidly. On the other hand, if a protein or functional group of proteins is no longer required, eukaryotes are able to rapidly decrease the concentrations of these proteins. While historically much attention and research has been devoted to how proteins are synthesized, the reverse process, i.e. how and when proteins are degraded, is not understood as well. This thesis is structured into three parts, addressing regulation of protein turnover on a molecular as well as on a system-wide level. In part one, to gain deeper insight into how eukaryotes regulate targeted protein degradation, the molecular mechanisms of regulation of Cullin-RING ubiquitin ligases (CRLs) by the COP9 signalosome (CSN), the deubiquitylating enzyme Ubp12p, and the cullin-associated protein Can1p were investigated in the fission yeast *Schizosaccharomyces pombe*. In a survey of eight F-box proteins, which confer target specificity to CRLs, the study uncovered the existence of variant F-box proteins lacking a critical proline residue required for efficient regulation by the CSN. The results suggest that distinctive features of the F-box motif specify the assembly of F-box proteins into CRL complexes thus destining them for regulation by the CSN through a mechanism, which can principally function independently of Can1p and Ubp12p. In part two, a large-scale biochemical approach was used to investigate whether E2 ubiquitin conjugating enzymes contribute to substrate selection. Because of the lack of a reliable method for the identification of specific ubiquitylation substrates, a new biochemical method termed SPASS allowing the study of specific E2 substrates was developed. SPASS was used to identify on a global scale ubiquitylation substrates of the *S. pombe* E2s Ubc7p and Ubc8p. Examination of the previously known and also the novel substrates revealed a potential new mechanism by which E2s confer substrate specificity utilizing heterodimerisation. Lastly in part three, to gain a system-wide insight on the regulation of protein expression in *S. pombe*, mRNA as well as protein levels of 1500 open reading frames were explored. For this purpose, a new label-free mass spectrometry based

method allowing relative quantification of protein levels was developed. The fission yeast protein data showed considerable correlations with mRNA levels and with the abundance of orthologous proteins in budding yeast. Functional pathway analysis indicated that the mRNA–protein correlation is strong for proteins involved in signalling and metabolic processes, but increasingly discordant for components of protein complexes, which clustered in groups with similar mRNA–protein ratios. Self-organizing map clustering of large-scale protein and mRNA data from fission and budding yeast revealed coordinate but not always concordant expression of components of functional pathways and protein complexes.

ZUSAMMENFASSUNG

Proteine sind molekulare Maschinen die den Grossteil der zellulären Vorgänge ausführen und damit allen Organismen Leben ermöglicht. In Zellen steht die vorhandene Menge eines jeden Genprodukts im Gleichgewicht von Synthese und Abbau. Um die Menge eines jeden einzelnen zellulären Proteins zu kontrollieren, können Eukaryonten dieses Gleichgewicht in Richtung Synthese oder Abbau verschieben. Falls eine Zelle ein bestimmtes Protein, oder eine funktionelle Gruppe an Proteinen zu einem bestimmten Zeitpunkt benötigt, muss sie die Menge dieser Proteine rasch erhöhen können und hat dazu Synthesewege evolviert. Falls ein Protein oder eine funktionelle Gruppe an Proteinen nicht mehr benötigt wird, können Eukaryonten diese Proteine sehr schnell abbauen. Wohingegen Proteinsynthese und Proteinabbau bei der Regulation von Zellfunktionen eine gleich wichtige Rolle spielen, sind aus historischen Gründen sind die Synthesewege besser erforscht als die Abbauwege. Daher gliedert sich diese Arbeit in drei Teile, welche die Regulation des Proteinumsatzes auf molekularer sowie auf systemweiter Ebene untersuchen. Der erste Teil setzt sich mit der molekularen Regulation von Cullin-RING Ubiquitinligasen (CRLs) durch das COP9 Signalosom (CSN), das deubiquityliernde Enzym Ubp12p und dem cullin-assoziirten Protein Can1p auseinander. Experimente wurden dazu in der Spalthefe *Schizosaccharomyces pombe* durchgeführt. Die Studie von acht F-box Proteinen, welche für die Spezifität von Ubiquitinligasen verantwortlich sind, deckte die Existenz einer Variante des F-box Motivs auf. Diese F-Box Variante, welcher ein spezifisches Prolin fehlt, kann nicht von dem Signalosom reguliert werden. Die Ergebnisse deuten an, dass die unterschiedliche Zusammensetzung des F-Box Motivs den Einbau der F-Box Proteine in CRLs und damit die Regulation durch das Signalosom bestimmt. Es wird gezeigt, dass dieser Mechanismus prinzipiell ohne Ubp12p und Can1p abläuft. Im zweiten Teil, wurde ein umfangreicher biochemischer Assay neu entwickelt, welcher für die Erforschung der Rolle von E2 ubiquitinkonjugierenden Enzymen bei der Substratauswahl angewandt wurde. Der SPASS benannte Assay wurde für die globale Identifikation von *S. Pombe* Ubc7p und Ubc8p Substraten angewandt. Die Untersuchungen der Ergebnisse, welche bekannte und neue E2 Substrate enthielten, deckten eine potenzielle neue Funktionsweise von E2s bezüglich der Substratspezifität auf, welche Heterodimerisierung andeutet. Um eine systemweite Einsicht in die Regulation von Proteinlevels in *S. pombe* zu bekommen, wurden die Expressionslevel

von 1500 Genprodukten auf mRNA- und Proteinebene untersucht. Dazu wurde eine neue, auf Massenspektrometrie beruhende Methode für die Quantifizierung von Proteinlevel entwickelt. Die Ergebnisse offenbarten eine starke Korrelation der Proteinlevel mit den jeweiligen mRNA mengen, sowie als auch mit den Expressionsleveln von orthologen Proteinen der Bäckerhefe *S. cerevisiae*. Eine funktionelle Pathwayanalyse ergab eine hohe Protein-mRNA Korrelation für Proteine der Kategorien Signalling und metabolische Prozesse, wohingegen eine abnehmende Koordination für Untereinheiten von Proteinkomplexen festgestellt wurden. Clusteranalyse von *S. pombe* und *S. cerevisiae* Gen- und Proteinexpression durch self-organizing Maps ergab, dass funktionelle Pathways und Proteinkomplexe koordiniert aber in den zwei Organismen nicht immer übereinstimmend exprimiert sind.

INTRODUCTION

Schizosaccharomyces pombe: A Model for Studies of Eukaryotic Genes

Several features of the fission yeast *Schizosaccharomyces pombe* make it exceptionally well suited for the study of eukaryotic gene expression. It is a relatively simple eukaryote that can be readily grown and manipulated in the laboratory, using a variety of highly developed and sophisticated methodologies. *Schizosaccharomyces pombe* cells share many molecular, genetic, and biochemical features with cells from multicellular organisms, making it a particularly useful model to study the structure, function, and regulation of genes from more complex species. Fission yeast has been used as a model eukaryote to study processes such as the cell cycle and cell morphology among many others. Recognized as a species over a century ago, the cone-shaped binary dividing organism (Figure I) started to become the subject of more intensive experimentation in the 1950s. It attracted interest from cell biologists because its cell division is more typical of most eukaryotes and is distinct from that of budding yeasts. Genetic and molecular biology methods in *S. pombe* are well developed as a result of several decades of research in this organism. Methods for strain maintenance, vegetative growth, genetic crosses and mutant isolation had mainly been worked out by the 1970s (Gutz et al., 1974). Marking the onset of molecular biology in this organism, DNA transformation was established two decades ago (Beach and Nurse, 1981). *S. pombe* is haploid throughout most of its life cycle. Two mating types exist: minus and plus. Laboratory strains in which natural mating-type switching is inactivated and thus are locked into one of the mating types were generated. The diploid phase is normally restricted to the zygote, which goes straight to meiosis without intervening mitoses (Figure II).

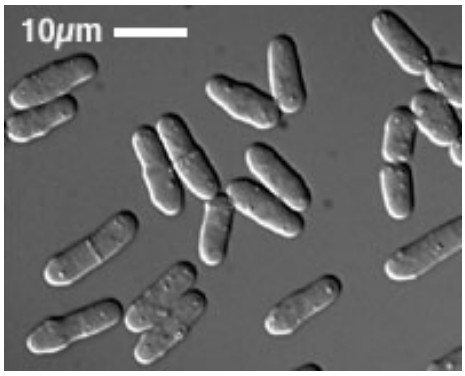


Figure I:**Light microscopic image of *S. pombe* cells (1000x)**

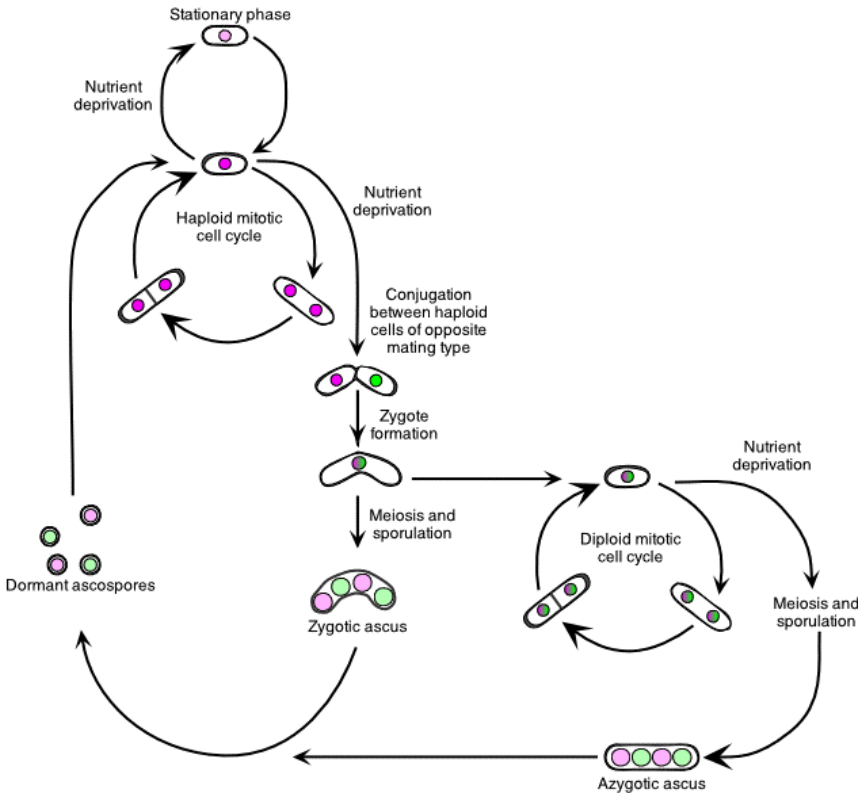


Figure II: **The S. *pombe* life cycle.** Adapted from MacNeill and Nurse, 1997. The magenta and green colors indicate the two different mating types (+ and -).

Recent calculations (Heckman et al., 2001) indicate that *S. pombe* separated from *Ascomycetes* more than 1,100 million years ago (Figure III). The full genome sequence which contains the smallest number of protein-coding genes yet recorded for a eukaryote (4,940) was published in 2002 (Wood et al., 2002). All the genetic information in *S. pombe* is contained in three chromosomes. Whereas the fraction of genes with introns is 5% in *S.*

cerevisiae, in *S. pombe* 43% of the genes have introns. The proportion of the genome that codes proteins in fission yeast is 60%, which corresponds to the 4,940 ORFs and indicates a lower gene density than in budding yeast. Nevertheless, the proportion of functional categories for proteins is essentially identical in budding and fission yeast (Wood et al., 2002). A comparison between the proteins that have been characterized in both fission and budding yeast showed that two thirds of *S. pombe* proteins have homologues in *S. cerevisiae*. There is a considerable number of proteins with a high level of conserved homology between the two yeasts and humans. Many of those genes are involved in human diseases. In addition, homologues to genes linked to human diseases have been characterized in *S. pombe* that are absent in *S. cerevisiae*.

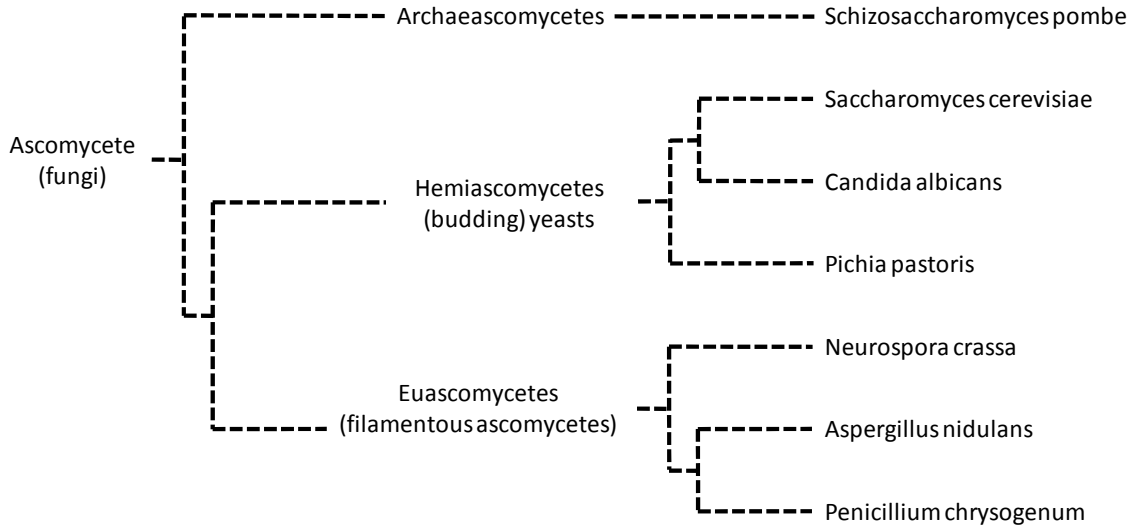


Figure III: **Phylogenetic relationship of selected fungi** (adapted from Franz Lang, University of Montreal)

The development of cDNA microarrays to measure gene expression profiles of entire organisms (Schena et al., 1995) revolutionized genetic analysis and allowed researchers to study gene expression globally with applications from understanding gene function in model organisms like yeasts to personalized medicine. For a long time, the standard method to establish a global picture of the protein content (proteome) of an organism has been protein two-dimensional gel electrophoresis (2D-PAGE). Given a method of assigning identities to the protein spots, it gives information about the amount and degree of posttranslational modification of the 500–1,000 most abundant proteins in a cell. New advances in mass spectrometry instrumentation allow the development and application of novel methods for the quantitative large scale measurement of protein expression, what should enable to draw

conclusions about the wiring of underlying molecular control mechanisms. The potential of global analyses of transcript and protein profiles in model organisms like fission yeast is becoming increasingly obvious with the appearance of "systems biology" tools that allow a more intuitive understanding of underlying regulation schemes.

The Ubiquitin-Proteasome System

The ubiquitin-proteasome system (UPS) represents the major pathway for intracellular protein degradation, a process in which an enzyme system tags unwanted proteins with several molecules of the 76-amino acid residue protein ubiquitin. The tagged proteins are then transported to the proteasome, a large multisubunit protease complex, where they are degraded (Ciechanover et al., 1984). In some circumstances, conjugation of ubiquitin to proteins can also regulate other functional aspects including localization and enzyme activity (reviewed in Schnell and Hicke, 2003). Many cellular processes are regulated by ubiquitin-mediated proteolysis including the cell cycle, DNA repair and transcription, protein quality control and the immune response among many others. Therefore, the number of enzymatic components of the UPS is immense. As an example, over 900 components of the UPS have been identified in *C. elegans*, indicating the importance of the process in cellular fate. Defects in the UPS have a causal role in many diseases, including neurodegenerative, autoimmune, and viral diseases as well as cancer (reviewed in Jiang and Beaudet, 2004).

The UPS can be separated into two steps: substrate ubiquitylation and substrate degradation. The dramatic impact ubiquitin attachment has on its targets – its life or death – led to the evolution of multiple layers of molecular regulation which are not completely resolved yet.

Ubiquitylation

The principles of protein ubiquitylation were established initially by Hershko, Ciechanover and Rose in the early 1980s and are being built upon to this day. The attachment of ubiquitin to a protein (the substrate) can be broken down into several steps (Pickart and Eddins, 2004). The C-terminus of ubiquitin is first activated by a ubiquitin-activating enzyme (E1). Through an ATP-dependent mechanism, ubiquitin is coupled to a cysteine side chain in E1, yielding a reactive E1~ubiquitin thioester intermediate. The activated ubiquitin is

subsequently passed to one of a number of distinct ubiquitin-conjugating enzymes (E2s) by transthiolation to a conserved cysteine of the E2. The E2 proteins catalyze substrate ubiquitylation in conjunction with a ubiquitin-protein ligase (E3) (Figure IV). Ubiquitin is covalently linked, forming an isopeptide bond between the terminal amino group of the target substrate lysine and the C-terminal glycine of ubiquitin. The process of ubiquitin conjugation is thought to be repeated in a way so that a following ubiquitin molecule is conjugates to one of the seven lysine residues (K6, K11, K27, K29, K33, K48 and K63) of the previous one (Hochstrasser, 2006). Once a poly-ubiquitin chain of a at least four moieties is formed, the substrate is recognized by the proteasome and subsequently degraded (Thrower et al., 2000). Contrary, attachment of a single ubiquitin molecule (monoubiquitylation) to a substrate, is not inducing proteasomal degradation, and has been implicated in processes like lysosomal sorting and trafficking, gene regulation and silencing among many others (reviewed in Schnell and Hicke, 2003).

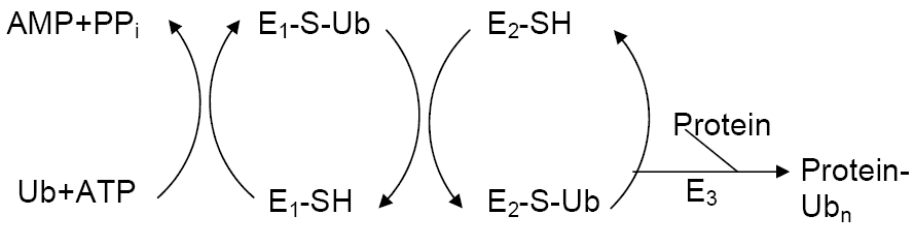


Figure IV: **Schematic outline of protein ubiquitylation reaction**. Adapted from Hershko et al., 1983.

Whereas only one E1 has been identified per organism, the existence of multiple, biochemically distinct E2s was initially demonstrated by Pickart and Rose (1985). We now know that mammalian cells contain many E2s and several hundred E3s (reviewed in Petroski and Deshaies, 2005). It became clear that the recognition of substrates as well as enzymatic ubiquitin conjugation must be precisely timed. Inappropriately ubiquitylated substrates would be degraded immediately making them unavailable to fulfill their normal functions, whereas too little ubiquitylation would lead to accumulation of proteins. Aberrant ubiquitylation can have dramatic consequences such as the development of diseases. The role of orchestrating substrate recognition and ubiquitin conjugation has been attributed to the E3 ubiquitin ligases (Hershko et al., 1983).

Ubiquitin Ligases

Two classes of E3s are conserved from yeast to man: homologous to the E6-AP carboxyl-terminus (HECT) domain ligases and really interesting new gene (RING) domain ligases. HECT domain E3s, which were originally characterized in studies with the human papillomavirus (HPV) (reviewed in Pickart, 2001), bind ubiquitin in a thioester bond utilizing a conserved cysteine residue before transfer onto a substrate. RING domain E3s do not form this catalytic thioester attachment between themselves and ubiquitin. It is thought that these E3s, with the help of their RING domains, provide the proximity for the substrate and the ubiquitin-charged E2, creating a chemical environment that allows the direct transfer of ubiquitin onto a lysine residue within the substrate, but the exact mechanism is unknown (Wu et al., 2003)

Some of the best-studied RING domain ubiquitin ligases are cullin-based ligases (CRLs) (Skowyra et al., 1997, Lyapina et al., 1998, Patton et al., 1998, Kominami et al., 1998). All CRLs share a similar modular complex topology (Figure V) (reviewed in Petroski and Deshaies, 2005) composed of a cullin core module, a C-terminally bound RING domain protein Rbx1p (Hrt1p, Roc1p) and varying adapters that recruit substrates to the N-terminal side of the cullin. The SCF complex is composed of the RING domain protein Rbx1p, the backbone protein Cul1p and variable F-Box proteins. The F-Box proteins are linked to Cul1p via an additional bridging protein Skp1p. The CRL3 complex is composed of the Cul3p backbone as well as the Rbx1p RING domain and variable Btb-domain substrate adapters (Geyer et al., 2003). Other cullin associated ligase complexes share a similar modular composition (Figure V). There are seven different cullins in humans (CUL1-7) and three in fission yeast (Cul1p, Cul3p and Cul4p).

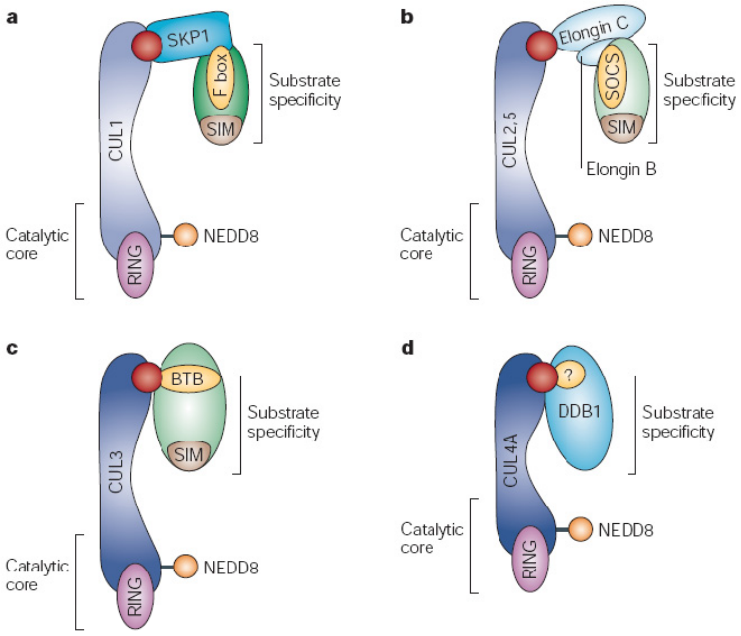


Figure V: **Modularity of cullin-RING ligases** (adapted from Petroski and Deshaies, 2005). a) CRL1/SCF complex. b) CRL2/5 complex. c) CRL3 complex. d) CRL4 complex.

Regulation of Protein Ubiquitylation

Studies performed mainly on CRLs have demonstrated that ubiquitin-dependent substrate degradation is subject to several levels of post-translational as well as transcriptional regulation. Posttranslational substrate modification is ordinarily required for recognition by CRLs (Petroski and Deshaies, 2005). At the level of the ligases, modification of cullins with the ubiquitin-related protein Nedd8 is essential for CRL activity in all organisms except budding yeast (reviewed in Pan et al., 2004). Further, the COP9 signalosome (CSN), a multiprotein complex, serves as an assembly and maintenance platform for CRLs, providing a favorable biochemical environment for exchange of adapters and thus for substrate specificity (Wolf et el., 2003, Wee et al., 2005). In addition, few cases are known where substrate adapter expression levels underlie transcriptional control, particularly in the adaptation of cells to environmental perturbations.

Objectives

The aim of this work was to deepen the understanding of how eukaryotes regulate protein turnover. Therefore, molecular studies on a prominent member of the RING-domain ubiquitin ligases, the SCF/CRL1 complex were conducted to investigate regulatory mechanisms of ubiquitylation. To gain a system-wide view on directed protein ubiquitylation, new biochemical tools allowing the comprehensive study of a multitude of proteins were developed and applied to study protein turnover in fission yeast.

References

Beach D, Nurse P (1981) High frequency transformation of the fission yeast Schizosaccharomyces pombe. Nature 290:140–142

Ciechanover A, Finley D, Varshavsky A. (1984) The ubiquitin-mediated proteolytic pathway and mechanisms of energy-dependent intracellular protein degradation. J Cell Biochem. 1984;24(1):27-53.

Geyer R, Wee S, Anderson S, Yates J, Wolf DA. (2003) BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. Mol Cell. 2003 Sep;12(3):783-90.

Gutz H, Heslot H, Leupold U, Loprieno N (1974) Schizosaccharomyces pombe. In: King RC (ed) Handbook of genetics, vol 1:

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001) Molecular evidence for the early colonization of land by fungi and plants. Science 293:1129–1133.

Hershko A, Heller H, Elias S, Ciechanover A. (1983) Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J Biol Chem. 1983 Jul 10;258(13):8206-14.

Hochstrasser M (2006) Lingering mysteries of ubiquitin-chain assembly. Cell. 2006 Jan 13;124(1):27-34.

Jiang Y.H., Beaudet A.L. (2004) Human disorders of ubiquitination and

Kominami K, Ochotorena I, Toda T. (1998) Two F-box/WD-repeat proteins Pop1 and Pop2 form hetero- and homo-complexes together with cullin-1 in the fission yeast SCF (Skp1-Cullin-1-F-box) ubiquitin ligase. Genes Cells. 1998 Nov;3(11):721-35.

Lyapina SA, Correll CC, Kipreos ET, Deshaies RJ. (1998) Human CUL1 forms an evolutionarily conserved ubiquitin ligase complex (SCF) with SKP1 and an F-box protein. Proc Natl Acad Sci U S A. 1998 Jun 23;95(13):7451-6.

MacNeill SA and Nurse P (1997) Cell cycle control in fission yeast. In The Molecular and Cellular Biology of the Yeast Saccharomyces. Pp. 697-763. Cold Spring Harbor Laboratory Press.

Pan, Z. Q., Kentsis, A., Dias, D. C., Yamoah, K. & Wu, K. (2004) Nedd8 on cullin: building an expressway to protein destruction. Oncogene 23, 1985–1997 .

Patton EE, Willems AR, Sa D, Kuras L, Thomas D, Craig KL, Tyers M. (1998) Cdc53 is a scaffold protein for multiple Cdc34/Skp1/F-box proteincomplexes that regulate cell division and methionine biosynthesis in yeast. Genes Dev. 1998 Mar 1;12(5):692-705. Erratum in: Genes Dev 1998 Oct 1;12(19):3144.

Petroski MD, Deshaies RJ. (2005) Function and regulation of cullin-RING ubiquitin ligases. Nat Rev Mol Cell Biol. 2005 Jan;6(1):9-20.

Pickart CM, Eddins MJ. (2004) Ubiquitin: structures, functions, mechanisms. Biochim Biophys Acta. 2004 Nov 29;1695(1-3):55-72.

Pickart, C.M. 2001. Mechanisms underlying ubiquitination. Annu Rev Biochem 70: 503-33

Pickart, C.M., and Rose, I.A. (1985) Functional heterogeneity of ubiquitin carrier proteins. J. Biol. Chem. 260, 1573-1581.

Schena M, Shalon D, Davis RW, Brown PO. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science. 1995 Oct 20;270(5235):467-70.

Schnell JD, Hicke L. (2003) Non-traditional functions of ubiquitin and ubiquitin-binding proteins. J Biol Chem. 2003 Sep 19;278(38):35857-60.

Skowyra D, Craig KL, Tyers M, Elledge SJ, Harper JW. (1997) F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. Cell. 1997 Oct 17;91(2):209-19.

Thrower, J.S., Hoffman, L., Rechsteiner, M., and Pickart, C.M. (2000). Recognition of the polyubiquitin proteolytic signal. EMBO J. 19, 94–102.

Wee S, Geyer RK, Toda T, Wolf DA. (2005) CSN facilitates Cullin-RING ubiquitin ligase function by counteracting autocatalytic adapter instability. Nat Cell Biol. 2005 Apr;7(4):387-91.

Wolf DA, Zhou C, Wee S. (2003) The COP9 signalosome: an assembly and maintenance platform for cullin ubiquitin ligases? Nat Cell Biol. 2003 Dec;5(12):1029-33

Wood V, Gwilliam R, Rajandream M-A, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson D, Quail MA, Rabbinowitsch E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schafer M, Muller-Auer S, Gabel C, Fuchs M, Fritzc C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dreano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sanchez M, del Rey F, Benito J, Dominguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerrutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P (2002) The genome sequence of Schizosaccharomyces pombe. Nature 415: 871–880

Wu, P.Y., Hanlon, M., Eddins, M., Tsui, C., Rogers, R.S., Jensen, J.P., Matunis, M.J., Weisman, A.M., Wolberger, C.P., and Pickart, C.M. (2003). A conserved catalytic residue in the ubiquitin-conjugating enzyme family. EMBO J. 22, 5241–5250.

CHAPTER I: F-Box-directed CRL Complex Assembly and Regulation by the CSN and CAND1

1.1. Background

CRLs represent an extensive class of multisubunit E3 ubiquitin ligases each consisting of a core module containing a member of the cullin family and the RING domain protein RBX1 (= HRT1, ROC1), which recruits E2 ubiquitin conjugating enzymes to the ligases (reviewed in (Petroski & Deshaies, 2005)). This core is joined by one of several hundred adapter proteins each of which appears to target a distinct array of substrates for ubiquitylation and proteasomal degradation. Whereas FBPs are tethered to the CUL1 core through the linker protein SKP1 to form SCF (or CRL1) complexes (Deshaies, 1999), CUL3 adapters are recruited into CRL3 complexes via their inherent BTB domains (Geyer et al., 2003; Pintard et al., 2003; Xu et al., 2003). Several members of both adapter families are unstable proteins, due to autoubiquitylation by the intrinsic ubiquitin ligase activity of their associated core modules (Galan & Peter, 1999; Geyer et al., 2003; Luke-Glaser et al., 2007; Rouillon et al., 2000; Wirbelauer et al., 2000; Zhou & Howley, 1998) .

CRLs are stimulated through modification of cullins with the ubiquitin-related peptide NEDD8 (reviewed in Pan et al., 2004). Cullin neddylation is reversed by the COP9 signalosome (CSN), a highly conserved protein complex that binds cullins (Lyapina et al., 2001; Schwechheimer et al., 2001; Zhou et al., 2001) and thereby exposes them to a deneddylating activity intrinsic to subunit 5 of the CSN (Cope et al., 2002). Consistent with neddylation being a stimulatory modification, purified CSN inhibits CRL activity *in vitro*. In addition, the CSN-associated deubiquitylating enzyme (DUB) Ubp12/USP15 acts to neutralize CRL activity when assayed *in vitro* (Groisman et al., 2003; Zhou et al., 2003).

These biochemical studies conflicted with genetic studies showing that CSN is required for efficient CRL-dependent substrate degradation *in vivo* (Reviewed in Cope & Deshaies, 2003; von Arnim, 2003; Wolf et al., 2003). This so-called CSN paradox was resolved by the demonstration that CSN's inhibitory enzymatic activities revealed *in vitro* serve to prevent the autocatalytic degradation of CRL substrate adapters thus promoting CRL activity *in vivo*. For example, the Cul3p adapter Btb3p is considerably less stable in fission yeast *csn* and *ubp12* mutants (Wee et al., 2005). In addition, several FBPs were since shown to be

destabilized in *csn* mutants of *S. pombe* (Zhou et al., 2003) and *N. crassa* (He et al., 2005), and in human CSN knockdown cells (Cope & Deshaies, 2006; Denti et al., 2006).

CAND1 is a highly conserved protein that binds to the unneddylated form of human and plant CUL1 and inhibits complex formation with SKP1-FBP modules (Hwang et al., 2003; Liu et al., 2002; Min et al., 2003; Oshikawa et al., 2003; Zheng et al., 2002a). By virtue of these properties, CAND1 inhibits CRL activity *in vitro*. However, CAND1 was also shown to be required for efficient CRL function *in vivo* (Chuang et al., 2004; Feng et al., 2004; Zheng et al., 2002a). This contradictory pattern is reiterating the CSN paradox, and it was therefore suggested that CAND1 and CSN participate in the same pathway of CRL adapter stabilization (Cope & Deshaies, 2003; He et al., 2005; Min et al., 2005), although experimental proof is still outstanding. It also remained unclear whether the model is broadly applicable to all FBPs. In the present report, we have examined these questions by determining the regulation of a panel of eight FBPs from fission yeast by CSN and CAND1.

1.2. Materials and Methods

Yeast Strains and Techniques

The *can1* and *ubp1* deletion strains and the epitope-tagged strains indicated in Table 1-1 were constructed by one-step gene replacement using PCR-generated fragments containing *ura4* or kanamycin cassettes (Bahler et al., 1998). The *ubp1 ubp12* double mutant was obtained by mating and confirmed by PCR. Strains containing Myc-tagged FBPs were provided by T. Toda and crossed into our wild-type background and into *csn5* and *can1* mutants. The *skp1-ts can1* mutant was created by mating, followed by verification of the recombinants by colony PCR. The strain was assayed for synthetic phenotypes by spotting serial dilutions exactly as described recently (Wee et al., 2005). N-terminally protein A tagged *can1* was constructed by exactly following the protocol from P. Werler (Werler et al., 2003).

Name	Genotype
M169/2	*leu1-32 ura4-d18 ade6-704 btb3.tev5xproA kan can1::ura4*
Swf418	*leu1-32 ura4-d18 ade6-704 btb3.tev5xproA kan pof1.13myc kan ubp12::ura4*
Swf416	*leu1-32 ura4-d18 ade6-704 btb3.tev5xproA kan pof1.13myc kan*
Swf294	*leu1-32 ura4-d18 ade6-704 btb3.tev5xproA kan*
M161/1	*leu1-32 ura4-d18 ade6-704 can1::ura4 h-*
M169/1	*leu1-32 ura4-d18 ade6-704 can1::ura4 h+*
M216/1	*leu1-32 ura4-d18 ade6-704 csn5.13myc kan can1::ura4*
C399/3	*leu1-32 ura4-d18 ade6-704 csn5.13myc kan*
G44	*leu1-32 ura4-d18 ade6-704 csn5::ura4*
M177/2	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan can1::ura4*
Swf326	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan csn3::ura4*
Swf296	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan csn4::ura4*
Swf297	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan csn4::ura4 csn5::ura5*
Swf389	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan csn5::ura4*
Swf320	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan ubp1::ura4*
Swf361	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan ubp1::ura4 ubp12::ura4*
Swf391	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan ubp12::ura4*
Swf387	*leu1-32 ura4-d18 ade6-704 pof1.13myc kan*
M184/2	*leu1-32 ura4-d18 ade6-704 pof10.13myc kan can1::ura4*
Swf329	*leu1-32 ura4-d18 ade6-704 pof10.13myc kan csn3::ura4*
Swf301	*leu1-32 ura4-d18 ade6-704 pof10.13myc kan csn4::ura4*
Swf365	*leu1-32 ura4-d18 ade6-704 pof10.13myc kan csn4::ura4 csn5::ura4*
Swf386	*leu1-32 ura4-d18 ade6-704 pof10.13myc kan csn5::ura4*
Swf384	*leu1-32 ura4-d18 ade6-704 pof10.13myc kan ubp12::ura4*
Swf279	*leu1-32 ura4-d18 ade6-704 pof10.13myc kan*
M276/1	*leu1-32 ura4-d18 ade6-704 pof10::ura4*
Swf379	*leu1-32 ura4-d18 ade6-704 pof12.13myc kan csn5::ura4*
Swf355	*leu1-32 ura4-d18 ade6-704 pof12.13myc kan*
M215/3	*leu1-32 ura4-d18 ade6-704 pof13.13myc kan csn5::ura4*
Swf387	*leu1-32 ura4-d18 ade6-704 pof13.13myc kan*

Swf395	*leu1-32 ura4-d18 ade6-704 pof3.13myc kan csn5::ura4*
Swf397	*leu1-32 ura4-d18 ade6-704 pof3.13myc kan ubp12::ura4*
Swf393	*leu1-32 ura4-d18 ade6-704 pof3.13myc kan*
Swf401	*leu1-32 ura4-d18 ade6-704 pof7.13myc kan csn5::ura4*
Swf399	*leu1-32 ura4-d18 ade6-704 pof7.13myc kan*
Swf377	*leu1-32 ura4-d18 ade6-704 pof8.13myc kan csn5::ura4*
Swf353	*leu1-32 ura4-d18 ade6-704 pof8.13myc kan*
Swf407	*leu1-32 ura4-d18 ade6-704 pof9.13myc kan csn5::ura4*
Swf409	*leu1-32 ura4-d18 ade6-704 pof9.13myc kan ubp12::ura4*
Swf405	*leu1-32 ura4-d18 ade6-704 pof9.13myc kan*
M141/15	*leu1-32 ura4-d18 ade6-704 proA.tev.can1*
DS448/2	*leu1-32 ura4-d18 ade6-704 h-*
DS448/1	*leu1-32 ura4-d18 ade6-704 h+*
M192/5	*skp1-A7 leu1-32 ura4-d18 ade6-704 can1::ura4*
Swf139	*skp1-A7 leu1-32 ura4-d18 ade6-704 csn5::ura4*
Swf131	*skp1-A7 leu1-32 ura4-d18 ade6-704*

Table 1-1: **Yeast strains used in this study**.

Plasmids

The plasmids expressing Csn5p JAMM mutants were described previously (Wee et al., 2005). Plasmids for FBP binding studies were prepared by amplifying the respective genes from *Schizosaccharomyces pombe* complementary DNA. PCR products were sequenced and then cloned into pREP plasmids, which drive the expression of amino-terminally Myc-epitope-tagged proteins from the thiamine-repressible *nmt1* promoter. FBP point mutants were constructed with the QuickChange site-directed mutagenesis method (Stratagene), following the manufacturer's instructions. Transformed cells were grown in Edinburgh minimal medium (EMM) lacking leucine for selection of positive transformants. Plasmids are listed in Table 1-2.

Name	Plasmid
C429/1	pRep3-myc-csn5
Sw58	pRep3-myc-csn5 H118A
M242/1	pRep3-myc-pof1
M242/3	pRep3-myc-pof1 P114S
M241/1	pRep3-myc-pof9
M241/3	pRep3-myc-pof9 P5S
M257/1	pRep3-myc-pof9 50-467
M209/6a	pRep3-myc-pof10
M209/3e	pRep3-myc-pof10 P35S
M242/5	pRep3-myc-pof12
M242/7	pRep3-myc-pof12 S15P
M193/1	pRep81-myc-pof1
M193/3	pRep81-myc-pof1 P114S

Table 1-2**: Plasmids used in this study**

Immunological Methods

Epitope-tagged proteins were detected by the monoclonal anti-Myc antibody 9E10 or with monoclonal anti-Protein A antibodies (Sigma). Cell lysates for immunoprecipitation were prepared as described (Zhou et al., 2001). Lysates were cleared by centrifugation, and proteins were precipitated with the respective antisera. Immunocomplexes were collected by binding to protein A beads, washed, and analyzed by immunoblotting as described (Zhou et al., 2001). Protein A-tagged proteins were precipitated using whole rabbit immunoglobulin absorbed to Dynabeads. Affinity-purified rabbit antisera against Cul1p, Cul3p, Skp1p, and Rbx1p were described before (Geyer et al., 2003; Seibert et al., 2002). For loading controls PSTAIR (Santa Cruz) antibodies were used at dilutions of 1:1.000.

Cycloheximide Chase Experiments

To measure the stability of adapter proteins, strains containing epitope-tagged versions of FBPs and Btb3p were grown to an OD_{600} of 1.0 in 50ml YES media. 100µg/ml cycloheximide was added, and cultures were incubated at 25°C. 10ml aliquots removed after the times indicated in the figures. Aliquots were diluted in 1mM NaN_3, in order to instantly kill the cells, harvested by centrifugation and frozen in liquid nitrogen. Once all samples of the chase period were obtained, cell lysates were prepared by standard bead lysis as described before (Wee et al., 2005), and equal amounts of proteins were analyzed by immunoblotting.

RT/PCR

Total cellular RNA was isolated by cell lysis in hot RNAzol (TelTest, Inc.) and three cycles each of vortexing in the presence of beads for 5 minutes and heating to 65^{o}C for 5 minutes. The RNA was extracted and precipitated according to the recommendations of the manufacturer. 2µg RNA was used in each RT/PCR reaction. Primer design and RT-PCR conditions were according to the manual supplied with the Platinum Quantitative RT-PCR Thermoscript kit (Invitrogen). The primer concentration for actin amplification was 0.25 µM. Other primers were used at 1 µM. *Pof13* expression was assayed using a iQ Multiplex Powermix enzyme kit (BioRad) following the supplied protocol. Triplicate reactions were analyzed with a iCycler iQ PCR detection system (BioRad).

1.3. Results

Differential Effect of CSN on FBP Levels

In previous studies we demonstrated that the stability of the CRL1 adapter Pop1p and the CRL3 adapter Btb3p is promoted by the CSN *in vivo* (Wee et al., 2005; Zhou et al., 2003). In addition, genetic interaction studies suggested that FBPs other than Pop1p are also subject to regulation by the CSN (Wee et al., 2005). The *S. pombe* genome encodes a minimum of 15 FBPs, but whether all are targets for regulation by the CSN is unknown.

We therefore compared the steady-state protein levels of eight FBPs in wild-type and *csn5* mutant cells. We utilized a panel of strains harbouring FBPs modified at their endogenous genomic loci with C-terminal Myc-epitope tags (Lehmann et al., 2004). As shown in Figure 1-1A Pof1p, 3p, 7p, 9p, and 10p steady-state levels were strongly reduced in *csn5* mutants, whereas Popf8p and Pof12p levels were unaffected. Pof13p was only discretely affected by *csn5* deletion (Fig. 1-1A, lanes 15, 16).

The levels of the CSN-regulated Pof1p and Pof10p were also diminished in *csn3* and *csn4* mutants (Supplementary Fig. 1-1). Conversely, downregulation of Pof10p in *csn5* mutants was complemented by providing wild-type *csn5* from a plasmid (Fig. 1-1B). This rescue failed in *csn4 csn5* double mutants (Fig. 1-1B). In addition, efficient rescue depended on the enzymatic function of the deneddylating enzyme Csn5p, since point mutants in the catalytic JAMM motif were unable to maintain Pof10p levels in *csn5* mutants (Fig, 1-1C). These results suggested that the cullin deneddylation function of the entire CSN complex is required to maintain the steady-state expression levels of CSN-sensitive FBPs.

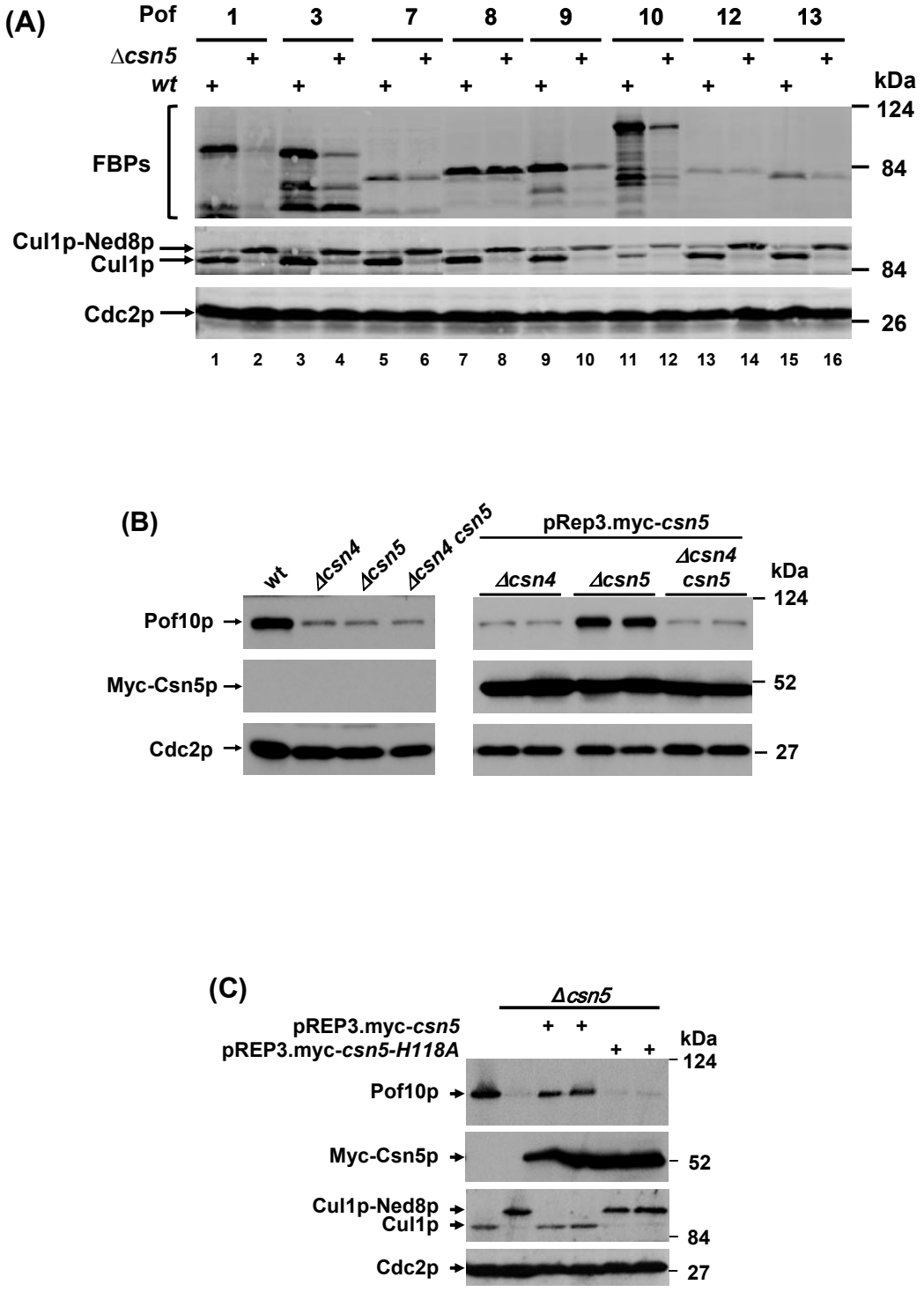


Fig. 1-1**: Differential Effect of CSN on FBP Levels**. (A) Steady-state levels of Myc-tagged FBPs in *csn5* deletion strains. (B) Left panel: Steady-state levels of Myc-tagged Pof10p in *csn4,* and *csn5* deletion strains. Right panel: Complementation of Pof10p levels. The indicated *csn4* and *csn5* deletion strains were transformed with a plasmid driving the expression of Myc-tagged Csn5p, and Pof10p-Myc levels were determined by immunoblotting. Duplicate strains are shown. (C) Dependence of the rescue of Pof10p on the JAMM motif. Myc-tagged wild-type Csn5p or the JAMM point mutant Csn5p-H118A were expressed in *csn5* deletion strains, and the level of Pof10p-Myc was determined by immunoblotting. Duplicate strains are shown.

CSN regulates FBP protein stability

The levels of the mRNAs encoding Pof1p and Pof10p were unchanged in *csn5* mutants as determined by semiquantitative RT-PCR (Supplementary Fig. 1-2), suggesting that CSN-sensitive FBPs might be regulated at the level of protein stability. Cycloheximide (CHX) chase experiments confirmed that the CSN-sensitive FBPs Pof1p, Pof3p, Pof7p, Pof9p, and Pof10p were considerably destabilized in *csn5* mutants (Fig. 1-2). No destabilization was observed for Pof8p and Pof12p. Pof13p stability was also unaffected in *csn5* mutants (Fig. 1-2) despite the minor decrease in Pof13p steady-state levels apparent in Fig. 1-1A. Since *pof13* mRNA was not downregulated in *csn5* mutants as determined by quantitative real time PCR (Supplementary Fig. 1-2B), decreased Pof13p protein levels in *csn5* mutants may reflect an unidentified role of CSN in facilitating Pof13p protein synthesis. In accordance with this notion, some CSN components have dual functions in proteolysis and protein synthesis (Luke-Glaser, 2007).

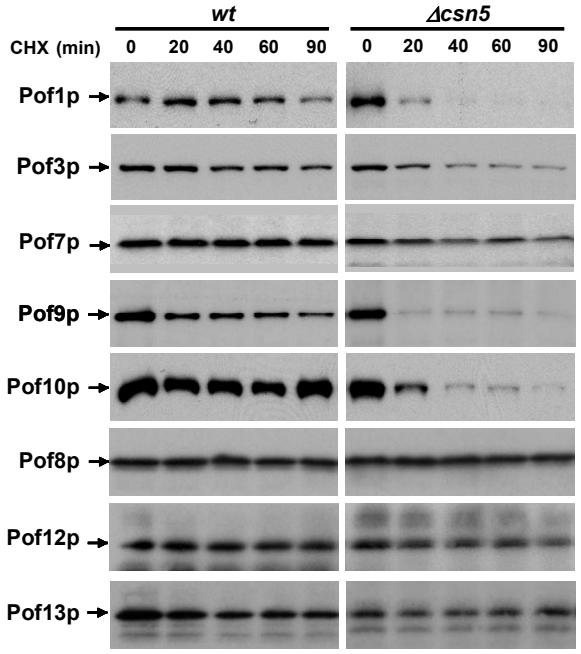


Figure 1-2: **Stability of FBPs in *csn5* mutants**. Strains expressing the indicated Myc-tagged FBPs in a wild-type or csn5 mutant background were employed in a cycloheximide chase experiment to determine their stability (see Materials and Methods).

The DUB Ubp12p Maintains the Stability of the CRL3 adapter Btb3p but not FBPs

We previously showed that the level of the Cul3p adapter Btb3p is strongly reduced in cells lacking CSN deneddylation activity (Wee et al., 2005). An even more pronounced downregulation was observed in cells deficient of the CSN-associated DUB Ubp12p, and

both effects were attributed to increased autocatalytic destruction of Btb3p by its associated CRL3 core module (Wee et al., 2005). Unlike with Btb3p, the steady-state levels of the CSN-regulated FBPs Pof1p, Pof3p, Pof9p, and Pof10p were only minimally affected in *ubp12* mutants when compared to *csn5* mutants (Fig. 1-3A), indicating that Ubp12p is not a major regulator of these FBPs.

To illustrate the differential effect of Ubp12p on Btb3p and FBPs more rigorously, we generated a *ubp12* deletion strain coexpressing protein A-tagged Btb3p and Myc-tagged Pof1p from their endogenous promoters. A CHX chase experiment revealed that Btb3p was drastically destabilized in *ubp12* mutants as described (Wee et al., 2005), whereas Pof1p was entirely stable in the very same cells (Fig. 1-3B).

To exclude the possibility that Ubp1p, a DUB sharing 48% amino acid similarity (32% identity) with Ubp12p and an overall identical domain structure and length (data not shown), maintained FBP stability in *ubp12* mutants, we compared Pof1p levels in *ubp1* and *ubp12* single mutants and in *ubp1 ubp12* double mutants. Pof1p levels were unchanged in either mutant (Fig. 1-3C). These data suggested that, unlike with Btb3p, the steady-state levels of CSN-sensitive FBPs were not affected by lack of CSN-associated DUB activity, although we cannot exclude minor destabilization as previously detected for Pop1p (Zhou et al., 2003).

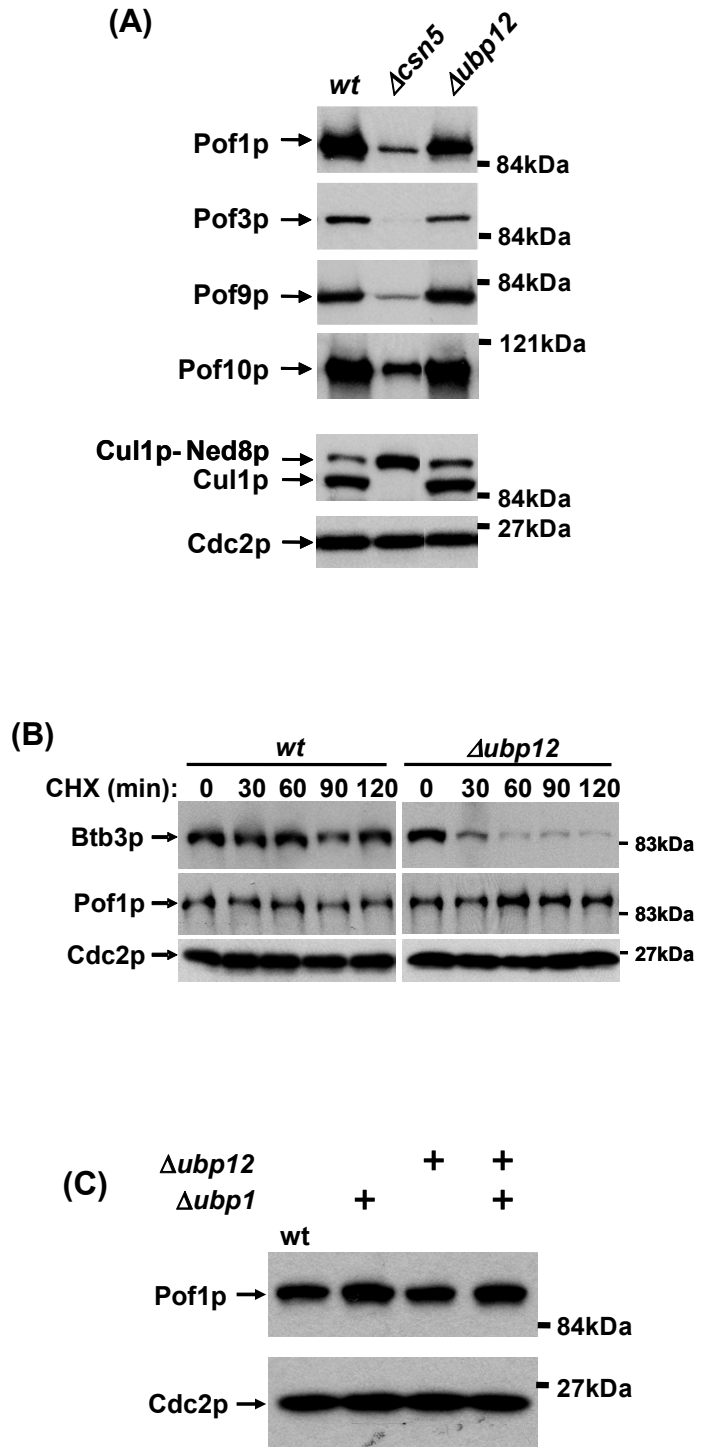


Figure 1-3: **The Ubp12p Maintains the Stability of the CRL3 adapter Btb3p but not FBPs**. (A) Steady-sate levels of CSN-sensitive FBPs in *csn5* and *ubp12* mutants. (B) A strain coexpressing protein A-tagged Btb3p and Myc-tagged Pof1p form their respective genomic loci was employed in a CHX chase experiment to determine the stability of the adapters. (C) Pof1p-Myc levels were determined in *ubp1* and *ubp12* single mutants, and in *ubp1 ubp12* double mutants.

CSN-Insensitive FBPs are Deficient in Forming Canonical CRL1 Complexes

To understand why the stability of only five of the eight FBPs tested was regulated by the CSN, we searched for structural features setting the two groups apart. Since the F-box is the only motif shared by all of these proteins, we performed a ClustalW alignment of the F-boxes of all 16 fission yeast FBPs. Human SKP2 and β-TRCP1 as well as budding yeast Cdc4p were also included in the analysis for reference. The alignment revealed that all CSN-insensitive FBPs missed a conserved proline residue at the beginning of the F-box motif (Fig. 1-4A). None of the other signature residues of the F-box motif segregated consistently with CSN regulation. Since the proline is one of the most highly conserved amino acid of the F-box motif across all species, we reasoned that it might be important for CRL complex formation.

To test this, Myc-tagged FBPs were immunoprecipated with Myc antibodies, followed by immunoblotting with Cul1p and Skp1p antibodies. Binding of Cul1p was only detected for the CSN-regulated FBPs Pof1p, 3p, 7p, 9p, and 10p (Fig. 1-4B). In contrast, Pof8p, 12p, and 13p showed no detectable interaction with Cul1p (Fig.1-4B). Skp1p exhibited a binding pattern closely resembling that of Cul1p, although a low level of Skp1p was also retrieved in immunoprecipitates of Pof8p, 12p, and 13p (Fig. 1-4B), suggesting that these FBPs are principally capable of interacting with Skp1p. Definitive evidence for this conjecture was provided by immunoprecipating the same cell lysates with Skp1p antisera, where all eight FBPs were found to efficiently interact with Skp1p (Fig. 1-4C). These results suggested that CSN-insensitive FBPs lacking the conserved proline residue in their F-boxes are impaired in binding Cul1p, and thus, in forming canonical CRL complexes.

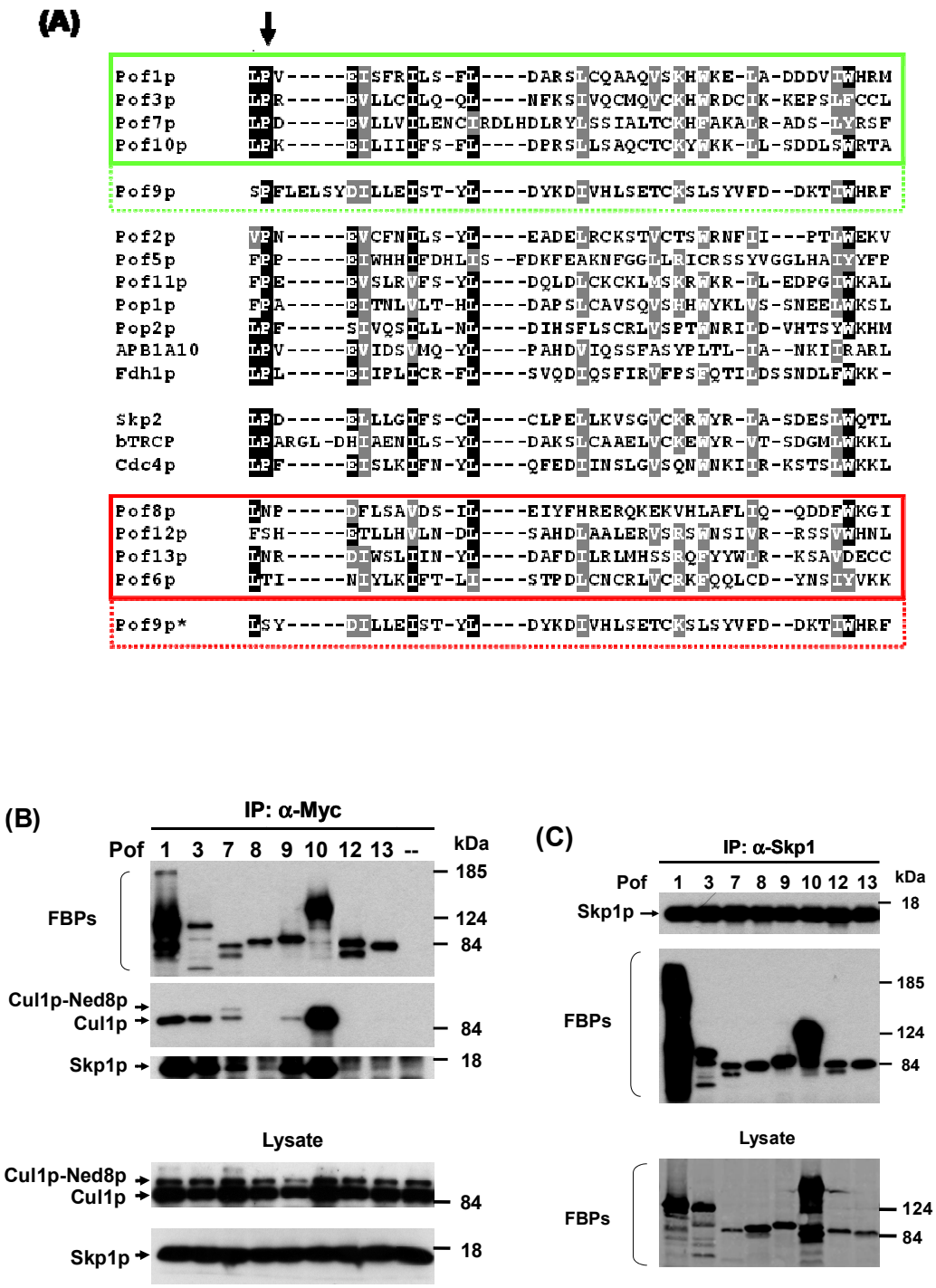


Figure 1-4: **CSN-insensitive FBPs are deficient in forming CRL complexes.** (A) ClustalW alignment of the F-boxes of the indicated proteins. The conserved proline residue is highlighted. (B) Interaction of FBPs with Cul1p and Skp1p. The indicated FBPs were immunoprecipitated with Myc antibodies, followed by immunoblotting with Cul1p and Skp1 sera. Total lysates are shown for reference of expression levels. The negative control (lane 9) represents cell lysate from a strain not expressing any Myc-tagged proteins. (C) The same lysates as in (B) were subjected to immunoprecipitation with Skp1p antisera, followed by immunoblotting with Myc antibodies to detect co-purification of Myc-tagged FBPs. The corresponding total lysate reference is shown.

The conserved proline residue determines CRL1 complex formation

To directly address the role of the proline residue, we constructed a point mutant of the CSN-regulated Pof1p, exchanging proline 114 for serine, which is found in the corresponding position in the F-box of the CSN-independent Pof12p (Fig. 1-4A). Similarly, we changed the conserved proline residues in Pof10p to serine. Myc-tagged wild-type and proline mutant FBPs were expressed from pRep81 plasmids driven by the low strength *nmt1* promoter, and binding to endogenous Cul1p was determined by co-immunoprecipitation. The proline mutant Pof1p and Pof10p proteins interacted with Cul1p much less efficiently than the respective wild-type proteins (Fig. 1-5A). For Pof1p, this effect was observed in both wild-type and *csn5* mutant backgrounds, whereas binding of Skp1p was not affected by the proline mutation (Supplementary Fig. 1-3A).

We next asked whether the proline was sufficient to target FBPs into a CRL1 complex. To this end, we changed serine 15 of the Pof12p F-box to proline and determined binding to Cul1p. As with endogenous Pof12p (see Fig. 1-4B), plasmid-derived wildtype Pof12p was inefficient in binding Cul1p (Fig. 1-5A). In contrast, proline-containing Pof12p bound Cul1p (Fig. 1-5A). This manipulation also increased the recruitment of Skp1p (Supplementary Fig. 1-3B). Thus, the F-box proline residue appears both required and sufficient to target FBPs into canonical CRL1 complexes.

The instability of Pof9p in *csn5* mutants (Fig. 1-2A) and the efficient binding of Pof9p to Cul1p and Skp1p (Fig. 1-4B) suggested that its F-box also contains the conserved proline. Indeed, under the parameters used, the ClustalW algorithm aligned proline 5, a residue near the beginning of the Pof9p F-box, with the conserved proline of other FBPs (Fig. 1-4A). The alignment also highlighted a short insertion following the proline that is shared by human β-TRCP but not other *S. pombe* FBPs (Fig. 1-4A). Remarkably, unlike with Pof1p and Pof10p, deletion of proline 5 did not reduce the binding of Pof9p to Cul1p but enhanced it instead (Fig. 1-5A). Manual editing of the sequence alignment revealed that Pof9p can also be aligned such that it features a serine in the position of the conserved proline, a manoeuvre that would place Pof9p into the group of CSN-independent FBPs (designated Pof9p* at the bottom of Fig. 1-4A).

To reconcile this apparent paradox, we considered the possibility that Pof9p may be targeted into CRL1 complexes independently of its F-box motif. Precedence for such a scenario was previously provided by our demonstration that the FBP Pop2p can be recruited into functional CRL1 complexes in an F-box independent manner through dimerization with another FBP, Pop1p (Seibert et al., 2002). Truncated Pof9p lacking the N-terminal F-box

retained efficient interaction with Cul1p similar in extent to the wildtype and the proline mutant proteins (Fig. 1-5B). This was equally apparent in immuno-purified CRL1 complexes prepared with Rbx1p antibodies and with antibodies directed against exogenously expressed Myc-Pof9p and its variants. In summary, these finding indicated that Pof9p, unlike most other FBPs, is targeted into CRL1 complexes and subjected to stability control by the CSN independently of the integrity of its F-box. Taken together, these data strongly suggest that proline-dependent recruitment of FBPs into CRL1 complexes targets FBPs for stabilization by the CSN pathway.

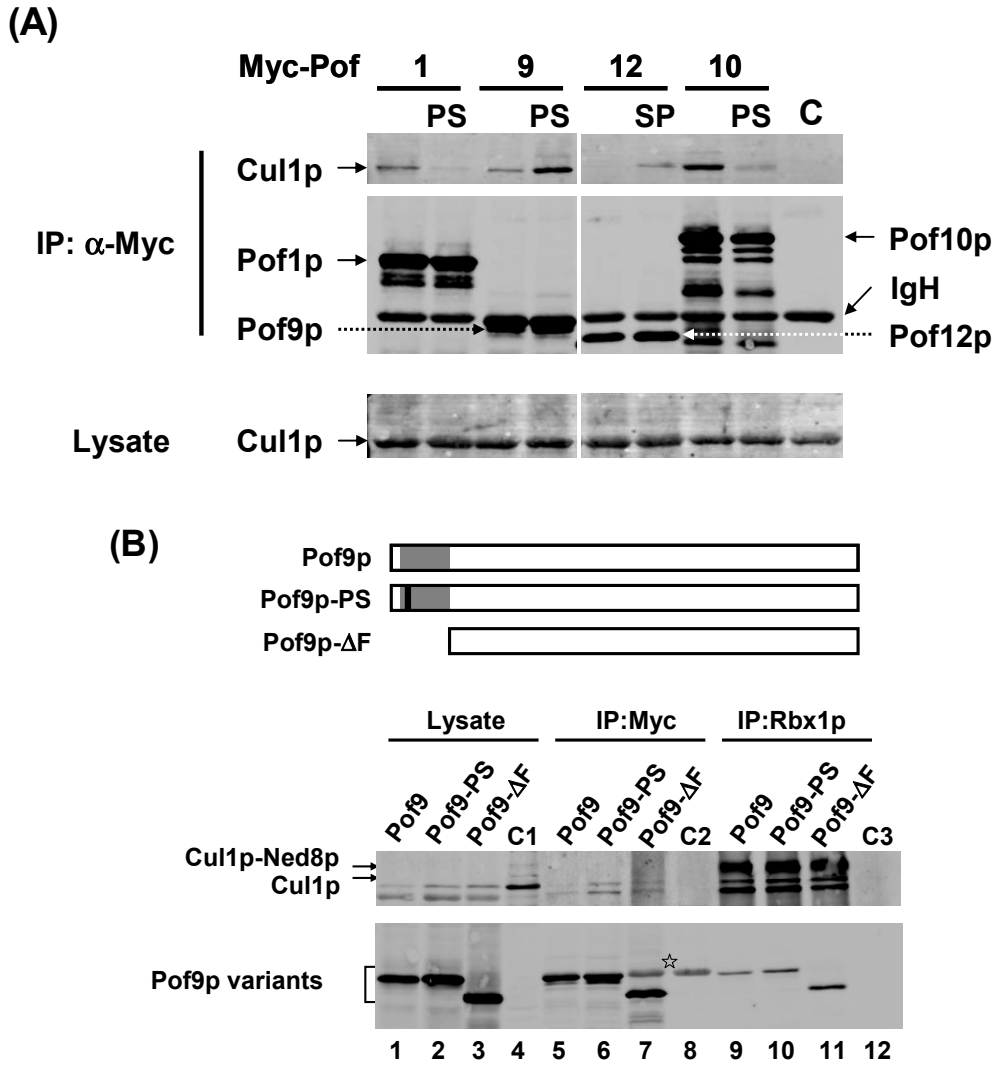


Figure 1-5: **The conserved proline residue determines CRL1 complex formation.** (A) Interaction of exogenous wildtype or mutated Pof1p, Pof1pP114S, Pof9p, Pof9pP5S, Pof12p, Pof12pS15P, Pof10p and Pof10pP35S with Cul1p. Myc-tagged FBPs were immunoprecipitated and copurification of Cul1p was assayed by immunoblotting with anti-Cul1p serum. (B) Interaction of Pof9p and Cul1p. Plasmid driven wildtype Pof9p, mutated Pof9pP5S, Pof9pΔF (Δ1-49) or *csn5* lysate lacking the plasmid were immunoprecipitated with Myc antibodies, followed by immunoblotting with Cul1p serum (lanes 5-9). The same lysates were subjected to immunoprecipitation with Rbx1p antisera, followed by immunoblotting with Myc antibodies to detect co-purification (lanes 9-11). Beads incubated in Pof9p expressed lysate without the addition of Rbx1p antibody served as negative control. The corresponding total lysate reference is shown in lanes 1-4.

CAND1 is not required for maintaining the stability of CSN-regulated FBPs

CAND1 was proposed to participate in the same process of CRL adapter stabilization as the CSN (Cope and Deshaies, 2003; He et al., 2005), although no experimental evidence was provided. To address this proposition in fission yeast, we turned our attention to the uncharacterized open reading frame SPAC1565.07c, which we named *can1*. This gene

encodes a protein with an overall similarity of 43% (22% identity) with human CAND1 over its entire length of 1220 amino acids. In addition, 25 of the 27 HEAT repeats found in human CAND1, which are responsible for its hallmark solenoid structure (Goldenberg et al., 2004), are predicted from the primary sequence of *S. pombe* Can1p (data not shown). These considerations suggested that Can1p is the putative orthologue of CAND1 from higher eukaryotes.

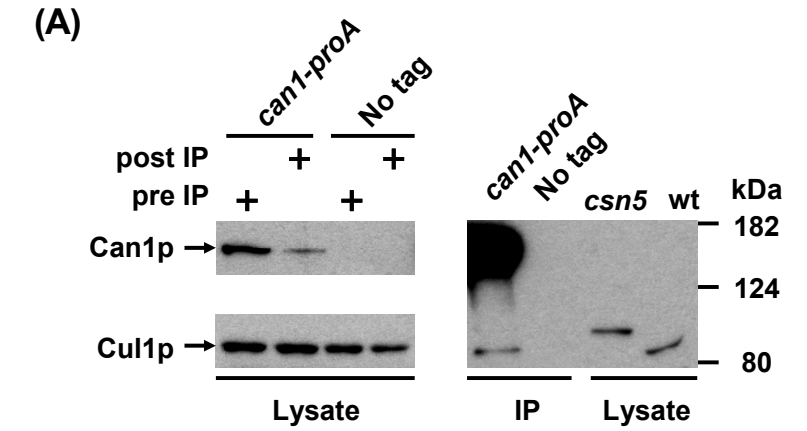
This notion was supported by the finding that Can1p modified with a single N-terminal protein A tag at the endogenous genomic locus co-immunoprecipitated the unneddylated form of Cul1p (Fig. 1-6A). Importantly, whereas >50% of ProA-Can1p was depleted from the cell lysate upon absorption to IgG resin, the bulk of Cul1p was retained in the lysate, indicating that only a minor fraction of Cul1p was in a stable complex with Can1p under steady-state conditions (Fig. 1-6A).

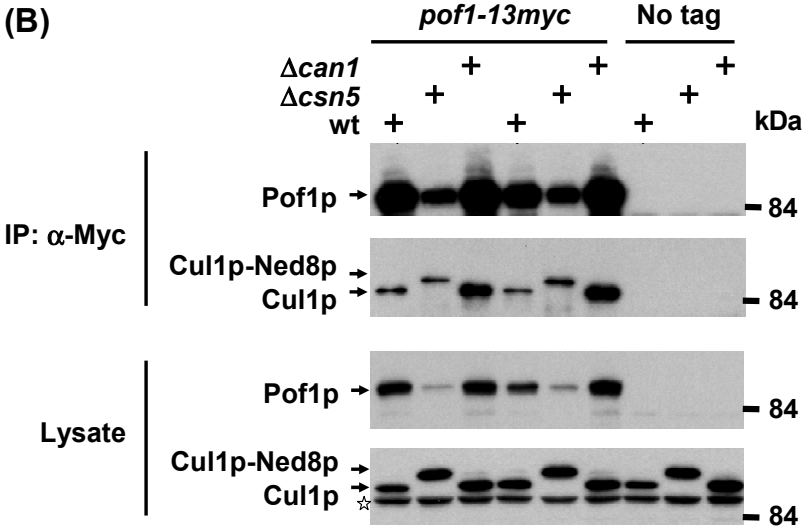
Like most *csn* and *ubp12* deletions strains (Mundt et al., 2002; Zhou et al., 2001; Zhou et al., 2003), haploid cells lacking *can1* were viable and did not exhibit any gross morphological or growth phenotypes (data not shown). Unlike *csn* mutants, however, *can1* mutants did not show accumulation of Cul1p in the neddylated state or downregulation of Pof1p (Fig. 1-6B). Nevertheless, binding of Pof1p to Cul1p was enhanced in *can1* mutants (Fig. 1-6B). This binding was also enhanced in *csn5* deletion strains, in particular when accounting for the low levels of Pof1p present in this mutant (Fig. 1-6B). Since neddylated Cul1p can not interact with CAND1 (Hwang et al., 2003; Liu et al., 2002; Min et al., 2003; Oshikawa et al., 2003; Zheng et al., 2002a), and since Cul1p is fully neddylated in *csn5* mutants, CSN deficiency appears to mimic CAND1 deficiency with respect to recruitment of F-box proteins to Cul1p.

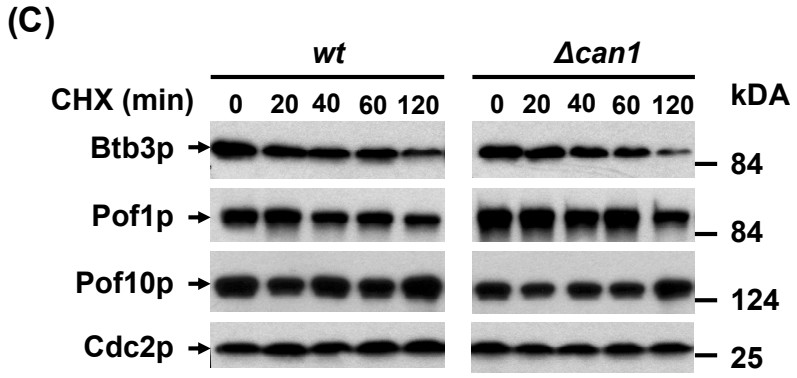
However, CAND1 deficiency does not phenocopy CSN deficiency with regards to control of FBP levels and stability. The downregulation of Pof1p steady state levels occurring in *csn5* mutants was not observed in *can1* mutants (Fig. 1-6B). In addition, the stability of the CSN-regulated Pof1p, Pof10p, and the CRL3 adapter Btb3p was unaffected (Fig. 1-6C). These findings indicated that CSN maintains CRL adapter stability independently of CAND1.

To further substantiate this conclusion, we performed a genetic assay to determine the impact of Can1p on CRL1 function more broadly. Using a strain containing a temperature-sensitive allele of *skp1*, which is specifically impaired in binding of FBPs (Lehmann et al., 2004), but not in binding of Cul1p (Wee et al., 2005), we previously showed that CSN becomes essential for viability when adapter recruitment to CRL core complexes is compromised. Whereas *skp1-ts csn5* mutants lost viability upon shift to the restrictive temperature as demonstrated before, *skp1-ts can1* double mutants were fully viable at 36.5 ^{0}C

(Fig. 1-6D). Thus, unlike CSN, Can1p is dispensable, even when adapter recruitment to CRL1 core complexes is compromised.


(A)
can1-proA
No tag
post IP
pre IP
Can1p
Cul1p
Lysate
can1-proA
No tag
csn5
wt
kDa
182
124
80
IP
Lysate


(B)
pof1-13myc
No tag
Δcan1
Δcsn5
wt
kDa
IP: α-Myc
Pof1p
Cul1p-Ned8p
Cul1p
84
Lysate
Pof1p
Cul1p-Ned8p
Cul1p
84


(C)
wt
Δcan1
CHX (min)
0 20 40 60 120
0 20 40 60 120
kDA
Btb3p
Pof1p
Pof10p
Cdc2p
84
84
124
25

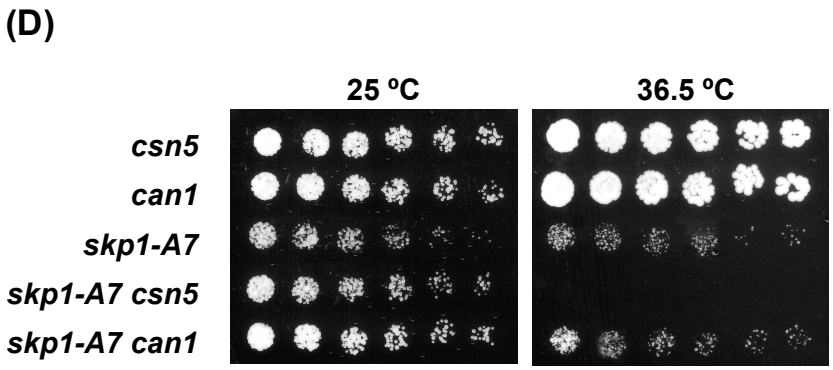


Figure. 1-6: **CAND1 is not Required for Maintaining the Stability of CSN-Regulated FBPs.** (A) Lysate from a strain expressing Can1p modified with a single N-terminal protein A tag at the endogenous genomic locus was absorbed to IgG beads, followed by immunoblotting with Cul1p sera. Total cell lysates before and after chromatography on IgG resin are shown to indicate the extent of the Can1p-Cul1p interaction (left panel). Wildtype and *csn5* lysate was blotted with Cul1p antiserum to indicate the unmodified and Ned8p-modified Cul1p band (lanes 3 and 4, right panel). (B) Interaction of Pof1p with Cul1p in wildtype, *csn5* and *can1* mutant strains. Duplicate strains are shown. Endogenously tagged Pof1p was immunoprecipitated with Myc antibodies, followed by immunoblotting with Cul1p antiserum. Total lysates are shown for reference of expression levels. The negative controls (lanes 7-9) represent cell lysate from strains not expressing any Myc-tagged proteins. (C) The stability of Pof1p-Myc, Pof10-Myc, and Btb3p-proA in *can1* mutants was determined by CHX chase. (D) Genetic interaction of *can1* and *csn5* with the temperature-sensitive allele *skp1-A7*. Serial dilutions of the indicated strains were spotted onto YES plates and incubated at the indicated temperatures. Whereas *csn5* shows synthetic interaction with *skp1-A7*, *can1* does not.

1.4. Discussion

F-box directed CRL assembly and control by the CSN

The results of this study add further credence to our model that CSN's cullin deneddylation activity revealed *in vitro* serves to maintain the stability of CRL adapters thus promoting CRL activity *in vivo* (Wolf et al., 2003; Zhou et al., 2003). CSN-mediated adapter stabilization was observed here for the FBPs Pof1p, 3p, 7p, 9p, and 10p, and for Pop1p and the CRL3 adapter Btb3p in our previous studies (Wee et al., 2005; Zhou et al., 2003). Other recent reports examining two FBPs in *N. crassa* (He et al., 2005), four human FBPs (Cope & Deshaies, 2006; Denti et al., 2006), and a *C. elegans* BTB protein (Luke-Glaser et al., 2007) have provided corroborating evidence.

However, not all FBPs are subject to this mode of regulation, a surprising finding, considering that all FBPs examined here share a canonical F-box motif and interact with Skp1p (Fig. 1-3) (Lehmann et al., 2004). Yet, the CSN-insensitive FBPs are not efficiently incorporated into CRL1 complexes, because they lack a critical proline residue in their F-boxes, which is required for binding of Cul1p, but not Skp1p. This proline, which is among the most highly conserved residues of the F-box motif is also missing from budding yeast Rcy1p, its fission yeast orthologue Pof6p, and from human Emi1, all of which were previously found to bind Skp1p, but not Cul1p (Galan et al., 2001; Hermand et al., 2003; Seol et al., 2001); and Peter K. Jackson, personal communication). Interestingly, unlike other human FBPs, Emi1 was recently shown not to require CSN for stability, although the reason for this discrepancy remained unknown (Cope & Deshaies, 2006).

Conversely, the CSN-regulated *N. crassa* FPBs FWD-1 and SCON-2, contain the proline. In addition, our point mutagenesis data clearly established the importance of the proline residue for binding of Pof1p and Pof10p to Cul1p (Fig. 4A, B). As shown by point mutagenesis of Pof12p, the conserved proline is not only required but sufficient for targeting FBPs into CRL complexes (Fig. 1-5A, Supplementary Fig. 1-3B).

These findings can be rationalized by the crystal structure of the human CUL1-RBX1-SKP1-SKP2 complex, which showed that the proline is the only conserved F-box residue within a cluster of three contiguous amino acids that makes side chain contacts with Cul1 (Zheng et al., 2002b). Despite its extensive interface with both CUL1 and the F-box, SKP1 is not sufficient to recruit FBPs into CRL complexes; instead, this process is specified by proline-dependent F-box-CUL1 interactions. An interesting possibility is that recruitment of

FBP-SKP1 dimers to the CRL1 core complex requires the induction of proline-dependent conformational changes in CUL1 perhaps involving binding-induced folding of the unstructured loops within the CUL1 N-terminus (Zheng et al., 2002b).

Based on these data, we propose that all FBPs that lack the critical proline residue are not engaged in canonical CRL1 complexes and are hence not subject to stabilization by the CSN. The same scenario would apply to 11 human FBPs that lack the proline. Conversely, CSN-mediated cullin deneddylation would assist in the assembly of all proline-containing FBPs into CRL1 complexes by shielding them from autocatalytic inactivation. As we have previously discussed, this mechanism provides a safe environment for the de novo assembly and maintenance of CRL complexes with labile FBPs (Wolf et al., 2003).

Role of CAND1 in FBP Regulation

CAND1 inhibits CRL activity *in vitro* (Hwang et al., 2003; Liu et al., 2002; Min et al., 2003; Oshikawa et al., 2003; Zheng et al., 2002a), but is required for full CRL activity *in vivo* (Chuang et al., 2004; Feng et al., 2004; Zheng et al., 2002a). This constellation reiterates the CSN paradox, and similar resolutions were invoked for both the CSN and CAND1 paradoxes. Specifically, CAND1 was suggested to disassemble CRL complexes by stable sequestration of Cul1 upon CSN-mediated deneddylation, thus leading to adapter stabilization as a consequence of escape from autocatalytic degradation (Cope & Deshaies, 2003; Cope & Deshaies, 2006; He et al., 2005; Zheng et al., 2002a).

At first sight, a model of adapter stabilization involving CAND1 seems attractive considering that inactivation of CSN is expected to phenocopy deletion of CAND1 because fully neddylated cullins such as present in *csn* mutants can not bind CAND1. Nevertheless, the results presented here contradict several key prediction of a model of adapter stabilization involving CAND1: First, we found that only a miniscule fraction of unneddylated Cul1p was in a stable complex with Can1p at endogenous steady-state expression levels, although the vast majority of Cul1p was in the unneddylated state (Fig. 1-6A). Conversely, unneddylated Cul1p readily interacted with Skp1p and five different FBPs (Fig. 4B). Both findings are inconsistent with quantitative sequestration of unneddylated Cul1 into neutral complexes with CAND1 as a mechanism for maintaining bulk FBP stability. Secondly, unlike inactivation of CSN, deletion of *can1* did not interfere with the stability of multiple FBPs or the Cul3p adapter Btb3p (Fig. 1-6C). Similarly, the FBP Col1 is not stabilized in plant CAND1 mutants

(Feng et al., 2004). Finally, unlike CSN, Can1p was not rendered essential when adapter recruitment to CRL1 core complexes was compromised (Fig. 1-6D).

Based on these findings, we conclude that CSN-dependent adapter stabilization principally functions independently of CAND1. CSN-mediated cullin deneddylation is sufficient to sustain this mode of regulation. While it is true that lack of Can1p increases FBP recruitment to Cul1p as shown for Pof1p (Fig. 1-6B), this does not coincide with autocatalytic Pof1p degradation (Fig. 1-6C), most likely because Cul1p engaged in these complexes is dominantly inhibited by CSN-mediated deneddylation or otherwise inactive. In *csn* mutants on the other hand, increased Pof1p recruitment to fully neddylated Cul1p leads to severe autocatalytic depletion (Fig. 1-6B).

If CAND1 does not maintain adapter stability, what might its function be in promoting CRL activity *in vivo*? Based on the low amount of Cul1p we found engaged in a stable complex with Can1p, a situation that is reiterated in *A. thaliana* (Feng et al., 2004), we envision that, contrary to current concept, Can1p forms only transient interactions with a subpopulation of Cul1p molecules. It is puzzling then why, in mammalian cells, CAND1 appears to form a stable complex with a substantial fraction of the total unneddylated CUL1 pool. It is of note that all identifications of stable cullin/CAND1 complexes relied on plasmid-based overexpression of cullins (Liu et al., 2002; Min et al., 2003; Min et al., 2005; Oshikawa et al., 2003). The only exception to this is the study by Zheng et al. where the endogenous CUL1/CAND1 complex was visualized in CUL1 immunoprecipitates from ^{35}S-methionine-labelled cells (Zheng et al., 2002a). In contrast, extensive mass spectrometric characterization of endogenous cullin complexes has either failed to identify CAND1 (Geyer et al., 2003; Higa et al., 2006; Liu et al., 2003) or revealed a similarly small fraction of the complex as seen here for Cul1p/Can1p (He et al., 2006). The only obvious commonality between the metabolic labelling and the ectopic overexpression experiments is the preferential detection of newly synthesized CUL1. Since cullins are generally stable proteins, their endogenous level of synthesis is expected to be relatively low. These findings suggest that, under normal conditions, CAND1 only associates with the small pool of newly synthesized and unneddylated CUL1.

Based on these considerations, we are proposing a new model for the reconciliation of the CAND1 paradox that is entirely decoupled from CSN and the cycles of neddylation and deneddylation proposed to underly CAND1 function. We envision that the partially folded N- and C-termini of newly synthesized, unneddylated CUL1 (and other cullins) are captured by CAND1. CAND1 then serves as a scaffold or brace for the proper folding of CUL1 into its

mature conformation, presumably involving chaperones. In support of these notions are the observations that cullins expressed in bacteria are generally insoluble even, if Rbx1 is co-expressed (unpublished observation), and that CAND1/cullin complexes derived by overexpression in mammalian cells copurify with substantial amounts of HSP70 and HSP60 (Oshikawa et al., 2003). CAND1-templated folding is further speculated to facilitate the association of CUL1 with RBX1 thus assembling the CRL core complex. Indeed, overexpression of CAND1 in mammalian cells was shown to discretely stimulate the recruitment of RBX1 to unneddylated CUL1, whereas knockdown of CAND1 had the opposite effect (Zheng et al., 2002a). Once the CUL1/RBX1 core complex is assembled, SKP1/FBP dimers can then displace CAND1 thus leading to the assembly of a complete, neddylation-competent CRL complex. In support of this notion, SKP1/SKP2 dimers (Bornstein et al., 2006) but not neddylating enzymes (Bornstein et al., 2006; Goldenberg et al., 2004; Hwang et al., 2003), are sufficient to release human CUL1 from CAND1 *in vitro*. This release was assumed to require critical FBP-CUL1 interactions (Bornstein et al., 2006), which we show are provided by the conserved proline residue of FBPs. Mature CRL complexes assembled in this way can then enter repeated cycles of neddylation and CSN-mediated deneddylation for long-term maintenance that no longer involve CAND1.

Upon induction of CAND1 deficiency, either by genetic mutation in yeast or by si-RNA-mediated knockdown in mammalian cells, CRL function would rely on the sizeable, folded, and stable pool of CUL1/RBX cores preexisting in the cells. This pool is only sluggishly replenished by low level synthesis of slowly folding CUL1. Lack of CAND1 would also allow a portion of slow folding naïve CUL1 to assemble prematurely with SKP1/FBP dimers (Fig. 1-6B) thus forming abortive CRL complexes. Both biochemical consequences of CAND1 deficiency are predicted to impair bulk cellular CRL function, although much less dramatically than deletion of the CSN or the neddylation system. This is true, for instance, for the auxin response and photomorphogenesis in *A. thaliana* (Chuang et al., 2004; Feng et al., 2004), for SCF^{SKP2}-mediated p27 degradation in mammalian cells (Zheng et al., 2002a), and for the synthetic effects with *skp1-ts* alleles in *S. pombe* (Fig. 1-6D). These quantitative differences in phenotype are readily explained in our model by the different pools of CUL1 proposed to be controlled by CSN and CAND1. Whereas CSN promotes the stability of the bulk steady-state pool of CRL complexes, CAND1 only affects the folding of a small portion of newly synthesized CUL1 and its assembly with RBX1. It is important to note, however, that we cannot entirely rule out the possibility that a small fraction of steady-state CUL1 is

occasionally reset into a similar naïve state as newly synthesized CUL1 consequently requiring CAND1 for reassembly with RBX1 and FBP-SKP1 dimers.

Role of Ubp12p in CRL adapter regulation

Like Can1p, the CSN-associated deubiquitylating enzyme Ubp12p was dispensable for the stability of the FBPs examined here (Fig. 1-3A, B). However, unlike Can1p, Ubp12p is essential in cells harbouring Skp1p ts-mutants that are impaired in adapter recruitment (Lehmann et al., 2004; Wee et al., 2005), suggesting that CSN-associated Ubp12p regulates at least a subset of CRL adapters. This notion is confirmed by our published observation that Ubp12p maintains the stability of the FBP Pop1p (Zhou et al., 2003) and, more dramatically, the CRL3 adapter Btb3p (Wee et al., 2005). Ubp12p therefore appears to have a differential effect on adapter stability, a conclusion we directly corroborated here by showing that Btb3p is destabilized in the same *ubp12* deleted cells, in which the FBP Pof1p is stable (Fig. 2B). Additional studies are required to address the mechanisms that destine a particular adapter for regulation by Ubp12p. On the other hand, Ubp12p may also regulate CRL activity by stabilizing components other than substrate adapters. Consistent with this notion are recent findings that CSN promotes the stability of CRL1 core subunits (He et al., 2005; Hetfeld et al., 2005; Wu et al., 2005), although we have not yet detected conditions under which similar regulation occurs in fission yeast.

1.5. References

Bahler J, Wu JQ, Longtine MS, Shah NG, McKenzie A, 3rd, Steever AB, Wach A, Philippsen P, Pringle JR (1998) Heterologous modules for efficient and versatile PCR-based gene targeting in Schizosaccharomyces pombe. Yeast 14(10): 943-951.

Bornstein G, Ganoth D, Hershko A (2006) Regulation of neddylation and deneddylation of cullin1 in SCFSkp2 ubiquitin ligase by F-box protein and substrate. Proceedings of the National Academy of Sciences 103(31): 11515-11520

Chuang H-w, Zhang W, Gray WM (2004) Arabidopsis ETA2, an Apparent Ortholog of the Human Cullin-Interacting Protein CAND1, Is Required for Auxin Responses Mediated by the SCFTIR1 Ubiquitin Ligase. Plant Cell 16(7): 1883-1897

Cope G, Suh GSB, Aravind L, Schwarz SE, Zipursky SL, Koonin EV, Deshaies RJ (2002) Role of Predicted Metalloprotease Motif of Jab1/Csn5 in Cleavage of NEDD8 from CUL1. Science 298: 608-611

Cope GA, Deshaies RJ (2003) COP9 signalosome: a multifunctional regulator of SCF and other cullin-based ubiquitin ligases. Cell 114(6): 663-671.

Cope GA, Deshaies RJ (2006) Targeted silencing of Jab1/Csn5 in human cells downregulates SCF activity through reduction of F-box protein levels. BMC Biochem 7(1): 1.

Denti S, Fernandez-Sanchez ME, Rogge L, Bianchi E (2006) The COP9 signalosome regulates Skp2 levels and proliferation of human cells. J Biol Chem 281(43): 32188-32196

Deshaies RJ (1999) SCF and Cullin/Ring H2-based ubiquitin ligases. Annu Rev Cell Dev Biol 15: 435-467

Feng S, Shen Y, Sullivan JA, Rubio V, Xiong Y, Sun T-p, Deng XW (2004) Arabidopsis CAND1, an Unmodified CUL1-Interacting Protein, Is Involved in Multiple Developmental Pathways Controlled by Ubiquitin/Proteasome-Mediated Protein Degradation. Plant Cell 16(7): 1870-1882

Galan JM, Peter M (1999) Ubiquitin-dependent degradation of multiple F-box proteins by an autocatalytic mechanism. Proc Natl Acad Sci U S A 96(16): 9124-9129.

Galan JM, Wiederkehr A, Seol JH, Haguenauer-Tsapis R, Deshaies RJ, Riezman H, Peter M (2001) Skp1p and the F-box protein Rcy1p form a non-SCF complex involved in recycling of the SNARE Snc1p in yeast. Mol Cell Biol 21(9): 3105-3117.

Geyer R, Wee S, Anderson S, Yates JRI, Wolf DA (2003) BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. Mol Cell 12: 783-790

Goldenberg SJ, Cascio TC, Shumway SD, Garbutt KC, Liu J, Xiong Y, Zheng N (2004) Structure of the Cand1-Cul1-Roc1 Complex Reveals Regulatory Mechanisms for the Assembly of the Multisubunit Cullin-Dependent Ubiquitin Ligases. Cell 119(4): 517-528

Groisman R, Polanowska J, Kuraoka I, Sawada J, Saijo M, Drapkin R, Kisselev AF, Tanaka K, Nakatani Y (2003) The ubiquitin ligase activity in the DDB2 and CSA complexes is

differentially regulated by the COP9 signalosome in response to DNA damage. Cell 113(3): 357-367.

Harrison C, Katayama S, Dhut S, Chen D, Jones N, Bahler J, Toda T (2005) SCF(Pof1)-ubiquitin and its target Zip1 transcription factor mediate cadmium response in fission yeast. Embo J 24(3): 599-610

He Q, Cheng P, He Q, Liu Y (2005) The COP9 signalosome regulates the Neurospora circadian clock by controlling the stability of the SCFFWD-1 complex. Genes Dev 19(13): 1518-1531

He YJ, McCall CM, Hu J, Zeng Y, Xiong Y (2006) DDB1 functions as a linker to recruit receptor WD40 proteins to CUL4-ROC1 ubiquitin ligases. Genes Dev 20(21): 2949-2954

Hermand D, Bamps S, Tafforeau L, Vandenhaute J, Makela TP (2003) Skp1 and the F-box Protein Pof6 Are Essential for Cell Separation in Fission Yeast. J Biol Chem 278(11): 9671-9677

Hetfeld BK, Helfrich A, Kapelari B, Scheel H, Hofmann K, Guterman A, Glickman M, Schade R, Kloetzel PM, Dubiel W (2005) The zinc finger of the CSN-associated deubiquitinating enzyme USP15 is essential to rescue the E3 ligase Rbx1. Curr Biol 15(13): 1217-1221

Higa LA, Wu M, Ye T, Kobayashi R, Sun H, Zhang H (2006) CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. Nat Cell Biol 8(11): 1277-1283

Hwang JW, Min KW, Tamura TA, Yoon JB (2003) TIP120A associates with unneddylated cullin 1 and regulates its neddylation. FEBS Lett 541(1-3): 102-108.

Lehmann A, Katayama S, Harrison C, Dhut S, Kitamura K, McDonald N, Toda T (2004) Molecular interactions of fission yeast Skp1 and its role in the DNA damage checkpoint. Genes Cells 9(5): 367-382

Liu C, Powell KA, Mundt K, Wu L, Carr AM, Caspari T (2003) Cop9/signalosome subunits and Pcu4 regulate ribonucleotide reductase by both checkpoint-dependent and -independent mechanisms. Genes Dev 17(9): 1130-1140.

Liu J, Furukawa M, Matsumoto T, Xiong Y (2002) NEDD8 modification of CUL1 dissociates p120(CAND1), an inhibitor of CUL1-SKP1 binding and SCF ligases. Mol Cell 10(6): 1511-1518.

Luke-Glaser S, Roy M, Larsen B, Le Bihan T, Metalnikov P, Tyers M, Peter M, Pintard L (2007) CIF-1, a Shared Subunit of the COP9/Signalosome and Eukaryotic Initiation Factor 3 Complexes, Regulates MEL-26 Levels in the Caenorhabditis elegans Embryo. Mol Cell Biol 27(12): 4526-4540

Lyapina S, Cope G, Shevchenko A, Serino G, Tsuge T, Zhou C, Wolf DA, Wei N, Shevchenko A, Deshaies RJ (2001) Promotion of NEDD8-CUL1 conjugate cleavage by COP9 signalosome. Science 292: 1382-1385

Min KW, Hwang JW, Lee JS, Park Y, Tamura TA, Yoon JB (2003) TIP120A Associates with Cullins and Modulates Ubiquitin Ligase Activity. J Biol Chem 278(18): 15905-15910.

Min KW, Kwon MJ, Park HS, Park Y, Yoon SK, Yoon JB (2005) CAND1 enhances deneddylation of CUL1 by COP9 signalosome. Biochem Biophys Res Commun 334(3): 867-874

Mundt KE, Liu C, Carr AM (2002) Deletion mutants in COP9/signalosome subunits in fission yeast Schizosaccharomyces pombe display distinct phenotypes. Mol Biol Cell 13(2): 493-502.

Oshikawa K, Matsumoto M, Yada M, Kamura T, Hatakeyama S, Nakayama KI (2003) Preferential interaction of TIP120A with Cul1 that is not modified by NEDD8 and not associated with Skp1. Biochem Biophys Res Commun 303(4): 1209-1216.

Pan ZQ, Kentsis A, Dias DC, Yamoah K, Wu K (2004) Nedd8 on cullin: building an expressway to protein destruction. Oncogene 23(11): 1985-1997

Petroski MD, Deshaies RJ (2005) Function and regulation of cullin-RING ubiquitin ligases. Nat Rev Mol Cell Biol 6(1): 9-20

Pintard L, Willis JH, Willems A, Johnson JL, Srayko M, Kurz T, Glaser S, Mains PE, Tyers M, Bowerman B, Peter M (2003) The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. Nature 425(6955): 311-316.

Rouillon A, Barbey R, Patton EE, Tyers M, Thomas D (2000) Feedback-regulated degradation of the transcriptional activator Met4 is triggered by the SCF(Met30)complex. Embo J 19(2): 282-294.

Schwechheimer C, Serino G, Callis J, Crosby WL, Lyapina S, Deshaies RJ, Gray WM, Estelle M, Deng XW (2001) Interactions of the COP9 signalosome with the E3 ubiquitin ligase SCFTIRI in mediating auxin response. Science 292(5520): 1379-1382.

Seibert V, Prohl C, Schoultz I, Rhee E, Lopez R, Abderazzaq K, Zhou C, Wolf DA (2002) Combinatorial diversity of fission yeast SCF ubiquitin ligases by homo- and heterooligomeric assemblies of the F-box proteins Pop1p and Pop2p. BMC Biochemistry 3: 22

Seol JH, Shevchenko A, Deshaies RJ (2001) Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. Nat Cell Biol 3(4): 384-391

von Arnim AG (2003) On again - off again: COP9 signalosome turns the key on protein degradation. Current Opinion in Plant Biology 6(6): 520-529

Wee S, Geyer RK, Toda T, Wolf DA (2005) CSN facilitates Cullin-RING ubiquitin ligase function by counteracting autocatalytic adapter instability. Nat Cell Biol 7(4): 387-391

Werler PJ, Hartsuiker E, Carr AM. (2003) A simple Cre-loxP method for chromosomal N-terminal tagging of essential and non-essential Schizosaccharomyces pombe genes. Gene. 2003 Jan 30;304:133-41.

Wirbelauer C, Sutterluty H, Blondel M, Gstaiger M, Peter M, Reymond F, Krek W (2000) The F-box protein Skp2 is a ubiquitylation target of a Cul1-based core ubiquitin ligase complex: evidence for a role of Cul1 in the suppression of Skp2 expression in quiescent fibroblasts. Embo J 19(20): 5362-5375.

Wolf DA, Zhou C, Wee S (2003) The COP9 signalosome: an assembly and maintenance platform for cullin ubiquitin ligases? Nat Cell Biol 5(12): 1029-1033.

Wu JT, Lin HC, Hu YC, Chien CT (2005) Neddylation and deneddylation regulate Cul1 and Cul3 protein accumulation. Nat Cell Biol 7(10): 1014-1020

Xu L, Wei Y, Reboul J, Vaglio P, Shin TH, Vidal M, Elledge SJ, Harper JW (2003) BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. Nature 425(6955): 316-321.

Zheng J, Yang X, Harrell JM, Ryzhikov S, Shim EH, Lykke-Andersen K, Wei N, Sun H, Kobayashi R, Zhang H (2002a) CAND1 binds to unneddylated CUL1 and regulates the formation of SCF ubiquitin E3 ligase complex. Mol Cell 10(6): 1519-1526.

Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, Wang P, Chu C, Koepp DM, Elledge SJ, Pagano M, Conaway RC, Conaway JW, Harper JW, Pavletich NP (2002b) Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. Nature 416(6882): 703-709.

Zhou C, Seibert V, Geyer R, Rhee E, Lyapina S, Cope G, Deshaies RJ, Wolf DA (2001) The fission yeast COP9/signalosome is involved in cullin modification by ubiquitin-related Ned8p. BMC Biochemistry 2: 7

Zhou C, Wee S, Rhee E, Naumann M, Dubiel W, Wolf DA (2003) Fission Yeast COP9/Signalosome Suppresses Cullin Activity through Recruitment of the Deubiquitylating Enzyme Ubp12p. Mol Cell 11(4): 927-938.

Zhou P, Howley PM (1998) Ubiquitination and degradation of the substrate recognition subunits of SCF ubiquitin-protein ligases. Mol Cell 2(5): 571-580

CHAPTER II: SPASS, a Biochemical Approach for the Identification of Substrates of Ubiquitin Conjugating Enzymes

2.1. Background

The function of protein ubiquitylation is to target substrate proteins to the proteasome for degradation (Hershko, 1983). Eukaryotic organisms evolved this system to rapidly degrade negative regulators, proteins that are not synthesized correctly, and proteins no longer required within the cell. The 26S proteasome, a multi-subunit protease, recognizes and degrades ubiquitylated substrates only when a polyubiquitin chain of a certain length is formed (Thrower et al., 2000). Ubiquitin chains are formed by the repeated attachment of ubiquitin molecules, which are linked via amide (isopeptide) linkages between the ε-amino group of one of seven lysines of one ubiquitin molecule to the C-terminal carboxyl group of the next ubiquitin in the chain (Chau et al., 1989). Ubiquitin conjugation is performed by a cascade of three major protein classes: ubiquitin activating enzymes (E1s), ubiquitin conjugating enzymes (E2s) and ubiquitin ligases (E3s). Utilizing this enzymatic cascade, a polyubiquitin chain is formed on lysine residues within the substrate protein.

Although the general principles of how ubiquitin molecules are activated and conjugated are fairly well understood (Pickart and Eddins, 2004), the exact mechanism of how a polyubiquitin chain is formed on substrate molecules is not clear (Hochstrasser, 2006). The C-terminus of ubiquitin is first activated by a ubiquitin-activating enzyme (E1). Through an ATP-dependent mechanism, ubiquitin is coupled to a cysteine side chain in E1, yielding a reactive E1~ubiquitin thioester intermediate. The activated ubiquitin is subsequently passed to one of a number of distinct ubiquitin-conjugating enzymes (E2s) by transthiolation to a conserved cysteine of the E2. The E2 proteins catalyze substrate ubiquitylation in conjunction with a ubiquitin-protein ligase (E3).

A large fraction of the proteome is ubiquitylated. A study in budding yeast (Peng et al., 2003) identified over 1000 different ubiquitylated proteins, representing roughly 17% of yeast proteins. The median half-life of a fraction of 3751 proteins in budding yeast was determined to be 43 min (Belle et al. 2006), suggesting that steady state levels of a large portion of proteins in yeast are regulated by constant protein synthesis and degradation. Furthermore the large number and variety of different components of the ubiquitin

conjugation enzyme cascade suggests there may be a multitude of ubiquitylation substrates in eukaryotes. Differentiation between the many substrates is thought to be achieved through interaction with specific E3 substrate adapters. It is thought that there are over 599 substrate adapters for Cullins, an E3 ligase, in humans and over 20 in fission yeast leading to a potentially huge number of substrates for Cullin based E3s alone (Petroski and Deshaies, 2005). A recent study has shown that the budding yeast E3 adaptor Grr1p can recruit more than 10 substrates to the ubiquitin machinery (Benanti et al., 2007), suggesting that multiple substrates can be ubiquitylated by the same E3 scaffold. In addition to the large variety of E3 substrate adaptors there are also many E2 ubiquitin conjugating enzymes (over 50 in humans, 16 in fission yeast) (Jiang and Beaudet, 2004). Whereas some E2s facilitate attachment of other ubiquitin-like proteins (Pickart and Eddins, 2004), the majority of E2s conjugate ubiquitin. It is not exactly clear why there is more than one ubiquitin conjugating enzyme, but it has been suggested that E2s selectively bind specific E3s and thereby confer specificity towards substrates. Another possibility is that ubiquitin conjugating enzymes form a variety of heterodimers which could confer an even greater combinatorial diversity allowing E3 and substrate selection.

The large number of potential ubiquitylating enzymes suggests that substrate-specific ubiquitylation plays an essential role in cellular regulation in eukaryotes. Despite the presence of presumably thousands of substrates and ubiquitin system components, the substrates targeted for degradation by particular E2 and E3 enzymes and have not been easily identified. One reason for the inability of matching substrates and ligases is that E3 recognition elements in substrates are often poorly defined. This minimizes the likelihood that bioinformatic approaches will lead to the identification of E2 or E3 substrates.

Systematic experimental approaches for the identification of specific ubiquitylation substrates can be approached in two ways. First, scientists have taken advantage of the attribute of substrates to accumulate within cells if a component of the ubiquitylation system is compromised. This requires a method of measuring the levels of thousands of proteins rapidly. Approaches used in the past were based on mass spectrometric or 2D-PAGE based protein quantification (Mayor et al. 2007), or on quantitative western blotting (Belle et al., 2006). Whereas proteomic protein quantification can never capture all proteins within a cell, quantitative immunoblotting requires either antibodies directed against every protein within an organism or a library of genomically epitope-tagged strains. Although a library of epitope-

tagged budding yeast strains is available, this method is currently not applicable for other organisms.

A second approach for the identification of specific ubiquitylation substrates are methods based on *in vitro* reconstituted ubiquitylation reactions. A ubiquitylation reaction composed of purified E1, E2, E3-complex and free ubiquitin is mixed with any purified protein and so tested for specific ubiquitylation activity upon ATP hydrolysis. The drawback of this approach is the need to clone and purify thousands of proteins one by one to be tested for specific ubiquitylation in conjunction with a specific set of E2 or E3 enzymes. Recent studies have found elegant ways to partially approach these problems. Gupta and colleagues (Gupta et al., 2007) use substrate microarrays to conduct the *in vitro* ubiquitylation reaction, circumventing the need of thousands of western blots. However, substrates still must be cloned and purified. Ayad and colleagues (Ayad et al., 2005) applied a similar method using a clone library which is *in vitro* transcribed and translated in pools of hundred proteins and used in an *in vitro* ubiquitylation reaction followed by autoradiography and mass spectrometry.

While some of the methods for identification of substrates are very elaborate, many drawbacks remain. Measuring changes in protein levels by mass spectrometry or immunoblotting does not demonstrate the ubiquitylation of the protein, but produces a candidate list requiring more detailed investigation. Identifying substrates in *in vitro* based enzymatic ubiquitylation reactions requires the cloning and purification of thousands of proteins which usually lack critical post-translational modifications necessary for ubiquitylation (reviewed in Petroski and Deshaies, 2005).

In this study, we present a method termed SPASS (**S**olid **P**hase **A**ssay for **S**ystematic Profiling of Ubiquitylation **S**ubstrates) allowing the identification of specific E2 substrates derived from a physiological context without the requirement of extensive protein purification or quantification. SPASS, was developed as a method which purifies and enriches the products of an *in vitro* ubiquitylation reaction controlled by specific ubiquitin conjugating enzymes. Bead-cross-linked ubiquitin is conjugated to an E2 of choice utilizing E1 catalyzed ubiquitin adenylation. As a pool of substrates, cell extract is used, containing all other necessary components of the ubiquitin system. Substrates linked to the ubiquitin beads are putative substrates of the chosen E2s and identified by mass spectrometry. We report a systematic study in which we identified putative *Schizosaccharomyces pombe* Ubc7p and Ubc8p substrates. Fission yeast Ubc7p has the highest sequence homology to Cdc34p, but

whether it functions as SCF E2 is not clear. Fission yeast Ubc8p has not been characterized yet, but it has the highest sequence homology to budding yeast and human Ubc8p. We include experiments demonstrating the biochemical feasibility of the screen as well as analysis and interpretation of the Ubc7p and Ubc8p results.

2.2. Materials and Methods

Chemicals and Materials

All chemicals were purchased from Sigma (St. Louis, MO) unless otherwise noted.

Protein detection by Western Blotting and Coomassie Staining

SDS-PAGE was performed using standard methods. Separated proteins were stained with Coomassie Brilliant Blue or transferred onto a PVDF membrane for immunoblotting (Immobilon-P; Millipore). Processing of membranes was performed as described previously (Wolf et al., 1999). 6xHis-tagged proteins were detected using a monoclonal anti-His antibody (ab18184, Abcam). Ubiquitin was detected with mouse monoclonal antibody P4D1 raised against ubiquitin (Santa Cruz). Mouse monoclonal antibodies were detected with polyclonal HRP-conjugate rabbit anti-mouse secondary antibody (Jackson).

Plasmids and Yeast Strains

The yeast strain DS448/2 (927 h- leu-1-32 ura4d-18) was used for all experiments. Plasmids used are listed in table 2-1:

ID	Name	Source
C210/1	pET28b-Ubc7	obtained from C. Zhou
C214/4	pET28b-Ubc8	obtained from C. Zhou
C210/2	pET28b-Ubc11	obtained from C. Zhou
C215/1	pET28b-Cdc34	obtained from C. Zhou
M172/1	pGex6p-Rsp5	obtained from J. Huibregtse
M172/2	pGex6p-Rsp5-CA	obtained from J. Huibregtse

Table 2-1: **Plasmids use in this study**

Purification of Recombinant Proteins

6x His-tagged *S. pombe* Ubc7p, Ubc8p, Ubc11p, and *S. cerevisiae* Cdc34p were purified using from pET28b plasmids using the *Escherichia coli* strain BL21 and immobilized metal affinity chromatography as described by the manufacturer (Clontech) Transformed bacteria were grown at 37^{o}C to an optical density (OD) of 0.6 at 595 nm in 1l of Luria broth (LB) and expression was induced by addition of 1mM isopropyl-b-1-thio-D-galactopyranoside (IPTG). After 3h of induction at room temperature (RT), the cells were harvested and lysed by

sonication in 30ml His-lysis buffer (phosphate buffered saline, 100mM NaCl, 1% Triton x-100, 10 mM imidazole, 10mM β=mercapto ethanol) and protease inhibitors (10µg/mL leupeptin, 10µg/mL pepstatin, 15µg/mL aprotinin, 1mM phenylmethylsulfonylfluoride (PMSF)). Lysates were clarified by centrifugation at 40,000g for 40min and His-tagged proteins were extracted from the clarified lysate by batch binding for 2h at 4°C using 3ml TALON super flow resin (Clontech). The resin was transferred to Flex Columns (Spectrum) and washed with 200ml His-lysis buffer and 10mM imidazole, followed by equilibration in 10ml 25mM Tris pH 8.0. His-tagged proteins were eluted in a buffer containing 10ml 25mM Tris pH 8.0, 200mM imidazole and 1mM DTT, brought to 50% glycerol and stored in aliquots at -80°C. The purified proteins were analyzed by 12% SDS-PAGE and Coomassie Blue staining and quantified using the D_c protein assay (BioRad) based on the Lowry method, resulting in 9.8mg/ml Cdc34p, 1.1mg/ml Ubc7p, 9.5mg/ml Ubc11 and 5.2mg/ml Ubc8p.

GST-tagged *S. cerevisiae* Rsp5p and Rsp5CAp were purified from pGex6p plasmids (obtained from Jon Huibregtse, University of Texas at Austin) using the *Escherichia coli* strain BL21 as described above, with the exception that 3ml of glutathione–Sepharose resin (Amersham Biosciences) was used instead of TALON resin, and equilibration was carried out in 25ml of PreScission cleavage buffer (PCB: 50mM Tris–HCl, pH 7.0, 150mM NaCl, 1mM TCEP, 10% glycerol) after the wash in 200ml lysis buffer. Rsp5p and Rsp5CAp were proteolytically cleaved from the GST tag by incubating the resin for 4h with 1ml of PCB containing 40U of PreScission protease (Amersham Biosciences) based on the manufacturer's instructions, aliquoted and stored at -80°C. The purified proteins were analyzed by 12% SDS-PAGE and Coomassie Blue staining and quantified using the D_c protein assay (BioRad) based on the Lowry method, resulting in 3.6mg/ml Rsp5p and 3.1 mg/ml Rsp5CAp.

In vitro Ubiquitylation Assays

All purified E2s were tested for ubiquitin conjugation activity in standard *in vitro* ubiquitylation assays based on a published protocol (Zhou et al., 2001). 5µg E2 was incubated with 0.1µg E1 (Ube1p, E-301, Boston Biochem), 0.8µM ubiquitin monomer (Sigma), reaction buffer (RX: 4mM MgAc, 1mM DTT, 1mM PMSF), one tenth of the volume of 10xATP regeneration system (10xATP-RS: 20 mM HEPES pH 7.4, 10 mM ATP, 300 mM creatine phosphate, 10 mM MgAc, 1.5 mg/ml creatine kinase, 10% glycerol) and the volume was adjusted to 20µl with HEPES buffer (20mM HEPES, pH 7.4, 100mM KAc, 1

mM DTT). After 2h incubation at 30°C, samples were analyzed by 12% SDS-PAGE followed by anti-his immunoblotting (see Supplement).

Activity of purified Rsp5p was tested by incubating the following reagents for 2h at 30°C: 300pmol Ubc11p, 100pmol Rsp5p, 10pmol E1 (Ube1p, E-301, Boston Biochem), 0.8µM ubiquitin monomer (Sigma), reaction buffer (RX), one tenth of the volume of 10xATP-RS and the volume was adjusted to 20µl with HEPES buffer. Negative controls were performed by replacing either Ubc11p with buffer or Rsp5p with Rsp5CAp. Samples were analyzed by 12% SDS-PAGE followed by anti-his immunoblotting (see Supplement).

Analytical SPASS Assays

The reverse phase transthiolation reactions were performed in a total volume of 30µl. Ubiquitin-agarose (U-400, Boston Biochem) was prepared by three washes in 30 volumes of HEPES buffer and diluted to the indicated concentrations. The indicated amounts of ubiquitin coupled to beads were adjusted to 10µl in HEPES buffer and incubated with 4µg E2, RX, 0.5µg E1 (E-201, Boston Biochem) and adjusted to 27µl with HEPES buffer. To start the reactions, 3µl 10xATP-RS (or 3µl HEPES buffer for the negative controls) was added and the reactions were placed in a 30°C heat block for the indicated times. Beads were pelleted for 1min at 5,000g using a table-top centrifuge and aliquots of the supernatant were analyzed by 12% SDS-PAGE followed by immunoblotting.

For the DUB cleavage assay, 100µg ubiquitin-beads were incubated with 20µg Cdc34p, 0.5µg E1, RX, 0.8mM ubiquitin and 3µl 10xATP-RS and adjusted to a total volume of 30µl with HEPES buffer. After 90min incubation at 30 °C, beads were washed 5 times in wash buffer (Tris pH 7.5, 250mM NaCl, 0.5% Triton x100) and the volume was adjusted to 20µl. The beads were incubated after the addition of 1µg/µl purified Otu1p deubiquitylating enzyme (obtained from Suneng Fu, Harvard University, Boston) for one additional hour at 30°C, pelleted at 5,000g with a table-top centrifuge and aliquots of the supernatant were analyzed by 10% and 15% SDS-PAGE and anti-His and anti-ubiquitin immunoblotting.

For the E2-E3 transthiolation assay, in a first step, 20µg *S. pombe* Ubc11p was incubated with 0.5µg E1, 100µg ubiquitin agarose, 10xATP-RS, RX and HEPES buffer in a total volume of 30µl for 60min at 30 °C. The beads were then washed 5 times in wash buffer (25mM Tris pH 7.5, 250mM NaCl, 0.5% Triton x100). The washed Ubc11p-ubiquitin-beads were split into three 10µl aliquots and incubated with 5µl wash buffer, buffer containing 10µg purified *S. cerevisiae* Rsp5CAp or 10µg purified *S. cerevisiae* Rsp5p respectively.

After 60min at 30°C, beads were pelleted at 5,000g in a table-top centrifuge and aliquots of the supernatant were analyzed by 12% SDS-PAGE and immunoblotting.

Preparative SPASS Assays

500µg E2 was incubated with 1mg ubiquitin-agarose equilibrated in HEPES buffer, 5µg E1, RX, 10xATP-RS and the volume was adjusted to 1ml with HEPES buffer. After 90min at 30°C, beads were gently washed 5 times in HEPES buffer containing 0.1% Tx100. Washed E2-conjugated beads and beads that were prepared in the same way without the addition of ATP were incubated with 10mg of fresh wild-type *S. pombe* (DS448/2) cell extract prepared as described below. After incubation for 1h at 30°C, beads were washed three times for 10min in HEPES buffer containing 0.1% Triton x-100, 8M Urea and twice for 10min in HEPES buffer, then equilibrated in Tris pH 7.5, 250mM NaCl, 0.5% Triton x-100. 1µg Otu1p was added to 90% slurry beads for 20h at 30°C. After DUB cleavage, the supernatant was transferred to a new tube and beads were washed once in 100µl analytical grade H_2O. The wash was united with the protein digest, resulting in 200µl samples which were precipitated by the addition of 66µl trichloroacetic acid (TCA) for 24 hours at 4°C. The protein pellets was washed once in -20°C pre-cooled acetone, air-dried and stored at -20°C until further analysis.

Preparation of S. pombe Cell Extract

S. pombe growth media and techniques were used as previously described (Moreno et al., 1991). In rich medium (YES: yeast extract, glucose, supplements) exponentially growing wild-type *S. pombe* cells (DS448/2) were washed once in ice-cold sterile water and then lysed by bead lysis and a FastPrep device (Bio 101) in a buffer containing 25mM HEPES pH 8.0, 1mM DTT, 10mM NEM, 1µg/ml MG132, 10mM NaF, 10µg/ml leupeptin, 10µg/ml pepstatin, 15µg/ml aprotinin, 1mM PMSF and 0.1% Triton x-100. Cell extracts were clarified by centrifugation for 15min at 4°C at 16,000g and used immediately in the SPASS assays.

Mass Spectrometry

Protein pellets were prepared for mass spectrometry by triple digest based on a protocol from the Yates Laboratory (The Scripps Research Institute, LaJolla CA): The protein pellets were reconstituted in resuspension buffer (8M Urea, 100mM Tris pH 8.5, 5mM Tris(2-carboxyethyl)phosphine (TCEP)) and incubated at 20°C for 30min. The solutions were

adjusted to 10mM iodoacetamide (IAM) and incubated for 30min at 20°C in the dark. Each sample was then split into three equal aliquots. The first aliquots were diluted to 2M Urea with 100mM Tris pH 8.5. 1M $CaCl_2$ was added to a final concentration of 2mM and modified Trypsin (Roche) was added to a final substrate to enzyme ratio of 0.01µg/µl. The Trypsin digests were incubated over night shaking at 37°C in the dark. The second aliquots were diluted to 2M Urea with 100mM Tris pH 8.5, and Elastase (Boehringer Mannheim) was added to a final substrate to enzyme ratio of 0.005µg/µl and incubated overnight shaking at 37°C protected from light. The third aliquots were diluted to 4M Urea with 100mM Tris pH 8.5, and Subtilysin (Boehringer Mannheim) was added to a final substrate to enzyme ratio of 0.004µg/µl and incubated overnight shaking at 37°C protected from light. The three aliquots of each sample were recombined and formic acid was added to a final concentration of 5%. Solid phase extraction with SPEC-PLUS PTC18 Cartridges (Ansys Technologies) was performed as follows: membranes were first washed with 50µl MeOH and then with 400µl 0.5% acetic acid. Peptide samples were loaded followed by four washes with 400µl 0.5% acetic acid. Peptides were eluted with 200µl 90% acetonitrile, 0.5% acetic acid and 5% acetonitrile. Both samples and controls were analyzed by MudPIT as previously described (Washburn et al., 2001). The MS data was searched against an *S. pombe* FASTA protein database using the SEQUEST tool version 2.7 and DTAselect (Tabb et al., 2002). The following cut-off parameters were used to filter the results:

Parameter	Criteria
Minimum +1 XCorr	1.8
Minimum +2 XCorr	2.5
Minimum +3 XCorr	3.5
Minimum DeltCN	0.08
Minimum charge state	1
Maximum charge state	3
Minimum ion proportion	0
Maximum Sp rank	1000
Minimum Sp score	-1
Modified peptide inclusion	Include
Tryptic status requirement	Any
Minimum peptides per locus	2

Table 2-2: **SEQUEST parameters**

Only proteins that had been identified by two or more peptides were included in the results.

2.3. Results and Discussion

Principles of SPASS

For this study, *S. pombe* Ubc7p and Ubc8p substrates were identified, using a strategy based on an *in vitro* ubiquitylation system. The SPASS (**S**olid **P**hase **A**ssay for **S**ystematic Profiling of Ubiquitylation **S**ubstrates) approach presented here is based on *in vitro* ubiquitylation reactions, taking advantage of the ability to directly control and measure specific protein ubiquitylation, while relying on a substrate pool derived from cells, thus circumventing the need of purifying every single protein in an organism. As discussed above, the main advantage of using a system reconstituted from purified proteins is the ability to directly monitor the attachment of ubiquitin moieties to a single substrate at a time. In the past, the main disadvantage of this approach was the necessity of cloning and purifying thousands of proteins serving as substrates, and the potential absence of key post-translational modifications of these substrates that are essential for ubiquitylation. Further, ubiquitylation only occurring under specialized physiological conditions could not be captured. The use of purified proteins as substrates also limited investigation to model systems like budding yeast for which extensive clone libraries exist. The SPASS approach has eliminated these problems.

SPASS consists of four basic steps (Figure 2-1). In the first step, a specific E2 is linked to N-terminally immobilized ubiquitin. Purified E1 is incubated with bead-immobilized ubiquitin, ATP, and purified E2. Upon completion of the reaction, the ubiquitin-beads are washed to remove free E1, free E2 and ATP. The resulting thiol-linked E2-ubiquitin beads are then incubated with *S. pombe* cell extract in a second step. Here, the ubiquitin conjugated E2 is allowed to interact with its specific E3 partners contained in the cell extract, and the ubiquitin is transferred to lysine residues within specific substrates. In a third step, the reaction is halted by chemical denaturation. Following rigorous wash steps and re-equilibration, substrates are eluted in a fourth step by protease cleavage utilizing the highly site specific deubiquitylating enzyme (DUB) Otu1p. Substrates are subsequently identified in the eluate using tandem mass spectrometry. Before application of this assay for the identification of *S. pombe* ubiquitylation substrates, we addressed feasibility of each step in detailed biochemical experiments.

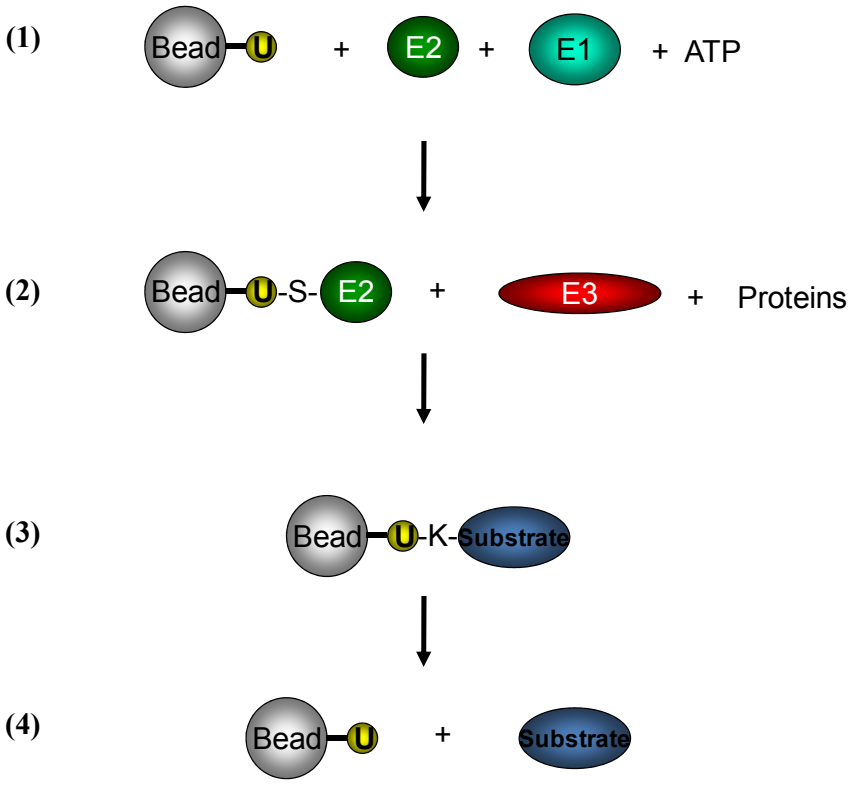


Figure 2-1: **Schematic outline of SPASS**

Ubiquitin Conjugating Enzymes form a Thioester with Immobilized Ubiquitin by Reversed Phase Transthiolation

The formation of a thioester linkage between ubiquitin and the active site cysteine of ubiquitin conjugating enzymes is essential for the occurrence of substrate ubiquitylation (Pickart and Eddins, 2004). From *in vitro* ubiquitylation studies (Raasi and Pickart, 2005) it is well established that incubation of purified E1, purified E2, ATP and free ubiquitin results in ubiquitin which is conjugated to the active-site cysteine residue within the E2. Whether this is also possible using immobilized ubiquitin was unknown. In order to investigate the formation of a thiol linkage between immobilized ubiquitin and ubiquitin conjugating enzymes, we performed a number of experiments based on the schematic outline in Figure 2-2. E1, E2, ATP and ubiquitin-beads were allowed to react (1). After incubation a sample was taken from the supernatant of the reaction (2). The amount of E2 depleted from the super natant was used as a measure of E2-ubiquitin conjugation.

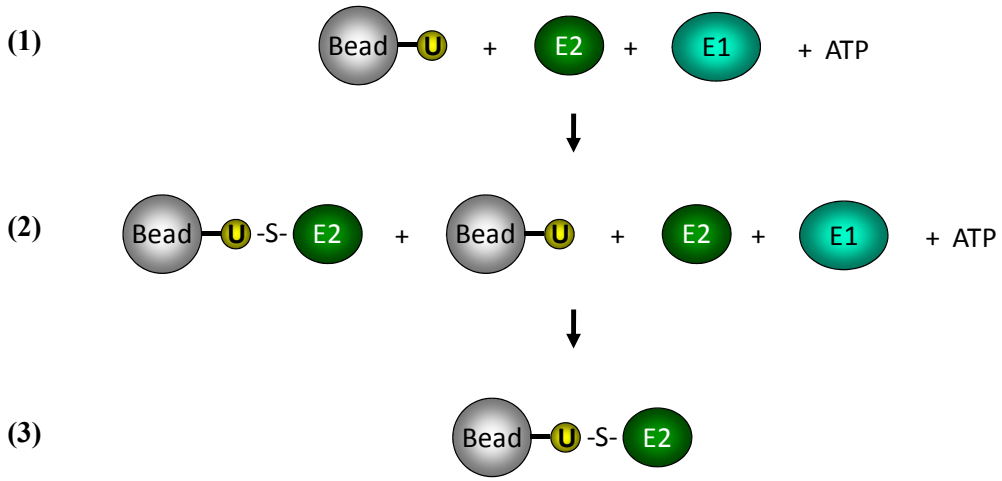


Figure 2-2: **Schematic outline of ubiquitin-bead-E2 thioester formation**

To determine whether E2 was conjugated to immobilized ubiquitin in an energy-dependent manner, purified *S. pombe* Ubc7p or *S. cerevisiae* Cdc34p were allowed to react with immobilized ubiquitin and purified E1 in the presence or absence of ATP in a time-course experiment, and samples from the supernatant were analyzed by immunoblotting (Figure 2-3). Although low affinity of E2 to ubiquitin was observed in absence of ATP, its addition dramatically accelerated the conjugation of E2 to immobilized ubiquitin. This energy-dependent E2 depletion strongly suggested that the reversed-phase ubiquitin-conjugation reaction took place.

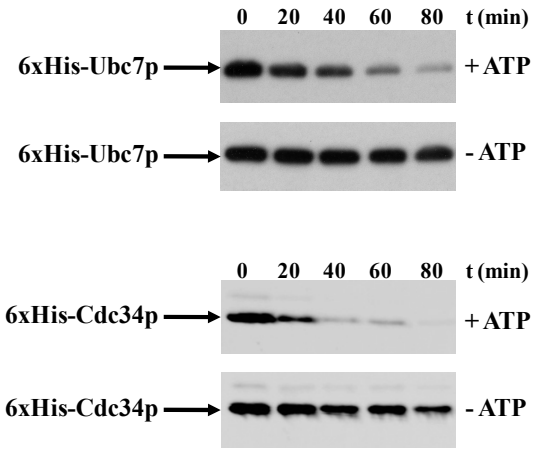


Figure 2-3: **E2s bind to ubiquitin-beads in an energy-dependent manner**. 4µg *S. p.* Ubc7p or 4µg *S. c.* Cdc34p were incubated with 0.5 µg E1 and 10ug ubiquitin-beads with or without ATP and an aliquots of the supernatants were analyzed by anti-His immuno blotting after the indicated reaction times.

To further investigate how E2 conjugation was influenced by the levels of available immobilized ubiquitin, we performed four reactions, where the amount of E1, E2 and ATP remained constant and the amount of ubiquitin linked to a fixed volume of beads increased. E2 levels were determined before the addition of ATP and after the reaction was carried out for 60min in the presence of various amounts of ubiquitin (Figure 2-4).

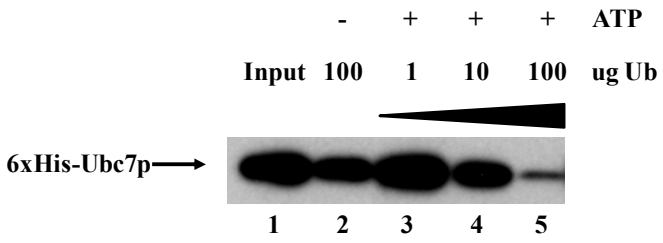


Figure 2-4: **ATP-dependent Ubc7p conjugation is limited by the availability of ubiquitin**: 10µg Ubc7p was incubated with 0.5µg E1, ATP and 1µg, 10µg or 100µg ubiquitin-beads and aliquots were analyzed after 60min by His immuno blotting.

A clear dose-dependent relationship was observed (lanes 3-5) in comparison to the ATP-negative sample (lane 2) and the starting signal (lane1), indicating the saturation of ubiquitin molecules with conjugated E2. Because 100µg ubiquitin were necessary to conjugate to 10µg Ubc7p, we concluded based on the molar ratios that 2.5% of the immobilized ubiquitin molecules were accessible to E2 conjugation.

In order to determine whether a thioester linkage was formed between the active site cysteine of the E2 and ubiquitin molecules, ubiquitin-beads which had been conjugated to purified *S. cerevisiae* 6xHis-Cdc34p were washed to remove free protein and a sample of the supernatant was analyzed by anti-His immunoblotting (Figure 2-5). Ubiquitin-beads which had been incubated with Cdc34p and E1 in the absence of ATP served as negative control. As expected no signal was observed under control conditions, demonstrating the absence of free E2 (lanes 1, 2) after the wash step. The washed Cdc34-conjugated ubiquitin-beads were then incubated with wash solution containing 50mM DTT and samples of the supernatant were analyzed. As expected, Cdc34p was recovered only from those ubiquitin beads that had been Cdc34-conjugated in the presence of ATP, but not in its absence (lanes 3, 4).

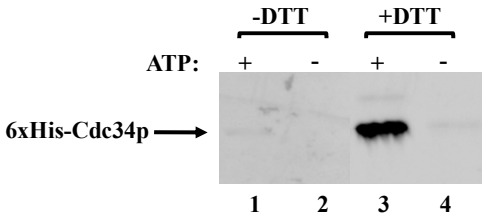


Figure 2-5: **Cdc34p forms a thioester with immobilized ubiquitin:** 4µg Cdc34p were incubated with 0.5µg E1, ATP and 10µg ubiquitin-beads. After 60 min, beads were washed three times in a buffer containing 8M Urea, 25mM Tris pH7.5 and samples were analyzed by anti-His immuno blotting (Lanes 1, 2). The beads were then incubated for 30min in the same buffer containing 50mM DTT and samples were analyzed (lanes 3, 4).

In summary, these results indicate that purified E2 enzyme can be covalently conjugated to N-terminally immobilized ubiquitin in such a way that the C-terminus of ubiquitin is linked to the active site cysteine residue of the E2. The conjugation is energy- and E1-dependent and exhibits dose-responsive behavior when the ubiquitin concentration is increased.

Immobilized Isopeptide Bond Linked Proteins are cleaved by the Deubiquitylating Enzyme Otu1p

Next, we determined whether an isopeptide bond could be formed between a lysine residue of a substrate and immobilized ubiquitin. We also assessed whether this isopeptide bond could be specifically cleaved by a deubiquitylating enzyme. As a model substrate we chose *S. cerevisiae* Cdc34p, which is known to have lysine–directed autoubiquitylation activity (Banerjee et al., 1993), and can therefore serve as both a ubiquitin conjugating enzyme and a substrate. The experiment is outlined in Figure 2-6. When Cdc34p is incubated with ubiquitin-beads, E1 ATP, and free ubiquitin, the E2 should polyubiquitylate itself on a lysine residue using the available free ubiquitin, but also incorporate ubiquitin molecules which are linked to beads. This incorporation of free and immobilized ubiquitin molecules should produce beads covalently connected to Cdc34p via a polyubiquitin chain made of isopeptide linkages. If purified deubiquitylation enzyme is added to washed polyubiquitin-linked Cdc34p-beads, it should only be possible to recover Cdc34p and mono-ubiquitin if they had been linked by isopeptide bonds between a substrate lysine and the di-glycine C-terminus of ubiquitin (Figure 2-7) (Wilkinson, 1997).

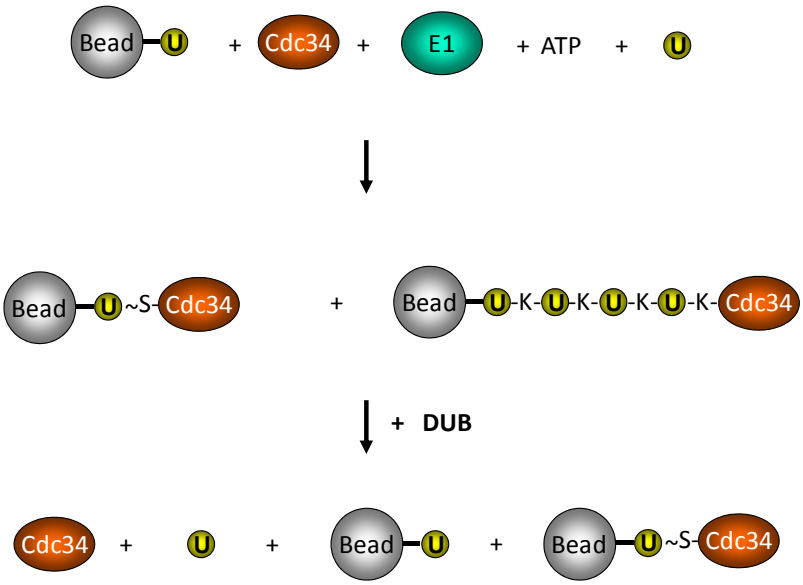


Figure 2-6: **Schematic outline of isopeptide formation using Cdc34p autoubiquitylation activity**

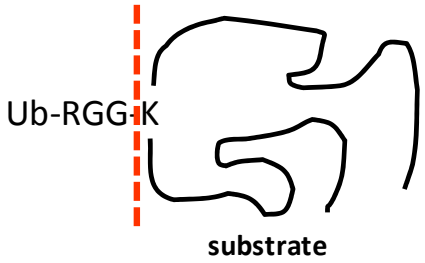


Figure 2-7: **Illustration of DUB isopeptide bond cleavage**. The red line indicates the cleavage between the C-terminal di-glycine residue of ubiquitin (left side) and the lysine residue of the substrate protein (right side).

To investigate the formation of isopeptide-linked Cdc34p, we incubated ubiquitin-beads with Cdc34p, E1, and free ubiquitin monomers with or without ATP (Figure 2-8). The formation of free poly-ubiquitylated Cdc34p was observed in the supernatant of the reaction (lane 2), as well as the formation of free high-molecular weight poly-ubiquitin chains (lane 4). A reduction in mono-ubiquitin moieties was also indicative of poly-ubiquitylation activity (lane 4). Beads were washed extensively after 90min incubation to remove any free protein and purified *S. pombe* Otu1p DUB protease was added to the beads to cleave substrates bound to the beads through isopeptide linked poly-ubiquitin chains.

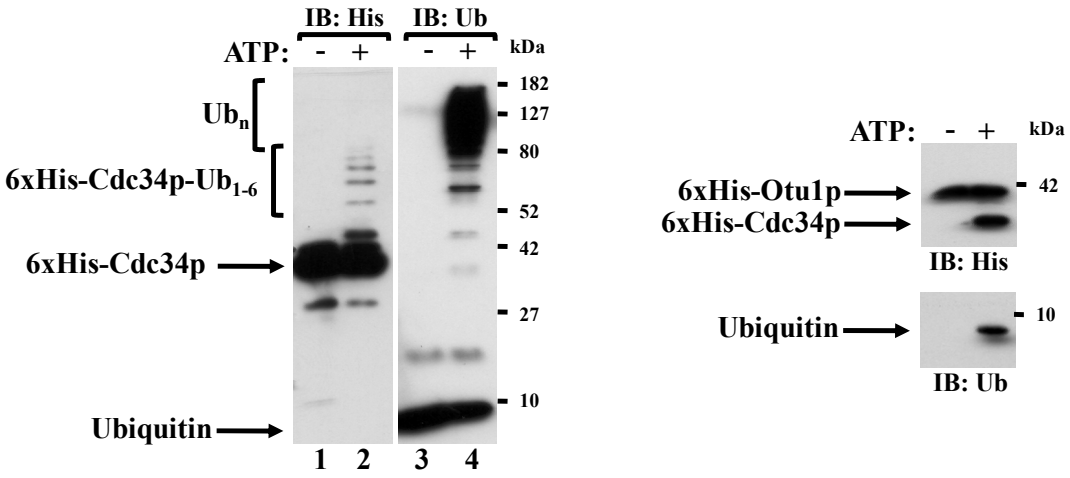


Figure 2-8: **Otu1p cleaves Cdc34p and ubiquitin-beads linked by isopeptide bonds.** Left panel: Supernatant of a reaction of ubiquitin-beads, free ubiquitin, E1, Cdc34p with (lane 2) or without (lane 1) ATP after 90min incubation at 30°C. The supernatants of the reactions have been blotted with anti-His (lanes 1, 2) or anti-ubiquitin (lanes 3, 4) antibodies. Right panel: Beads of the reactions with and without ATP have been washed 5 times in buffer containing 25mM Tris pH 7.5, 250mM NaCl, 0.5% Triton x100 and then incubated 1h at 37°C after the addition of 1µg of purified Otu1p protease. Supernatants were analyzed by anti-his and anti-ubiquitin immunoblotting.

After two hours of incubation, Cdc34p and mono-ubiquitin were recovered only from Cdc34p-ubiquitin-beads that had been prepared in the presence of ATP (Figure 2-8, right panel). Beads incubated with the proteins in the absence of ATP did not result in cleavage of E2 and ubiquitin, suggesting the formation of isopeptide linkages.

Only Catalytically Active Rsp5p displaces immobilized thioester linked E2

In order to investigate whether E2 which had been conjugated via thioester linkage to immobilized ubiquitin could still interact with an E3 ubiquitin ligase, we used the *S. cerevisiae* homologous of the E6-associated protein (E6AP) C-terminus (HECT) -domain ligase Rsp5p in an experiment that is outlined in Figure 2-9.

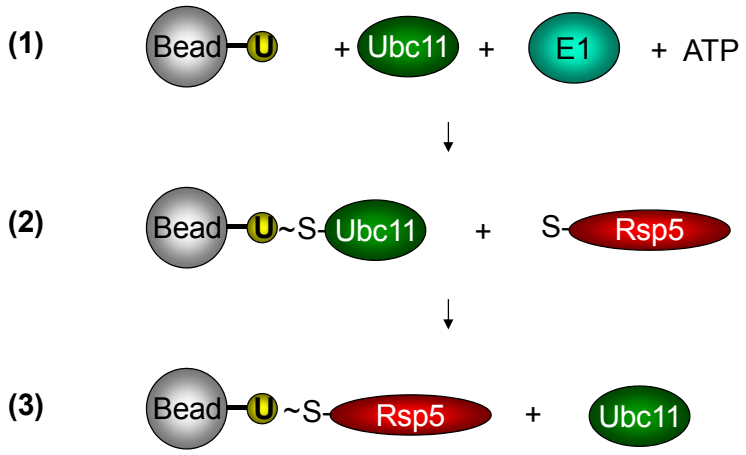


Figure 2-9: **Schematic outline of E2-HECT E3 transthiolation**

S. pombe Ubc11p was conjugated to ubiquitin-beads by incubation with E1 and ATP (1). The Ubc11p-conjugated ubiquitin-beads were then extensively washed to remove any ATP and free protein and incubated with either purified wild-type Rsp5p or purified mutant Rsp5CAp where the active site cysteine 777 residue was replaced by alanine (2) (Wang et al., 1999). Free Ubc11p was recovered in the supernatant only after beads were incubated with Rsp5p, but not Rsp5CAp (3), indicating successful transthiolation (Figure 2-10). This result indicates that E2 which is thioesterified with immobilized ubiquitin can be released in exchange for the active site cysteine of a HECT-domain ubiquitin ligase in an ATP-independent reaction.

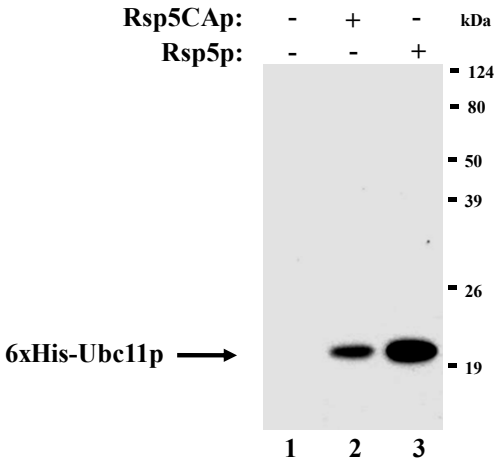


Figure 2-10: **Rsp5p but not Rsp5CAp forms a thioester with ubiquitin-beads in an energy-independent manner, releasing pre-conjugated E2.** *S. pombe* Ubc11p was incubated with E1, ubiquitin-beads and ATP for 60min at 30°C. The beads were then washed 5 times in wash buffer and a sample was taken from the super natant (lane 1). The washed Ubc11p-ubiquitin-beads were split into two aliquots and incubated with buffer containing 10µg purified untagged Rsp5CAp (lane 2) or 10µg purified untagged Rsp5p (lane 3) for 60min at 30°C. The supernatants of the reactions were analyzed by anti-His immunoblotting.

SPASS of S. pombe Ubc7p and Ubc8p

We used full deployment of SPASS to identify putative *S. pombe* Ubc7p and Ubc8p substrates. Cell extract, prepared from vegetative exponentially growing wild-type *S. pombe* cells was incubated with washed Ubc7p or Ubc8p pre-conjugated ubiquitin-beads as described in detail in the methods section. Equal amounts of ubiquitin-beads that were incubated with each of the E2s in the absence of ATP served as negative controls. Following the reaction, beads were washed and denatured, leaving only covalently linked proteins. Linked proteins were then cleaved off the beads using Otu1p and analyzed by tandem mass spectrometry. Proteins present in both ATP-positive samples and the ATP-negative controls were considered to be non-specific interactions and were disregarded.

Of the 166 putative substrates identified in the Ubc7p sample, 77 were present also in the negative control, including Otu1p and Ubc7p. Similarly, 160 of the 226 putative Ubc8p substrates were present in the negative control including Otu1p and Ubc8p (see Supplementary Table 2-1). Whereas Otu1p was present in equal quantities based on spectrum counts (Liu et al., 2004) in samples and controls, as would be expected since it was added in equal amounts to both samples, significant differences in the amounts of Ubc7p and Ubc8p were observed in samples incubated in the presence of ATP. Ubc7p was 55 times more abundant in the samples prepared in the presence of ATP, and Ubc8p was 12 times more

abundant in the ATP-positive sample, compared to the respective controls. The overabundance of Ubc7p and Ubc8p in the ATP-positive sample underlines the successfully carried out enzymatic ubiquitin-conjugation reaction and is likely due to autoubiquitylation (Chen et al., 1991). The relatively small amount of Ubc7p and Ubc8p present in the ATP-negative sample could be due to non-covalently bound protein that could not be removed in the denaturing wash steps. All other non-specifically bound proteins (75 for Ubc7p, and 158 for Ubc8p) were present in equal amounts in samples and controls based on quantification by spectrum counts (Figure 2-11).

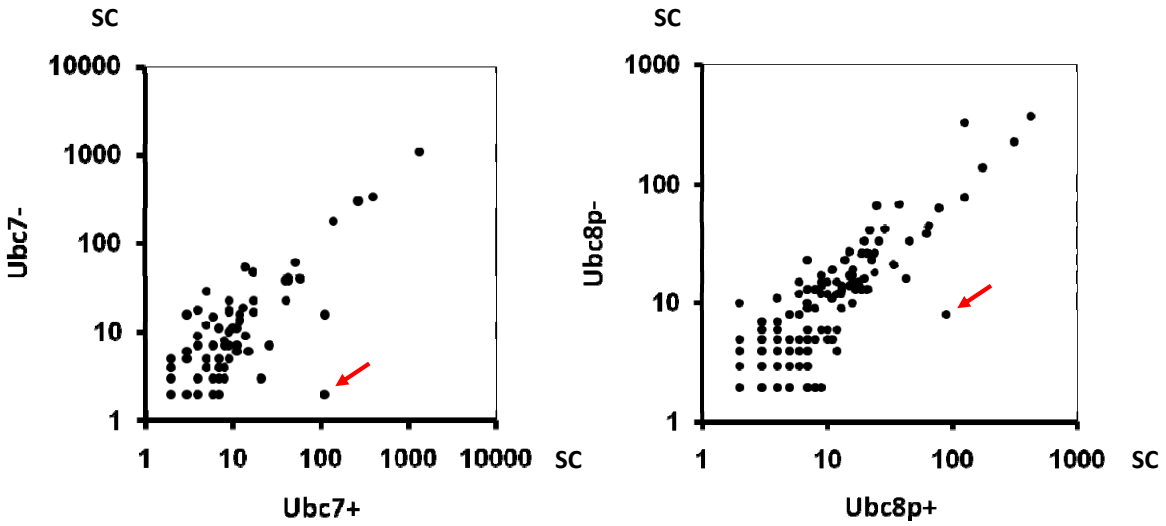


Figure 2-11: **Non-specific proteins are present in samples and controls in equal quantities**. Spectrum counts (SC) for proteins identified in ATP-positive and ATP-negative reactions were plotted. Red arrows indicate significantly higher levels of recovered Ubc7p or Ubc8p in ATP+ samples.

The majority of proteins were identified as non-specific in all four samples and were generally highly abundant proteins such as translation elongation factors, ubiquitin or actin (Figure 2-12). Proteins found in all four samples could have been attached to free unoccupied bead-linked ubiquitin utilizing an E2-E3 cascade present in the cell extract. This would explain how these proteins could have been recovered by DUB-cleavage which is specific to ubiquitin-substrate isopeptide linkages. The second explanation would be that certain very abundant proteins were not completely removed from beads, even after stringent wash steps. The second explanation seems to be likely since those proteins were well known non-specific contaminants in large scale affinity purifications (Krogan, 2006).

Proteins identified exclusively in the Ubc7p sample and Ubc7p control but in neither the Ubc8p sample nor the Ubc8p control, and vice versa, are most likely specific Ubc7p or Ubc8p contaminants respectively. This explanation can be rationalized by the finding that no protein was identified exclusively in either of the negative controls.

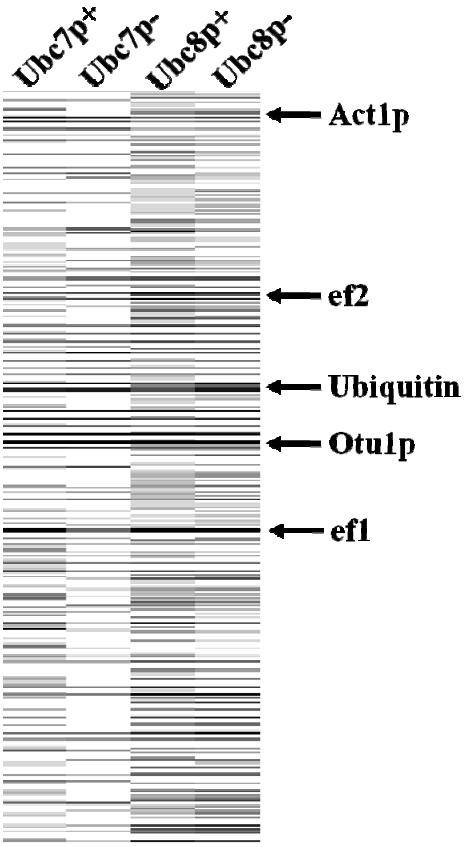


Figure 2-12: **Virtual gel of SPASS-identified proteins**. Spectrum counts of proteins identified in samples (+) and controls (-) are displayed in shades of grey, where darker shades indicate higher abundance and white reflects absence of a protein. High-level contaminants are indicated as examples.

55 proteins that were exclusively identified in the ATP-positive Ubc7p sample were regarded as putative Ubc7p substrates. (Table 2-3). Interestingly, three HECT-domain E3 ubiquitin ligases were among the proteins: the *S. pombe* Rsp5p ortholog Pub1p, the putative Ufd4p ortholog SPAC12B10.01c which had not been identified yet in fission yeast, and Ptr1p. Although HECT-domain E3s utilize their catalytic cysteine for ubiquitin ligase activity, it has been proposed that Rsp5p and other HECT-domain E3s assemble ubiquitin polymers on a lysine residue (Verdecia et al., 2003, Mouchantaf et al., 2006, Nuber et al., 1998) through autoubiquitylation, providing a valid explanation for the detection of these

proteins in DUB-cleaved sample. The finding that these HECT-domain ubiquitin ligases were identified exclusively in the Ubc7p sample suggests that specifically *S. pombe* Ubc7p, but not Ubc8p serves these ligases. In addition, four proteins previously known to be Rsp5p substrates in budding yeast were identified. These included the three uncharacterized fission yeast proteins SPBC660.06, SPAC1002.12c, and SPCC1223.07c, which are putative orthologs to Wwm1p (Kus et al., 2005, Gupta et al., 2007), Uga2p (Kus et al., 2005) and Dps1p (Gupta et al., 2007) respectively, and *S. pombe/S. cerevisiae* Tps1p. Most importantly, Wwm1p and Dps1p were shown to belong to a group of Rsp5p substrates that do not to bind to the E3 (Gupta et al., 2007), suggesting that the identification of these substrates was in fact a result of reversed phase ubiquitylation.

26 of the 51 proteins identified (51%) were shown previously to be ubiquitylated in budding yeast (Peng et al., 2003, Hitchcock et al., 2003), whereas the other 25 had not yet been identified as substrates. As determined by gene ontology mapping, 10 of the proteins identified were involved in vesicular transport and seven in translation, processes known to be extensively regulated through ubiquitylation (Saleki et al., 1997, Urbe, 2005, Mukherjee et al., 2006, Hicke et al., 2003).

Locus	Name	Product	Ortholog	Name
SPBC582.08	SPBC582.08	alanine aminotransferase	YDR111C	ALT2
SPAC4G9.09c	arg11	N-acetyl-gamma-glutamyl-phosphate reductase/acetylglutamate kinase	YER069W	ARG5,6
SPAC631.02	SPAC631.02	bromodomain protein	YDL070W	BDF2
SPAC26A3.05	chc1	clathrin heavy chain Chc1	YGL206C	CHC1
SPBC31F10.11c	cwf4	complexed with Cdc5 protein Cwf4	YLR117C	CLF1
SPAC3C7.11c	cnx1	calnexin	YAL058W	CNE1
SPBC56F2.09c	arg5	arginine specific carbamoyl-phosphate synthase Arg5	YOR303W	CPA1
SPCC1223.07c	SPCC1223.07c	aspartate-tRNA ligase	YLL018C	DPS1
SPAC23A1.12c	SPAC23A1.12c	phenylalanine-tRNA ligase beta subunit	YLR060W	FRS1
SPAC3G9.06	frs2	phenylalanine-tRNA ligase alpha subunit Frs2	YFL022C	FRS2
SPBC12C2.11	SPBC12C2.11	glutamine-fructose-6-phosphate transaminase	YKL104C	GFA1
SPBC26H8.06	grx4	glutaredoxin Grx4	YDR098C	GRX3
SPBC4F6.17c	SPBC4F6.17c	mitochondrial matrix chaperone Hsp78	YDR258C	HSP78
SPAC17G8.06c	SPAC17G8.06c	dihydroxy-acid dehydratase	YJR016C	ILV3
SPAC22A12.15c	bip1	BiP	YJL034W	KAR2
SPAC25G10.07c	cut7	kinesin-like protein Cut7	YBL063W	KIP1
SPAC56F8.10	met9	methylenetetrahydrofolate reductase Met9	YGL125W	MET13
SPAC139.01c	SPAC139.01c	nuclease, XP-G family	YNL085W	MKT1
SPAC1142.04	SPAC1142.04	Noc2p-Noc3p complex subunit Noc2 family	YOR206W	NOC2
SPAC23D3.06c	nup146	nucleoporin Nup146	YDR192C	NUP42
SPCC757.09c	rnc1	RNA-binding protein that suppresses calcineurin deletion Rnc1	YBR233W	PBP2
SPAPB2B4.04c	SPAPB2B4.04c	P-type ATPase, calcium transporting Pmc1	YGL006W	PMC1
SPAC23G3.11	rpn6	19S proteasome regulatory subunit Rpn6	YDL097C	RPN6
SPBC16C6.07c	rpt1	19S proteasome regulatory subunit Rpt1	YKL145W	RPT1
SPCC1682.16	rpt4	19S proteasome regulatory subunit Rpt4	YOR259C	RPT4
SPAC3A11.12c	rpt5	19S proteasome regulatory subunit Rpt5	YOR117W	RPT5
SPBC23G7.12c	rpt6	19S proteasome regulatory subunit Rpt6	YGL048C	RPT6
SPAC11G7.02	pub1	ubiquitin-protein ligase E3	YER125W	RSP5
SPBC31F10.06c	sar1	ADP-ribosylation factor Sar1	YPL218W	SAR1
SPBC24C6.05	sec28	coatomer epsilon subunit	YIL076W	SEC28
SPAC4D7.01c	sec71	Sec7 domain	YDR170C	SEC7
SPAC1F12.07	SPAC1F12.07	phosphoserine aminotransferase	YOR184W	SER1
SPAC29A4.15	SPAC29A4.15	serine-tRNA ligase	YDR023W	SES1
SPAC19D5.04	ptr1	ubiquitin-protein ligase E3	YDR457W	TOM1
SPAC328.03	tps1	alpha,alpha-trehalose-phosphate synthase [UDP-forming]	YBR126C	TPS1
SPBC1539.09c	trp1	anthranilate synthase component II	YKL211C	TRP3
SPAC22F8.05	SPAC22F8.05	alpha,alpha-trehalose-phosphate synthase	YML100W	TSL1
SPCC1672.05c	SPCC1672.05c	tyrosine-tRNA ligase	YGR185C	TYS1
SPBC119.02	Ubc4	ubiquitin conjugating enzyme Ubc4	YBR082C	UBC4/5
SPAC12B10.01c	SPAC12B10.01c	ubiquitin-protein ligase E3	YKL010C	UFD4
SPAC1002.12c	SPAC1002.12c	succinate-semialdehyde dehydrogenase	YBR006W	UGA2
SPBC16C6.02c	vps1302	chorein homolog	YLL040C	VPS13
SPBC660.06	SPBC660.06	conserved fungal protein	YFL010C	WWM1
SPAC7D4.12c	SPAC7D4.12c	DUF1212 family protein	YJL107C	YJL107C
SPBC18E5.01	SPBC18E5.01	cycloisomerase 2 family	NONE	
SPCC16A11.16c	SPCC16A11.16c	ARM1 family	NONE	
SPBC13G1.02	SPBC13G1.02	mannose-1-phosphate guanyltransferase	NONE	
SPCC11E10.09c	SPCC11E10.09c	alpha-amylase homolog	NONE	
SPAC2E1P3.04	SPAC2E1P3.04	copper amine oxidase	NONE	
SPAPB2C8.01	SPAPB2C8.01	glycoprotein	NONE	
SPBC1289.16c	SPBC1289.16c	copper amine oxidase	NONE	
known ubiquitylation	known Rsp5p substrate	HECT-domain E3		

Table 2-3: **Putative S. pombe Ubc7p substrates identified by SPASS**

One of the few known Ubc8p substrates in budding yeast, fructose-1, 6-bisphosphatase (FBPase) (Schule et al., 2000) was not identified among the 58 putative *S. pombe* Ubc8p substrates (Table 2-4). This is not surprising, since *S. cerevisiae* FBPase is only degraded when cells are shifted from a non-fermentable carbon source to a glucose-containing medium. Interestingly, Aro1p and Lys1p, two highly conserved proteins involved in amino acid metabolism, were among the putative Ubc8p substrates. Aro1p and Lys2p, the respective budding yeast homologues have recently been shown to be destabilized by the SCF-Grr1p complex a RING-domain E3 ubiquitin ligase (Benanti et al., 2007). SCF-Grr1p has been proposed to link cell cycle control and nutrient sensing in yeast. Whereas its cell cycle regulating activity utilizes Cdc34p to degrade G1 cyclins (Barral et al., 1995), this finding indicates possible Ubc8p involvement in ubiquitylation of metabolic proteins in both yeasts. To date no substrates of the *S. pombe* Grr1p ortholog Pof2p have been identified.

Further, two putative fission yeast RING-domain E3 ubiquitin ligases were identified in the Ubc8p sample: The uncharacterized protein SPCC4G3.12c which shares similarity with the mouse Rnf12/RLIM E3 ligase, and a subunit of the widely conserved CCR4-Not transcriptional repressor complex Ccr4p (Albert et al., 2002). Another protein identified among the Ubc8p substrates was Edg2p, a subunit of the NAC (nascent polypeptide-associated complex) in mammals or EGD (enhancer of Gal4p DNA binding) complex in yeast. Egd2p is a known substrate of the CCR4-Not complex (Panasenko et al., 2006). Degradation of human histone deacetylase HDAC2 was shown to be dependent on RLIM and Ubc8p (Kramer et al., 2003). The identification of the RLIM ortholog SPCC4G3.12c in the Ubc8p sample suggests a possible involvement of Ubc8p in histone modification in *S. pombe.*

28 of the 58 proteins identified (47%) were shown to be ubiquitylated in budding yeast before (Peng et al., 2003, Hitchcock et al., 2003), considerably more than the proteome average of 15%, whereas the other 30 had not yet been identified as substrates.

7 ribosomal proteins were among the putative Ubc8p substrates. Although ribosomal proteins have been traditionally regarded as non-specific contaminants in affinity purifications, evidence has accumulated in the literature over the past years that many ribosomal proteins have ribosome-independent functions unrelated to protein synthesis (Naora, 1999, Zimmermann,. 2003). The 40S ribosomal subunit S3 (rps101) identified here as a putative Ubc8p substrate, is a protein that participates in the processing of DNA damage, functioning as a general base-damage endonuclease (Kim et al., 1995) and was shown previously to be regulated via ubiquitin-dependent proteolysis (Kim et al., 2006). Similarly, the subunit S20, identified here, has been shown to be ubiquitylated on a lysine residue (Peng

et al., 2003). Further, recent findings showed strong ubiquitylation of the subunit L28 during the yeast cell cycle (Spence et al., 2000). 12 of the proteins identified were involved in mRNA metabolic processes, 10 in DNA metabolic processes, and 9 in response to stress, as determined by gene ontology mapping.

Only one of the proteins identified in the Ubc7p dataset was also identified among the 58 putative Ubc8p substrates. Interestingly, the identity of that protein was determined to be Ubc4p. This is not surprising, since most E2s behave *in vitro* as monomers, but both *in vitro* and *in vivo* analyses have suggested an ability to form homo- and hetero dimers, particularly if the E2 is charged with ubiquitin (Varelas et al., 2003, Chen et al. 1993, Pickart and Eddins, 2004). Although the exact molecular mechanism of how polyubiquitin chains are synthesized has been a mystery (reviewed by Hochstrasser, 2006), the identification of Ubc4p exclusively in the Ubc7p and Ubc8p samples which were thioesterified with ubiquitin, suggests a mechanism involving heterodimeric E2 action. A situation where Ubc4p would serve as a general E2 and its binding partners Ubc7p or Ubc8p provide specificity towards ligases or substrates would be plausible.

Interestingly, budding yeast Rsp5p and the CCR4-Not complex have been shown to have E3 activity with Ubc4p *in vitro* (Kee et al., 2007; Laribee et al., 2007). The fact that the *S. pombe* Rsp5p ortholog, as well as several of its putative substrates were identified exclusively in the Ubc7p sample, and the CCR4-Not complex subunit and a known substrate were identified in the Ubc8p sample, supports a model in which Ubc4p would cooperate with another E2 to acquire E3 and thus substrate specificity. No other protein was identified in both samples, a finding that supports this model.

The detection of previously described substrates identified in this study suggests that the SPASS approach is capable of discovering proteins ubiquitylated by specific E2s *in vivo* and that the proteins in the data sets are likely candidates of Ubc7p and Ubc8p specific ubiquitylation.

Locus	Name	Product	Ortholog	Name
SPBC405.01	ade1	phosphoribosylamine-glycine ligase	YGL234W	ADE5,7
SPCC825.01	SPCC825.01	ribosome biogenesis ATPase, Arb family	YER036C	ARB1
SPAC1834.02	aro1	pentafunctional aromatic polypeptide Aro1	YDR127W	ARO1
SPAC24H6.10c	SPAC24H6.10c	phospho-2-dehydro-3-deoxyheptonate aldolase	YBR249C	ARO4
SPAC869.11	SPAC869.11	amino acid permease, unknown 6	YBR068C	BAP2
SPAC23A1.17	SPAC23A1.17	WIP homolog	YJL020C	BBC1
SPAC9.03c	brr2	U5 snRNP complex subunit Brr2	YER172C	BRR2
SPCC970.03	SPCC970.03	cytochrome b5 reductase	YIL043C	CBR1
SPCC31H12.08c	ccr4	CCR4-Not complex subunit Ccr4	YAL021C	CCR4
SPBC25H2.12c	cct7	chaperonin-containing T-complex eta subunit Cct7	YJL111W	CCT7
SPCC757.07c	ctt1	catalase	YDR256C	CTA1
SPBC17D1.06	dbp3	ATP-dependent RNA helicase Dbp3	YGL078C	DBP3
SPBC25H2.05	egd2	nascent polypeptide-associated complex alpha subunit Egd2	YHR193C	EGD2
SPBC3D6.13c	SPBC3D6.13c	protein disulfide isomerase	YIL005W	EPS1
SPBC215.09c	erg10	acetyl-CoA C-acetyltransferase Erg10	YPL028W	ERG10
SPAC13A11.02c	erg11	sterol 14-demethylase	YHR007C	ERG11
SPCC1494.10	SPCC1494.10	transcription factor	YER109C	FLO8
SPAC56F8.03	SPAC56F8.03	translation initiation factor IF2	YAL035W	FUN12
SPCC550.06c	hsp10	mitochondrial heat shock protein Hsp10	YOR020C	HSP10
SPBC8D2.06	SPBC8D2.06	isoleucine-tRNA ligase	YBL076C	ILS1
SPAC167.01	ppk4	serine/threonine protein kinase Ppk4	YHR079C	IRE1
SPAP7G5.04c	lys1	aminoadipate-semialdehyde dehydrogenase	YBR115C	LYS2
SPAC17H9.14c	SPAC17H9.14c	protein disulfide isomerase	YOL088C	MPD2
SPBC11G11.03	SPBC11G11.03	60S acidic ribosomal protein	YKL009W	MRT4
SPAC1B3.13	SPAC1B3.13	U3 snoRNP-associated protein Nan1	YPL126W	NAN1
SPAC1805.04	nup132	nucleoporin Nup132	YKR082W	NUP133
SPAC27F1.07	SPAC27F1.07	dolichyl-diphospho-oligosaccharide-protein glycosyltransferase	YJL002C	OST1
SPAC1786.02	SPAC1786.02	phospholipase	YMR006C	PLB2
SPBC336.04	cdc6	DNA polymerase delta catalytic subunit Cdc6	YDL102W	POL3
SPAC4A8.14	prs1	ribose-phosphate pyrophosphokinase Prs1	YBL068W	PRS4
SPBC839.13c	rpl1601	60S ribosomal protein L13/L16	YIL133C	RPL16A
SPAC23A1.11	rpl1602	60S ribosomal protein L13/L16	YIL133C	RPL16A
SPBC11C11.07	rpl1801	60S ribosomal protein L18	YNL301C	RPL18B
SPAC11E3.15	rpl22	60S ribosomal protein L22	YFL034C-A	RPL22B
SPBC1685.02c	rps1202	40S ribosomal protein S12	YOR369C	RPS12
SPCC962.04	rps1201	40S ribosomal protein S12	YOR369C	RPS12
SPBC839.05c	rps1701	40S ribosomal protein S17	YDR447C	RPS17B
SPCC24B10.09	rps1702	40S ribosomal protein S17	YDR447C	RPS17B
SPAC13G6.02c	rps101	40S ribosomal protein S3	YLR441C	RPS1A
SPCC576.09	rps20	40S ribosomal protein S20	YHL015W	RPS20
SPBC1711.07	SPBC1711.07	WD repeat protein Rrb1	YMR131C	RRB1
SPCC550.14	vgl1	vigilin	YJL080C	SCP160
SPBC215.15	sec13	COPII-coated vesicle component Sec13	YLR208W	SEC13
SPAC57A7.10c	sec21	coatomer gamma subunit Sec21	YNL287W	SEC21
SPAC1F7.01c	spt6	transcription elongation factor Spt6	YGR116W	SPT6
SPAC17G6.14c	uap56	ATP-dependent RNA helicase Uap56	YDL084W	SUB2
SPBC12D12.03	cct1	chaperonin-containing T-complex alpha subunit Cct1	YDR212W	TCP1
SPCC23B6.03c	tel1	ATM checkpoint kinase	YBL088C	TEL1
SPAC4D7.05	sum1	translation initiation factor eIF3i	YMR146C	TIF34
SPCC1919.09	tif6	translation initiation factor eIF6	YPR016C	TIF6
SPAC1F5.11c	SPAC1F5.11c	phosphatidylinositol kinase	YHR099W	TRA1
SPAC3F10.02c	trk1	potassium ion transporter Trk1	YJL129C	TRK1
SPACUNK4.16c	SPACUNK4.16c	alpha,alpha-trehalose-phosphate synthase	YML100W	TSL1
SPBC119.02	ubc4	ubiquitin conjugating enzyme Ubc4	YBR082C	UBC4, 5
SPCC1494.05c	ubp12	ubiquitin C-terminal hydrolase Ubp12	YJL197W	UBP12
SPAC15A10.04c	zpr1	zinc finger protein Zpr1	YGR211W	ZPR1
SPAC821.05	SPAC821.05	translation initiation factor eIF3h	NONE	
SPCC4G3.12c	SPCC4G3.12c	RING-domain ubiquitin-protein ligase E3	NONE	
known SCF substrate	known ubiquitylation	known CCR4-Not substrate	RING-domain E3	

Table 2-4: **Putative S. pombe Ubc8p substrates identified by SPASS**

2.4. Conclusions

In this study, we demonstrated that SPASS can be used to identify, on a global scale, ubiquitylation substrates of the *S. pombe* E2s Ubc7p and Ubc8p. A number of biochemical assays was used to validate the details of the assay before application in a large scale study. 51 putative Ubc7p substrates and 58 putative Ubc8p substrates were identified. Since only a handful of ubiquitylation substrates are known in fission yeast, we evaluated the results based on whether the proteins were known substrates in other organisms. To our surprise, we identified several HECT-domain and RING-domain E3s among the substrates of both E2s, suggesting extensive E3 autoubiquitylation. Interestingly, Ubc7p ubiquitylated exclusively HECT-domain ligases, whereas Ubc8p ubiquitylated two RING-domain ligases. Consistently, we also identified known substrates of those E3s in each of the respective Ubc samples. We further identified known substrates of other E3 ligases as well as previously unknown substrates. Overall, about 50% of putative substrates identified here have been shown to be ubiquitylated before, a significant enrichment over the 15% proteome average, providing a useful resource for further investigation.

It has not been clear what role E2s play in the process of providing specificity in terms of ligase or substrate selection. Our data indicates that Ubc4p interacts with and is possibly auto-ubiquitylated by both, Ubc7p and Ubc8p. Conversely, none of the putative substrates was identified in both E2 preparations, suggesting that specificity towards E3 and substrates is mediated by Ubc7p or Ubc8p. One explanation for this finding could be that Ubc4p requires a second E2 to form a heterodimer *in vivo* in order to be catalytically active in conjunction with specific E3s. A ubiquitin ligase, the SCF-Grr1p complex may be able to be catalytically active with more than one E2. It functions with Cdc34p in ubiquitylation of cell cycle regulators and – supported by the identification of two putative SCF-Grr1p substrates in this study – with Ubc8p. This finding would indicate that besides the known requirement of substrate modification, differential E2 selection may provide an additional layer of substrate specificity.

To gain more insight into how the ubiquitin proteasome system is regulated *in vivo*, the SPASS system may prove useful in future studies of cellular ubiquitylation. Further studies could identify substrates of all E2s within an organism. This more comprehensive view would provide detailed insight into substrate and E3 selection. SPASS could be utilized

studies investigating the E2 catalytic mechanism and could also be applied to understanding the formation of E2 heterodimers.

Our approach should support the screening of other ubiquitin systems both in humans and in lower organisms. Furthermore, once proteins are covalently linked to the ubiquitin-beads, deubiquitylating enzymes could be screened comprehensively for substrate specificity.

2.5. References

Albert TK, Hanzawa H, Legtenberg YI, de Ruwe MJ, van den Heuvel FA, Collart MA, Boelens R, Timmers HT. (2002) Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. EMBO J. 2002 Feb 1;21(3):355-64.

Ayad NG, Rankin S, Ooi D, Rape M, Kirschner MW. (2005) Identification of ubiquitin ligase substrates by *in vitro* expression cloning. Methods Enzymol. 2005;399:404-14.

Banerjee A, Gregori L, Xu Y, Chau V. (1993) The bacterially expressed yeast CDC34 gene product can undergo autoubiquitination to form a multiubiquitin chain-linked protein. J Biol Chem. 1993 Mar 15;268(8):5668-75.

Barral, Y., Jentsch, S. & Mann, C. (1995) G1 cyclin turnover and nutrient uptake are controlled by a common pathway in yeast. Genes Dev. 9, 399–409 (1995).

Belle A, Tanay A, Bitincka L, Shamir R, O'Shea EK. (2006) Quantification of protein half-lives in the budding yeast proteome. Proc Natl Acad Sci U S A. 2006 Aug 29;103(35):13004-9.

Benanti JA, Cheung SK, Brady MC, Toczyski DP. (2007) A proteomic screen reveals SCFGrr1 targets that regulate the glycolytic-gluconeogenic switch. Nat Cell Biol. 2007 Oct;9(10):1184-91.

Chau, V., Tobias, J.W., Bachmair, A., Marriott, D., Ecker, D.J., Gonda, D.K., and Varshavsky, A. (1989). A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science 243, 1576–1583.

Chen ZJ, Niles EG, Pickart CM. (1991) Isolation of a cDNA encoding a mammalian multiubiquitinating enzyme (E225K) and overexpression of the functional enzyme in Escherichia coli. J Biol Chem. 1991 Aug 25;266(24):15698-704.

Chen, P., Johnson, P., Sommer, T., Jentsch, S., and Hochstrasser, M. (1993). Multiple ubiquitin-conjugating enzymes participate in the *in vivo* degradation of the yeast MAT alpha 2 repressor. Cell 74, 357–369.

Gupta R, Kus B, Fladd C, Wasmuth J, Tonikian R, Sidhu S, Krogan NJ, Parkinson J, Rotin D (2007) Ubiquitination screen using protein microarrays for comprehensive identification of Rsp5 substrates in yeast. Mol Syst Biol. 2007;3:116.

Hershko A. (1883) Ubiquitin: roles in protein modification and breakdown. Cell. 1983 Aug;34(1):11-2.

Hicke L, Dunn R. (2003) Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. Annu Rev Cell Dev Biol. 2003;19:141-72.

Hitchcock AL, Auld K, Gygi SP, Silver PA. (2003) A subset of membrane-associated proteins is ubiquitinated in response to mutations in the endoplasmic reticulum degradation machinery. Proc Natl Acad Sci U S A. 2003 Oct 28;100(22):12735-40.

Hochstrasser M (2006) Lingering mysteries of ubiquitin-chain assembly. Cell. 2006 Jan 13;124(1):27-34.

Horak J, Wolf DH. (2005) The ubiquitin ligase SCF(Grr1) is required for Gal2p degradation in the yeast Saccharomyces cerevisiae. Biochem Biophys Res Commun. 2005 Oct 7;335(4):1185-90.

Jiang YH, Beaudet al. (2004) Human disorders of ubiquitination and proteasomal degradation. Curr Opin Pediatr 16: 419–426.

Kee Y, Huibregtse JM. (2007) Regulation of catalytic activities of HECT ubiquitin ligases. Biochem Biophys Res Commun. 2007 Mar 9;354(2):329-33.

Kim J, Chubatsu LS, Admon A, Stahl J, Fellous R, Linn S. (1995) Implication of mammalian ribosomal protein S3 in the processing of DNA damage. J Biol Chem. 1995 Jun 9;270(23):13620-9.

Kim TS, Jang CY, Kim HD, Lee JY, Ahn BY, Kim J. (2006) Interaction of Hsp90 with ribosomal proteins protects from ubiquitination and proteasome-dependent degradation. Mol Biol Cell. 2006 Feb;17(2):824-33.

Kramer OH, Zhu P, Ostendorff HP, Golebiewski M, Tiefenbach J, Peters MA, Brill B, Groner B, Bach I, Heinzel T, Gottlicher M. (2003) The histone deacetylase inhibitor valproic acid selectively induces proteasomal degradation of HDAC2. EMBO J. 2003 Jul 1;22(13):3411-20.

Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, Li J, Pu S, Datta N, Tikuisis AP, Punna T, Peregrin-Alvarez JM, Shales M, Zhang X, Davey M, Robinson MD, Paccanaro A, Bray JE, Sheung A, Beattie B, Richards DP, Canadien V, Lalev A, Mena F, Wong P, Starostine A, Canete MM, Vlasblom J, Wu S, Orsi C, Collins SR, Chandran S, Haw R, Rilstone JJ, Gandi K, Thompson NJ, Musso G, St Onge P, Ghanny S, Lam MH, Butland G, Altaf-Ul AM, Kanaya S, Shilatifard A, O'Shea E, Weissman JS, Ingles CJ, Hughes TR, Parkinson J, Gerstein M, Wodak SJ, Emili A, Greenblatt JF. (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. Nature. 2006 Mar 30;440(7084):637-43. Epub 2006 Mar 22.

Kus B, Gajadhar A, Stanger K, Cho R, Sun W, Rouleau N, Lee T, Chan D, Wolting C, Edwards A, Bosse R, Rotin D. (2005) A high throughput screen to identify substrates for the ubiquitin ligase Rsp5. J Biol Chem. 2005 Aug 19;280(33):29470-8.

Laribee RN, Shibata Y, Mersman DP, Collins SR, Kemmeren P, Roguev A, Weissman JS, Briggs SD, Krogan NJ, Strahl BD. (2007) CCR4/NOT complex associates with the proteasome and regulates histone methylation. Proc Natl Acad Sci U S A. 2007 Apr 3;104(14):5836-41.

Liu H, Sadygov RG, Yates JR 3rd. (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem. 2004 Jul 15;76(14):4193-201.

Mayor T, Graumann J, Bryan J, Maccoss MJ, Deshaies RJ. (2007) Quantitative profiling of ubiquitylated proteins reveals proteasome substrates and the substrate repertoire influenced by the Rpn10 receptor pathway. Mol Cell Proteomics. 2007 Jul 20

Moreno S, Klar A, Nurse P (1991) Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. Methods Enzymol. 1991;194:795-823.

Mouchantaf R, Azakir BA, McPherson PS, Millard SM, Wood SA, Angers A. (2006) The ubiquitin ligase itch is auto-ubiquitylated *in vivo* and *in vitro* but is protected from degradation by interacting with the deubiquitylating enzyme FAM/USP9X.

Mukherjee S, Tessema M, Wandinger-Ness A. (2006) Vesicular trafficking of tyrosine kinase receptors and associated proteins in the regulation of signaling and vascular function. Circ Res. 2006 Mar 31;98(6):743-56.

Naora H. (1999) Involvement of ribosomal proteins in regulating cell growth and apoptosis: translational modulation or recruitment for extraribosomal activity? Immunol Cell Biol. 1999 Jun;77(3):197-205.

Nuber U, Schwarz SE, Scheffner M. (1998) The ubiquitin-protein ligase E6-associated protein (E6-AP) serves as its own substrate. Eur J Biochem. 1998 Jun 15;254(3):643-9.

Panasenko O, Landrieux E, Feuermann M, Finka A, Paquet N, Collart MA. (2006) The yeast Ccr4-Not complex controls ubiquitination of the nascent-associated polypeptide (NAC-EGD) complex. J Biol Chem. 2006 Oct 20;281(42):31389-98.

Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP. (2003) A proteomics approach to understanding protein ubiquitination. Nat Biotechnol. 2003 Aug;21(8):921-6. Epub 2003 Jul 20.

Petroski MD, Deshaies RJ. (2005) Function and regulation of cullin-RING ubiquitin ligases. Nat Rev Mol Cell Biol. 2005 Jan;6(1):9-20.

Pickart CM, Eddins MJ. (2004) Ubiquitin: structures, functions, mechanisms. Biochim Biophys Acta. 2004 Nov 29;1695(1-3):55-72.

Raasi S, Pickart CM. (2005) Ubiquitin chain synthesis. Methods Mol Biol. 2005;301:47-55.

Saleki R, Jia Z, Karagiannis J, Young PG. (1997) Tolerance of low pH in Schizosaccharomyces pombe requires a functioning pub1 ubiquitin ligase. Mol Gen Genet. 1997 May 20;254(5):520-8.

Schule T, Rose M, Entian KD, Thumm M, Wolf DH. (2000) Ubc8p functions in catabolite degradation of fructose-1, 6-bisphosphatase in yeast. EMBO J. 2000 May 15;19(10):2161-7.

Skowyra D, Craig KL, Tyers M, Elledge SJ, Harper JW. (1997) F-box proteins are receptors that recruit phosphorylated substrates to the SCF ubiquitin-ligase complex. Cell. 1997 Oct 17;91(2):209-19.

Skowyra D, Koepp DM, Kamura T, Conrad MN, Conaway RC, Conaway JW, Elledge SJ, Harper JW. (1999) Reconstitution of G1 cyclin ubiquitination with complexes containing SCFGrr1 and Rbx1. Science. 1999 Apr 23;284(5414):662-5.

Spence J, Gali RR, Dittmar G, Sherman F, Karin M, Finley D. (2000) Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. Cell. 2000 Jul 7;102(1):67-76.

Tabb DL, McDonald WH, Yates JR 3rd. (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J Proteome Res. 2002 Jan-Feb;1(1):21-6.

Thrower, J.S., Hoffman, L., Rechsteiner, M., and Pickart, C.M. (2000) Recognition of the polyubiquitin proteolytic signal. EMBO J. 19, 94–102

Urbé S. (2005) Ubiquitin and endocytic protein sorting. Essays Biochem. 2005;41:81-98.

Varelas, X., Ptak, C., and Ellison, M.J. (2003). Cdc34 self-association is facilitated by ubiquitin thiolester formation and is required for its catalytic activity. Mol. Cell. Biol. 23, 5388–5400.

Verdecia MA, Joazeiro CA, Wells NJ, Ferrer JL, Bowman ME, Hunter T, Noel JP. (2003) Conformational flexibility underlies ubiquitin ligation mediated by the WWP1 HECT domain E3 ligase.

Wang G, Yang J, Huibregtse JM. (1999) Functional domains of the Rsp5 ubiquitin-protein ligase. Mol Cell Biol. 1999 Jan;19(1):342-52.

Washburn MP, Wolters D, Yates JR 3rd. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol. 2001 Mar;19(3):242-7.

Wilkinson KD (1997) Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. FASEB J. 1997 Dec;11(14):1245-56.

Wolf DA, McKeon F, Jackson PK (1999) F-box/WD-repeat proteins Pop1p and Sud1p/Pop2p form complexes that bind and direct the proteolysis of Cdc18p. Curr. Biol. 1999, 9:373-376

Zhou C, Seibert V, Geyer R, Rhee E, Lyapina S, Cope G, Deshaies R.J, Wolf DA (2001) The fission yeast COP9/signalosome is involved in cullin modification by ubiquitin-related Ned8p. BMC Biochemistry 2:7

Zimmermann RA. (2004) The double life of ribosomal proteins. Cell. 2003 Oct 17;115(2):130-2.

CHAPTER III: Comparative Proteomic and Transcriptomic Profiling of the Fission Yeast *Schizosaccharomyces pombe*

3.1. Background

Schizosaccharomyces pombe is a unicellular archiascomycete fungus displaying many properties of more complex eukaryotes. It has been estimated that fission yeast diverged from budding yeast 1100 million years ago (Heckman et al., 2001), thus accounting for their considerable divergence in genome organization (Wood et al., 2002). Despite differences in the number of genes, the number of introns, and centromere size, basic cellular processes are highly conserved between the two yeasts with 3600 proteins being predicted or confirmed orthologs (Wood, 2006). However, it is still unclear to what extent mechanisms of gene expression in fission yeast overlap with those in budding yeast.

S. pombe is a well-established model organism for the study of cell-cycle regulation, cytokinesis, DNA repair and recombination, and checkpoint pathways, but only 1500 of its predicted 4900 genes and proteins have been experimentally characterized. Although mRNA profiling has begun to address functional aspects of the fission yeast genome (Mata et al., 2002; Chen et al., 2003; Rustici et al., 2004; Oliva et al., 2005; Peng et al., 2005; Marguerat et al., 2006), the notion was expressed that mRNA levels are only a partial reflection of the functional state of an organism (Greenbaum et al., 2003). It is widely accepted that a comprehensive understanding of the genomic information will require, besides other strategies, means of analyzing quantitative differences in protein expression on a proteome-wide scale (Anderson et al., 2000; Bakhtiar and Tse, 2000; Yates, 2000).

Several quantitative methods, including ICAT (Gygi et al., 1999), iTRAQ (Ross et al., 2004), stable isotope labelling (Ong et al., 2002; Washburn et al., 2002), AQUA (Gerber et al., 2003), spectral sampling (Liu et al., 2004; Kislinger et al., 2006), protein abundance indexing (Ishihama et al., 2005), and whole-genome ORF epitope tagging (Ghaemmaghami et al., 2003; Matsuyama et al., 2006), have been employed for proteomic analyses of model organisms, in particular budding yeast. All of these techniques have their intrinsic strengths and limitations, including the bias of mass spectrometry-based methods toward proteins of medium to high abundance, and the potential for interference of epitope tags with endogenous protein function, expression, and localization. In addition, epitope tagging can

only interrogate putative known ORFs and is only applicable to organisms that are readily amenable to genetic manipulation in a high-throughput format. Mass spectrometry, on the contrary, can potentially identify new proteins and is broadly applicable to any proteome for which a corresponding genome sequence is available.

Weighing the advantages and disadvantages of currently available methods, we have embarked on a mass spectrometry-based approach for relative quantitation of unmodified fission yeast proteins. In addition, we have compared mRNA and protein expression profiles in fission yeast and budding yeast to assess the overall protein–mRNA correlation in these related organisms.

3.2. Materials and Methods

Chemicals and materials

Porcine sequencing grade modified trypsin was obtain from Promega (Madison, WI). HPLC-grade purity deionized water was purchased from Fisher Scientific (Pittsburgh, PA). All other chemicals were from Sigma-Aldrich (St. Louis, MO), unless otherwise noted.

Determination of protein concentration

Protein concentration of the lysate and all resulting fractions was measured using a Lowry-type RC DC assay according to manufacturer's protocol (Biorad, Hercules, CA).

Preparation of fission yeast cell lysate

S. pombe cells (DS 448/2=927 h- leu-1-32 ura4d-18) were grown in 50 ml YES to mid-log phase (OD595=0.68). Cells were washed in STOP buffer (150 mM NaCl, 10 mM EDTA, 50 mM NaF, 1 mM NaN_3) and lysed in 450 µl buffer containing 7.7 M urea, 2.2 M thiourea, 0.55% CHAPS, 10 mM Tris (pH 8.5), 200 mM DTT and protease inhibitors by bead lysis in a Fastprep device (Bio 101). The cell homogenate was cleared by centrifugation and the bead lysis was repeated once with the pellet of insoluble debris. The two homogenates were pooled (950 µl) and incubated at room temperature (RT) for 30 min. A volume of 5.2 µl of 99% N,N-dimethylacrylamide (Sigma) was added, followed by another incubation at RT for 30 min after which 10 µl 2 M DTT was added for 5 min at RT. The homogenate was cleared by centrifugation for 15 min at 14 000 g, resulting in a denatured, reduced, and alkylated sample with a concentration of 10 mg/ml.

Sample prefractionation by isoelectric focusing (IEF) with a ZOOM device (Invitrogen)

1.5 mg of concentrated lysate was diluted to 0.4 mg/ml in 3.6 ml LB. 40 µl carrier ampholytes (Invitrogen) and 40 µl fresh 2 M DTT were added. The diluted sample was loaded into the 5 chambers (pH 3-4.6, 4.6-5.4, 5.4-6.2, 6.2-7, 7-10) of a liquid phase isoelectric focusing device (ZOOM IEF Fractionator, Invitrogen), and isoelectric focusing was carried out at RT in three steps (100 V, 200 V, 600 V) at 2 mA current limit and 2 W power limit. The current at each step was allowed to drop to 0.2 mA before switching to the next higher voltage. The isoelectric focusing was finished when a final current of 0.2 mA at 600 V was reached. The five fractions were collected from the chambers and subjected to

chloroform-methanol precipitation as described in Ref. (Wessel and Flugge, 1984). The protein pellets were resuspended in 1 M urea, 25 mM ammonium bicarbonate, pH 8.6, and digested in two repetitive incubation steps with 10 µg sequencing grade trypsin (Promega) each for 24 h at 37 ^{o}C. The resulting tryptic digests were subjected to solid phase extraction using C18 cartridges (Varian, Palo Alto, CA), according to the manufacturer's instructions. Samples were concentrated in a SpeedVac (Thermo, CA) and resuspended in 100 µl of aqueous solvent containing 5% formic acid, 2% acetonitrile. Samples were analyzed by 1D nano-LC tandem mass spectrometry as described below.

Sample prefractionation with a Multicompartment Electrolizer (MCE)

150 µl (1.5 mg) of the lysate was used for two MCE experiments. Each sample was diluted to 15.5 ml with 7 M urea, 2 M thiourea, 0.5% CHAPS. 5 ml of sample was applied to the pH 5-6.5 chambers of two MCEs configured with isoelectric membranes of pH 3, 5, 6.5, 8, and 11. Each electrode chamber was filled with 5 ml of 7 M urea, 2 M thiourea. Separation chambers were filled with 5 ml of 7 M urea, 2 M thiourea, 0.5% CHAPS. Isoelectric focusing was performed at 18 ^{o}C for 9.27 ± 0.05 kVh according to manufacturer's recommendations (Proteome Systems, Woburn, MA). Fractions were collected, concentrated by ultrafiltration in VIVASPIN 20 ml centrifugal filtration devices (Vivascience, Edgewood, NY) with a 10 kDa molecular weight cut-off (MWCO), and digested with 10 µg of sequencing grade trypsin (Promega) each, for 24 h at 37 ^{o}C. Tryptic peptides resulting from the two MCE fractionation experiments were subjected to 1D nano-LC MS/MS as described below.

Sample prefractionation by IEF on immobilized pH gradient (IPG) strips

50 µl (500 µg) cell lysate was suspended in 350 µl CHAPS buffer (7M urea, 2M thiourea, 2% CHAPS, 2% ampholytes 3-10, 65 mM DTT, 0.1% bromophenol blue). 400 µl were loaded onto an 18 cm immobilized pH 3-10 nonlinear gradient strip (Amersham, Piscataway, NJ) and passively rehydrated for 16 hours. The strip was focused to 100,000 Vh in a Genomic Solutions Investigator IEF device, equilibrated in 10 ml equilibration buffer I (6 M urea, 375 mM Tris/HCl pH 7.4, 2% SDS, 2% glycerol, 2% DTT), followed by 10 ml equilibration buffer II (6 M urea, 375 mM Tris/HCl pH 7.4, 2% SDS, 2% glycerol, 2% iodoacetamide). The strip was washed ten times with alternating 5 min long cycles of 40% methanol, 10% acetic acid and deionized water. The IPG strip was cut into 48 pieces of equal length, placed into 1.5 ml micro centrifuge tubes, and subjected to in-gel digestion with 1 µg

of sequencing grade trypsin (Promega) each for 24 h at 37 ^{o}C in 0.5 ml of 25 mM ammonium bicarbonate, pH 8.6. Protein digests were collected; gel slices were washed once with 0.2 ml of 25 mM ammonium bicarbonate, pH 8.6 and twice with 0.2 ml of 5% formic acid, 50% acetonitrile. Samples were analyzed by 1D nano-LC tandem mass spectrometry as described below.

Sample prefractionation by 1D-PAGE

100 µl (1 mg) of cell lysate was boiled for 5 min in an equal volume of 2x Laemmli buffer and separated on a 12% SDS-poly-acrylamide gel (15 cm x 15 cm x 1.5 mm). The gel was stained with Coomassie Brilliant Blue and destained in 25% methanol, 7.5% acetic acid. The gel was sliced horizontally into 12 strips with roughly similar protein content as estimated from the staining intensity. The gel strips were washed three times in 10 ml of 25 mM ammonium bicarbonate 50% acetonitrile, followed by dehydration in 10 ml of HPLC-grade acetonitrile. The gel strips were fully dried in a vacuum concentrator (SpeedVac) and rehydrated in 5 ml 50 mM ammonium bicarbonate containing 1 µg/ml trypsin, followed by incubation for 24 h at 37 ^{o}C. Protein digests were collected and gel strips were washed once with 10 ml of 25 mM ammonium bicarbonate, pH 8.6 and twice with 10 ml of 5% formic acid, 50% acetonitrile. The combined peptide solutions were concentrated to a volume of ~0.5 ml in a SpeedVac. The resulting digests were subjected to solid phase extraction using C18 cartridges (Varian) according to the manufacturer's protocol, concentrated in a SpeedVac, and resuspended in 100 µl aqueous solvent containing 5% formic acid, 2% acetonitrile. Samples were analyzed by 1D nano-LC tandem mass spectrometry as described below.

Prefractionation on strong anion exchange (SAX) membrane adsorber spin columns

Two Vivapure Q Mini spin columns were equilibrated with 100 mM sodium bicarbonate buffer, pH 8.8 by loading 400 µl onto the column and spinning at 2000g for 5 minutes. 100 µl lysate (1 mg protein) was bound to the equilibrated columns and spun at 2000g for 5 minutes. After protein binding, the column was washed twice with 400 µl of 100 mM sodium bicarbonate, pH 8.8. Fraction 1 was eluted by applying 400 µl of 100 mM sodium bicarbonate, pH 8.8 containing 100 mM NaCl to the column and spinning at 2000g for 5 minutes. Subsequent elutions were performed with 400 µl 200 mM sodium bicarbonate buffer, pH 8.8, 300 mM (fraction 3), 400 mM (fraction 4), 500 mM (fraction 5), and 1 M NaCl

(fraction 6), respectively. All centrifugation steps were performed at 4 ºC. Fractions were desalted using 3 ml C18 reversed-phase solid phase extraction columns (Varian) and concentrated to a volume of 100 µl each. Organic solvent (methanol) was removed using a centrifugal evaporator (SpeedVac), and the samples were resuspended in 500 µl of 20 mM ammonium bicarbonate buffer, pH 8.6. Samples were analyzed by 1D nano-LC tandem mass spectrometry as described below.

On-Line 2-D LC ESI-MS/MS analysis

For the analysis of total unfractionated cell lysate by LC MS, the lysate was digested in solution. 150 µl (1.5 mg) lysate was subjected to chloroform-methanol precipitation and the precipitate was resuspended in 1 M Urea 25 mM ammonium bicarbonate, pH 8.6. Trypsin was added at a ratio of 1:50 relative to the protein amount and the sample was incubated for 12 hrs at 37 ^{o}C under agitation. Upon digestion, the sample was subjected to solid-phase extraction on a 3 ml C18 reversed-phase column (Varian, Palo Alto, CA), and concentrated to a volume of 100 µl. Organic solvent (methanol) was evaporated in a SpedVac, and the sample was resuspended in 100 µl of aqueous solvent containing 5% formic acid 2% acetonitrile.

100 µl of protein digest was injected onto a 300 µm i.d. 5 cm strong cation exchange (SCX) column (Dionex, Sunnyavle, CA), followed by elution with twelve salt plug injections (5 mM – 500 mM ammonium formate, pH 3.0). Columns used were a C18 solid phase extraction (SPE "trapping") column (300 µm i.d. 5 mm, Dionex) and a self-packed 75 µm i.d. 15 cm nano-LC reversed-phase fused silica PicoFrit column (stationary phase: Magic C18AQ, 3 µm, 100 Å (Michrom Bioresources, Auburn, CA); column: PicoFrit, 15 µm i.d. pulled tip (New Objective, Woburn, MA)). The columns were connected through a 10-port Valco valve (custom LCQ plumbing configuration). Thirteen fractions resulting from twelve salt plug injections onto the SCX column and the flowthrough fraction, were desalted on the "trapping" column, and subsequently eluted onto the nano-LC column, where fractions were separated with a linear gradient of acetonitrile in 0.1% formic acid. The eluent was introduced into the LCQ Deca XP Plus mass spectrometer by nanoelectrospray. Two hour-long gradients (2% B to 40% B in 110 min and 40% B to 95% B in 10 minutes; where solvent A was 2% acetonitrile, 0.1% formic acid, and solvent B was 5% isopropanol, 0.1% formic acid, 85% acetonitrile) were used for the reversed phase separation of the first four fractions. The remaining fractions were separated with one hour-long gradients. One full MS

scan between 400 and 1800 m/z and five full MS/MS scans of the most intense ions were acquired in data-dependent MS/MS scanning experiments. The scan time was set to 50 µs and up to three scans were accumulated. The temperature of the heated capillary was set to 220 ^{o}C. A spray voltage of 1.95 kV was used in all experiments.

Protein identification by 1D nano-LC tandem mass spectrometry

Columns used: C18 solid phase extraction "trapping" column (300 µm i.d. 5 mm, Dionex) and a self-packed nano-LC reversed-phase fused silica PicoFrit column (75 µm i.d. 15 cm, 15 µm i.d. pulled tip (NewObjective); stationary phase: Magic C18AQ, 3 µm, 100 Å (Michrom Bioresources)). Each protein digest (20-100 µl) was loaded onto the SPE column connected on-line with the nano-LC column through the 10-port Valco valve (custom LCQ plumbing configuration). The sample was desalted on the "trapping" column, eluted onto and separated on the nano-LC column with a one-hour linear gradient of acetonitrile in 0.1% formic acid. The eluent was introduced into the LCQ Deca XP Plus mass spectrometer by nanoelectrospray.

MS data processing

The MS data ".raw" files acquired by the mass spectrometer were converted to ".dta" files using Bioworks 3.2 (ThermoElectron, CA). The ".dta" files were submitted for database search on a Sorcerer search engine (Sage-N Research, Thermo Electron, CA) using the SEQUEST algorithm (Eng et al., 1994). The search was performed against a combined forward and reverse *S. pombe* FASTA protein database containing 10,149 protein entries constructed using a Pearl script kindly provided by Dr. T. Rejtar. Methionine oxidation (+15.9994 atomic mass units (amu)) and cysteine alkylation (+58.0446 amu with acrylamide) were set as differential modifications. No static modifications or differential posttranslational modifications were employed in the first round of searching. A peptide mass tolerance equal to 1.5 amu and a fragment ion mass tolerance equal to 0.8 amu were used in all searches. Monoisotopic mass type, fully tryptic peptide termini, and up to 2 missed cleavages were used in all searches. The SEQUEST output (".out" files) was analyzed using DTASelect software (kindly provided by Dr. J. Yates IIIrd, The Scripps Research Institute, CA). A second SEQUEST search was performed on LC MS data derived from 2D LC MS experiments and 1D PAGE prefractionation followed by 1D LC MS. A composite forward-reverse protein database translated from the gene database (Sanger Center, UK) containing all

putative open reading frames (ORFs) encoded by the *S. pombe* genome was used for this search. All proteins over 25 amino acids translated in three reading frames without filtering for internal and overlapping ORFs were considered as putative proteins (the database was kindly provided by Dr. V. Wood, Sanger Center, UK). Methionine oxidation (+15.9994 amu); cysteine alkylation (+58.0446 amu with acrylamide); phosphorylation of serine, threonine and tyrosine (+79.9663); and lysine ubiquitylation (+114.14) were set as differential modifications. The SEQUEST output was analyzed using DTASelect software. Peptide matches with posttranslational modifications were examined and filtered manually using rigorous evaluation of SEQUEST scores (XCorr $\geq$ 2.45, 2.65, 3.75 for charge states +1, +2, +3, respectively; ΔCn $\geq$ 0.1; Sp $\leq$ 5; RSp $\geq$ 250; and ion proportion ratio $\geq$0.4) in combination with detailed visual inspection of tandem mass spectra. SEQUEST matches corresponding to peptides ubiquitylated on C-terminal lysine residues were removed despite acceptable scores. Up to four posttranslational modifications per peptide were allowed. All peptide matches are contained in Supplementary Data File 3-1.

The balance between reliability and sensitivity of the protein identification data was set by adjusting the estimated false positive peptide identification rate (FPPeR) to 1% using DTASelect and sequence-based filtering as described elsewhere (Resing et al., 2004). The FPPeR was calculated as the number of peptide matches from a "reverse" database divided by the total number of "forward" protein matches, in percents (similar to Refs. (Peng et al., 2003; Qian et al., 2005)). For every dataset, filtering parameters of DTASelect software were set such that the specified FPPeR was obtained in every search. The application of adjusted primary scores (XCorr), relative scores (ΔCn and Sp, RSp), and the minimum ion proportion score was essential to accomplish an acceptable error rate without compromising sensitivity significantly. Duplicate peptide matches were purged on the basis of Xcorr with the use of DTASelect software to eliminate redundancy caused by homologous proteins and protein isoforms. When multiple similar proteins were identified, only the entry with the highest score was included on the ID list.

Merging MS datasets for spectral sampling

SEQUEST results for multiple 1D LC MS/MS runs resulting from the same upfront fractionation method were merged and validated using DTASelect software as above with the modifications described here. SEQUEST results for multiple 1D LC MS/MS runs or multiple online 2D LC MS/MS experiments were merged and compared using DTAContrast software

and DTASelect output files. DTAContrast allowed assembling and comparing SEQUEST results derived from multiple runs and prefractionation techniques with identification of overlapping and non-overlapping protein entries. While duplicate peptide matches were purged on the basis of Xcorr with the use of DTASelect to eliminate redundancy for a single fractionation method, duplicate peptide matches were kept when merging and validating datasets derived from different prefractionation techniques. Additional validation of merged datasets was performed using DTASelect software with a target FPPeR value ≤1%. Similar to the technique used by Kislinger et al. (Kislinger et al., 2006), cumulative spectral counts obtained by the various prefractionation techniques were scored for each protein in the merged dataset.

cDNA mircoarray analysis

Two 50 ml fission yeast cultures were harvested at mid-log phase (OD595 = 0.6). Cells were washed in a buffer containing 150 mM NaCl, 10 mM EDTA, 50 mM NaF and 1 mM NaN_3, and subject to RNA isolation as follows: 5 ml RNAzol (Invitrogen) preheated to 65 oC and 1 ml silica beads (BioSpec) were added to the cell pellets and the tubes were subject to three cycles of 2 min vortexing followed by 5 min heating to 65 oC. 500 µl chlorophorm was added, followed by 5 min incubation on ice and centrifugation for 30 min at 5000g. 2.5 ml of the aqueous layer were removed, followed by addition of 2 ml isopropanol and incubation on ice for 20 min. The samples were spun at 5000g for 30 min and the liquid was removed by vacuum aspiration. The resulting pellets were washed with cold 80% ethanol, dried and resuspended in RNAse-free water, resulting in samples of 400 µg RNA at 5 µg/µl. 100 µg total RNA of each sample was further purified using an RNAeasy kit (Qiagen) following the manufacturer supplied protocol, resulting in 60 µg purified RNA at a concentration of 0.9 µg/µl. Aminoallyl nucleotide incorporation (aa-dUTP) via first strand cDNA synthesis followed by coupling of the aminoallyl groups to either Cyanine 3 or 5 (Cy3/Cy5) fluorescent molecules was carried out following the standard operating procedure (SOP # M004) of the Institute for Genomic Research (http://www.tigr.org/).

Labeled cDNAs were hybridized onto glass slide microarrays at the Stony Brook University Microarray Facility. Microarrays contained 4988 predicted ORFs and transcripts as annotated by the Sanger Center and described in detail in reference (Oliva et al., 2005). Median pixel intensities for each spot were corrected by local background subtraction and intensities obtained from the two independently prepared samples were averaged. The

microarray dataset was named "Pombe-mRNA" and was deposited with the public ArrayExpress (www.ebi.ac.uk/arrayexpress/) database.

Statistical analysis

Spectral count modeling - In order to apply likelihood-based goodness-of-fit criteria for model comparison in a setting with spectral counts as outcomes, negative binomial log-linear regression was used. The negative binomial distribution is commonly used to model counts that are overdispersed in comparison with Poisson outcomes (e.g. Ref. (Land et al., 1996; Thurston et al., 2000)). It is defined as follows: Y has a negative binomial distribution with mean parameter μ and dispersion parameter σ if

$$P(Y = y) = \frac{\Gamma(y + \sigma^{-1})}{\Gamma(\sigma^{-1})\Gamma(y+1)} \left(\frac{1}{1+\mu\sigma} \right)^{1/\sigma} \left(\frac{\mu}{\mu + \sigma^{-1}} \right).$$

The dispersion parameter σ represents overdispersion with respect to a Poisson variable, and if $\sigma = 0$, then Y has a Poisson distribution. As with Poisson regression, the assumed relationship between Y and a vector of covariates ($x_1, \ldots, x_J$) is specified log-linearly as

$$\mu(x) = \exp\left(\beta_0 + \sum_{j=1}^{J} \beta_j x_j \right).$$

Note that the vector may include the same covariate raised to different powers. Figure 3-1 suggests a quadratic relationship between the logarithm of the number of tryptic peptides, assuming no miscleavages, and ln(μ). Note that the quadratic model fit the data better than a linear model; this implies that a nonmonotonic relationship fits better than a monotonic relationship. Similar quadratic relationships were evident between other measures of protein size (number of amino acids) and ln(μ). Consequently, five quadratic models were fitted as follows:

$$\mu(x) = \exp\left\{ \beta_0 + \beta_L (x - m) + \beta_Q (x - m)^2 \right\}$$

where x is given in Table 3-1 and m was the observed mean of x.

Model	Model description	Model specification	LR Stat	P Value	AIC
0	No adjustment	$\beta_L = \beta_Q = x = 0$			13931.88
1	Adjustment by number amino acids	$x = \log_2$ (length of protein in amino acids)	20.83	<0.0001	13915.04
2	Adjustment by number of tryptic peptides, assuming 0 miscleavages	$x = \log_2$ (number of tryptic peptides corresponding to protein)	32.4	<0.0001	13903.48
3	Adjustment by number of tryptic peptides, assuming no more than 1 miscleavage	$x = \log_2$ (number of tryptic peptides corresponding to protein)	34.21	<0.0001	13901.67
4	Adjustment by number of tryptic peptides, assuming no more than 2 miscleavages	$x = \log_2$ (number of tryptic peptides corresponding to protein)	27.84	<0.0001	13908.03

Table 3-1: Quadratic models for adjusting spectral counts to protein size

We used maximum likelihood to fit model parameters β_0 and σ for Model 0 and to fit model parameters β_0, β_L, β_Q and σ for Models 1-4. For Models 1-4, the likelihood ratio test statistic G_i for comparing Model i to Model 0 is computed as $G_i = 2(L_i - L_0)$, where L_j is the maximized log-likelihood assuming Model j. Using standard likelihood theory, under the null hypothesis that Model 0 adequately describes the data, G_i has an approximate chi-square distribution with 2 degrees-of-freedom, from which a P-value is computed. A larger value of G_i represents stronger evidence against the adequacy of Model 0 compared with Model i. Since the likelihood ratio test is valid only for nested models, we used AIC (Akaike, 1974) to compare Models 1-4 with each other. AIC, a measure of goodness-of-fit, is computed as $AIC_j = 2\,p_j - 2\,L_j$, where p_j is the number of model parameters in Model j ($p_j = 2$ for Model 0 and $p_j = 4$ for Models 1-4). Smaller values of AIC suggest better fit.

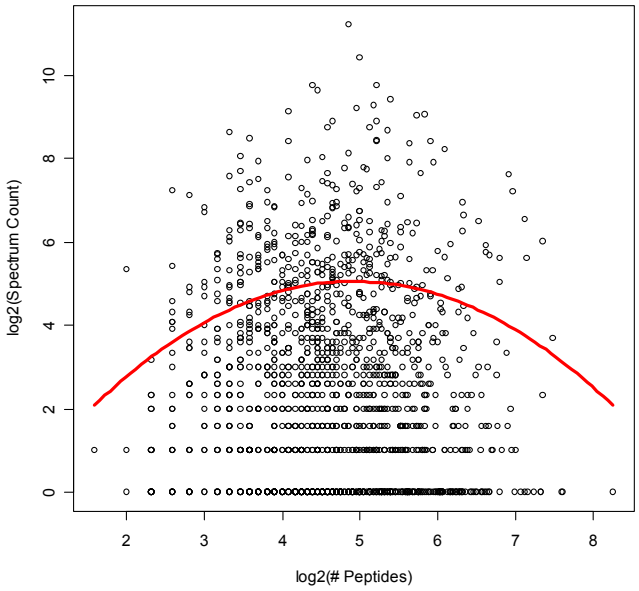


Figure 3-1: Quadratic relationship between the logarithm of the number of tryptic peptides and spectral counts, assuming no miscleavage.

Model 3 had the smallest AIC and largest likelihood ratio test statistic (Table 3-1), suggesting that the optimal adjustment is achieved by considering the number of fully trytpic peptides, assuming one miscleavage. On the basis of the parameters fit to Model 3, an adjusted spectral count was calculated as

$$\text{Adjusted Spectral Count} = \text{Spectral Count} / \exp\left\{\hat{\beta}_L (x-m) + \hat{\beta}_Q (x-m)^2\right\},$$

where x is defined in Table 3-1 for Model 3. Assuming Model 3 is correct, $\beta_0 = \hat{\beta}_0$, $\beta_L = \hat{\beta}_L$, and $\beta_Q = \hat{\beta}_Q$, the mean Adjusted Spectral Count (averaged over all observed proteins) is exp(β_0). Thus, Adjusted Spectral Count is constructed so that the result is independent of protein size as measured by the number of tryptic peptides.

While several ribosomal proteins are encoded by more than a single gene, many of these isoforms were positively identified based on peptides with small variations in sequence. In cases where a single peptide identified two or three ribosomal isoforms, these peptides were added to the quantitation of ribosomal proteins by averaging their ASCs across all isoforms.

Variability testing - To assess the variability of the adjusted spectral count measurements we used a random intercept model. The adjustment model was fitted using log2 (adjusted SC) as the response. Analyses were done using standard software for linear-mixed-effects models (lme in the nlme package of R) The confidence intervals, obtained using the delta method for

computing variance, show precision of the estimate, but are considered approximations, since they assume that the responses are lognormal. A lognormal distribution is deemed a reasonable approximation as apparent from the residual histograms (data not shown).

According to this analysis, reproducibility is very high for three independent 2D LC MS/MS data sets (R^2 = 0.85, Table 3-2). Reproducibility is lower, but still significant, in a joint analysis of all 2D LC, 1D PAGE, and IEF datasets (R^2 = 0.64; Table 3-2).

	On-Line 2D LC[a]		1D PAGE vs. IEF vs. 2D LC[b]	
	Raw SC	Adjusted SC	Raw SC	Adjusted SC
R^2	0.85 (0.83, 0.87)	0.85 (0.82, 0.87)	0.61 (0.57, 0.66)	0.64 (0.59, 0.68)

a Data from three repeat 2D LC-MS/MS runs were included
b Data from all 1D PAGE-based, IEF-based, and 2D LC-based experiments were included

Table 3-2: Variability tests for spectrum counts

Protein-protein and protein-mRNA correlations – The following datasets were used for fission yeast protein and mRNA expression: The Adjusted Spectral Count (Pombe-ASC) data and the cDNA microarray data (Pombe-mRNA) determined in the present study. For correlation with *S. cerevisiae* orthologs, two datasets were used. The first was derived from published abundance ranked 2D LC ESI MS/MS data similar to the data we have generated here (Liu et al., 2004), but with the adjustment of spectral counts to protein size (= Cerevisiae-ASC dataset). This set contained 472 *S. cerevisiae* orthologs of the fission yeast proteins detected in this study. The second list was assembled from the published absolute quantitation data derived from whole genome ORF tagging with the TAP epitope (Cerevisiae-TAP dataset; Ref. (Ghaemmaghami et al., 2003)). This dataset contained 1033 orthologs. Finally, Pombe-ASC was compared with recent relative protein quantitation data provided by reverse protein array analysis of ~4900 fission yeast strains each overexpressing a single epitope tagged protein from an integrated plasmid carrying the nmt1 promoter (Pombe-ORFeome dataset, Ref. (Matsuyama et al., 2006). This dataset contained 1305 proteins that were also represented in the Pombe-ASC dataset.

Pearson-, and where indicated, Spearman rank correlation coefficients were computed between Pombe-ASC and each of Pombe-mRNA, Cerevisiae-TAP, Cerevisiae-ASC, and Pombe-ORFeome.

Assignment of ortholog proteins - In order to determine the evolutionary conservation of each identified protein, we categorized the data into three major classes. 60 proteins

identified in this study were represented on a list of 592 S. pombe-specific proteins (kindly provided by Dr. J Bähler; Sanger Center). A list of "yeast-specific proteins" (ORFs with orthologs in *S. pombe* and *S. cerevisiae* but not in *C. elegans* and/or *D. melanogaster*) and the "core set" (ORFs with orthologs in *S. pombe, S. cerevisiae, C. elegans* and *D. melanogaster*) were constructed using the InParanoid tool (Remm et al., 2001) (http://inparanoid.cgb.ki.se/). High confidence orthologs of *S. pombe* proteins were searched for in budding yeast, worm and fly in version 4.0 of InParanoid. Only proteins obtaining a score of 1.0 were subsequently used. The set of proteins conserved exclusively in yeast contained 917 proteins out of which 291 were identified, and the set of core proteins contained 1548 proteins out of which 731 were identified in this study. The remaining uncategorized group of 383 identified proteins contained either low confidence orthologs, orthologs only present in worm or fly, or proteins completely absent from *S. cerevisiae*, *C. elegans* or *D. melanogaster*, but present in lower organisms. Membership of all identified proteins in any of these groups is indicated in Supplementary Data File 3-2.

SOM clustering - From the Cerevisiae-ASC dataset (Liu et al., 2004) and published mRNA data derived from cDNA microarray analysis of wild-type *S. cerevisiae* (=Cerevisiae-mRNA) (Gasch et al., 2001), 445 protein-mRNA pairs were extracted that had corresponding ortholog pairs in the fission yeast datasets established in this study. All four mRNA and protein expression datasets were preprocessed by log-transformation and subsequent standardization. Each of the variables log2 (Pombe-ASC), log2 (Pombe-mRNA), log2 (Cerevisiae-ASC), and log2 (Cerevisiae-mRNA) were standardized by subtracting their medians and dividing by their standard deviations. Standardized variables were subsequently subject to unsupervised cluster analysis using EXPANDER (Shamir et al., 2005). A self organizing map (SOM) algorithm with a grid width and length of 4 x 4 was instructed to build, in 1,000,000 iterations, 16 clusters with an average homogeneity of 0.81 and an average separation of -0.048. Clustering results were visualized using TreeView (Eisen et al., 1998), with blue color indicating low values, yellow indicating high values and shades in between indicating intermediate values. Each of the 16 clusters was interrogated for significant ($p < 0.05$) overrepresentation of Gene Ontology (GO) attributes using FuncAssociate (Berriz et al., 2003) with the entire dataset (445 S. pombe ORFs) set as the "universe of genes" as reference. ORFs within a cluster belonging to an overrepresented GO attribute were extracted and mean expression profiles were constructed with MS-Excel. When a completely randomized dataset was subjected to the same clustering analysis (SOM

with 16 clusters), none of the clusters had any GO attributes overrepresented with $p < 0.05$ (data not shown).

3.3. Results

Analysis of the S. pombe proteome by multidimensional prefractionation and LC ESI MS/MS

We devised the extensive multidimensional biochemical prefractionation scheme outlined in Figure 3-2, starting with total cell lysate from wild-type fission yeast cells growing vegetatively in mid-log phase in rich media. Aliquots of the lysate were fractionated by preparative isoelectric focusing (IEF) on immobilized pH gradients, or in two different liquid-phase formats, by one-dimensional (1D) gel electrophoresis, and by strong ion-exchange chromatography in a spin column format (Doud et al., 2004), followed by analysis of individual fractions by 1D liquid chromatography coupled with electrospray ionization tandem mass spectrometry (LC ESI MS/MS). In parallel, total fission yeast lysate was subjected to on-line 2D LC ESI MS/MS (='MudPIT'; Washburn et al., 2001), upon in-solution digestion into tryptic peptides. Altogether, 3 million mass spectra were collected and rigorously searched against the fission yeast protein database using the SEQUEST algorithm (Eng et al., 1994). Mass spectra were matched to 12 413 nonredundant peptides (Supplementary Data File 3-1), resulting in the identification of 1465 proteins (Supplementary Data File 3-2) with a predicted false-positive peptide identification rate of 1.05%, as determined by searching against a combined forward and reverse protein database (Peng et al., 2003a). The identified proteins cover 29.5% of the predicted fission yeast proteome. To our knowledge, this represents the highest percent coverage of native, unmodified proteins reported to date for any eukaryotic proteome. We also confirmed 40 predicted sequence orphans as well as five hypothetical proteins, and identified three new proteins, which were listed as dubious ORFs (SPAC13G6.13, SPBC354.04) or pseudogenes (SPBC16E9.16c) in the *S. pombe* genome database.

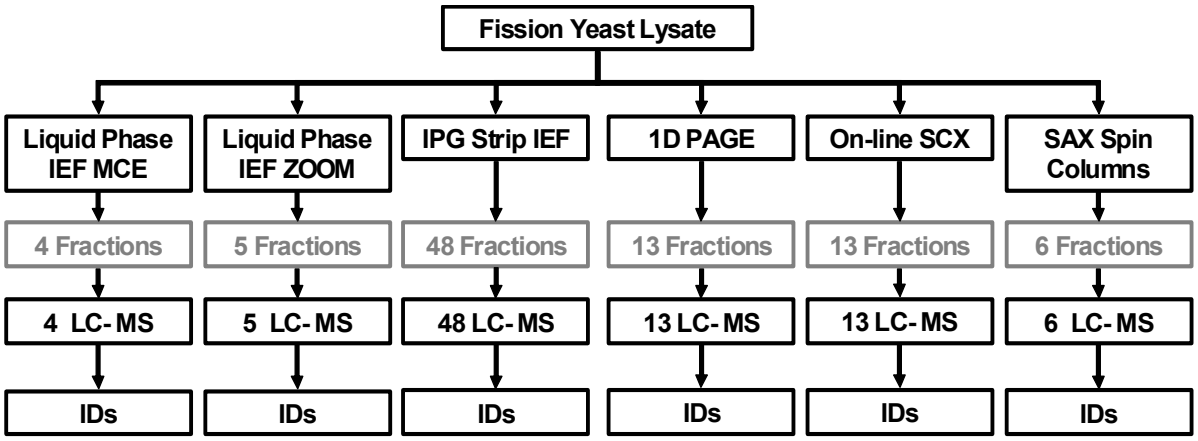


Figure 3-2: Flow chart of sample prefractionation. IEF=isoelectric focusing; ZOOM, MCE (multicompartment electrolizer)=liquid-phase IEF devices, IPG=immobilized pH gradient strips, SAX=strong anion exchange, LC-MS=liquid chromatography and mass spectrometry.

Although the individual prefractionation techniques contributed to the total protein count to different extents (Figure 3-3), the extensive scale of the combined approaches identified a list of proteins that was representative of the whole proteome across the entire range of molecular weights and isoelectric points (Figures 3-4 and 3-5). Most major Gene Ontology (GO) attributes for *S. pombe* were represented, indicating that our study broadly sampled across cell functions (Supplementary Data File 3-2). For example, we identified 132 of 141 ribosomal proteins and all subunits of the 26S proteasome and the CCT chaperonin complex. We also identified all enzymes of the cysteine, glutamate, glycine, isoleucine, leucine, proline, threonine, valine, aspartate, adenine and aromatic amino-acid biosynthesis pathways as well as 45 kinases (23% of all kinases predicted from the genome sequence), 20 predicted transcriptional regulators (14%), and 21 mitochondrial proteins (15%).

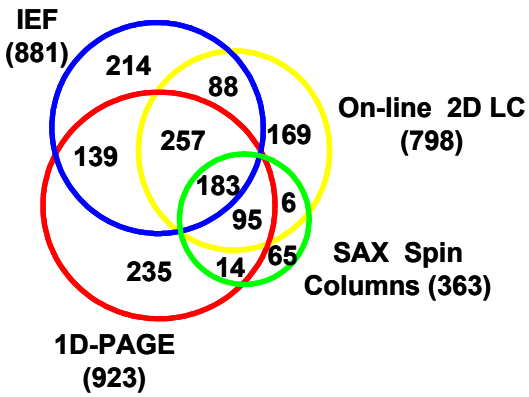


Figure 3-3: Summary of the number of proteins identified with each prefractionation method. The overlap between fractions is indicated.

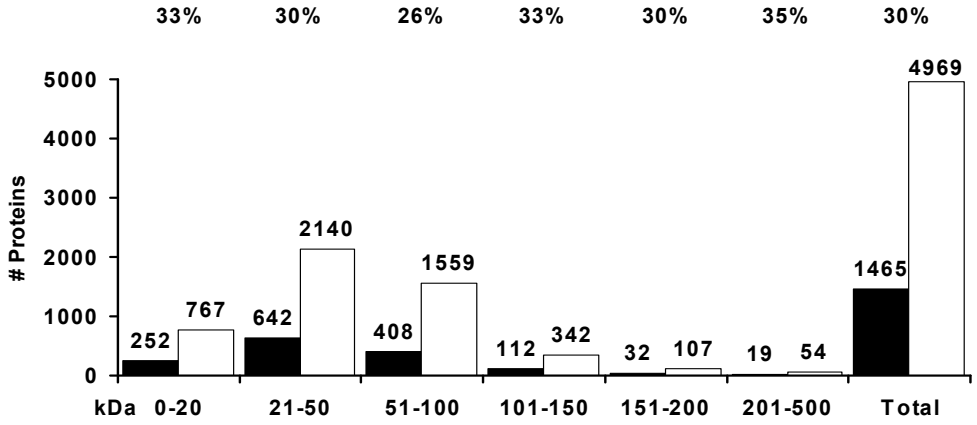


Figure 3-4: Molecular weight distribution of the identified proteins compared to the theoretical proteome. (Solid bars: identified proteins, white bars: theoretical possible proteins)

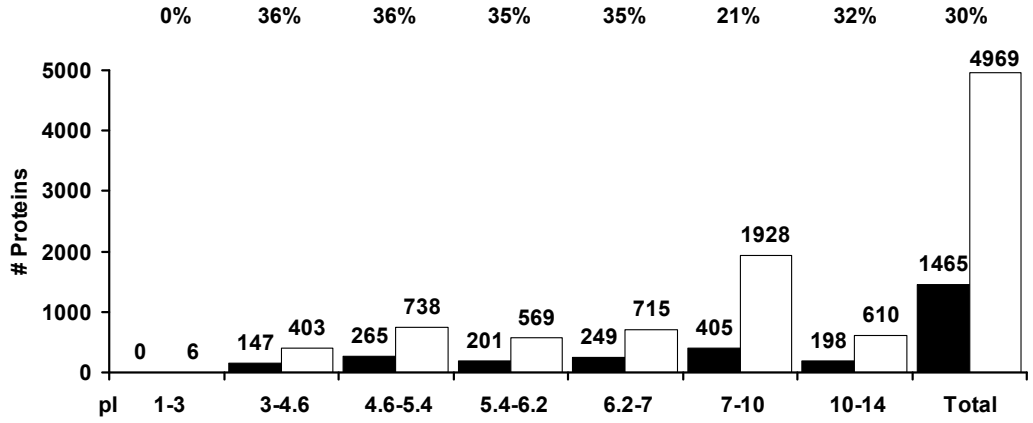


Figure 3-5: Isoelectric point distribution of the identified proteins compared to the theoretical proteome. (Solid bars: identified proteins, white bars: theoretical possible proteins)

More detailed analysis revealed equal identification rates for essential and non-essential proteins (both 36%; Figure 3-6) based on 187 proteins present in our data set for which information on essentiality is available in fission yeast (83 essential, 104 nonessential genes/proteins). Similarly, yeast-specific proteins were represented at the same rate as the entire proteome (30%; Figure 3-6). Metazoan 'core' proteins (proteins common to *S. pombe*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster*; see methods section), were overrepresented (47%; Figure 3-6), a finding that is consistent with their higher mRNA levels (Mata and Bahler, 2003). In contrast, we undersampled proteins containing predicted transmembrane domains (14%) and *S. pombe*-specific proteins (10%; Figure 3-6). Although not all membrane proteins may be equally amenable to extraction

under our sample preparation conditions, the underrepresentation of *S. pombe*-specific proteins is mostly due to their specialized functions in the sexual differentiation pathway (data not shown), which are not expressed in vegetatively growing cells used here.

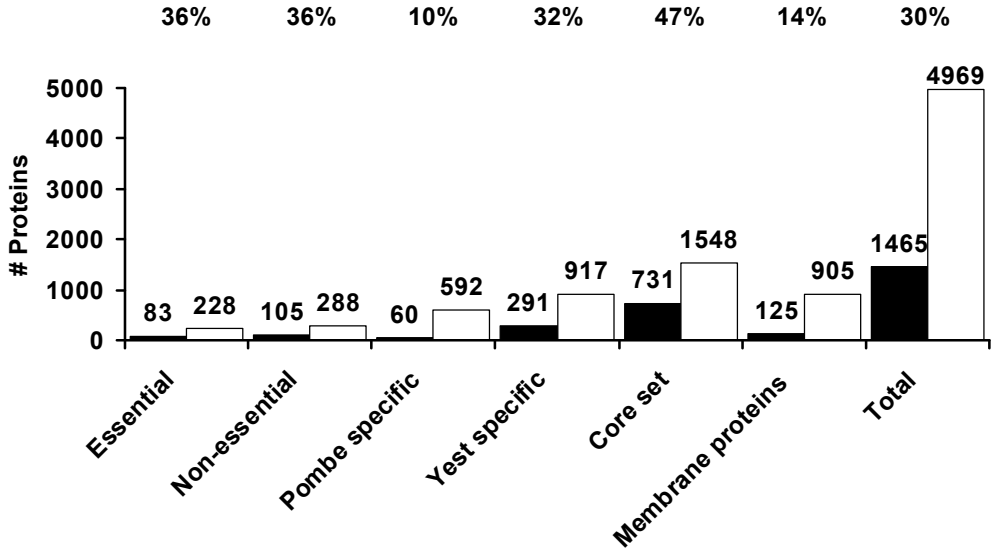


Figure 3-6: Fractions of proteins identified that belong to the indicated categories. (Solod bars: identified proteins, white bars: theoretical possible proteins)

Label-free relative quantitation of S. pombe proteins

To quantitatively rank the identified proteins relative to each other, we used spectral counts. Spectral counts represent the number of nonredundant mass spectra identifying the same protein. Whereas spectral counts are predicted to increase linearly with protein abundance (Liu et al., 2004), this relationship is amended by protein size, with larger proteins having a statistically higher probability of being detected. The relationship is further modified by the sequence-dependent number of peptides produced by the tryptic cleavage. Finally, an allowance for up to three enzymatic miscleavages is often granted during the SEQUEST database search, thus further distorting the theoretical linear relationship between spectral count and protein abundance.

To apply an appropriate adjustment of spectral counts to a measure of protein size, we compared goodness-of-fit statistics applied to negative binomial regression models to determine which of the above parameters (number of amino acids, number of tryptic peptides, miscleavages) figured most prominently. The models revealed that adjustment to

the number of tryptic peptides with one miscleavage resulted in the most optimal fit statistics for the experimental LC ESI MS/MS data (Figure 3-1 and Table 3-1).

Based on adjusted spectral counts (ASCs), we assembled a ranked list of all 1465 proteins identified (Supplementary Data File 3-2). This quantitative ranking reflects the abundance of each protein relative to all others and their quantitative distances. The ranked list was validated by comparing it to absolute quantitation data established for a series of 27 cytokinesis-related fission yeast proteins (Wu and Pollard, 2005). While these absolute measurements rely on epitope tagging, the tagged alleles were extensively validated for functionality under various conditions and in various genetic backgrounds, thus suggesting that tagging did not interfere with normal protein expression (Wu and Pollard, 2005). Of the 27 cytokinesis proteins, 10 were represented on our list. Plotting our ASC data versus the absolute quantitation data revealed a close correlation (rP=0.98; Figure 3-7), suggesting that ASCs provide a good approximation of relative protein abundance.

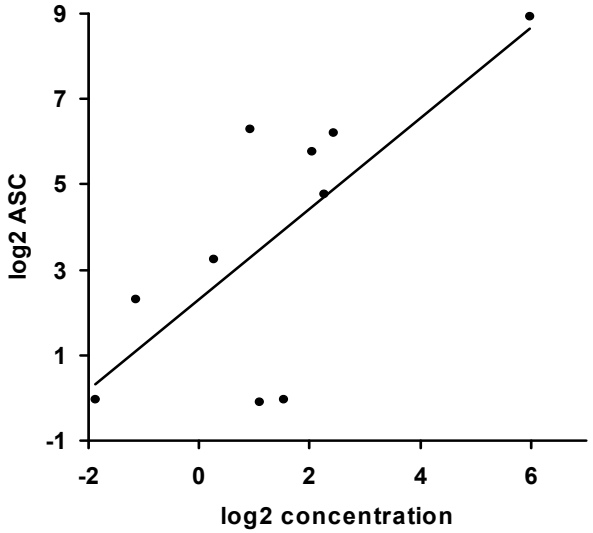


Figure 3-7: Correlation of published absolute quantitation data for several cytokinesis proteins with their corresponding ASCs.

The range of ASCs spanned more than three orders of magnitude (Figure 3-8). The mean ASC was 68.0, whereas the median was 14.6, indicating that the vast majority of the 1465 proteins identified are of relatively low abundance compared to a small number of hyperabundant proteins (Figure 3-8). The group of the 30 most abundant proteins (ASC between 584 and 4269) contained proteins of which all but three have orthologs in budding

yeast that were also detected by whole-genome TAP tagging (Ghaemmaghami et al., 2003). This group includes eight glycolytic enzymes, six enzymes involved in biosynthetic pathways, seven translation factors, five heat-shock proteins, as well as two thioredoxin peroxidases (Supplementary Data File 3-2). The most abundant fission yeast protein is Eno101, a subunit of the phosphopyruvate hydratase complex (ASC=4269), followed by phosphoglycerate kinase (Pgk1, ASC=2301) as a distant second.

The group of the 30 least abundant proteins detected (ASC=0.93–0.95) contains a variety of enzymes involved in RNA metabolism (two helicases, Argonaute 1, two RNA-binding proteins) and ubiquitin-mediated proteolysis, two SH3 domain proteins, three kinases, as well as eight proteins of unknown function (Supplementary Data File 3-2). Notably, 10 out of these 30 proteins do not have orthologs in budding yeast. In addition, seven out of those 20 that do have orthologs did not give signals in the TAP-tagging approach (Ghaemmaghami et al., 2003).

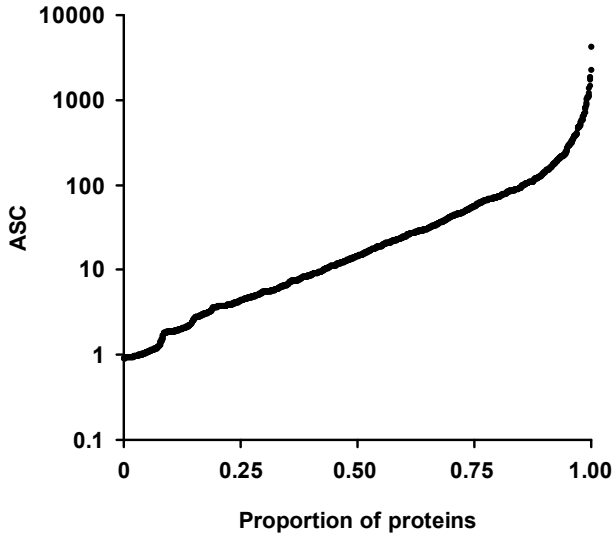


Figure 3-8: ASCs for each of the 1465 identified proteins plotted on a log scale.

Our quantitative data also indicated that the median abundance of metazoan core proteins (ASC=24.2) is significantly higher than that of all proteins detected (ASC=14.6, $P<0.05$), whereas the abundance of *S. pombe*-specific proteins is considerably lower (ASC=5.5; Figure 3-9). This finding is consistent with the higher representation of core proteins in our data set (Figure 3-6) and with their higher mRNA levels as reported previously (Mata and Bahler, 2003). In addition, essential proteins are considerably more abundant (median ASC=12.6)

than non-essential proteins (ASC=7.5). This finding can be rationalized by the enrichment of highly expressed core proteins in the set of essential proteins (Supplementary Data File 3-2).

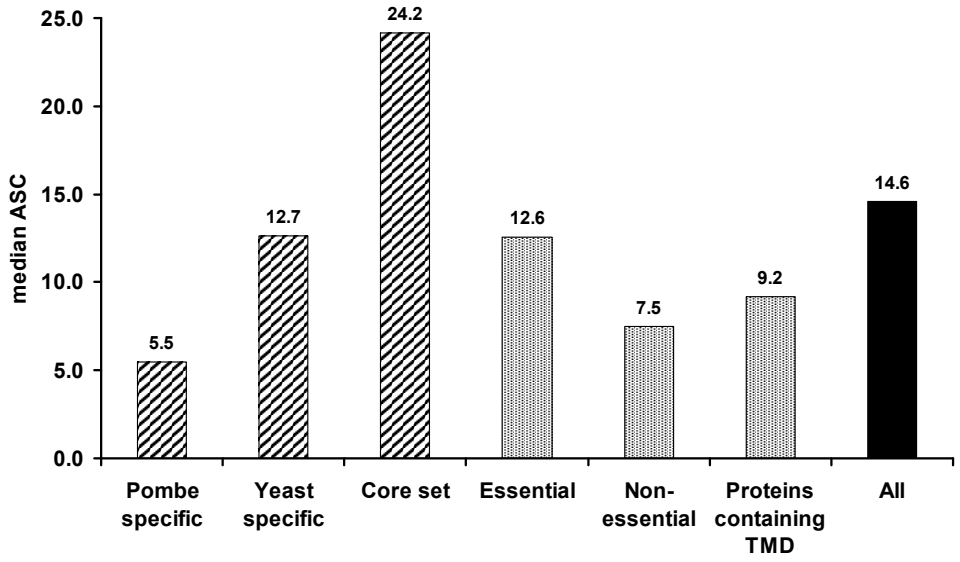


Figure 3-9: Median ASCs for proteins belonging to the indicated categories. All values were statistically different at $p<0.05$ (TMD=transmembrane domain).

Analysis of 10 protein complexes for which we identified greater than 80% of their known or predicted subunits, and which are involved in a large variety of cellular processes, indicated that the translation initiation factor eIF4 is the most abundant protein complex in *S. pombe* (median ASC=85.7; Figure 3-10). eIF4 is similar in abundance to the ribosome (ASC=70.7), but three- to four-fold more abundant than eIF2 (ASC=21.7) and eIF3 (ASC=32.0), two other translation initiation factor complexes (Figure 3-10). Although, during the process of translation initiation, all of these eIFs are known to join a stoichiometric 43S initiation complex, it is thought that eIF2 and eIF3, but not eIF4, dissociate from the mRNA upon successful scanning for the initiator AUG codon (Gebauer and Hentze, 2004). Our finding that eIF2 and eIF3 are considerably less abundant than eIF4 and the ribosome therefore, underpins the concept that the former eIFs are only transiently involved during the initiation reaction, whereas the cap-binding eIF4 complex and the ribosome stay on the mRNA during translation. Our data also indicate that the protein synthesis machinery (ribosome, eIFs) and the protein folding and degradation machinery (CCT chaperonin, proteasome) are among the most abundant molecular modules in fission yeast and perhaps other eukaryotes (Figure3-10).

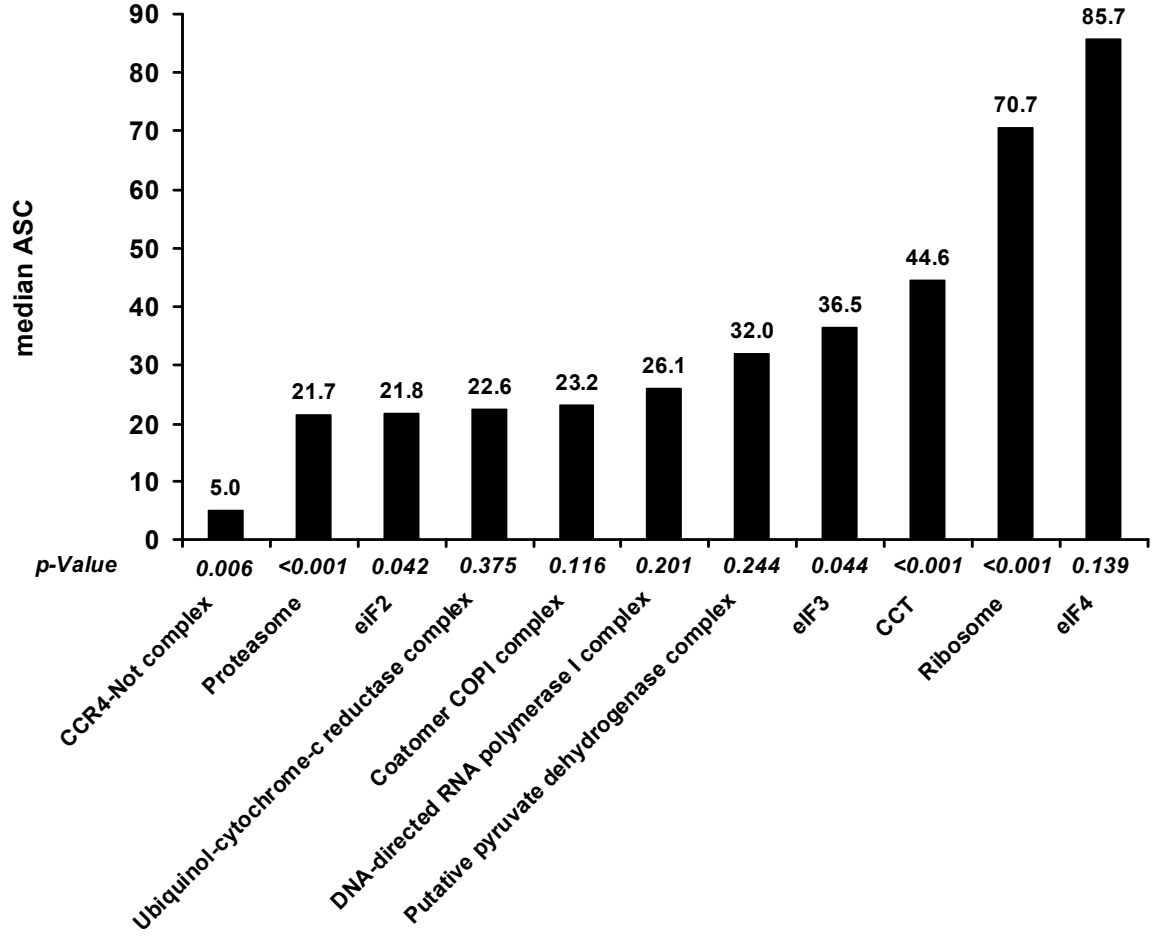


Figure 3-10: Median ASCs for subunits of the indicated protein complexes. For protein complexes with few subunits, P-values are not always <0.05 owing to some outliers.

Comparison of S. pombe proteome data with S. cerevisiae

We compared the abundance ranked list of *S. pombe* proteins with similar lists of S. cerevisiae proteins. This was carried out for the subset of proteins that have known or predicted orthologs in budding yeast (1285 of 1465 proteins based on ortholog mapping information in *S. pombe* GeneDB (www.genedb.org/genedb/pombe/index.jsp). Two data sets of S. cerevisiae proteins were used. The first was derived from published 2D LC ESI MS/MS data (Liu et al., 2004) that we subjected to our adjustment of spectral counts to the number of tryptic peptides (=Cerevisiae-ASC data set). This set contained 473 pairs of orthologous proteins that were detected in both studies. The second list was assembled from the absolute quantitation data derived from whole-genome ORF tagging with the TAP epitope (Cerevisiae-TAP data set; Ghaemmaghami et al., 2003). This data set contained 1033 orthologs, 252 fewer than the theoretically possible 1285, because 20% of the native fission

yeast proteins we detected by 2D LC ESI MS/MS could not be quantified when TAP tagged in budding yeast (Ghaemmaghami et al., 2003). For example, our data set contained all 32 subunits of the 26S proteasome (Finley et al., 1998; Supplementary Data File 3-4), whereas only 25 of these subunits were detected in the ORF tagging approach (Ghaemmaghami et al., 2003). Similarly, we identified 94% of all cytosolic ribosomal subunits, whereas only 76% were identified by TAP tagging in budding yeast (Supplementary Data File 3-4).

Both budding yeast data sets correlated with the fission yeast protein list as indicated by Pearson correlation coefficients of 0.56 and 0.45 and Spearman rank correlation coefficients rs of 0.55 and 0.42, respectively (Figure 3-11; Figure 3-12). Notably, our data showed an overall stronger correlation with the budding yeast 2D LC ESI MS/MS data presented by Liu et al. (2004) (=Cerevisiae-ASC). This finding reinforces previously expressed notions regarding the limitations of comparing mass spectrometry-based proteomics data to absolute quantitation based on ORF tagging (Liu et al., 2004). However, organism-specific differences in protein expression are also expected to distort the correlation (see Figure 3-20 – 3-25).

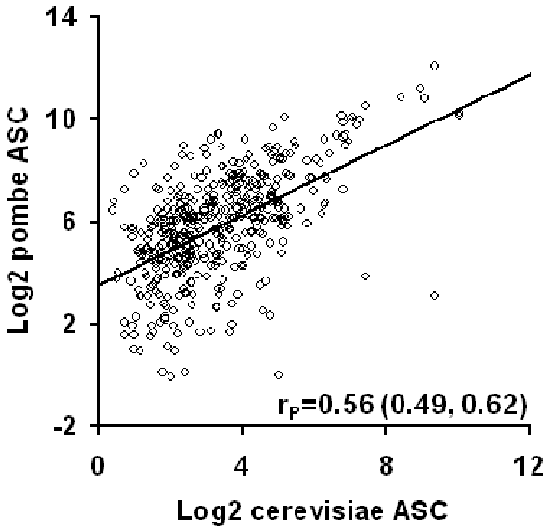


Figure 3-11: Correlation between the Pombe-ASC and the *Cerevisiae*-ASC data sets.

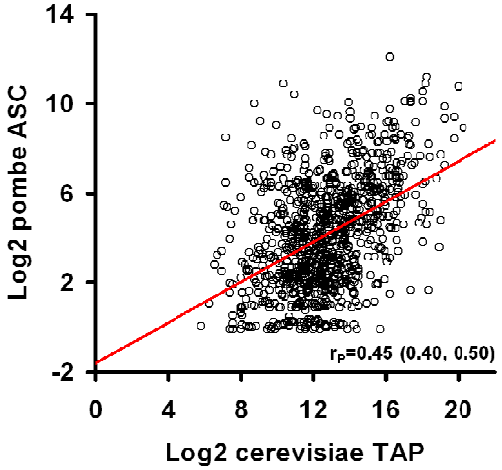


Figure 3-12: Linear regression and Pearson correlation coefficient representing the relationship between protein abundance in S. pombe and S. cerevisiae. Cerevisiae-TAP refers to the whole genome TAP tagging dataset (Ghaemmaghami et al., 2003).

Nonetheless, our LC ESI MS/MS data showed a remarkable overlap with the Cerevisiae-TAP data set in the relative frequency distribution of the detected proteins across the entire dynamic range (Figure 3-13). For example, 88% of the 1033 budding yeast proteins, for which we have identified the fission yeast orthologs, are present at under 50 000 molecules/cell, 62% are under 10 000 molecules/cell, and 11% are under 1000 molecules/cell. This finding suggests that the dynamic range of multidimensional prefractionation and LC ESI MS/MS analysis is not necessarily inferior to that of the wholeORF tagging approach.

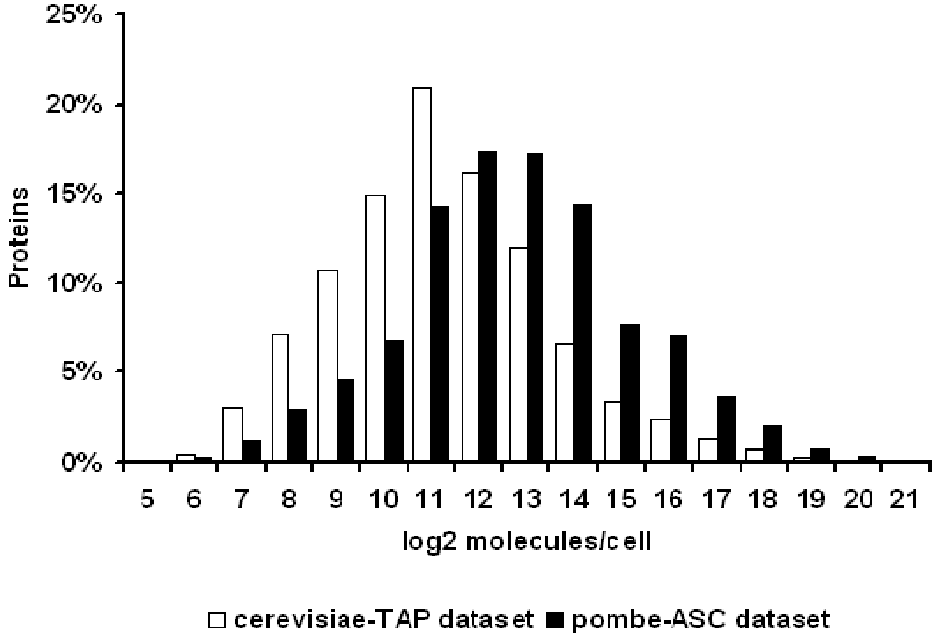


Figure 3-13: Numbers of fission yeast proteins detected were plotted against the absolute amount of their budding yeast orthologues as determined in (Ghaemmaghami et al., 2003). The budding yeast data are shown as a reference.

Correlation of protein and mRNA levels in fission yeast

We next determined the overall correlation of our protein data set with mRNA abundance as estimated by cDNA microarray analysis. Total RNA was prepared from the same *S. pombe* strain maintained under identical growth conditions as used for the proteomic analyses, followed by hybridization onto *S. pombe* cDNA microarrays (Oliva et al., 2005; Zhou et al., 2005). Background subtracted hybridization values averaged from three parallel experiments (see Supplementary Data File 3-2) were used to estimate mRNA abundance. Although it is clear that the hybridization values obtained on cDNA microarrays are influenced by factors other than mRNA abundance (probe length, GC content, etc.), these variations are relatively minor with probes longer than 500 bp as used here (Lyne et al., 2003). Similarly, Mata and Bahler (2003) have previously used absolute hybridization signals as approximate measures of mRNA levels in fission yeast.

The comparison of 1367 protein–mRNA pairs for which data were obtained (Supplementary Data File 3-2) revealed a Spearman rank correlation coefficient (rS) of 0.61 and a Pearson correlation coefficient (rP) of 0.58 (Figure 3-14), indicating a substantial correlation between mRNA and protein abundance in fission yeast. The extent of correlation is very similar in budding yeast as determined with the whole-genome TAP tagging data set

(rS=0.57, Ghaemmaghami et al., 2003), and by an independent re-evaluation of additional large-scale budding yeast data sets (rP=0.66; Greenbaum et al., 2003).

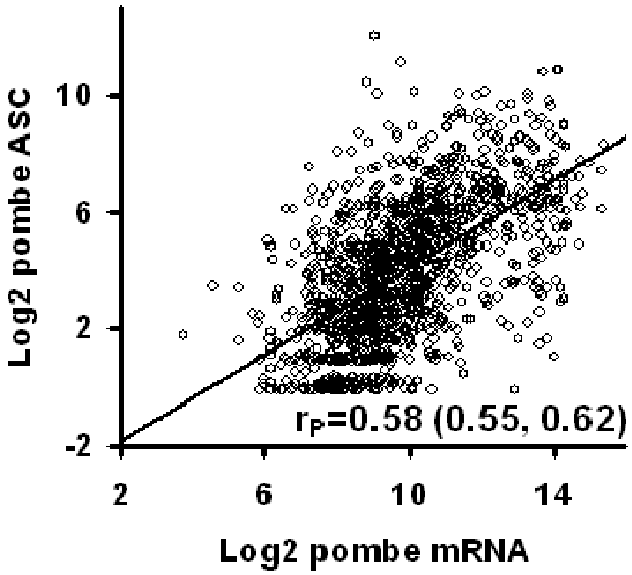


Figure 3-14: Scatter plot representing the relationship between mRNA and protein (Pombe-ASC). The Pearson correlation coefficient is indicated.

The mean mRNA intensity of proteins detected in our multidimensional analysis was 2462, whereas for undetected proteins the number was 420 (Figure 3-15). This comparison confirmed the expectation that mass spectrometry-based proteomics has a bias towards detecting proteins encoded by highly expressed mRNAs. However, a significant portion of low-abundance mRNAs may encode proteins that never accumulate in the vegetative state. Consistent with this notion is the demonstration that 1033 genes are induced more than four-fold during nitrogen starvation and meiosis (Mata et al., 2002). The actual vegetative translatome may therefore be devoid of many of the proteins encoded by such developmentally regulated mRNAs. These proteins may also escape detection by ORF tagging and immunoblotting, thus explaining why the dynamic range of our LC ESI MS/MS analysis was comparable to the whole-genome ORF tagging approach employed in budding yeast (Ghaemmaghami et al., 2003; Figure 3-13).

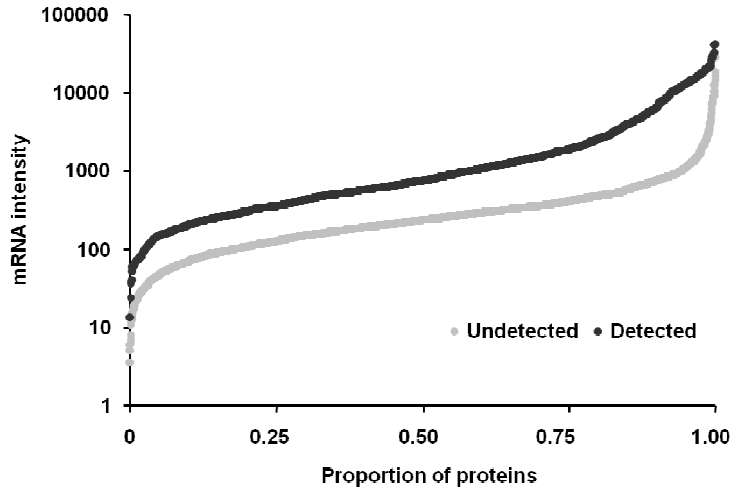


Figure 3-15: mRNA fluorescence intensities for the proteins detected and undetected in this study.

Functional pathway analysis

Although the overall protein–mRNA correlation is surprisingly high, we wondered whether this correlation is maintained throughout specific functional pathways, protein families, and multisubunit protein complexes. We calculated the Pearson correlation coefficients for several subclasses of protein–mRNA pairs that were highly represented in our data set (see Supplementary Data Files 3-3 and 3-4 for individual proteins). Whereas a high coefficient was obtained for kinases (rP=0.80; Figure 3-16), the correlation was weak for transporters (rP=0.21) and the unfolded protein response (UPR) pathway (rP=0.12), and moderately strong for glycolytic enzymes (rP=0.36) and transcription factors (rP=0.42). Correlations similar to those observed for all proteins (rP=0.58) were found for the categories amino-acid biosynthesis (rP=0.63), signal transduction (rP=0.61), protein translation (rP=0.5), stress response (rP=0.58), and cell-cycle regulation (rP=0.67; Figure 3-16).

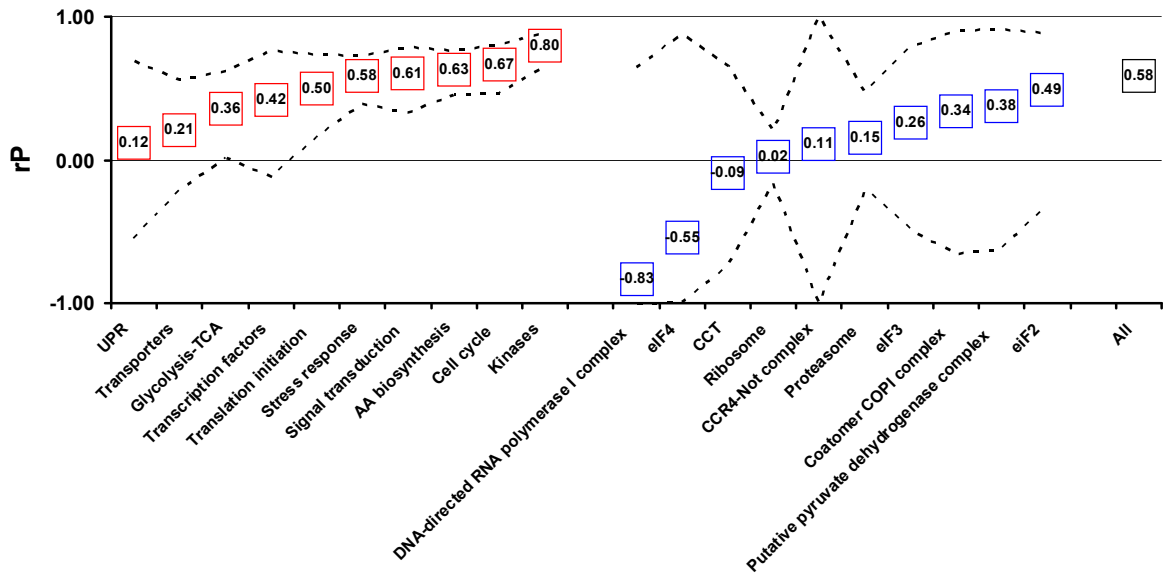


Figure 3-16: Protein–mRNA correlation coefficients for proteins belonging to the indicated pathways, protein families, and complexes. Dashed lines indicate 95% confidence intervals (AA=amino acid, UPR=unfolded protein response, TCA=tricarboxylic acid cycle).

For the majority of multisubunit protein complexes, very low or even negative correlation coefficients were obtained (Figure 3-16). Previous bioinformatics studies have suggested that a high protein–mRNA correlation (i.e. the higher the mRNA, the higher the protein) as observed here for kinases and cell-cycle components reflects control of protein abundance primarily at the level of mRNA synthesis, whereas poor correlation is indicative of post-transcriptional control (Greenbaum et al., 2003). By extension, negative correlations indicate extensive control at the post-transcriptional level (i.e. the higher the mRNA, the lower the protein and vice versa). The subunits of presumed stoichiometric protein complexes such as the 80S ribosome, the 26S proteasome, and the CCT complex would therefore be controlled substantially at the post-transcriptional level.

The poor protein–mRNA correlation for complexes would be expected, if their subunits were coordinately regulated. For example, if all subunits of a protein complex had exactly equal protein and mRNA levels, say 5.0 units and 1.0 unit, respectively, then all data points would coincide at the very same coordinates of a protein versus mRNA plot (x=5; y=1; protein–mRNA ratio=5). Consequently, the protein–mRNA correlation would be zero for the subunits of this protein complex.

Indeed, we noticed that the protein and mRNA data points for many protein complexes were not randomly scattered over the entire data map, but tended to cluster together. To comprehensively illustrate this, we determined the protein–mRNA ratio individually for

every protein in a given pathway, family, or complex, and compared it to the entire data set. Individual ratios of functional pathway components were used to determine their location and relative distance on the ratio distribution curve of the entire data set of 1367 protein–mRNA pairs. This reference curve indicates the extent and orientation of the deviations of all observed ratios from the median ratio, which was arbitrarily set to 1.0. The partitioning of pathway components along this curve thus informs about the degree to which they cluster around certain protein–mRNA ratios and their distances from the median. The graphical representation of clustering effects was enhanced by displaying the data points for specific pathway components at equal distance laid over the reference curve, thus causing informative phase shifts of the curves.

This analysis revealed strong deviations from the reference curve for several protein complexes, suggesting more consistent protein–mRNA ratios for individual subunits than observed for all proteins. Ribosomal subunits clustered with relatively higher levels of mRNA than protein (Figure 3-17; Supplementary Data File 3-5), whereas the shape of the ratio distribution curve for eIF3, the COP1 complex, and several other protein complexes (Supplementary Data File 3-5) indicated clustering around the median ratio (Figure 3-17). This differential distribution was even more pronounced for the eight subunits of the CCT complex (Figure 3-17). In other words, all eight subunits of the CCT complex displayed highly similar protein–mRNA ratios, and therefore appear to be coordinately regulated at the mRNA and protein levels. Thus, although the protein–mRNA correlations were low for multisubunit protein complexes, clustering of their protein–mRNA ratios around similar values indicated coordinate regulation of complex subunits (Table 3-3). Although this regulation could principally occur at any level, the low protein–mRNA correlation suggests a substantial contribution of post-transcriptional mechanisms (Greenbaum et al., 2003). Notably, the UPR pathway showed a similar pattern in correlation and ratio distribution (Figure 3-16 and 3-17), perhaps suggesting that components of this pathway are also present in stoichiometric amounts.

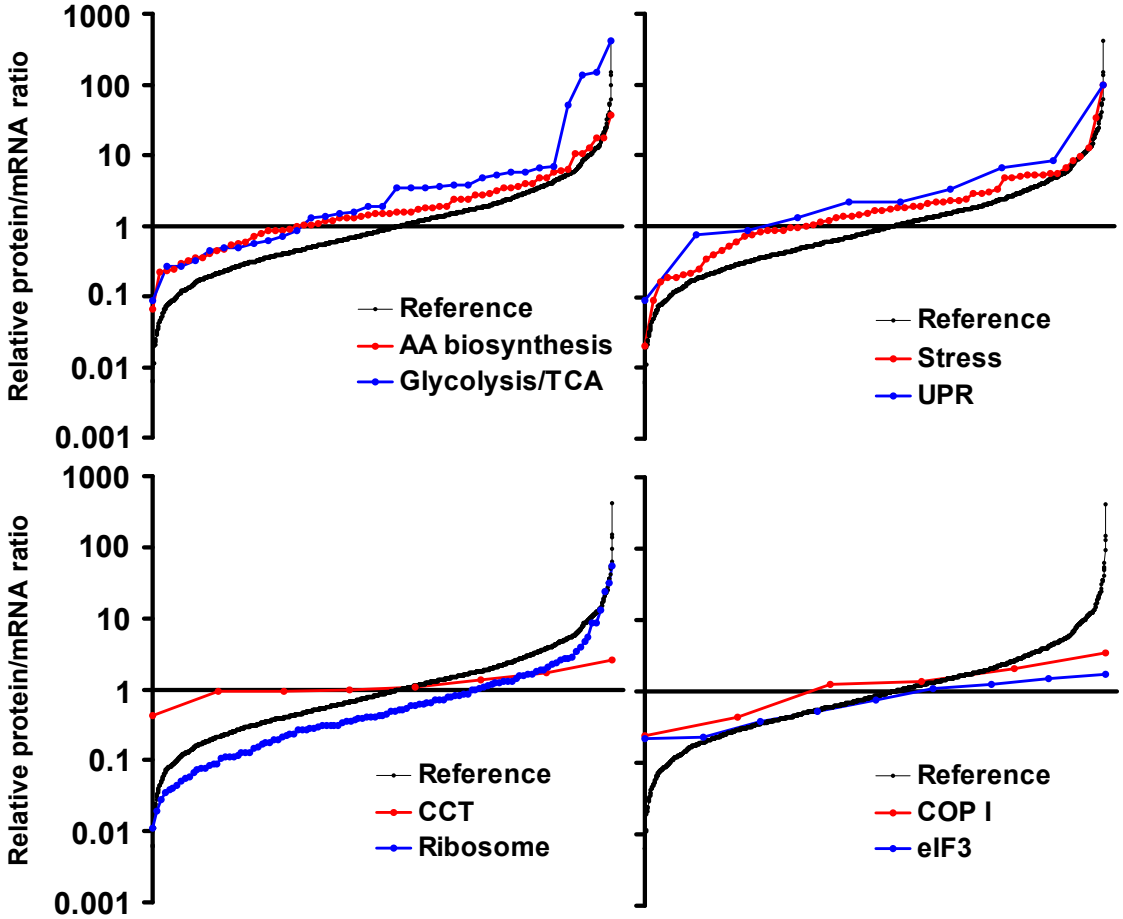


Figure 3-17: Protein–mRNA ratios for individual members of the indicated pathways, protein families, or complexes. The data are displayed relatively to median centered ratios of the entire data set of 1381 mRNA–protein pairs (black graphs).

The reverse scenario, clustering of protein–mRNA ratios around similar values, but relatively high protein–mRNA correlation, was observed for the stress response pathway as well as for glycolysis and amino-acid biosynthesis (Figure 3-16 and 3-17). This pattern indicated that protein and mRNA expression varied widely among the members of these groups (Table 3-3). This might reflect the fact that proteins involved in hierarchical signal transduction cascades or linear and circular metabolic pathways do not necessarily cooperate in stoichiometric amounts. Rather signal amplification and the specific activities of metabolic enzymes may govern the varying levels of protein required for these functions.

Most other pathways and protein families showed a considerable overlap of protein–mRNA ratios with the reference curve, indicating no clustering. Among those were entities with low (transporters; Figure 3-16) and high (kinases; Figure 3-16) protein–mRNA

correlation. For these remaining cases, high protein–mRNA correlations would suggest control primarily at the transcriptional level, whereas low correlations would indicate extensive post-transcriptional control (Table 3-3) (Greenbaum et al., 2003).

		Protein – mRNA ratio	
		Clustering	No clustering
Protein – mRNA correlation	High	Methabolic and signal transduction pathways Non-coordinate expression	Protein families (kinases), cell cycle Transcriptional control
	Low	Multisubunit protein complexes; UPR Coordinate regulation	Protein families (transporters) Post transcriptional control

Table 3-3: Protein–mRNA relationships

Protein and mRNA relationship as a correlate of post-translational modifications

Although no specific enrichment strategies were employed, rigorous interrogation of our peptide data sets obtained by mass spectrometry provided high confidence indications for post-translational modification (PTM) of 53 peptides, which were matched to 51 proteins. A total of 40 proteins contained at least one peptide that was phosphorylated on either serine, threonine, or tyrosine (Table 3-4). The set of phosphoproteins was enriched for protein kinases (15% versus 1.6% in the entire proteome), a finding that is consistent with the known propensity of these enzymes to autophosphorylate and/or be part of kinase cascades. The budding yeast orthologs of eight of these proteins were previously shown to be phosphorylated by methods other than mass spectrometry. In one case, acetyl coenzyme-A carboxylase, the serine phosphorylation site we mapped in fission yeast exactly corresponds to the same position where the budding yeast protein was found to be modified (Ficarro et al., 2002).

For another set of 11 proteins, we mapped the precise sites of modification with the diglycine moieties created upon trypsin digestion of ubiquitylated lysines (Table 3-4). Independent evidence for ubiquitylation of the budding yeast ortholog of one of these proteins was provided previously (Peng et al., 2003b). Five proteins contained both phosphorylated and ubiquitylated peptides (Table 3-4), a finding that is consistent with the well-established connection between phosphorylation and ubiquitylation (Karin and Ben-Neriah, 2000).

The median abundance of phosphorylated (ASC=3.58) and ubiquitylated (ASC=2.84) proteins was considerably lower than the abundance of all 1465 proteins in the data set (ASC=14.6; Figure 3-18). Ubiquitylated proteins also showed a stark dissociation of median mRNA levels, which were relatively high (633 versus 757 in the entire data set), from protein levels, which were very low (2.84 versus 14.6; Figure 3-18). This finding indicates that extensive proteolytic control of these proteins through the ubiquitin–proteasome pathway may be dominant over their relatively high mRNA expression levels. This conclusion was further strengthened by comparing individual protein–mRNA ratios of ubiquitylated proteins to median adjusted ratios for the entire data set. This analysis revealed clustering of ubiquitylated proteins with relatively higher mRNA than protein levels, whereas phosphoproteins showed a distribution largely congruent with the reference curve (Figure 3-19).

ID	Name	Product	Phosphopeptide
SPAC1834.01	sup45	translation release factor eRF1	K.FHT#EALAELLES#DQR.F
SPCC24B10.21	tpi1	triosephosphate isomerase	R.RT#IFKES#DEFVADK.T
SPAC16E8.10c	SPAC16E8.10c	mitochondrial ribosomal protein subunit S7	K.AKAEKIVAT#ALS#IIQK.E
SPBC16G5.07c	SPBC16G5.07c	prohibitin	R.FS#RILT#PGVAFLAPIIDK.I
SPAC18B11.11	SPAC18B11.11	GTPase activating protein	K.VLS#EWLT#DLFTIIDDM*R.A
SPAC23G3.12c	SPAC23G3.12c	serine protease	R.Y#VEVCGAKFHNLSYQLAR.Q
			K.K@GT#ALVLDKDKGLAVT#S#R.S
SPAC20H4.09	SPAC20H4.09	ATP-dependent RNA helicase	R.T#LS#T#DLLLGVLK.R
SPAPB2B4.04c	SPAPB2B4.04c	P-type calcium ATPase	R.T#EGQAT#PLQLRLS#R.V
SPBC342.02	SPBC342.02	glutaminyl-tRNA synthetase	R.LMFLPDPIKVTLENLDDS#Y#R.E
SPCC1450.11c	cek1	serine/threonine protein kinase Cek1	K.PENLLIS#QNGHLK.L
SPAC56E4.04c	cut6	acetyl-CoA carboxylase	R.LQS#VSDLSWYVNK.T
SPBC839.15c	ef1a-c	translation elongation factor EF-1 alpha	K.MVPS#KPMCVEAFTDYAPLGR.F
SPAC18B11.04	ncs1	related to neuronal calcium sensor Ncs1	K.NKDGQLTLEEFCEGS#KR.D
SPBC651.01c	nog1	GTP binding protein Nog1	R.EGYYDS#DQEIEDADEEEVLEK.A
SPAC10F6.09c	psm3	mitotic cohesin complex subunit Psm3	K.S#KVALELQSSQLSRQIEFSK.K
SPCC962.04	rps1201	40S ribosomal protein S12	R.QAHLCVLCES#CDQEAYVK.L
SPCC417.08	tef3	translation elongation factor 3	R.FKLRKYLGNMS#EFVK.K
SPAC17C9.03	tif471	translation initiation factor eIF4G	R.SGSQVSDQVVESPNSSTLS#PR.N
SPAC1006.09	win1	MAP kinase kinase kinase Win1	R.LSDNELAS#FVK.E
SPCC4B3.11c	SPCC4B3.11c	BolA domain	K.S#K@AFQGKNTLAQHR.L
SPCC16C4.02c	SPCC16C4.02c	sequence orphan	K.NLSSATVILS#NLLK.A
SPAC4G8.09	SPAC4G8.09	leucine-tRNA ligase	K.VQLSYQKM*S#K.S
SPBC27B12.08	SPBC27B12.08	AP-1 accessory protein	K.VVS#LMIELLENLTAVNDPK.L
SPAC25A8.01c	SPAC25A8.01c	fun thirty related protein Fft3	K.KS#QVLDALPKKTR.I
SPBC8D2.06	SPBC8D2.06	isoleucine-tRNA ligase	K.NVIVS#GLVMAEDGKKM*SK@R.L
SPAC343.07	mug28	meiotically upregulated gene Mug28	K.VHDKENAFAEATGTSILS#S#K.A
SPBC146.14c	sec26	coatomer beta subunit	R.AS#LGEVPILAS#EEQLLK.D
SPBC29A3.09c	SPBC29A3.09c	AAA family ATPase	K.ELEELS#KDQTADQAIS#R.R
SPAC1635.01	SPAC1635.01	voltage-dependent anion-selective channel	K.Y#ALDKDT#FVK.G
SPBC1683.07	mal1	alpha-glucosidase Mal1	R.TPM*HWDSSPNGGFT#K.A
SPBC1861.09	ppk22	serine/threonine protein kinase	-.MARET#EFNDK.S
SPAC29E6.03c	uso1	armadillo repeat protein	K.LT#KQLDDIK@NQFGIISSK.N
SPBC947.10	SPBC947.10	zinc finger protein	K.RAFSEIKNAT#FLNIPER.V
SPAPB18E9.02c	ppk18	serine/threonine protein kinase Ppk18	K.QKTELAT#FT#TY#K.E /
			K.QKTELAT#FTT#Y#K.E
SPAC30D11.09	cwf19	complexed with Cdc5 protein Cwf19	R.KYGQNYEY#AKQIAK.D
SPAC343.01c	erg8	phosphomevalonate kinase	K.GY#ASTTTLDDKCGTVRVK.S
SPAC4G8.05	ppk14	serine/threonine protein kinase	K.SGK@FY#AM*KVLSKQEM*IK.R
SPBC3F6.05	rga1	GTPase activating protein	K.NSGAIY#DKNDGTQK.G
SPAC11E3.12	SPAC11E3.12	conserved eukaryotic protein	K.IY#GVNTKEKLVDIM*EALTQK.K
SPAC2F3.13c	SPAC2F3.13c	queuine tRNA-ribosyltransferase	R.ELVAWILLQLY#VYIKEHGK.E

ID	Name	Product	Ubiquitylated peptide
SPAC4G8.05	ppk14	serine/threonine protein kinase	K.SGK@FY#AMKVLSKQEMIK.R
SPAC589.10c	SPAC589.10c	ribosomal-ubiquitin fusion protein	R.TLSDYNIQK@ESTLHLVLR.L
SPBC28F2.10c	kap1	kinesin-associated protein	K.IGSSATSGSFPVIKSLMDK@R.S
SPAC23G3.12c	SPAC23G3.12c	serine protease	K.K@GT#ALVLDKDKGLAVT#S#R.S
SPCC736.11	ago1	argonaute	K.NK@SDGDRNGNPLPGTIIEK.H
			K.LT#KQLDDIK@NQFGIISSK.N
SPCC4B3.11c	SPCC4B3.11c	BolA domain	K.S#K@AFQGKNTLAQHR.L
SPBC354.01	gtp1	GTP binding protein Gtp1	R.LARLPK@SVVISCNMK.L
SPAC1420.02c	cct5	chaperonin-containing T-complex epsilon subunit	K.EKFQEMIK@HVK.D
SPCP1E11.11	SPCP1E11.11	RNA-binding protein	K.VASKLIVIIK@K.Y
SPAC3G6.04	rnp24	RNA-binding protein Rnp24	R.FNDAESLGQEDKPNFK@RAR.K
SPBC8D2.06	SPBC8D2.06	isoleucine-tRNA ligase	K.NVIVS#GLVMAEDGKKMSK@R.L

# denotes phosphorylation
@ denotes ubiquitylation
* denotes oxidation

Table 3-4: Post-translationally modified peptides

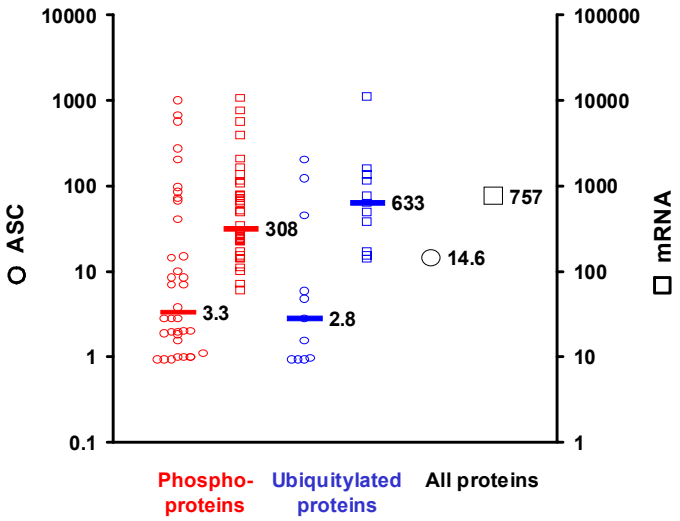


Figure 3-18: mRNA levels and ASCs for 40 phosphorylated and 11 ubiquitylated proteins as compared to the entire data set. Medians are indicated by the vertical bars.

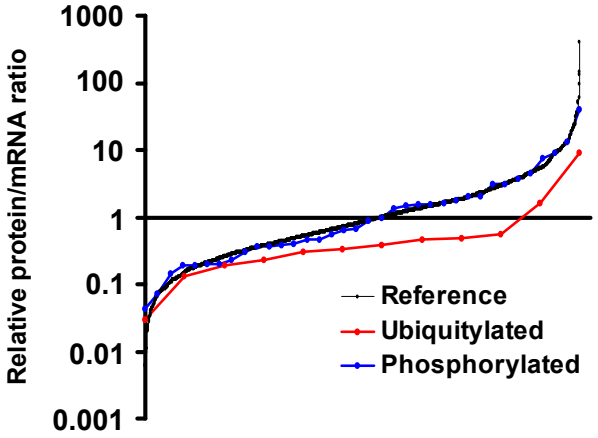


Figure 3-19: Individual protein–mRNA ratios for phosphorylated and ubiquitylated proteins relative to median centered ratios of the entire data set.

Steady-state proteome and transcriptome comparison of S. pombe and S. cerevisiae

The generation of quantitative fission yeast protein and mRNA data sets and the availability of corresponding data sets for budding yeast enabled the first large-scale comparison of mRNA and protein levels of two eukaryotic organisms. For this, we used the Cerevisiae-MS data (Liu et al., 2004) with adjustment of spectral counts to the number of tryptic peptides and published mRNA data derived from cDNA microarray analysis of wild-

type S. cerevisiae grown under conditions comparable to those of our fission yeast strains (Gasch et al., 2001). As the raw values of the four data sets were on different scales, they were log-transformed and standardized (see Methods section). As a result, each data set contained a continuum of mRNA and protein values ranging from high to low abundance for 445 distinct entities common to all four data sets. A self-organizing map (SOM) algorithm was used to arrange the four data sets into distinct clusters (see Methods section). The algorithm was instructed to assemble 16 clusters, because this number achieved good performance in reproducibility (data not shown), average cluster homogeneity (0.81), and separation (-0.048).

The SOM revealed many similarities in the mRNA and protein abundance patterns in the two yeasts, but also marked differences. The most frequent patterns represented roughly equal mRNA and protein levels in both organisms (clusters 3, 4, 7, 9, 10, and 13; Figure 3-20). In addition, one pattern was indicative of concordantly low mRNA and high protein abundance in both yeasts (cluster 6), whereas cluster 15 showed the opposite pattern. Among the discordant patterns were those with higher mRNA and protein levels in *S. pombe* (cluster 1), as well as various patterns where either mRNA or protein levels found in one yeast deviated from what was found in the other (clusters 2, 5, 8, 11, 12, 14, and 16).

The clusters were further interrogated for overrepresented *S. pombe* GO terms using the FuncAssociate tool (Berriz et al., 2003). In total, seven nonredundant GO attributes were found significantly (P0.0005) overrepresented in the clusters (Figure 3-20). As noise is a notorious feature of large-scale functional genomics data, the biological significance of these patterns will require further validation by more targeted experiments. However, as not a single GO attribute was enriched in SOM clusters derived from a random data set under identical conditions (data not shown), our data suggest that many pathways and complex subunits are coordinately, albeit not necessarily concordantly regulated in both fission and budding yeasts. For example, 6/13 components of the microtubule cytoskeleton organization GO category present in our data sets were coordinately and concordantly regulated in both yeasts (cluster 10; Figure 3-21). In contrast, ATPases and entities involved in chromatin remodelling and intracellular transport were coordinately, but discordantly regulated with mRNA levels being low in budding yeast (cluster 14, Figure 3-22 – 3-24).

Although 48 out of 121 fission yeast ribosomal subunits present in all data sets were coordinately regulated, they partitioned into two distinct clusters (3 and 16, Figure 3-20). Both clusters indicated that fission yeast ribosomal protein mRNAs are typically higher than the subunits they encode (Figures 3-25 and 3-22). Notably, higher mRNA than protein levels

were previously reported also in human cells (Ishihama et al., 2005). Coordinate post-transcriptional regulation of ribosomal proteins in both budding and fission yeasts was already observed in previous reports (Washburn et al., 2003; Bachand et al., 2006). Ribosomal proteins are known to be subject to extensive transcriptional and post-transcriptional control as indicated by short mRNA half-lives (Li et al., 1999) and extensive translational regulation (Meyuhas, 2000; Bachand et al., 2006). Although presumably serving to provide stoichiometric amounts of complex subunits, such control might also ensure the excess availability of individual ribosomal subunits that fulfill extraribosomal functions (Wool, 1996), a repertoire, that may vary from one organism to another.

Overall, our comparison reinvigorates the conclusion gained from previous functional genomics studies that similarities in the control of gene expression in the two yeasts are less pronounced than expected from genome comparisons (Mata et al., 2002; Rustici et al., 2004; Oliva et al., 2005). Only a remarkably small fraction of transcriptomic changes during cell-cycle progression (Rustici et al., 2004; Oliva et al., 2005) and sexual differentiation (Mata et al., 2002) is shared among the two yeasts. True organism-specific differences are therefore likely to underlie the moderate overall correlation in protein abundance in the two yeasts (Figure 3-11 and Figure 3-12) as well as the different patterns of mRNA and protein expression revealed here by SOM clustering (Figure 3-20—3-25).

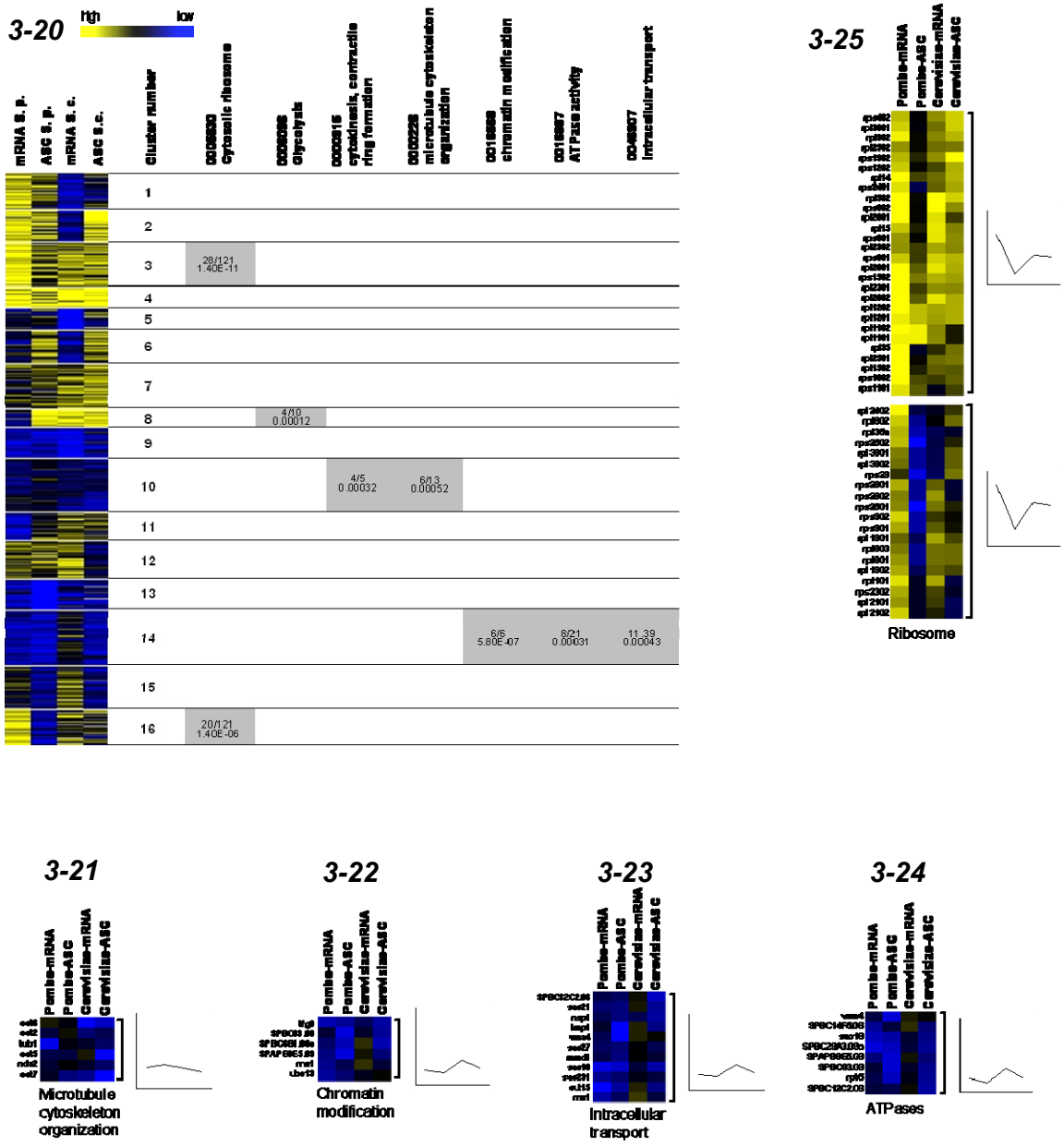


Figure 3-20 – 3-25: Comparison of proteome and transcriptome data from S. pombe and S. cerevisiae. (3-21) Self-organizing map cluster analysis of the fission and budding yeast mRNA and ASC protein data sets. The table on the right shows GO terms overrepresented in the various clusters and the P-values of enrichment. Also indicated is the number of proteins with a particular GO attribute enriched in each cluster over the total number with this attribute present in the entire data set. The names of the genes and proteins in the individual clusters are listed in Supplementary Data File 3-6. (3-22 – 3-26) Detailed view of subclusters containing (3-22) microtubule cytoskeleton organization (GO: 0000226), (3-23) chromatin modification components (GO: 0016568), (3-24) components involved in intracellular transport (GO: 0046907), (3-25) ATPases (GO: 0016887), and (3-26) ribosomal proteins (GO: 0005830). The graphs next to the heat maps indicate the mean variations in signal intensities.

3.4. Discussion

Shotgun proteomics employing multidimensional prefractionation and tandem mass spectrometry, aided by mathematical modelling of spectral count information, enabled a label-free relative quantitation of 30% of the theoretical fission yeast proteome corresponding to an estimated 50% of the entire vegetative translatome. Whereas Eno101, a subunit of the phosphopyruvate hydratase complex, was revealed as the single most abundant protein, the translation initiation factor eIF4 represents the most abundant protein complex. Highly abundant proteins also included the core set of proteins conserved in metazoans. Among the least abundant proteins observed in this study were *S. pombe*-specific proteins, a series of nonessential proteins, as well as proteins modified by phosphorylation and ubiquitylation. Whereas there was a positive overall correlation between protein and mRNA abundance in fission yeast similar to what was observed in other organisms, simple correlations proved insufficient to asses regulatory patterns of gene expression. Contrasting individual protein–mRNA ratios to the ratio distribution curve representing all entities suggested common schemes of control for subunits of protein complexes, unstable ubiquitylated proteins, and several functional pathways. The first large-scale comparison of mRNA and protein abundance in two related eukaryotic model organisms indicated frequently coordinate, but rarely concordant regulation, an observation that further underscored the marked differences in gene expression in the two yeasts noted previously (Mata et al., 2002; Rustici et al., 2004; Oliva et al., 2005). The data presented should become a valuable resource for the fission yeast community as well as researchers mining comprehensive gene expression data sets for systems biology.

3.5. References

Anderson NL, Matheson AD, Steiner S (2000) Proteomics: applications in basic and applied biology. Curr Opin Biotechnol 11: 408–412

Bachand F, Lackner DH, Bahler J, Silver PA (2006) Autoregulation of ribosome biosynthesis by a translational response in fission yeast. Mol Cell Biol 26: 1731–1742

Bakhtiar R, Tse FL (2000) Biological mass spectrometry: a primer. Mutagenesis 15: 415–430

Berriz GF, King OD, Bryant B, Sander C, Roth FP (2003) Characterizing gene sets with FuncAssociate. Bioinformatics 19: 2502–2504

Berriz, G.F., King, O.D., Bryant, B., Sander, C. and Roth, F.P. (2003) Characterizing gene sets with FuncAssociate. Bioinformatics, 19, 2502-2504.

Chen D, Toone WM, Mata J, Lyne R, Burns G, Kivinen K, Brazma A, Jones N, Bahler J (2003) Global transcriptional responses of fission yeast to environmental stress. Mol Biol Cell 14: 214–229

Doud MK, Schmidt MW, Hines D, Naumann C, Kocourek A, Kashani-Poor N, Zeidler R, Wolf DA (2004) Rapid prefractionation of complex protein lysates with centrifugal membrane adsorber units improves the resolving power of 2D-PAGE-based proteome analysis. BMC Genomics 5: 25

Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A., 95, 14863-14868.

Eng JK, McCormack AL, Yates III JR (1994) An approach to correlate MS/MS data to amino acid sequences in a protein database. Am Soc Mass Spectrom 5: 976–989

Eng, J.K., McCormack , A.L. and Yates, J.R., 3rd. (1994) An approach to correlate MS/MS data to amino acid sequences in a protein database. American Soc. Mass Spectrom., 5, 976-989.

Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, Shabanowitz J, Hunt DF, White FM (2002) Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae. Nat Biotechnol 20: 301–305

Finley D, Tanaka K, Mann C, Feldmann H, Hochstrasser M, Vierstra R, Johnston S, Hampton R, Haber J, McCusker J, Silver P, Frontali L, Thorsness P, Varshavsky A, Byers B, Madura K, Reed SI, Wolf D, Jentsch S, Sommer T, Baumeister W, Goldberg A, Fried V, Rubin DM, Toh-e A (1998) Unified nomenclature for subunits of the Saccharomyces cerevisiae proteasome regulatory particle. Trends Biochem Sci 23: 244–245

Gasch AP, Huang M, Metzner S, Botstein D, Elledge SJ, Brown PO (2001) Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. Mol Biol Cell 12: 2987–3003

Gasch, A.P., Huang, M., Metzner, S., Botstein, D., Elledge, S.J. and Brown, P.O. (2001) Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. Mol Biol Cell, 12, 2987-3003.

Gebauer F, Hentze MW (2004) Molecular mechanisms of translational control. Nat Rev Mol Cell Biol 5: 827–835

Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci USA 100: 6940–6945

Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea EK, Weissman JS (2003) Global analysis of protein expression in yeast. Nature 425: 737–741

Ghaemmaghami, S., Huh, W.K., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K. and Weissman, J.S. (2003) Global analysis of protein expression in yeast. Nature, 425, 737-741.

Greenbaum D, Colangelo C, Williams K, Gerstein M (2003) Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol 4: 117

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17: 994–999

Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB (2001) molecular evidence for the early colonization of land by fungi and plants. Science 293: 1129–1133

Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4: 1265–1272

Karin M, Ben-Neriah Y (2000) Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. Annu Rev Immunol 18: 621–663

Kislinger T, Cox B, Kannan A, Chung C, Hu P, Ignatchenko A, Scott MS, Gramolini AO, Morris Q, Hallett MT, Rossant J, Hughes TR, Frey B, Emili A (2006) Global survey of organ and organelle protein expression in mouse: combined proteomic and transcriptomic profiling. Cell 125: 173–186

Kislinger, T., Cox, B., Kannan, A., Chung, C., Hu, P., Ignatchenko, A., Scott, M.S., Gramolini, A.O., Morris, Q., Hallett, M.T., Rossant, J., Hughes, T.R., Frey, B. and Emili, A. (2006) Global Survey of Organ and Organelle Protein Expression in Mouse: Combined Proteomic and Transcriptomic Profiling. Cell, 125, 173-186.

Land, K.C., McCall, P.L. and Nagin, D.S. (1996) A comparison of Poisson, negative binomial, and semiparametric mixed Poisson regression models with empirical applications to criminal careers data. Sociological Methods and Research, 24, 387-442.

Li B, Nierras CR, Warner JR (1999) Transcriptional elements involved in the repression of ribosomal protein synthesis. Mol Cell Biol 19: 5393–5404

Liu H, Sadygov RG, Yates III JR (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem 76: 4193–4201

Liu, H., Sadygov, R.G. and Yates, J.R., 3rd. (2004) A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem, 76, 4193-4201.

Lyne R, Burns G, Mata J, Penkett CJ, Rustici G, Chen D, Langford C, Vetrie D, Bahler J (2003) Whole-genome microarrays of fission yeast: characteristics, accuracy, reproducibility, and processing of array data. BMC Genomics 4: 27

Marguerat S, Jensen TS, de Lichtenberg U, Wilhelm BT, Jensen LJ, Bahler J (2006) The more the merrier: comparative analysis of microarray studies on cell cycle-regulated genes in fission yeast. Yeast 23: 261–277

Mata J, Bahler J (2003) Correlations between gene expression and gene conservation in fission yeast. Genome Res 13: 2686–2690 [E-pub 2003 Nov 12]

Mata J, Lyne R, Burns G, Bahler J (2002) The transcriptional program of meiosis and sporulation in fission yeast. Nat Genet 32: 143–147

Matsuyama A, Arai R, Yashiroda Y, Shirai A, Kamata A, Sekido S, Kobayashi Y, Hashimoto A, Hamamoto M, Hiraoka Y, Horinouchi S, Yoshida M (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. Nat Biotechnol 24: 841–847, [E-pub 2006 Jun 25]

Matsuyama, A., Arai, R., Yashiroda, Y., Shirai, A., Kamata, A., Sekido, S., Kobayashi, Y., Hashimoto, A., Hamamoto, M., Hiraoka, Y., Horinouchi, S. and Yoshida, M. (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. Nat Biotechnol., 24, 841-847. Epub 2006 Jun 2025.

Meyuhas O (2000) Synthesis of the translational apparatus is regulated at the translational level. Eur J Biochem 267: 6321–6330

Oliva A, Rosebrock A, Ferrezuelo F, Pyne S, Chen H, Skiena S, Futcher B, Leatherwood J (2005) The cell cycle-regulated genes of Schizosaccharomyces pombe. PLoS Biol 3: e225

Oliva, A., Rosebrock, A., Ferrezuelo, F., Pyne, S., Chen, H., Skiena, S., Futcher, B. and Leatherwood, J. (2005) The cell cycle-regulated genes of Schizosaccharomyces pombe. PLoS Biology, 3.

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1: 376–386

Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP (2003a) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res 2: 43–50

Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP (2003b) A proteomics approach to understanding protein ubiquitination. Nat Biotechnol 21: 921–926

Peng X, Karuturi RK, Miller LD, Lin K, Jia Y, Kondu P, Wang L, Wong LS, Liu ET, Balasubramanian MK, Liu J (2005) Identification of cell cycle-regulated genes in fission yeast. Mol Biol Cell 16: 1026–1042, [E-pub 2004 Dec 22]

Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J. and Gygi, S.P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. J Proteome Res, 2, 43-50.

Qian, W.J., Liu, T., Monroe, M.E., Strittmatter, E.F., Jacobs, J.M., Kangas, L.J., Petritis, K., Camp, D.G., 2nd and Smith, R.D. (2005) Probability-based evaluation of peptide and protein identifications from tandem mass spectrometry and SEQUEST analysis: the human proteome. J Proteome Res, 4, 53-62.

Remm, M., Storm, C.E. and Sonnhammer, E.L. (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. J Mol Biol., 314, 1041-1052.

Resing, K.A., Meyer-Arendt, K., Mendoza, A.M., Aveline-Wolf, L.D., Jonscher, K.R., Pierce, K.G., Old, W.M., Cheung, H.T., Russell, S., Wattawa, J.L., Goehle, G.R., Knight, R.D. and Ahn, N.G. (2004) Improving reproducibility and sensitivity in identifying human proteins by shotgun proteomics. Anal Chem., 76, 3556-3568.

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ (2004) Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3: 1154–1169, [E-pub 2004 Sep 22]

Rustici G, Mata J, Kivinen K, Lio P, Penkett CJ, Burns G, Hayles J, Brazma A, Nurse P, Bahler J (2004) Periodic gene expression program of the fission yeast cell cycle. Nat Genet 36: 809–817

Shamir, R., Maron-Katz, A., Tanay, A., Linhart, C., Steinfeld, I., Sharan, R., Shiloh, Y. and Elkon, R. (2005) EXPANDER--an integrative program suite for microarray data analysis. BMC Bioinformatics, 6, 232.

Thurston, S.W., Wand, M.P. and Wiencke, J.K. (2000) Negative binomial additive models. Biometrics, 56, 139-144.

Washburn MP, Koller A, Oshiro G, Ulaszek RR, Plouffe D, Deciu C, Winzeler E, Yates III JR (2003) Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. Proc Natl Acad Sci USA 100: 3107–3112

Washburn MP, Ulaszek R, Deciu C, Schieltz DM, Yates III JR (2002) Analysis of quantitative proteomic data generated via multidimensional protein identification technology. Anal Chem 74: 1650–1657

Washburn MP, Wolters D, Yates III JR (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotechnol 19: 242–247

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

Wood V (2006) Schizosaccharomyces pombe comparative genomics, from sequence to systems. In Comparative Genomics using fungi as models. Topics in Current Genetics, Sunnerhagen P and Piskur J (eds) Vol. 15, pp. 233–285. Heidelberg: Springer Verlag

Wood V, Gwilliam R, Rajandream M-A, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S, Basham D, Bowman S, Brooks K, Brown D, Brown S, Chillingworth T, Churcher C, Collins M, Connor R, Cronin A, Davis P, Feltwell T, Fraser A, Gentles S, Goble A, Hamlin N, Harris D, Hidalgo J, Hodgson G, Holroyd S, Hornsby T, Howarth S, Huckle EJ, Hunt S, Jagels K, James K, Jones L, Jones M, Leather S, McDonald S, McLean J, Mooney P, Moule S, Mungall K, Murphy L, Niblett D, Odell C, Oliver K, O'Neil S, Pearson

D, Quail MA, Rabbinowitsch E, Rutherford K, Rutter S, Saunders D, Seeger K, Sharp S, Skelton J, Simmonds M, Squares R, Squares S, Stevens K, Taylor K, Taylor RG, Tivey A, Walsh S, Warren T, Whitehead S, Woodward J, Volckaert G, Aert R, Robben J, Grymonprez B, Weltjens I, Vanstreels E, Rieger M, Schafer M, Muller-Auer S, Gabel C, Fuchs M, Fritzc C, Holzer E, Moestl D, Hilbert H, Borzym K, Langer I, Beck A, Lehrach H, Reinhardt R, Pohl TM, Eger P, Zimmermann W, Wedler H, Wambutt R, Purnelle B, Goffeau A, Cadieu E, Dreano S, Gloux S, Lelaure V, Mottier S, Galibert F, Aves SJ, Xiang Z, Hunt C, Moore K, Hurst SM, Lucas M, Rochet M, Gaillardin C, Tallada VA, Garzon A, Thode G, Daga RR, Cruzado L, Jimenez J, Sanchez M, del Rey F, Benito J, Dominguez A, Revuelta JL, Moreno S, Armstrong J, Forsburg SL, Cerrutti L, Lowe T, McCombie WR, Paulsen I, Potashkin J, Shpakovski GV, Ussery D, Barrell BG, Nurse P (2002) The genome sequence of Schizosaccharomyces pombe. Nature 415: 871–880

Wool IG (1996) Extraribosomal functions of ribosomal proteins. Trends Biochem Sci 21: 164–165

Wu JQ, Pollard TD (2005) Counting cytokinesis proteins globally and locally in fission yeast. Science 310: 310–314

Wu, J.Q. and Pollard, T.D. (2005) Counting cytokinesis proteins globally and locally in fission yeast. Science, 310, 310-314.

Yates JR (2000) Mass spectrometry. From genomics to proteomics. Trends Genet 16: 5–8

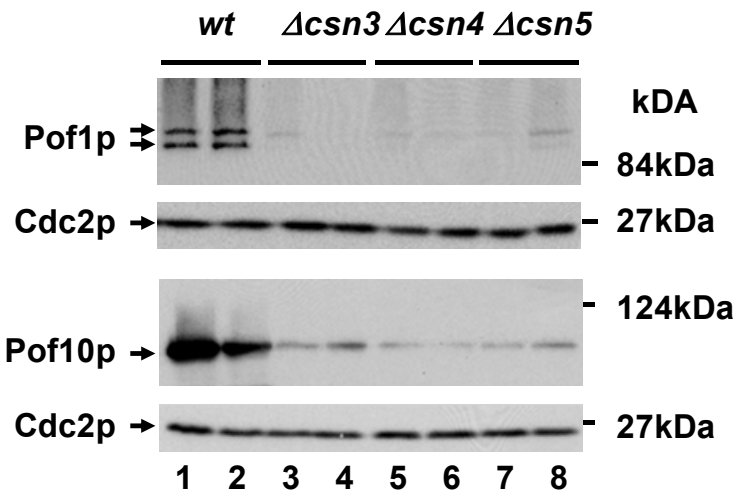
Zhou C, Arslan F, Wee S, Krishnan S, Ivanov A, Oliva A, Leatherwood J, Wolf DA (2005) PCI proteins eIF3e and eIF3m define distinct translation initiation factor 3 complexes. BMC Biol 3: 14

CONCLUSIONS

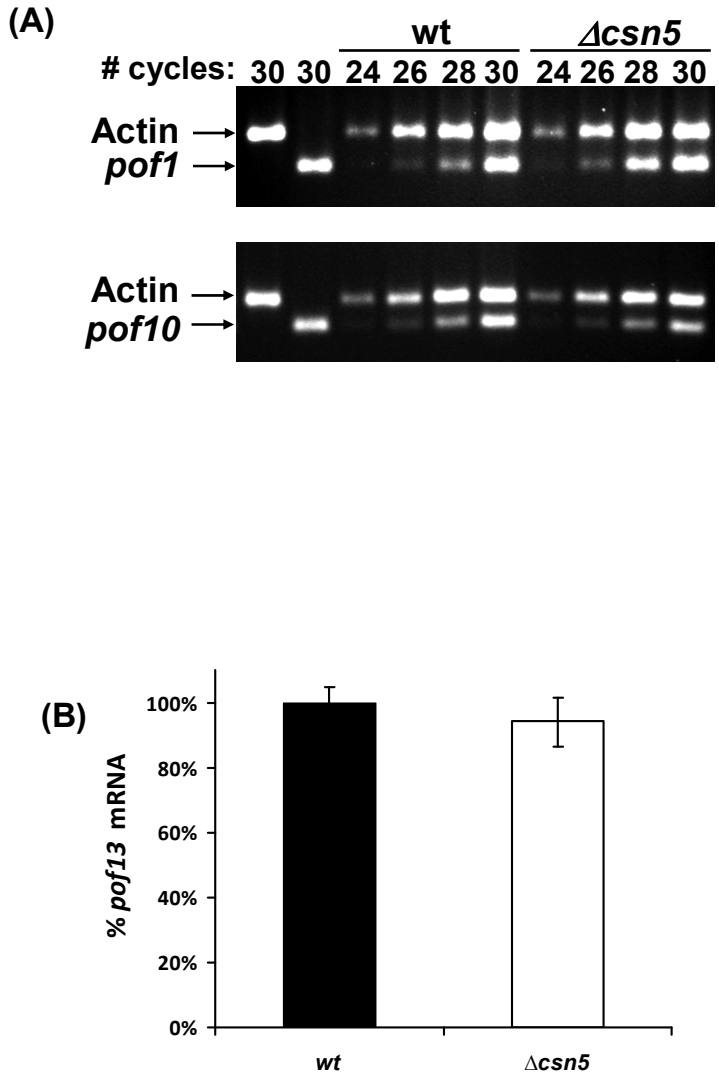
The regulation of protein levels in the cell is the final and arguably the most important step in control of cellular function. The most fundamental cellular process, cell division, is regulated by the ubiquitin-proteasome system through timely degradation of cyclins. Regulation of growth factor receptors by the UPS is essential for normal growth and differentiation of cells, and malfunction of this process contributes to development of many types of cancer. Diverse specialized cellular functions, for example neuronal synapse formation and neurotransmitter release, are regulated by the UPS and malfunction of the ubiquitin ligase pathway is implicated in Parkinson's disease. Comparison of closely related species, such as humans and chimpanzees which are 98% identical at the genomic level and thus express highly similar proteins, would suggest that divergence in protein sequence is simply not sufficient to account for the vast phenotypic differences between related species. The interpretation of highly similar genomes into very different organisms is thought to depend at least in part on regulation of protein levels through tightly controlled synthesis and degradation. This dissertation demonstrates that functional orthologs in *S. pombe* and *S. cerevisiae* have in fact widely different expression levels on a genome-wide scale. Small changes in protein sequences may lead to much greater changes in protein interaction networks, thereby amplifying a minor molecular change into a significant difference in cellular function. In part one of this thesis, evidence is provided that through evolution, as little as one amino acid change in the F-box of ubiquitin ligase substrate adapters is sufficient to abolish or change its function dramatically. Transient changes in protein stability and activity can have far-reaching functional effects through complex interactions with other proteins in networks. Deciphering these networks is an intriguing and complex problem, to which further system-wide studies on the protein level may be able to provide solutions.

SUPPLEMENT

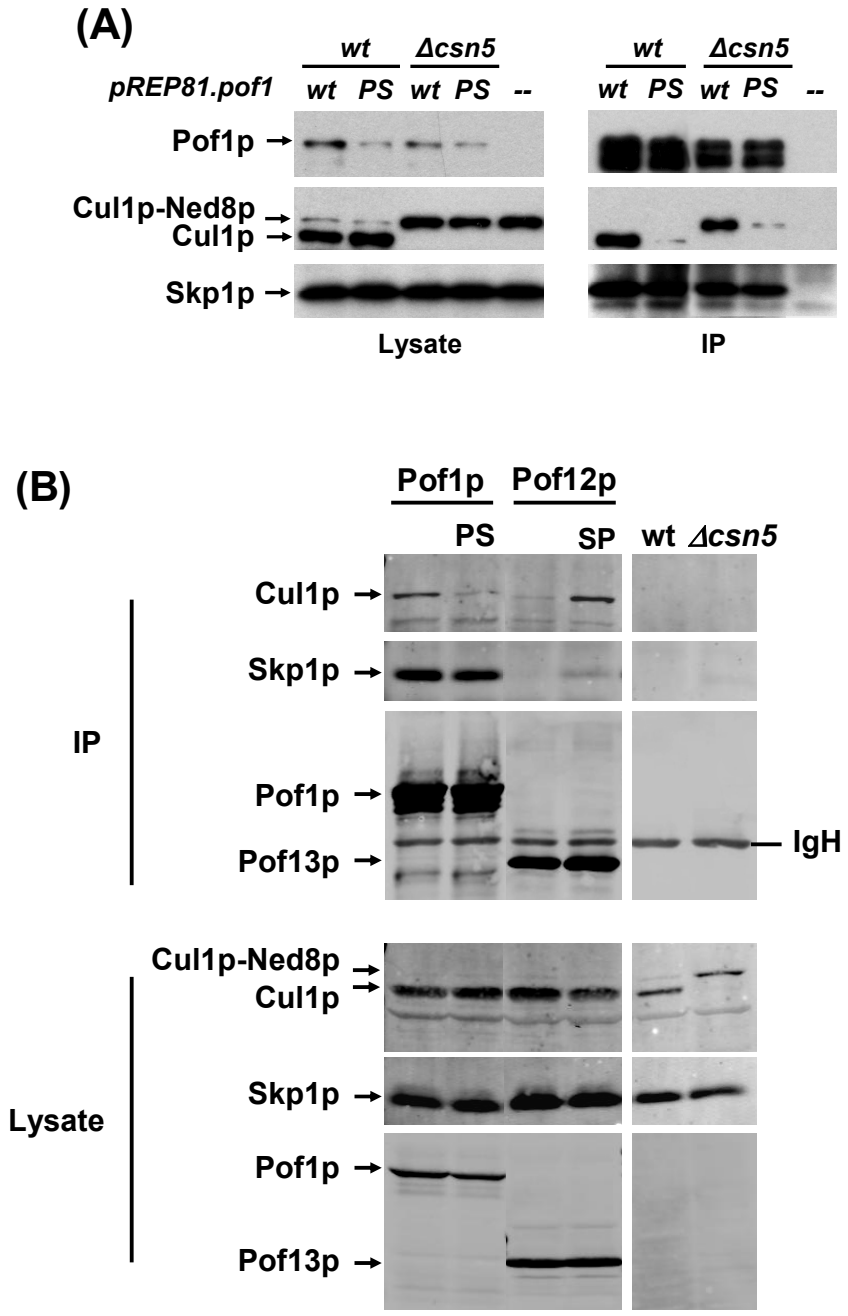
Supplement to Chapter 1



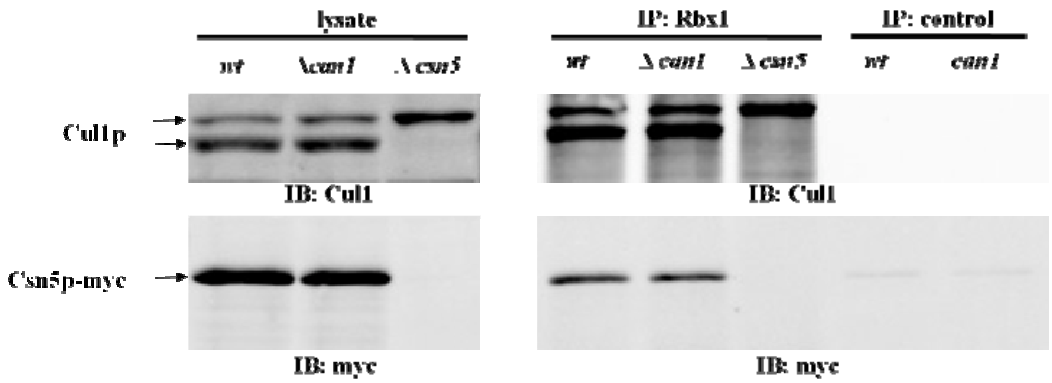
Supplementary Figure 1-1: **Differential Effect of CSN on FBP Levels.** Steady-state levels of Myc-tagged Pof1p and Pof10p in *csn3*, *csn4*, and *csn5* deletion strains. Duplicate strains are shown for each mutant. Note that Pof1p is known to sometimes migrate as a duplet on SDS gels (Harrison et al, 2005).



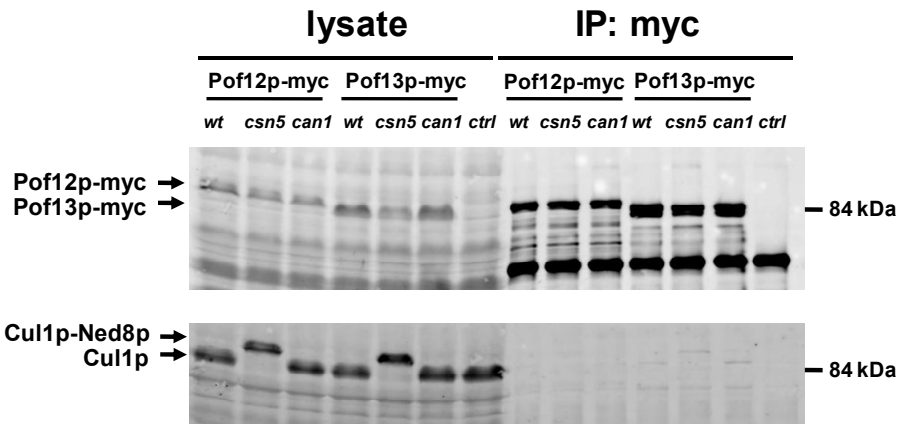
Supplementary Figure 1-2: **Adapter mRNA levels are not influenced by CSN**. (A) Total mRNA was isolated from the indicated strains and used in (A) RT/PCR reactions with primers specific for btb3 and actin as a loading control. Samples were removed after the number of PCR cycles indicated. (B) qRT/PCR reactions, assaying *pof13* expression in triplicates in an iCycler iQ (BioRad). Normalization was performed using the expression of 18S rDNA as an internal control.



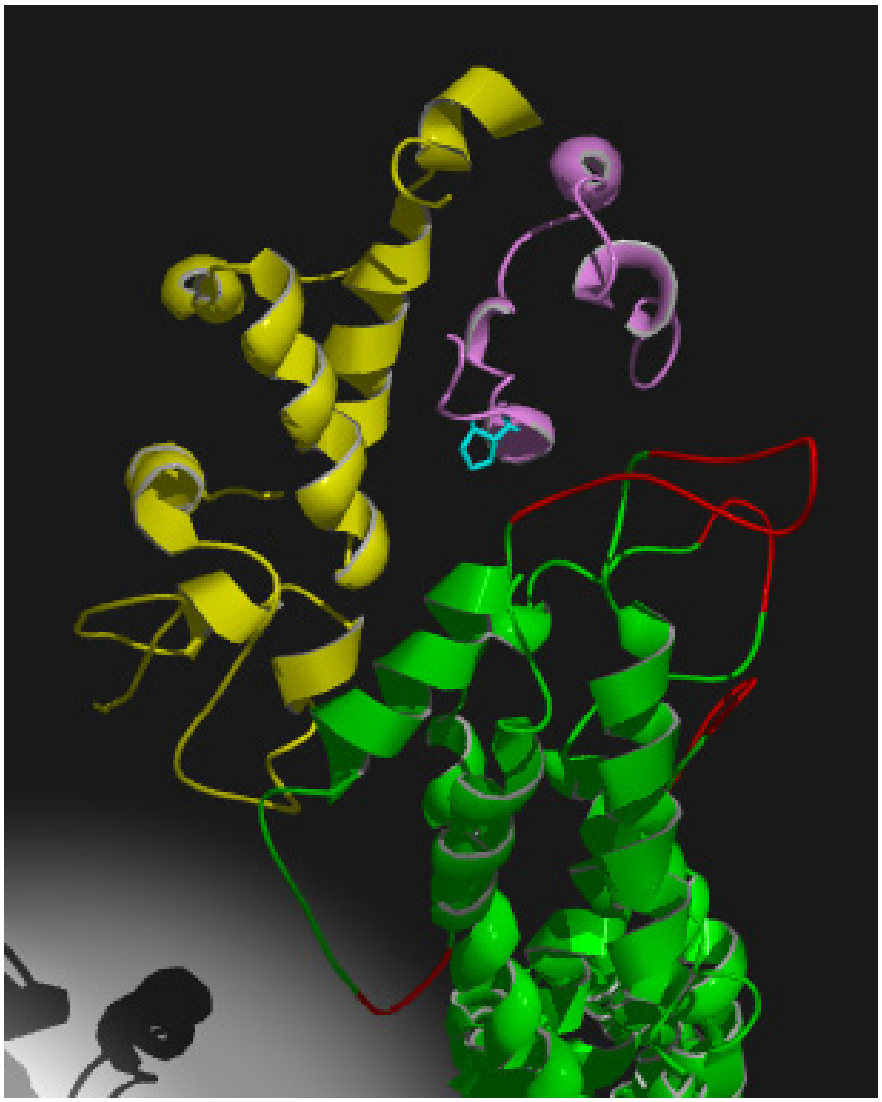
Supplementary Figure 1-3: **The Conserved Proline Residue Determines FBP binding to Cul1p.** (A) Interaction of plasmid derived wild-type Pof1p and Pof1pP114S with Cul1p and Skp1p. Myc-tagged Pof1p was immunoprecipitated and copurification of Cul1p and Skp1p was assayed by immunoblotting with the respective antisera. (B) Interaction of plasmid derived wildtype and mutated FBPs with Cul1p and Skp1p. Myc-tagged FBPs were immunoprecipitated and copurification of Cul1p and Skp1p was assayed by immunoblotting with the respective antisera.



Supplementary Figure 1-4**: Can1p has no influence on the interaction of Cul1p with the CSN**. Interaction of endogenously Myc-tagged Csn5p with Cul1p. Rbx1p was immunoprecipitated in wildtype, *can1* and *csn5* mutants and copurification of Cul1p and Csn5p was assayed by immunoblotting with Cul1p and antiserum or monoclonal Myc antibody.

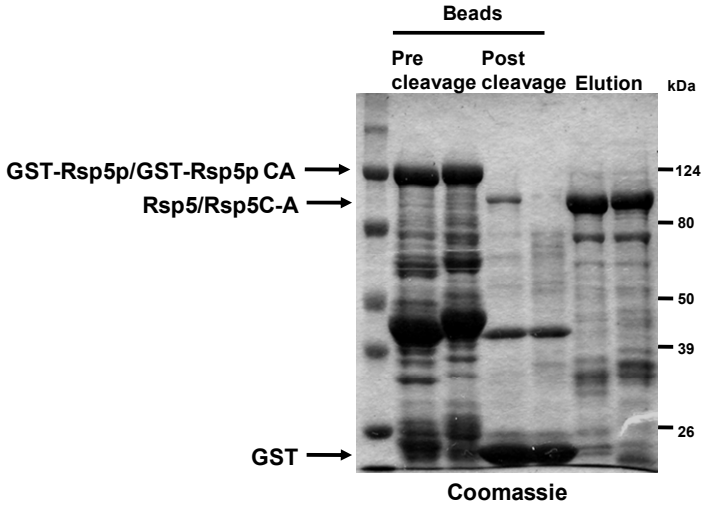


Supplementary Figure 1-5: **CSN and Can1p do not affect the affinity of Cul1p to proline-less FBPs.** Interaction of endogenously Myc-tagged FBPs with Cul1p. FBPs were immunoprecipitated in wildtype, *can1* and *csn5* mutants and copurification of Cul1p and was assayed by immunoblotting with Cul1p and antiserum.

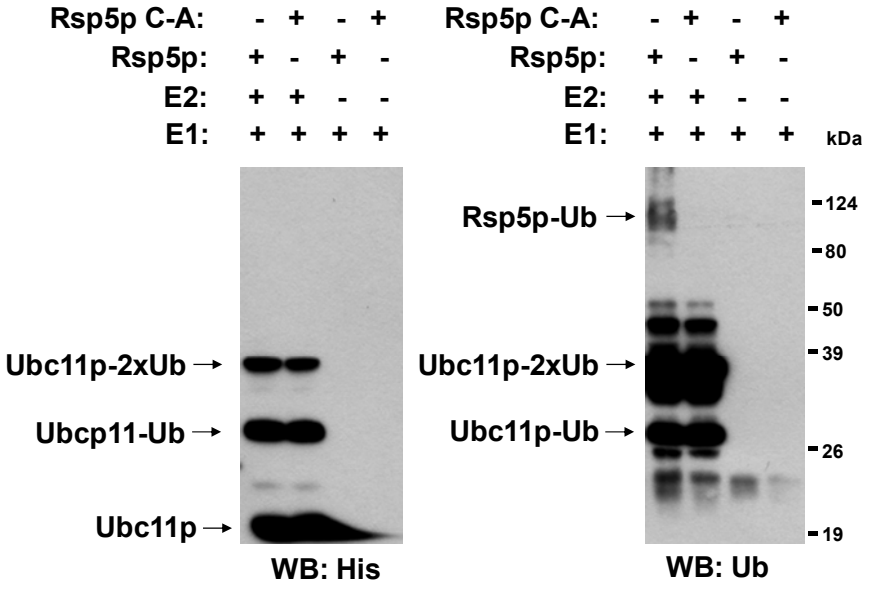


Supplementary Figure 1-6: **Model of the fission yeast Sic1p (yellow) Cul1p (green) F-Box (pink) complex**. The position of the conserved F-Box proline residue is indicated in blue.

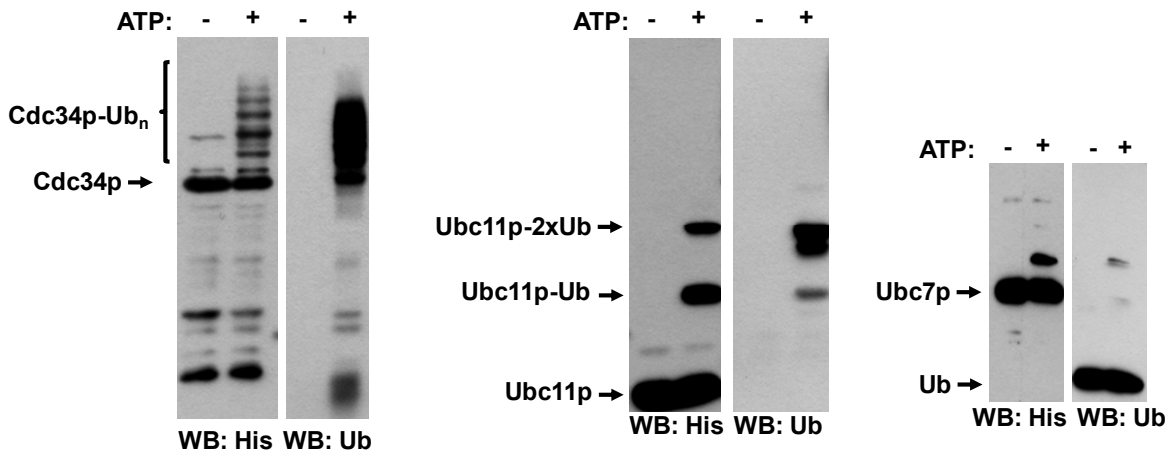
Supplement to Chapter 2



Supplementary Figure 2-1: **Purification of untagged Rsp5p and Rsp5CAp.** Details are described in the methods section. Lanes 1, 2: bead bound GST-Rsp5p and Rsp5pCA. Lanes 3, 4: Beads after protease cleavage. Lanes 5, 6: Cleaved Rsp5p and Rsp5CAp.



Supplementary figure 2-2: **In vitro ubiquitylation activity of untagged Rsp5p but not Rsp5pCA.** In vitro ubiquitylation assay as described in the methods section.



Supplementary Figure 2-3: **In vitro ubiquitin conjugation activity of the indicated E2s.** Assays are described in the methods section

Supplementary Table 2-1: All proteins identified. The numbers after the locus IDs denoe the spectrum count.

ID	Ubc7+	Ubc7-	Ubc8+	Ubc8-		ID	Ubc7+	Ubc7-	Ubc8+	Ubc8-
SPBC17G9.06c	4	2		3		SPAC9.09			4	5
SPCC1672.11c	2	2				SPAC56F8.10	6			
SPAC3F10.03	6	3		2		SPBP19A11.03c	8	8		2
SPAC24C9.12c	7	4	10	6		SPBC16A3.15c	6	2	2	
SPAC22A12.16	5	5	18	15		SPAC1805.04			2	
SPBC1773.10c	2	5	6			SPAC23D3.06c	2			
SPAC10F6.01c	8	7	34	21		SPCC16A11.10c			4	4
SPBC3B8.03	26	7	17	13		SPAC24C9.14	1320	1061	419	369
SPAC1006.07	14	9	11	11		SPAC57A7.04c			13	9
SPAC57A7.12	4	9	4			SPAC26F1.03			3	3
SPAC1F8.07c	9	10	125	78		SPBC16H5.02	17	17	9	14
SPCC584.01c	12	13	46	33		SPBC14F5.04c			21	26
SPBC1734.11	4	18	10	5		SPAC1A6.04c	4		4	2
SPBPB2B2.06c	5	29				SPAC1071.10c			13	12
SPAC589.10c	52	60	25	66		SPCC1020.01c			3	7
SPBPJ4664.04	2			5		SPAC1556.07	5	4		2
SPBC1703.07	5		16	10		SPAC167.01			2	
SPCC1322.04	6		6	12		SPAC4A8.14			2	
SPBC16H5.08c	5		4	5		SPAC19D5.04	2			
SPAC10F6.03c	12		3	3		SPBC1604.21c	8	4		8
SPBC17A3.04c	3		2	2		SPAC11G7.02	2			
SPBP4H10.15	2		4			SPAC4H3.10c	43	41	20	33
SPCC1223.07c	13					SPAC8E11.02c	9	7	5	2
SPAC1F12.07	8					SPAC17A2.13c	5	4		
SPAC1002.12c	5					SPCC757.09c	2			
SPBC18E5.01	3					SPBC18E5.04			9	6
SPCC16A11.16c	3					SPAP7G5.05			9	6
SPAC29A4.15	3					SPCC1183.08c		9	4	2
SPBC13G1.02	3					SPAC26A3.07c	3			
SPAC139.01c	3					SPBC17G9.10	3			
SPBC12C2.11	3					SPCC16C4.13c		2	6	8

SPBC4F6.17c	3					SPCC31H12.04c		2	6	8
SPAC23A1.12c	3					SPAC664.05	7		22	41
SPBC660.06	2					SPCC576.11	2		4	11
SPCC11E10.09c	2					SPAC1783.08c	2		4	11
SPAC1142.04	2					SPBC839.13c			5	
SPCC1672.05c	2					SPAC23A1.11			5	
SPAC17G8.06c	2					SPBC2F12.04	4		9	15
SPAC7D4.12c	2					SPCC364.03	5		10	15
SPBC582.08	2					SPBC11C11.07			2	
SPAC631.02	2					SPBC365.03c	3		5	4
SPAC2E1P3.04	2					SPAC959.08	3		7	4
SPAPB2C8.01	2					SPAC11E3.15			2	
SPBC1289.16c	2					SPAC3G9.03	2			3
SPAC22F8.05	2					SPCC1322.11	2			3
SPAC12B10.01c	2					SPBC685.07c			4	7
SPAPB2B4.04c	2					SPCC74.05			3	4
SPAC24H6.10c			3			SPBC776.11			4	7
SPAC56F8.03			3			SPCC5E4.07	2		5	8
SPACUNK4.16c			3			SPAC9G1.03c			2	3
SPBC11G11.03			2			SPAC17A5.03	6	7	12	15
SPCC970.03			2			SPAPB8E5.06c	6	7	12	15
SPAC17H9.14c			2			SPAC23A1.08c			2	4
SPAC821.05			2			SPCC1322.15			2	4
SPAC27F1.07			2			SPBC1711.06	5		9	12
SPAC869.11			2			SPBP8B7.03c	7	4	16	17
SPAC1B3.13			2			SPAC1687.06c			12	4
SPBC1711.07			2			SPAC3H5.12c	6	15	24	26
SPBC3D6.13c			2			SPBC11C11.09c	6	15	24	26
SPAC1786.02			2			SPCC622.18		3	3	6
SPCC825.01			2			SPAC3H5.07	2			2
SPCC4G3.12c			2			SPBC29A3.04		7	16	15
SPCC1494.10			2			SPAC1F7.13c			9	14
SPBC8D2.06			2			SPBC2F12.07c			9	14
SPAC23A1.17			2			SPBC839.04			9	14
SPAC1F5.11c			2			SPAC4G9.16c			12	12
SPCC622.12c			15	27		SPCC613.06			8	9
SPBC660.16			16	19		SPBC16G5.01	11	7		
SPBC839.16			18	15		SPBC17D11.07c	4	3		
SPAC29A4.02c			15	14		SPAC23G3.11	4			
SPAC589.06c			7	13		SPBC582.07c	3	2		
SPBC19C7.06			16	10		SPCC1682.10	39	38	4	3
SPAC1635.01			3	7		SPCC18.14c	9	17	7	10
SPBC17G9.03c			4	6		SPAC644.15		11	11	5
SPCC1281.06c			8	5		SPBC3B9.13c		10	10	5
SPCC794.07			2	5		SPCP1E11.09c		8	12	6
SPAC26F1.13c			3	5		SPBP8B7.06	12	14	24	18
SPAC890.04c			2	4		SPBC23G7.15c		5	5	
SPBC16G5.05c			4	4		SPAC1071.08	10	11	21	13
SPAC30C2.04			6	3		SPAC31G5.17c			13	14
SPAC9E9.06c			4	3		SPBP22H7.08			13	13
SPCC191.02c			7	3		SPAC13G6.02c			2	
SPAC17A5.15c		2	6	3		SPAC22H12.04c			2	2
SPAC9E9.09c			9	2		SPCC962.04			10	
SPAC19G12.08			7	2		SPBC1685.02c			7	
SPAC25H1.08c			3	2		SPCC1393.03	2			
SPBC1711.05			3	2		SPAC1071.07c	2			
SPCC962.01			4	2		SPBC839.05c			3	
SPAC23C4.06c			2	2		SPCC24B10.09			3	
SPBC32H8.12c	58	40	14	23		SPCC576.08c			2	3

SPBC405.01			2			SPCC576.09			3	
SPCPB16A4.03c			6	5		SPAC8C9.08			10	5
SPAC6F12.10c	7	2	8	2		SPAC328.10c			10	5
SPCC13B11.01	2	3	10	12		SPAC13G6.07c			2	3
SPBC11C11.04c			2	2		SPAC18G6.14c			3	6
SPBC4F6.18c	3			5		SPBC16C6.07c	4			
SPAC4G9.09c	2					SPBC4.07c	3	5		
SPBC56F2.09c	4					SPCC576.10c	4	7		
SPAC1834.02			5			SPCC1682.16	4			
SPAC222.12c			2	10		SPAC3A11.12c	2			
SPAC22A12.15c	2					SPBC23G7.12c	4			
SPAC9.03c			4			SPBC14F5.05c	21	3	43	16
SPCC31H12.08c			2			SPBC31F10.06c	2			
SPBC12D12.03			3			SPAC1556.02c			6	15
SPAC1420.02c			5	3		SPBC215.15			2	
SPBC25H2.12c			2			SPAC57A7.10c			2	
SPAC23H4.09			4	7		SPBC16C6.13c		3	5	4
SPBC336.04			2			SPBC24C6.05	2			
SPAC26A3.05	6					SPAC4D7.01c	2			
SPAC3C7.11c	2					SPBC1709.05	9	23		
SPAC6B12.15			20	16		SPAC1F7.01c			3	
SPCC757.07c			2			SPAC13G7.02c	5	12	7	9
SPAC56E4.04c	12	16	4	5		SPCC1739.13	17	47	26	33
SPAC25G10.07c	2					SPAC664.11	3		11	19
SPBC31F10.11c	2					SPBC27.08c			3	5
SPBP8B7.16c			3	2		SPAC4D7.05			2	
SPBC17D1.06			2			SPCC1795.11	2		5	3
SPCC1223.08c	11		4	5		SPBC32F12.11	388	338	174	140
SPAC1093.06c	2		5	3		SPCC417.08	7	3	63	38
SPAC1002.09c			2	4		SPCC23B6.03c			2	
SPBC428.02c	15		7	5		SPBC25H2.02	9	5	13	9
SPCC794.09c	111	16	309	228		SPBC336.10c		3	2	3
SPAC23A1.10	111	16	309	228		SPCC1919.09			2	
SPBC839.15c	111	16	309	228		SPAC328.03	3			
SPAC513.01c	17	23	65	44		SPCC576.03c	7	11	7	23
SPCP31B10.07	17	23	65	44		SPAC3F10.02c			3	
SPBC25H2.05			2			SPBC1539.09c	5			
SPBC1815.01	3	6	3	7		SPAC19A8.15	7	4	7	6
SPBC215.09c			2			SPBC3F6.03	3	16	4	
SPAC13A11.02c			2			SPAC7D4.07c	14	54	15	17
SPAC926.09c	11	11	38	68		SPBC800.05c	4			2
SPAC4A8.11c	2	4	29	42		SPAC17G6.14c			2	
SPBC19C2.07		3	21	26		SPBP16F5.04	109	2		
SPBC14C8.03			6	3		SPBC119.02	3		6	
SPAC3G9.06	2					SPBC211.07c			89	8
SPAC23D3.04c	3		2	5		SPAC11G7.04	52	60	25	66
SPBC354.12	141	179	79	63		SPAC6G10.11c	52	60	25	66
SPBC26H8.06	3					SPBC337.08c	264	304	125	330
SPBC2F12.14c	11	6	6	4		SPCC1494.05c			3	
SPBC2G5.06c			7	3		SPBC6B1.06c	13	19		
SPCC550.06c			3			SPAC27F1.03c	2		19	26
SPAC12G12.04	5	5	2			SPAC1805.12c	52	60	25	66
SPAC926.04c	4		8	13		SPAC22G7.06c	40	23	7	9
SPAC6G10.08			5	5		SPAC57A10.12c			6	5
SPBP35G2.07	3		7	6		SPAC1002.19	15	6		
SPBC56F2.12	43	38	23	23		SPAC1002.17c	8	3		
SPBC14F5.03c	6	2				SPCC550.14			2	
SPBC18H10.02			7	4		SPAC343.05			2	4
SPAC9E9.03	6		19	13		SPAC7D4.10	2			2

SPBC3E7.16c			5	3		SPAC637.05c			9	17
SPBC887.04c	9	18		3		SPBC16C6.02c	2			
SPAP7G5.04c			3			SPBC1703.10	7	5		
SPAC227.18			5	4		SPAC15A10.04c			3	
SPBC1105.02c	5			3		SPBC1778.01c			3	4

Supplementary Table 2-1

Supplement to Chapter 3

Supplementary Data Files 1-6 can be accessed by web:

Supplementary Data File 3-1: http://www.nature.com/msb/journal/v3/n1/extref/msb4100117-s2.xls
Supplementary Data File 3-2: http://www.nature.com/msb/journal/v3/n1/extref/msb4100117-s3.xls
Supplementary Data File 3-3: http://www.nature.com/msb/journal/v3/n1/extref/msb4100117-s4.xls
Supplementary Data File 3-4: http://www.nature.com/msb/journal/v3/n1/extref/msb4100117-s5.xls
Supplementary Data File 3-5: http://www.nature.com/msb/journal/v3/n1/extref/msb4100117-s6.xls
Supplementary Data File 3-6: http://www.nature.com/msb/journal/v3/n1/extref/msb4100117-s7.xls

ACKNOWKEDGEMENTS

I am grateful to Dieter A. Wolf my mentor, who taught me all I know about *S. pombe* and more over the past few years. I am also grateful to Dieter H. Wolf, who has been my co-advisor in letting me conduct this thesis. I appreciate my best colleagues, Rory Geyer and Susan Wee, who introduced me not only to the American way of life but also many basic concepts of molecular biology and biochemistry. I am grateful to all other member of Wolf lab, I had a great time working with them. I am thankful to Susan Wee who originally started the F-Box project and contributed an enormous collection of yeast strains. I'm also thankful to Kay Hoffmann who was a great help with bioinformatics analysis of F-Boxes. Thanks to Charles Hoffmann in solving the mystery of 5-FOA counter selection and Takashi Toda for providing yeast strains. Further, I'm thankful to Ian McLeoud and John Yatesas for providing resources and expertise in proteomics analysis of SPASS samples as well as Jon Huibregtse for providing plasmids. I am also grateful for fantastic discussion and ideas concerning this project with Cecile Pickart, Brenda Schulman, Dan Finley, Ray Deshaies and Joan and Ron Conaway. I would also like to thank Andreas Houseman and Alexander Ivanov for helping me conducting the proteomics study as well as Janet Leatherwood for microarray analysis. I'm also thankful to Vivian Wood for supported access to *S. pombe* genome data. Last but not least I'm thankful to Elizabeth Heilig for critically reading this document.

EIDESSTATTLICHE ERKLÄRUNG

Hiermit versichere ich, dass ich diese Arbeit selbst verfasst und dabei keine andern als die angegebenen Quellen und Hilfsmittel verwendet habe.

Stuttgart, den 29.10.2007

Michael Schmidt