Characterization of Ubiquitin-mediated Protein Quality Control in the Nucleus

Eric Kyle Fredrickson

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2012

Reading Committee:
Richard Gardner, Chair
Chris Hague
Peter Brzovic

Program Authorized to Offer Degree:
Pharmacology

University of Washington

Abstract

Characterization of Ubiquitin-mediated Protein Quality Control in the Nucleus

Eric Kyle Fredrickson

Chair of Supervisory Committee:

Assistant Professor Richard Gardner

Department of Pharmacology

Due to the importance of proper protein folding, cells have evolved a collection of protein quality control (PQC) mechanisms that maintain overall cellular protein homeostasis. One type of PQC is ubiquitin-proteasome mediated degradation, where misfolded proteins are eliminated from the cell. In the yeast nucleus, San1, a ubiquitin-protein ligase, mediates the degradation of misfolded proteins. The data within this thesis reveals the mechanism by which San1 recognizes misfolded proteins, the abnormal feature San1 recognizes in misfolded proteins and how San1 maintains its stability. First off, San1 directly binds misfolded proteins using N- and C- terminal domains that are largely disordered. This disorder likely allows San1 to adopt different conformations necessary to bind differently shaped misfolded proteins. Additionally, the requirement of molecular chaperones for San1-mediated degradation differs among classes of substrates. Next, the data demonstrates San1 targets exposed hydrophobicity in a diverse collection of misfolded proteins and that recognition can occur with as few as five contiguous hydrophobic residues. Additionally, San1 preferentially targets the hydrophobic residues most likely to lead to protein aggregation (Phe and Ile residues). Further supporting a preference for aggregation prone regions, San1 targets the same hydrophobic regions in misfolded proteins that cause toxicity and aggregation. Lastly, San1 maintains its own stability by lacking Lys residues and highly hydrophobic regions within its disordered N- and C- terminal domains. Addition of either Lys residues or regions of hydrophobicity to these disordered regions led to San1 autodegradation. These features of San1 are highly conserved among fungal homologs, suggesting their importance in maintaining function. Overall, this work elucidates how and why San1 functions as a central regulator of nuclear protein homeostasis.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

PQC Protein quality control

ER Endoplasmic reticulum

ERAD ER associated degradation

GBD Gal4 DNA-binding domain

GAD Gal4 activation domain

IDP Intrinsically disordered protein

BSA Bovine serum albumin

K-D Kyte-Doolittle

ANS 1-Anilinonaphthalene-8-sulfonate

NM Nuclear membrane

SUMEB SDS, urea, MOPS, EDTA, bromophenol blue

NLS Nuclear localization signal

GFP Green fluorescent protein

LIST OF FIGURES

Figure Number

PREFACE

A large portion of the data and text in chapters 1, 2 and 4 of this dissertation is derived from the following publications:

Fredrickson, E. F. and R. G. Gardner. 2012. Selective degradation of abnormal proteins by ubiquitin-mediated protein quality control. *Seminars in Cell and Developmental Biology* 22: in press.

Rosenbaum JC*, Fredrickson EK*, Oeser ML, Garrett-Engele CM, Locke MN, Richardson LA, Nelson ZW, Hetrick ED, Milac TI, Gottschling DE, Gardner RG. Disorder targets misorder in nuclear quality control degradation: a disordered ubiquitin ligase directly recognizes its misfolded substrates. Molecular Cell (2011); 41:93-106.

Fredrickson EK*, Rosenbaum JC*, Locke MN, Milac TI, Gardner RG. Exposed hydrophobicity is a key determinant of nuclear quality control degradation. Molecular Biology of the Cell (2011); 22.

ACKNOWLEDGEMENTS

First off, I want to thank my advisor, Richard Gardner. I was a spring rotation student and he already had three graduate students as a new professor. Luckily, he rolled the dice and let me join the lab. Without this gamble, my graduate career would be vastly different and I cannot thank him enough for it. My favorite traits of Richs' are his tenacity and joy for science. I swear they are contagious. I think I know much of what it takes to be a successful scientist and I owe that to Rich. Lastly, I am grateful for his encouragement and commitment to make me a better scientist.

I also want to thank the present and past members of the Gardner lab for making a great lab environment. Especially Pam Gallagher, Carrie Garrett-Engele, Catherine Konopka, Melissa Locke, Michelle Oeser, Ben Reed, Lauren Richardson and Joel Rosenbaum. A special thanks is due to Joel because we were co-first authors on two papers. It was a great pleasure working with him.

I also want to thank my father, Jim. He is a scientist at a national lab and was a de facto second mentor during my graduate career. No matter how things were going in lab, I could always chat about science and get his advice. I admire his success and unflappability. I owe part of my growth as a scientist to him. Thanks Dad.

Throughout the results sections I will be presenting data collected by others. Chapters II and IV would make little sense without their data and I am grateful for their contributions. In the text I will list them by first name, but here is a list of the figures they contributed (either in full or in part) and their full name:

Rich Gardner: Figure 2.21, 4.30
Joel Rosenbaum: Figure 2.3, 2.4, 2.6, 4.8, 4.9, 4.10, 4.22, 4.23, 4.25, 4.31, 4.33, 4.34
Melissa Locke: Figure 2.11, 2.12, 4.5, 4.6
Michelle Oesar: Figure 2.5, 2.22; Carrie Garrett-Engle: Figure 2.17, 2.18, 2.23, 2.24
Tom Milac: Figure 4.2; Zara Nelson: Figure 2.19

DEDICATION

To my wife, Molly. Her steadfast love and support is amazing.

CHAPTER I: BACKGROUND AND SIGNIFICANCE

To function properly, cells rely on proteins successfully accomplishing specific actions. Fundamental to protein action is the acquisition of a protein's 3-dimensional structure, and thus the proper folding of a cell's protein cohort is critical for cells. However, protein folding often goes awry, either through environmental stress, such as heat, or protein synthesis errors, such as genetic mutation. If not managed properly misfolded proteins can acquire toxic aggregation-prone states that lead to cellular dysfunction. Highlighting the importance of proper protein folding are 35+ human disorders characterized by protein aggregates, which include Parkinson's, Kennedy's and Alzheimer's (Skovronsky et al., 2006). The nucleus appears to be particularly susceptible to protein aggregation, as a majority of aggregation disorders involve aggregates appearing in the nucleus (Woulfe, 2007).

Due to the central importance of protein folding, cells have evolved a collection of protein quality control (PQC) mechanisms that maintain overall cellular protein homeostasis, or proteostasis (Balch et al., 2008). PQC systems can be divided into those that function as either primary or secondary PQC defenses. The cell's primary PQC defenses are directly involved in repairing or removing misfolded proteins. Repair systems are chiefly composed of protein chaperones, whereas removal systems are principally involved in proteolytic destruction either by the proteasome or via autophagy. In many cases, the PQC repair and removal machinery function together in a triage hierarchy that has the potential to determine if a misfolded protein is salvageable and then direct the PQC action towards either repair or removal (Wickner et al., 1999). In contrast, the cell's secondary PQC defenses are adaptive transcriptional responses that balance the primary PQC capacities with the extent of the cellular burden caused by misfolded proteins. They can also reduce global translation as a way to attenuate the production of misfolded proteins. In eukaryotes, PQC systems typically mitigate protein misfolding in a compartment-specific way, with each subcellular compartment housing a distinct set of PQC repair, removal, and adaptive capabilities.

There is now a considerable wealth of information on the different types of cellular PQC, which I will not cover in its entirety in this introduction. I refer the readers to excellent reviews for chaperone-mediated folding and repair and for PQC adaptive

stress responses (Akerfelt et al., 2010; de la Torre-Ruiz et al., 2010; Hartl et al., 2011; Ron and Walter, 2007; Voisine et al., 2010). Herein this chapter, I will introduce select ubiquitin-dependent PQC removal systems that operate in each eukaryotic cellular compartment (Figure 1.1).

Ubiquitin-mediated proteasomal degradation

Protein degradation by the ubiquitin-proteasome system has two primary purposes in the cell: 1) the temporal or spatial regulation of normal proteins, and 2) the removal of misfolded proteins. For each type of degradation, a specific subset of proteins must first be uniquely distinguished from the global pool of cellular proteins and subsequently ubiquitinated. Ubiquitination is canonically achieved via an enzymatic cascade wherein a ubiquitin-protein ligase (E3) partners with a specific ubiquitin-conjugating enzyme (E2) that has been charged with ubiquitin by a ubiquitin-activating enzyme (E1) (Glickman and Ciechanover, 2002). The ubiquitin-protein ligase typically confers substrate specificity within each ubiquitination cascade, either by possessing intrinsic substrate-binding domains or by recruiting auxiliary proteins that impart substrate specificity.

One critical aspect of ubiquitination is the ability of a ubiquitin-protein ligase to distinguish its substrates from other proteins. In regulated degradation of normal proteins, ubiquitin-protein ligases typically recognize degrons, which are small specific linear amino acid sequences located within substrates (Ravid and Hochstrasser, 2008). The degron recognized by the ligase often varies based on the auxiliary proteins that are bound to the ligase. For example, cullin RING ubiquitin-protein ligases (CRLs) are able to recognize many different types of degrons through the directed binding of distinct adapter F-box proteins (Petroski and Deshaies, 2005). The interaction of a particular ubiquitin-protein ligase with different substrate-binding auxiliary proteins allows it to target a larger number of protein substrates. Importantly, these ubiquitin-protein ligases are still only able to target a limited number of substrates due to this sequence specificity requirement.

In contrast to the regulated degradation of normal proteins, substrate recognition in PQC degradation is unlikely to be achieved via the recognition of linear sequence-specific degrons in misfolded substrates for two key reasons. First, any protein has the

capacity to misfold and most proteins in each cellular compartment share little, if any sequence homology. Thus, two different misfolded proteins will not likely possess the same sequence-specific degron. Second, a key purpose of PQC degradation is to destroy structurally abnormal proteins that share the same linear sequence with their normal counterparts. Thus, the features of misfolding recognized by PQC ubiquitin-protein ligases must transcend linear sequence. In the subsequent sections of this chapter, I will introduce the unique set of challenges PQC degradation encounters as well as specific PQC ubiquitin-protein ligases in each cellular compartment.

Endoplasmic reticulum PQC

The unique properties of the endoplasmic reticulum (ER) present numerous challenges for PQC ubiquitin-protein ligases in substrate recognition. First, *de novo* protein folding occurs in the ER and ER PQC ubiquitin-protein ligases must be capable of differentiating between nascent polypeptides that are in the process of folding and proteins that have become misfolded. In addition, the ER is a membrane-bound organelle where structural lesions may be present in transmembrane segments or in regions located on the luminal or cytoplasmic side of the membrane. Thus, ER PQC ubiquitin-protein ligases must have location-specific recognition mechanisms that can sense where the lesions are relative to the ER membrane. Another complication in ER PQC substrate recognition is the presence of disulfide bonds and glycosylation moieties in resident proteins. These posttranslational modifications must be queried for defects by PQC machinery. Lastly, ubiquitin and proteasomes are not present in the ER lumen, so ubiquitination and proteasome degradation can occur only on the cytoplasmic side of the ER membrane. Accordingly, the ER PQC ubiquitin-protein ligases must have a means to recognize substrates on the luminal side of the ER membrane while directing substrate ubiquitination on the cytoplasmic side of the ER membrane.

The first ubiquitin-protein ligase found to play a role in ER PQC degradation is the yeast protein Hrd1/Der3 (Bordallo et al., 1998; Hampton et al., 1996). Hrd1 is an integral ER-membrane protein containing a transmembrane domain that traverses the ER membrane six times and a cytoplasmic RING domain that mediates the transfer of ubiquitin to its substrates (Deak and Wolf, 2001; Gardner et al., 2000; Hampton et al., 1996). Hrd1 specifically targets misfolded ER proteins that present lesions in the ER lumen, known as ER-associated degradation-luminal (ERAD-L), or ER membrane,

referred to as ER-associated degradation-membrane (ERAD-M) (Carvalho et al., 2006). Several mammalian homologs of Hrd1 have been identified – gp78, hHrd1 (synovilin), and Rfp2 (Fang et al., 2001; Kaneko et al., 2002; Kikkert et al., 2004; Lerner et al., 2007; Nadav et al., 2003), with their roles in ER PQC degradation explored to varying degrees.

Another PQC ubiquitin-protein ligase in the ER is Doa10, which was initially identified as the ubiquitin-protein ligase involved in degradation of the MATα2 transcription factor (Swanson et al., 2001). Doa10 has been subsequently shown to target misfolded proteins in the ER (Huyer et al., 2004), specifically with lesions on the cytoplasmic side of the ER, called ER-associated degradation-cytoplasm (ERAD-C) (Carvalho et al., 2006). Doa10 contains a transmembrane domain that traverses the ER membrane fourteen times and a cytoplasmic RING domain (Kreft et al., 2006). In mammalian cells, TEB4 is described as a homolog of Doa10 due to its similar membrane topology (Kreft et al., 2006). TEB4 has recently been shown to be involved in the degradation of the ER-resident type 2 deiodinase (Zavacki et al., 2009).

Cytoplasmic PQC

The cytoplasmic environment presents challenges to PQC ubiquitin-protein ligases that are both similar to and distinct from the ER. Similar to the ER, *de novo* protein folding occurs in the cytoplasm and cytoplasmic PQC ubiquitin-protein ligases must have the ability to differentiate between misfolded proteins and nascent unfolded polypeptides in the process of folding. Unlike the ER, cytoplasmic PQC does not have to contend with multiple classes of substrate lesions that present themselves differently in relation to a membrane. It is possible, however, that there might be specific distinct regions within the cytoplasm in which misfolded proteins behave differently (Kaganovich et al., 2008), and these regions could require different ubiquitin-protein ligases.

The first identified ubiquitin-protein ligase in the cytoplasm that is involved in cytoplasmic PQC degradation is the mammalian protein CHIP. CHIP contains a U-box domain necessary for its ubiquitin-protein ligase activity via interaction with the ubiquitin-conjugating enzyme UbcH5 (Ballinger et al., 1999; Connell et al., 2001; Jiang et al., 2001), and a tetratricopeptide repeat (TPR) domain that is essential for CHIP's interactions with Hsp70 and Hsp90 chaperones (Ballinger et al., 1999; Connell et al., 2001). In *S. cerevisiae* there is no identified homolog to CHIP. Rather, the RING-domain

ubiquitin-protein ligase Ubr1 appears to mediate cytoplasmic PQC degradation (Eisele and Wolf, 2008; Heck et al., 2010; Nillegoda et al., 2010; Prasad et al., 2010). Originally, Ubr1 was characterized for its role in N-end rule degradation, in which certain residues at the N-terminus of proteins serve as degrons (Varshavsky, 2011). Two specific regions in Ubr1 were identified that direct the ubiquitination of substrates containing N-terminal residues of either type 1 (Arg, Lys, or His) or type 2 (Leu, Ile, Phe, Trp, or Tyr) (Varshavsky, 2011). It was recently found that Ubr1 also mediates the PQC degradation of misfolded cytoplasmic proteins (Eisele and Wolf, 2008; Heck et al., 2010; Nillegoda et al., 2010; Prasad et al., 2010) and this is independent of its role in the N-end rule pathway (Heck et al., 2010; Nillegoda et al., 2010). Additionally, yeast Ubr2 has been shown to mediate cytoplasmic PQC degradation, (Nillegoda et al., 2010) as well as human Ubr1 (Sultana et al., 2011).

Mitochondrial PQC

Mitochondria present a specific set of challenges for PQC degradation. One challenge is that mitochondria are bound by both inner and outer membranes that divide the mitochondria into distinct subcompartments. This means that separate PQC degradation pathways will be required to recognize misfolded proteins in each subcompartment. Similar to the ER, ubiquitin and proteasomes are not located inside the mitochondria. Because of this, inner mitochondria PQC degradation appears to be independent of the proteasome, instead relying on AAA-ATPase proteases similar to bacterial systems (Kaser and Langer, 2000). Another challenge for PQC degradation in the mitochondria is that there is a continual production of reactive oxygen species (ROS) as a byproduct of ATP production. An environment with a high level of ROS can lead to increased protein oxidation and misfolding, which the mitochondrial PQC degradation machinery will need to mange robustly.

Two ubiquitin-protein ligases have been identified in mammalian cells that associate with the outer mitochondrial membrane and participate in PQC degradation. First is Parkin, which contains a RING-in between RING domain that is required for its ubiquitin-protein ligase activity (Shimura et al., 2000; Zhang et al., 2000). Inactivation of Parkin is a major cause of juvenile Parkinson's disease (Kitada et al., 1998), but the exact role in PQC is unknown. Secondly is MITOL (mitochondrial ubiquitin-protein ligase), which contains a multi-spanning membrane domain that passes through the

outer mitochondrial membrane four times (Yonashiro et al., 2006). In addition, MITOL possesses a PHD-variant of the RING domain that is responsible for its ligase activity (Yonashiro et al., 2006). To date, there is no identified homolog to either MITOL or Parkin in yeast that functions in mitochondrial PQC.

Nuclear PQC

Unlike the cytoplasm and ER, *de novo* protein folding does not occur in the nucleus. Instead, nuclear proteins are typically translated and folded in the cytoplasm and imported into the nucleus. Thus, PQC degradation systems designed to detect errors in nascent protein folding will likely be absent in the nucleus. The nucleus does have different subdomains such as chromatin, the nucleoplasm, the nuclear membrane, and the nucleolus. Thus, different PQC ubiquitin-protein ligases might be required to manage misfolded proteins that arise in each of these different subnuclear regions.

In *S. cerevisiae*, the nuclear RING ubiquitin-protein ligase San1 mediates the PQC degradation of mutant or misfolded nuclear proteins (Gardner et al., 2005). San1 does not target normal versions of the same proteins (Gardner et al., 2005), establishing a specific role for San1 in PQC degradation. While no mammalian homologs of San1 are known, a *S. pombe* homolog was recently identified and its function in PQC degradation established (Matsuo et al., 2011).

For any PQC degradation pathway, two major questions exist. First, how is the recognition of abnormal proteins mediated? Second, what feature in abnormal proteins is recognized by PQC machinery? As I sought answers to these questions, it became apparent how San1 remains a stable protein is another important aspect of nuclear PQC. The data in the following chapters answer these questions and give rise to new questions necessary to further our understanding of nuclear PQC degradation.

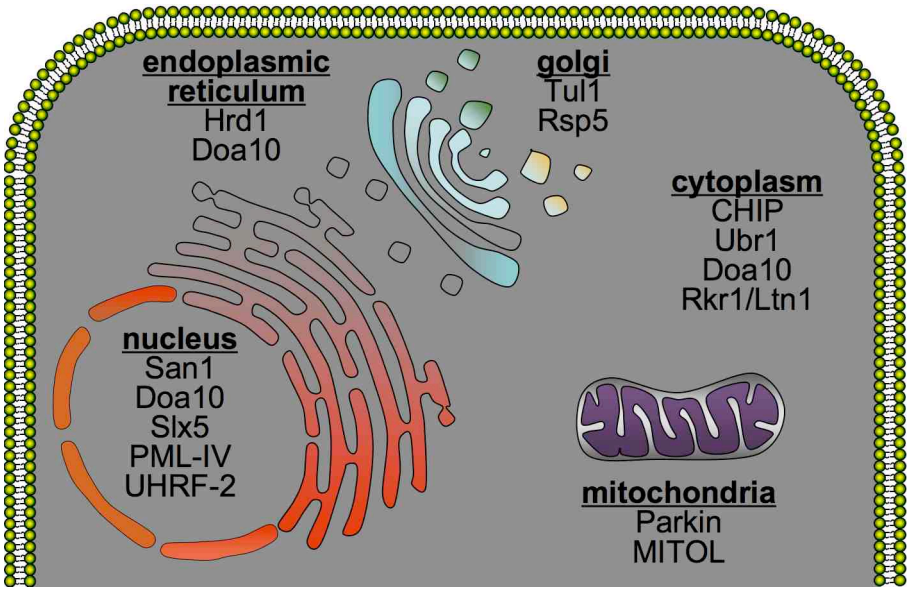


Figure 1.1: PQC ubiquitin-protein ligases in the cell. Schematic representation of the cell with PQC ubiquitin-protein ligases listed in the appropriate cellular compartment. Yeast PQC ligases are listed as the main example, except in the cases where there is no yeast homolog.

CHAPTER II: SAN1 SUBSTRATE RECOGNITION MECHANISM

Introduction on PQC substrate recognition

Any protein has the capability to misfold and each distinct misfolded protein must be recognized by a limited number of PQC degradation systems. As discussed in the previous chapter, each cellular compartment has a unique set of challenges for PQC degradation machinery. These challenges can dictate how recognition of misfolded proteins occurs. For example, in the cytoplasm where *de novo* protein synthesis occurs, substrate recognition must distinguish between nascent unfolded polypeptides and abnormal proteins that have lost their structure. In contrast, a post synthesis compartment such as the nucleus does not encounter nascent polypeptides and thus may have different substrate recognition mechanisms. Here I will highlight what is known about substrate recognition for each PQC ubiquitin-protein ligase introduced in the previous chapter before discussing our experiments characterizing how San1 recognizes substrates.

With the ER PQC ubiquitin-protein ligase Hrd1, substrate recognition varies depending on the type of misfolded protein. However, there are a core group of proteins required to mediate ubiquitination of all Hrd1 substrates. One member of the core Hrd1 complex involved in substrate recognition is Hrd3 (SEL1L in mammals (Lilley and Ploegh, 2005; Mueller et al., 2006)), which is predominantly luminal with a single transmembrane span and a small cytoplasmic region (Gardner et al., 2000; Hampton et al., 1996; Plemper et al., 1999). Hrd3 directly binds misfolded proteins in the ER lumen (Gauss et al., 2006). Due to direct substrate binding, it seems likely that Hrd3 plays a role in Hrd1 substrate recognition. However, Hrd3 also regulates Hrd1 autoubiquitination and stability (Gardner et al., 2000), which may explain the requirement of Hrd3 for substrate degradation. With ERAD-M substrates, recognition appears to be performed directly by Hrd1 (Sato et al., 2009). Mutational analysis of Hrd1's transmembrane domain demonstrated that the degradation of Hrd1's ERAD-M substrates could be disrupted without affecting degradation of other substrates (Sato et al., 2009), suggesting Hrd1 employs multiple substrate recognition mechanisms.

Degradation of ERAD-L substrates requires the Hrd1 core complex and additional cofactors. Hrd3 may be involved in ERAD-L substrate recognition because it is

essential for their degradation and directly interacts with them. Another Hrd1 cofactor required in the degradation of ERAD-L substrates that could recognize substrates is the Hsp70 chaperone, Kar2 (also known as BiP) (Plemper et al., 1997). Kar2 likely couples to the core Hrd1 complex via interaction with a tetratricopeptide repeat (TPR) domain in Hrd3. It is unclear if Kar2 directly binds ERAD-L substrates, or performs functions outside of substrate recognition.

Recognition of misfolded ER proteins often also depends on posttranslational modifications. For instance, certain N-glycan moieties on ERAD-L substrates lead to degradation after their recognition by additional Hrd1 complex factors, such as the lectin Yos9 (Bhamidipati et al., 2005; Buschhorn et al., 2004; Kim et al., 2005; Szathmary et al., 2005) (OS-9 in mammals (Christianson et al., 2008). Yos9 is coupled to the Hrd1 complex via interactions with Hrd3 and Kar2 (Christianson et al., 2008; Denic et al., 2006; Gauss et al., 2006). N-glycan interaction with Yos9 depends on Yos9's mannose-6-phosphate receptor function and is required for degradation of glycolsylated ERAD-L substrates (Hosokawa et al., 2009; Quan et al., 2008; Szathmary et al., 2005). Interestingly, Yos9 is involved in the degradation of certain non-glycosylated substrates (Benitez et al., 2011; Jaenicke et al., 2011), suggesting it might have additional substrate recognition roles.

With the other ER PQC ubiquitin-protein ligase, Doa10, substrate degradation also requires a protein complex. Members of this protein complex that recognize substrates are unclear. Recognition of certain ERAD-C substrates additionally requires Hsp70 (of the Ssa class) and Hsp40 chaperones (Huyer et al., 2004). It remains an open question if these chaperones are responsible for recognizing Doa10 substrates or perform tasks other then substrate recognition.

In the cytoplasm, CHIP ubiquitinates misfolded proteins that are clients of HSP70 and HSP90 chaperones (Murata et al., 2001). Thus substrate recognition is mediated by the chaperones. While CHIP interacts with both Hsp70 and Hsp90 chaperones, data suggests that CHIP has a preference for ubiquitinating Hsp70-bound proteins (Stankiewicz et al., 2010). Although substrate recognition by CHIP is heavily dependent on chaperone recognition, CHIP itself also appears capable of binding misfolded proteins directly (Rosser et al., 2007).

Also in the cytoplasm, Ubr1-mediated PQC degradation requires Hsp70 and Hsp110 chaperones (Heck et al., 2010; Nillegoda et al., 2010; Park et al., 2007; Prasad et al., 2010). Hsp70 functions to keep substrates soluble (Park et al., 2007), but it is not yet clear if these chaperones function in substrate recognition, similar to CHIP or if Ubr1 directly binds misfolded proteins. In support of direct interaction with substrates, Ubr1 is able to ubiquitinate a thermally denatured substrate *in vitro* without the aid of a chaperone (Nillegoda et al., 2010). Addition of Hsp70 increases the Ubr1-dependent ubiquitination of the denatured substrate and robust ubiquitination only occurs when Ubr1 is added during thermal denaturation (Nillegoda et al., 2010). Additional work will be necessary to clarify the role of chaperones and how substrates are recognized in Ubr1-mediated PQC ubiquitination.

To understand substrate recognition in nuclear PQC degradation, we focused on how San1 recognizes its substrates. San1 does not contain any known chaperone-binding motifs, like CHIP, although it might bind substrates through other adaptor proteins. Alternatively, San1 might possess one or more domains that directly bind its substrates. These results demonstrate San1 directly targets substrates using substrate recognition modules imbedded in highly disordered regions outside of San1's RING domain.

San1 directly interacts with substrates

To query how San1 recognizes substrates, we needed a reliable substrate interaction assay. Our first choice for an interaction assay was the two-hybrid assay (Fields and Song, 1989). One reason we chose the two-hybrid is its high throughput capability. For the 2-hybrid to report San1's PQC function accurately, San1 must interact with mutant proteins that it degrades, but not with their normal counterparts. Rich previously found that San1 specifically targets the missense mutant proteins Sir4-9, Cdc68-1, and Cdc13-1 for degradation whereas normal Sir4, Cdc68, and Cdc13 are not degraded (Gardner et al., 2005). We decided to use these substrates to test the ability of the two-hybrid to faithfully reproduce San1's selectivity. To test the interaction between San1 and substrate, Rich fused the Gal4 DNA-binding domain (GBD) to a ubiquitination-deficient version of San1, which contains the C279S substitution in the RING domain (Gardner et al., 2005). I fused the Gal4 activation domain (GAD) to either normal or

mutant Sir4, Cdc68, and Cdc13. I co-expressed the GAD fusions with GBD-San1^{C279S} in cells containing the *HIS3* gene under the control of a Gal4-inducible promoter (Figure 2.1). These cells require a successful 2-hybrid interaction for transcription of the *HIS3* gene and colony growth on media lacking histidine. As anticipated if the two-hybrid reproduced San1 selectivity, I observed an interaction between San1 and mutant Sir4-9, Cdc68-1, and Cdc13-1 but not with normal Sir4, Cdc68, and Cdc13 (Figure 2.2).

While the two-hybrid is evidence of direct interaction between San1 and its substrates, we wanted an assay that would query substrate interaction as well as identify any binding partners of San1 in an unbiased manner. For this, we chose a co-immunoprecipitation (coIP) tandem mass spectrometry approach coupled with *in vivo* formaldehyde crosslinking to identify both stable interactors (such as binding partners) and transient interactors (such as substrates). Rich previously appended a 3xHSV epitope tag to natively expressed San1 in a *sir4-9* strain and found it was functional for Sir4-9 degradation (Gardner et al., 2005). Joel used this strain to immunoprecipitate San1 from cells expressing the known San1 substrate Sir4-9. Joel also used a congenic *sir4-9* strain with untagged San1 as a negative control. All coIPs were performed in triplicate with independent cultures. Joel categorized a protein as interacting with San1 if the summed spectral counts for that protein across the tagged replicates exceeded the summed spectral counts across the untagged replicates by ≥3-fold, and if the spectral counts in the tagged replicates were statistically different from the untagged replicates as measured by a t-test *($p \leq 0.05$)*.

Mass spectrometry identified 328 total proteins, and 25 met the enrichment criteria. We chose to list the proteins with ≥10 summed spectral counts in the tagged strain with their p-values (Figure 2.3). The San1 substrate Sir4-9 had the greatest number of spectral counts in the tagged coIP replicates and showed the greatest enrichment over the untagged replicates. Joel did not observe enrichment for any chaperone in *sir4-9* cells (Figure 2.3, bottom). He also did not identify any other protein with a function consistent with PQC; 8 of the 10 other enriched proteins are nucleolar and function in ribosome biogenesis. Perhaps San1 regulates a subset of nucleolar proteins or, more likely, these proteins are prone to misfolding under the growth conditions used. Because Sir4-9 had the greatest enrichment in the coIP and no other protein identified fit the role of an intermediary, the interaction between San1 and Sir4-9 is likely direct.

We next determined if San1 could directly ubiquitinate a misfolded substrate using an *in vitro* ubiquitination assay in which we generated the misfolded substrate through heat or chemical denaturation. Previous *in vitro* ubiquitination assays using CHIP required chaperones for ubiquitination of denatured luciferase (Murata et al., 2001), indicating the assay is a good proxy for direct or indirect substrate interaction. Joel performed the assay using recombinant San1 purified from *E. coli,* ubiquitin cascade components (Uba1, UbcH5a, and ubiquitin), ATP and luciferase that was either native or denatured by preincubation in 4M urea or at 42˚C. In the presence of all San1-pathway components, he detected a ladder of ubiquitinated species with denatured luciferase and not with native luciferase (Figure 2.4). The laddering did not require chaperones, providing further evidence that San1 binds substrates directly.

We also attempted the *in vitro* ubiquitination assay using purified forms of San1's native yeast substrates. We found, however, that most San1 substrates aggregate during purification from *E. coli*. To circumvent this problem, Michelle developed an "*in coli"* ubiquitination assay in which she co-expressed either San1 yeast substrate Cdc68-1 or Cdc13-1 with San1, Uba1, Cdc34 (San1's E2), and ubiquitin in *E. coli* – a prokaryote that lacks the ubiquitin-proteasome machinery and thus any protein that would function with San1. In this way, San1 has the opportunity to access each substrate as it is made and before it has a chance to aggregate. Using this technique, Michelle found that San1 ubiquitinated Cdc68-1 and Cdc13-1 without the aid of any yeast protein other than the required ubiquitination cascade factors (Figure 2.5). Taken together, the result of each assay supports the conclusion that San1 directly binds its misfolded substrates.

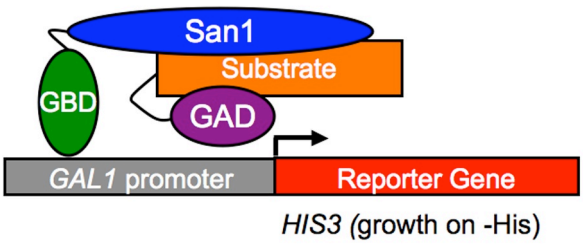


Figure 2.1: Two-hybrid experimental schematic. San1^{C279S} is fused to the GBD and substrates are fused to the GAD. Interaction forms an active Gal4 transcription factor and drives expression of HIS3.

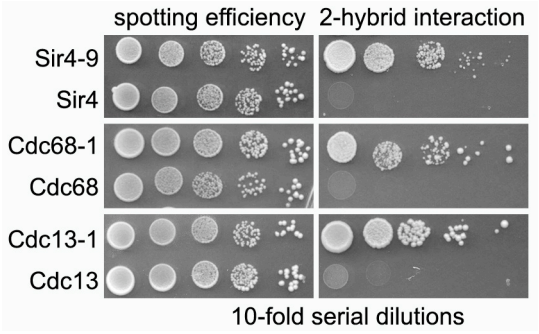


Figure 2.2: Two-hybrid interaction between San1 and its missense substrates. Cells expressing GBD-San1^{C279S} and GAD-substrate fusions were spotted onto media plus or minus histidine to measure spotting efficiency and two-hybrid interaction.

Protein	3HSV-San1 replicates				San1 replicates				ratio	p value
	1	2	3	sum	1	2	3	sum		
Sir4-9	43	53	56	152	0	1	0	1	152.0	0.00288
San1	48	36	44	128	1	0	0	1	128.0	0.00324
Rpa135	10	15	18	43	1	4	0	5	8.6	0.00853
Mdn1	8	6	6	20	0	0	0	0	>20	0.00492
Lys21	5	5	5	15	0	1	0	1	15.0	0.00253
Nop58	4	5	6	15	2	3	0	5	3.0	0.02095
Rrp5	4	7	4	15	0	1	0	1	15.0	0.01628
Rrp12	5	4	5	14	0	0	0	0	>14	0.00253
Dbp2	4	6	3	13	0	1	0	1	13.0	0.01633
Sub2	3	5	5	13	2	1	1	4	3.3	0.01429
Dbp3	4	3	4	11	0	0	0	0	>11	0.00408
Nug1	4	3	4	11	0	1	0	1	11.0	0.00105
Chaperones										
Ssa1/2	30	39	40	109	54	43	30	127	0.9	0.24634
Hsp82	12	18	20	50	19	23	15	57	0.9	0.26127
Hsp104	3	4	4	11	3	3	1	7	1.6	0.08671
Sse1	1	4	7	12	8	10	3	21	0.6	0.16598
Ydj1	1	3	3	7	1	2	0	3	2.3	0.10326

Figure 2.3: San1 proteomics show enriched substrates, not chaperones. 3HSV-tagged San1 was immunoprecipitated and analyzed by mass spectrometry. Proteins with ≥ 3-fold enrichment compared to untagged control are shown in the top panel. Chaperones known to function in PQC degradation are shown at the bottom. Numbers are spectral counts for each protein in each replicate.

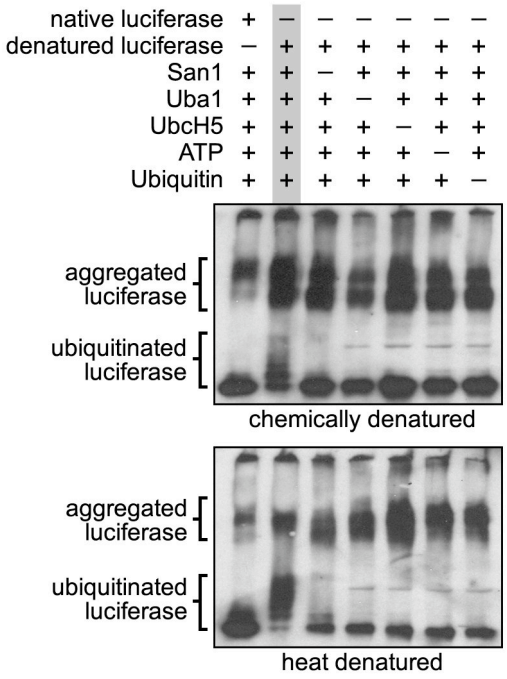


Figure 2.4: San1 does not require chaperones in an *In vitro* ubiquitination assay. Chemical or heat denatured luciferase was added to a reaction containing the San1 ubiquitination cascade. The resulting reactions were probed on a western blot with anti-luciferase antibody. Aggregated or ubiquitinated species are noted on the left. Annotations above the lanes indicate reagents present in the reaction.

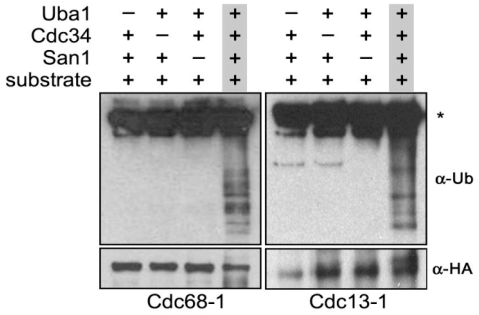


Figure 2.5: San1 pathway selectivity reconstituted in *E. coli.* 2xHA-tagged substrates were immunoprepcipated from *E. coli* cells expressing the San1 ubiquitination cascade and analyzed by western blot. Blots were probed with anti-ubiquitin antibodies to assess substrate ubiquitination and anti-HA antibodies to assess substrate immunoprecipitation. Annotations above the lanes indicate components expressed. The antibody band is marked by an asterisk.

San1 is largely unstructured

To understand how San1 might directly bind its substrates, we examined San1's sequence and domain composition. The only region of San1, a 610 residue protein, with a known function is the RING domain, which spans from residue 165 to residue 280. The RING domain confers ubiquitin ligase activity through interaction with a ubiquitin conjugating enzyme (Dasgupta et al., 2004; Gardner et al., 2005). Because the RING domain is the active site of an E3, we deemed it unlikely to also be the region for substrate interaction. The regions N- and C-terminal to the RING domain do not contain any known domains that could inform us of substrate binding or another function. In addition to lacking identifiable domains or motifs, the N- and C- terminal regions of San1 are predicted to contain little secondary structure using JPRED (Cole et al., 2008) (Figure 2.6).

Due to the lack of identifiable structural features in San1 outside of the RING domain, we predicted San1 might be an intrinsically disordered protein. To test our hypothesis, we applied multiple protein disorder prediction methods to San1's sequence. We first used the Predictor of Naturally Disordered Regions (PONDR) VL-XT algorithm, which predicts disorder based on sequence attributes typically found in regions that are absent in x-ray and NMR structures (Romero et al., 2001). Confirming our hypothesis, PONDR predicted San1 to contain large regions of disorder (Figure 2.7). The disorder within the RING domain contains the nuclear localization signal, which needs to be accessible for nuclear import. Other disorder prediction methods, such as FoldIndex (Prilusky et al., 2005) and IUPred (Dosztanyi et al., 2005) confirmed the PONDR prediction of disorder in San1 with a similar topology given by PONDR (Figure 2.8).

Because of the consistent prediction of greater than 60% disorder in San1, I decided to experimentally test for disorder in active San1. I first used size exclusion chromatography to examine if San1 had characteristics of a disordered protein. Size exclusion chromatography is used to characterize disorder because intrinsically disordered proteins have larger hydrodynamic volumes than structured proteins, which causes them to migrate at a rate indicative of a higher molecular weight than their actual mass (Receveur-Brechot et al., 2006). It was previously found that San1 purified from yeast migrates at ~500 kDa in gel filtration rather than its calculated size of ~66 kDa and it was proposed this was due to San1 existing in a multi-protein complex in yeast

(Dasgupta et al., 2004). Due to the disorder prediction and previous size exclusion data, I repeated the gel filtration assay using San1 purified from *E. coli* under non-denaturing conditions and found that San1 migrated at ~500 kDa, while the similarly sized bovine serum albumin (BSA) migrated at ~75 kDa (Figure 2.9). Because of the absence of yeast proteins, this result indicates San1 is either disordered or runs as an oligomer. To examine if San1 is a multimeric protein, I co-expressed differently tagged versions of San1 in *E. coli* and found that neither version co-purified with the other (data not shown), indicating that San1 does not self-associate. Also supporting the lack of multimers, Rich and I both found San1 does not interact with itself in a two-hybrid (data not shown). It is possible that San1 oligomerizes transiently and future experiments will need to address this possibility.

Due to the chance that San1 oligomerizes, I wanted to examine the lack of structure in San1 with another assay. For this, I used limited proteolysis, which takes advantage of the fact that protease cleavage sites are usually inaccessible in structured regions but readily accessible in regions of disorder (Fontana et al., 2004). I exposed *E. coli* purified San1 and BSA to trypsin or thermolysin for 0, 2, 10, or 45 minutes and analyzed their proteolysis by SDS-PAGE followed by silver staining. I found San1 was completely digested at the earliest time points whereas BSA resisted digestion throughout the time course of protease treatment (Figure 2.10). I also observed that intact San1 migrated at ~100 kDa on the acrylamide gel, larger than its calculated ~66 kDa mass. Intrinsically disordered proteins often show this behavior due to a low content of hydrophobic residues (Tompa, 2002).

Lastly, I analyzed San1's secondary structure by circular dichroism (CD), which is a technique used to determine the α-helical, β-sheet, and random coil content of proteins (Greenfield, 2006). The CD spectrum of San1 had a minimum close to 200 nm (Figure 2.11, blue), which is characteristic of a protein with mostly random coil content and is typical for intrinsically disordered proteins (Receveur-Brechot et al., 2006). By contrast, the CD spectrum of BSA had minima at 208 and 222 nm, which is characteristic of a protein with mostly α-helical content (Figure 2.11, yellow). The San1 spectrum had a small valley at ~224 nm indicating San1 possesses some α -helical content. This is consistent with the secondary structure prediction using Jpred (Cole et al., 2008), which predicts San1 possesses ~15% α-helical and ~5% β-sheet content (Figure 2.6).

Altogether, the experimental data confirmed the theoretical predictions that San1 is intrinsically disordered.

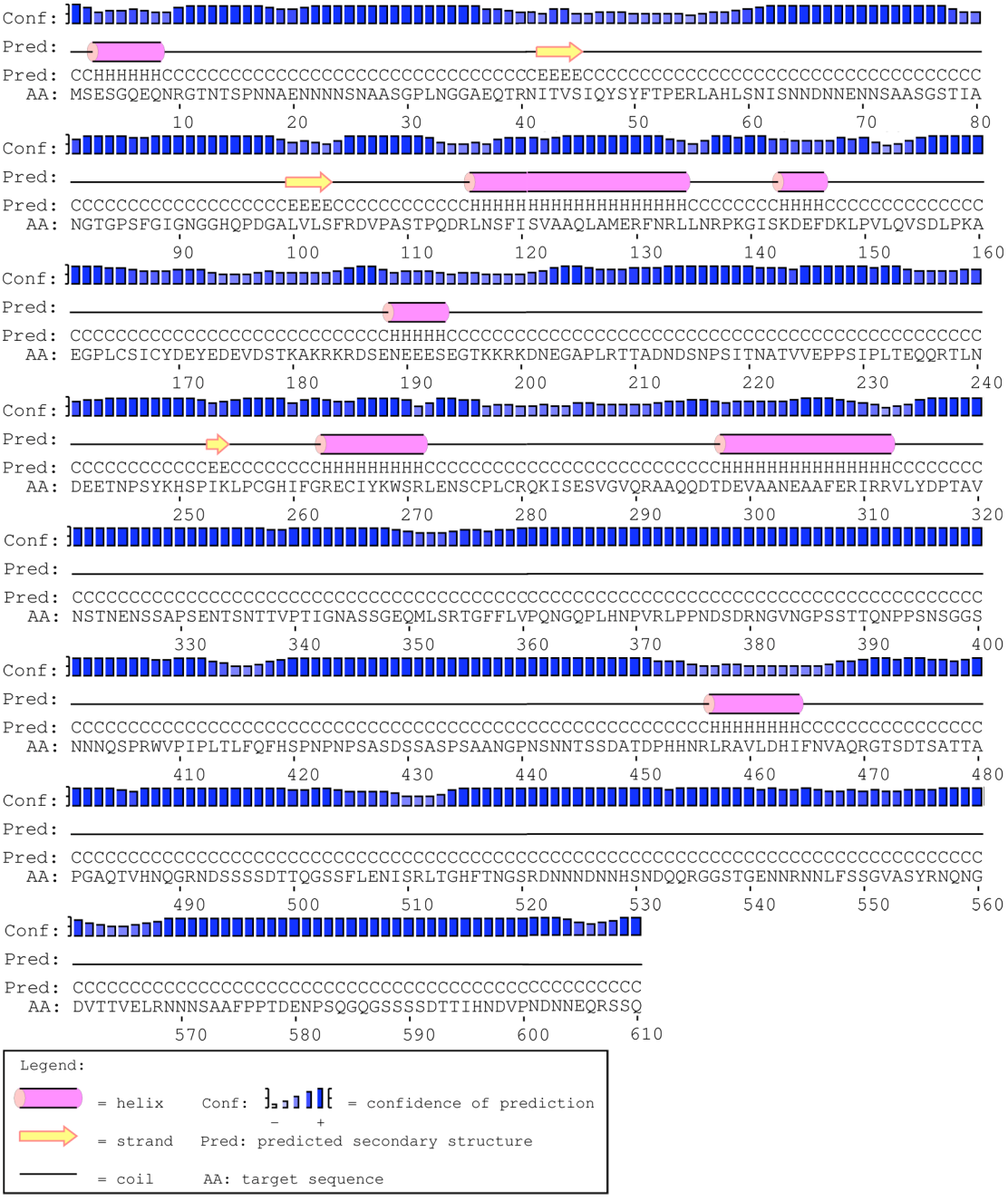


Figure 2.6: Predicted secondary structure of San1. Secondary structure predicted by Jpred (http://www.compbio.dundee.ac.uk/www-jpred/) is indicated below each residue in San1's sequence.

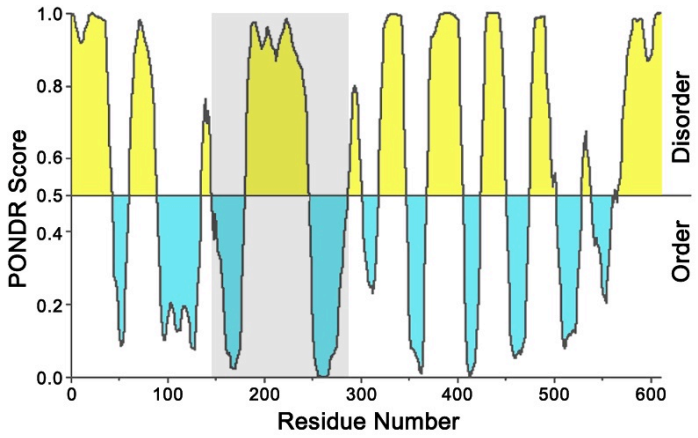


Figure 2.7: Disorder prediction of San1 via PONDR. Disorder, as predicted by PONDR (http://www.pondr.com/), is plotted on the y-axis across San1's sequence. Yellow indicates disorder and blue indicates order. The gray box marks the RING domain.

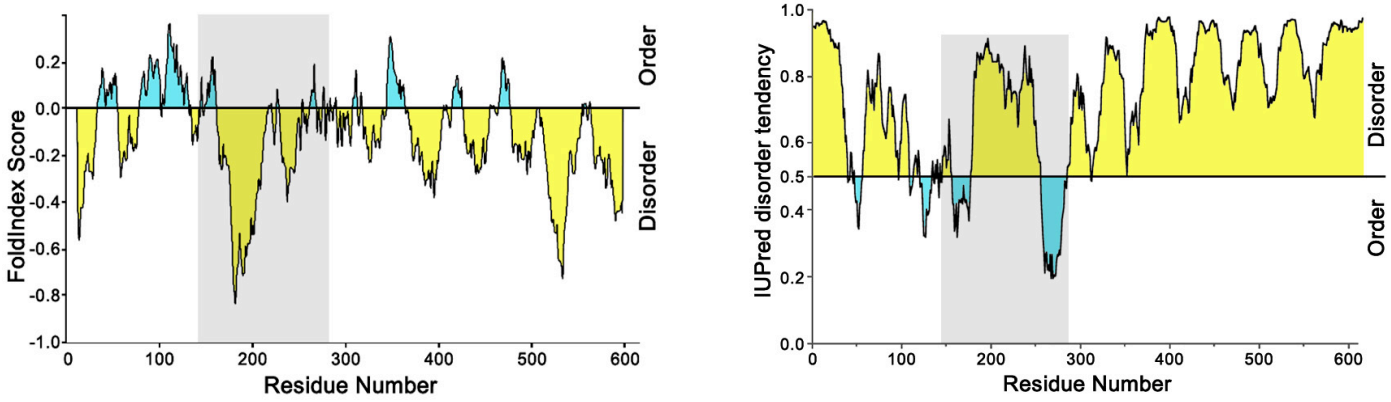


Figure 2.8: FoldIndex and IUPred predictions of San1 disorder. FoldIndex (http://bip.weizmann.ac.il/fldbin/findex) and IUPred (http://iupred.enzim.hu/index.html) also predict San1 to be disordered. Yellow indicates disorder and blue indicates order. The gray box marks the RING domain.

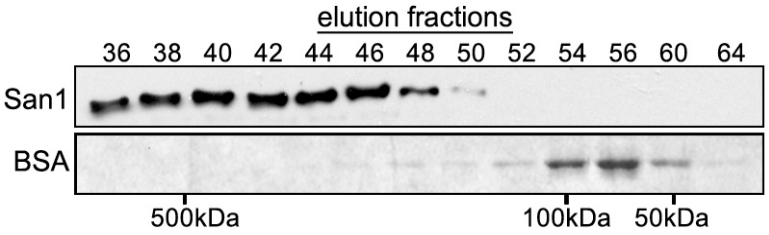


Figure 2.9: Gel filtration of recombinant San1 and BSA. Recombinant San1 or BSA was loaded onto a Superdex 200 column and eluted with 50mM NaCl, 15mM Na_2HPO_4, pH7.3. In the top panel, a western blot of San1 was probed with anti-HSV antibodies. At bottom, a gel containing BSA was stained with Coomassie.

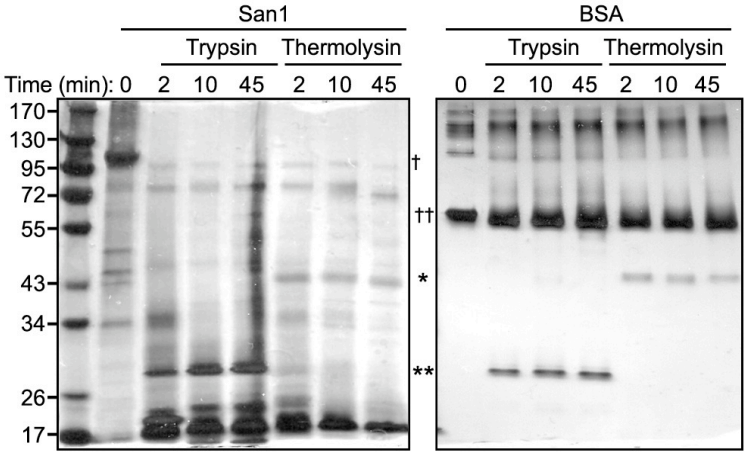


Figure 2.10: Limited proteolysis of recombinant San1 and BSA. Purified San1 or BSA was incubated with trypsin or thermolysin for the indicated times. Proteins were separated by SDS-PAGE and visualized by silver stain. The locations of trypsin (*), thermolysin (**), San1 (†), and BSA (††) are marked.

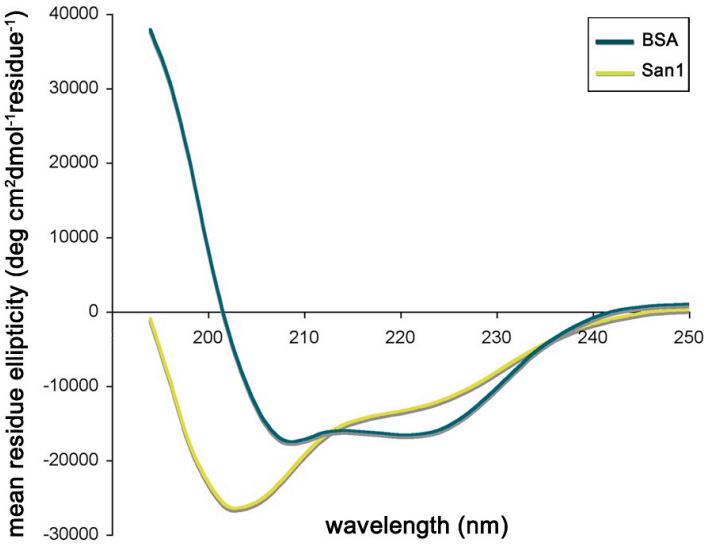


Figure 2.11: CD spectra of recombinant San1 and BSA. CD spectra of San1 (yellow) and BSA (blue) were recorded at 0.2 mg/ml for San1 and 0.15 mg/ml for BSA in 50mM NaCl and 15mM Na_2PO_4, pH7.3 at 25˚C.

San1 contains binding sites within its disordered N- & C-terminal regions

Many proteins contain large regions of disorder and this lack of structure gives proteins specific properties. One such property is allowing a protein to have a wide range of conformational states, which confers the ability to bind many differently shaped molecules (Dunker et al., 2008). San1's intrinsic disorder in the N- and C-terminal regions could provide San1 the conformational flexibility necessary to bind the broad variety of misfolded nuclear proteins. While previously uncharacterized in PQC degradation, there is precedent for this type of recognition in repair PQC mediated by chaperones. For example, intrinsically disordered regions in in the N-terminus of small heat shock proteins (sHSPs) are required for the binding of misfolded substrates (Haslbeck et al., 2004; Jaya et al., 2009; Reichmann et al., 2012; Stromer et al., 2004).

Because sHSPs use disorder to bind misfolded proteins, we hypothesized that San1 uses its disordered N- and C-terminal regions to bind its misfolded substrates. If this was the case with San1, we reasoned that San1's intrinsic disorder should be maintained in San1 homologs. We therefore examined San1's homologs from different *Saccharomyces* species. To test their ability to complement *S. cerevisiae SAN1* function, Melissa placed each homolog *SAN1* into a *san1*Δ *S. cerevisiae* strain and examined the degradation of Sir4-9 and temperature sensitivity of *cdc68-1*. She observed that *S. bayanus, S. mikatae* and *S. paradoxus SAN1* genes each function equivalently to *S. cerevisiae SAN1* in both degradation (Figure 2.12) and temperature sensitivity (Figure 2.13). Each homolog also shared the same predicted disorder topology as *S. cerevisiae* San1 (Figure 2.14), indicating disorder is likely important for San1 function.

Another property of disordered proteins is that they often contain small ordered regions, which are important for binding of other proteins (Dunker et al., 2008). We predicted that if San1 contains small ordered regions that are important for substrate interaction, these regions would be conserved. To test our hypothesis, we compared *S. cerevisiae's* San1 sequence with the homologs' sequences using ClustalW (Thompson et al., 1994). We observed discrete identical segments distributed along San1's length that suggest conserved function (Figure 2.15, yellow highlighting). We also noticed that the pattern of alternating identical and non-identical segments in the ClustalW alignment mirrored the pattern of alternating order and disorder predicted in San1's PONDR VL-XT topology. PONDR VL-XT predicts binding sites as dips in disorder (Romero et al., 2001),

and such dips are readily observed in San1's PONDR VL-XT profile (Figure 2.7, blue regions). These dips map to the identical segments identified by ClustalW. Using ANCHOR, another predictor of binding sites within disordered regions (Dosztanyi et al., 2009), we obtained a similar result as with PONDR with the predicted binding sites mapping to the identical segments (Figure 2.16). Altogether, the conservation and predications suggest San1 contains binding sites within its disordered N and C termini.

To test our hypothesis, we decided to conduct a systematic deletion analysis of San1 and use the 2-hybrid assay to identify actual substrate-binding sites in San1. Prior to conducting the analysis, we envisioned two models for how San1 engages substrates. The first model is based on the conventional view of protein-protein interactions. In this model, San1 binds all substrates identically using the same defined substrate-binding site, similar to how Hsp70 chaperones bind substrates (Mayer and Bukau, 2005). If this model were correct, particular deletions in San1 would affect each substrate's interactions with San1 equally and identically. The second model combines San1's disorder and binding-site predictions with ideas of how the plasticity of disorder might be used to accommodate the binding of many differently shaped molecules. In this model, San1 binds each substrate differently using a distinct combination of substrate-recognition sites dispersed throughout its disordered regions. If this model were correct, deletions in San1's disordered regions would have varying effects on each individual substrate's interactions.

To identify substrate-binding sites and resolve which model is correct, Carrie made 20 small deletions in San1 that spanned the length of each small identical and non-identical segment identified by ClustalW. Carrie examined the steady-state protein levels of each GBD-San1 deletion mutant to ensure that any changes in interactions were not due to reduced San1 levels (Figure 2.17). I then assessed how each deletion affected San1's ability to interact with Sir4-9, Cdc68-1, and Cdc13-1 in the 2-hybrid assay. To best visualize the results of the 2-hybrid interaction data, we used a heat map where the green color indicates the degree to which each San1 deletion affected the substrate interaction (Figure 2.18). As expected, the RING domain deletion (residues 165-285) had little effect on substrate interactions consistent with its primary function being interaction with the ubiquitin conjugating enzyme. In agreement with the second proposed model, no single interaction profile emerged for San1-substrate interactions;

multiple deletions in the N- and C-terminal disordered regions differentially affected each substrate interaction with San1 having a distinct interaction profile with each substrate. The identical, ordered segments were generally more important for substrate interactions particularly those spanning residues 295-319, 340-368, 402-422, 454-473, and 545-567. Their overall importance is consistent with these regions serving as conserved substrate-recognition modules.

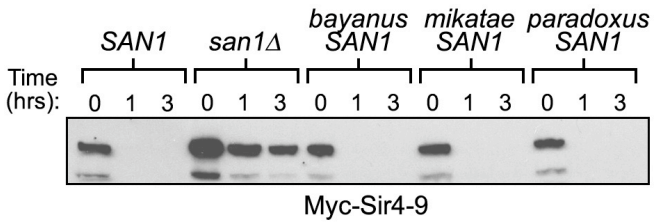


Figure 2.12: San1 homolog complementation of Sir4-9 degradation. Cycloheximide-chase assay of cells expressing Myc-Sir4-9 and San1 homologs was performed to assess complementation. Western blots were probed with anti-Myc antibodies. Time after cycloheximide addition is indicated above each lane.

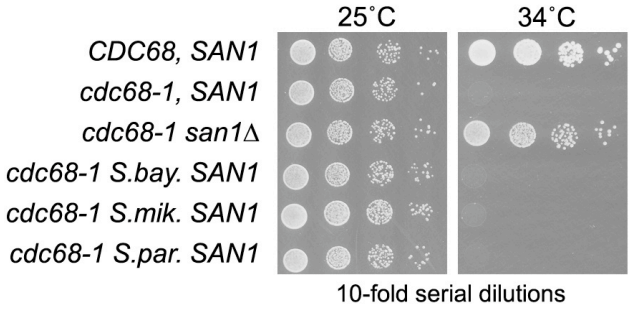


Figure 2.13: San1 homolog complementation of *cdc68-1* temperature sensitivity. *cdc68-1* cells expressing San1 homologs were grown to test complementation. Ten-fold dilutions of cells were incubated at permissive (25˚C) and restrictive (34˚C) temperatures.

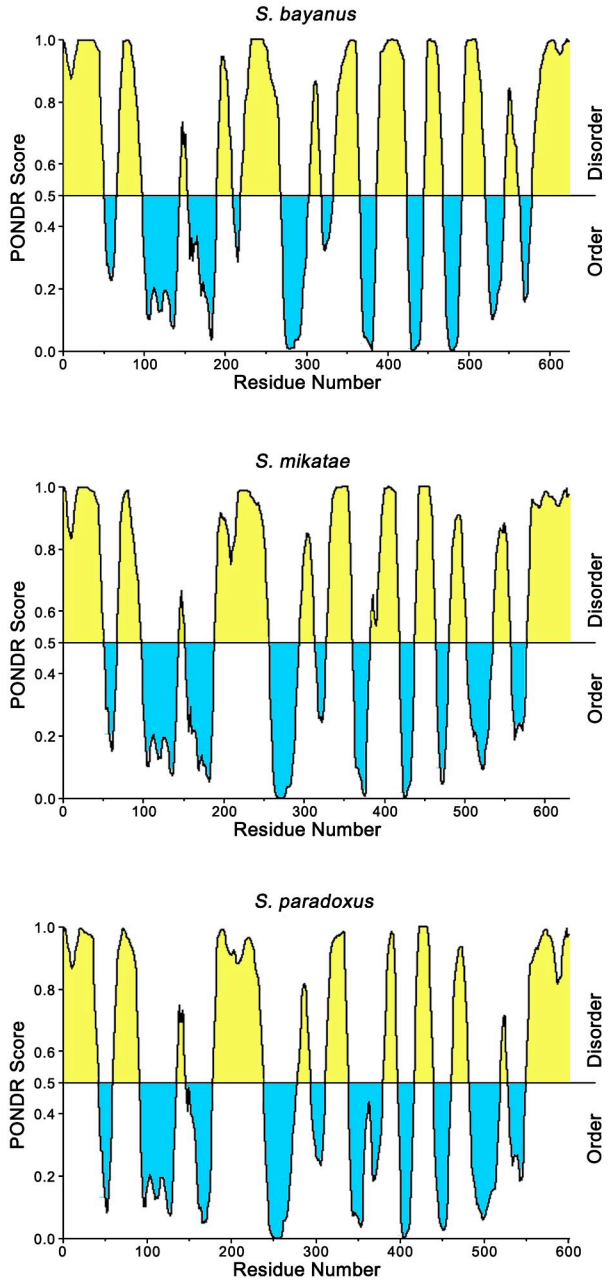


Figure 2.14: Disorder prediction of homologs via PONDR. Disorder, as predicted by PONDR (http://www.pondr.com/), is plotted on the y-axis across the San1 sequence from *S. bayanus, S. paradoxus*, and *S. mikatae*. Yellow indicates disorder and blue indicates order.

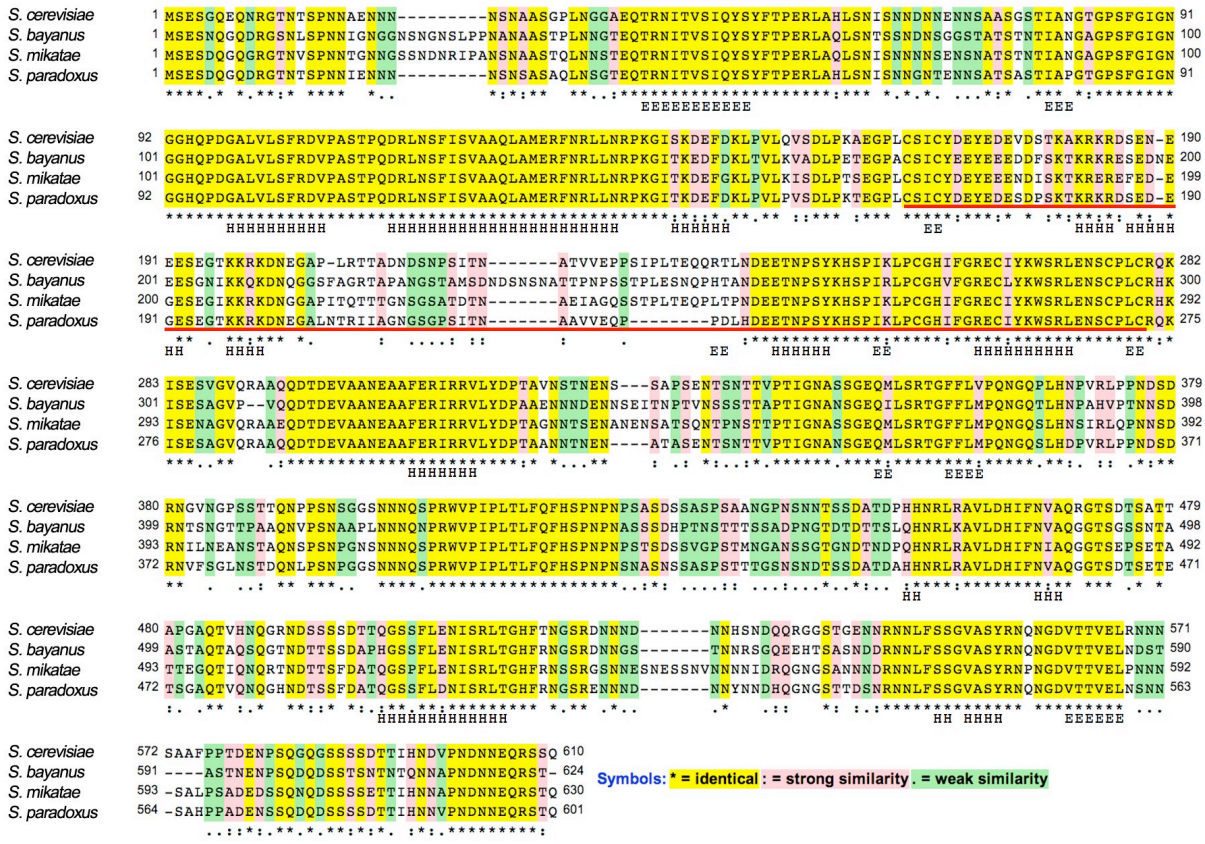


Figure 2.15: CLUSTALW alignment of San1 and its homologs. ClustalW alignment of *S. cerevisiae* San1 with its homologs from related *Saccharomyces* species was performed using Fungal Alignment viewer in the Yeast Genome Database (http://www.yeastgenome.org/cgi-bin/FUNGI/showAlign). The red bar underlines the RING domain.

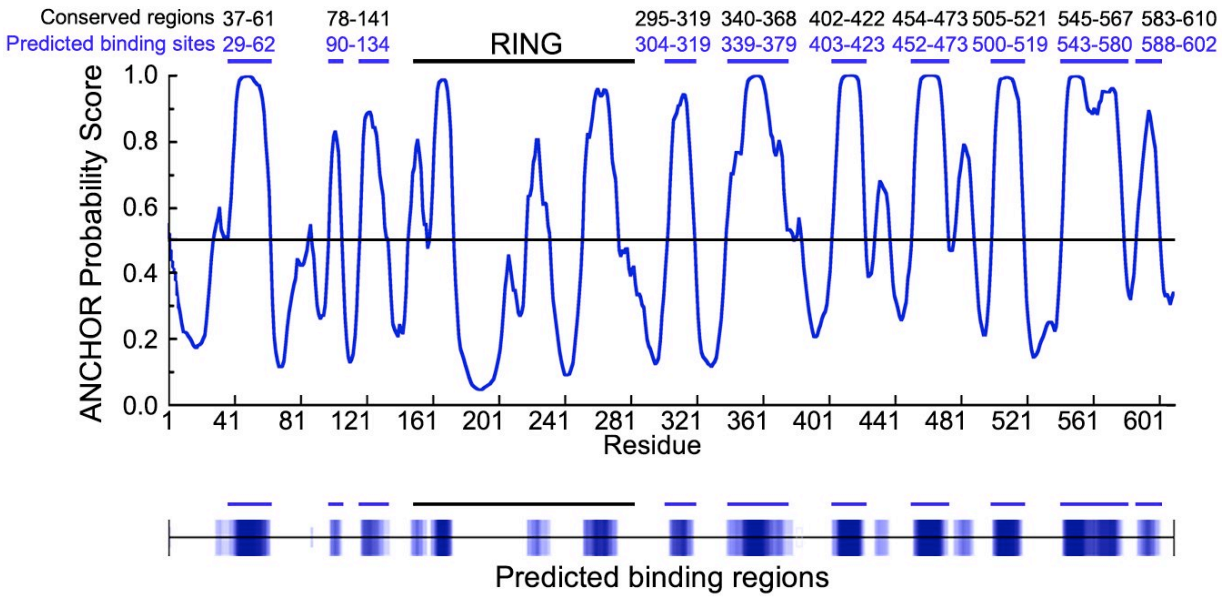


Figure 2.16: Prediction of San1 binding sites via ANCHOR. ANCHOR prediction of binding sites (http://anchor.enzim.hu/) within San1. Blue bars at the top mark the predicted binding sites. Black bars marked conserved regions.

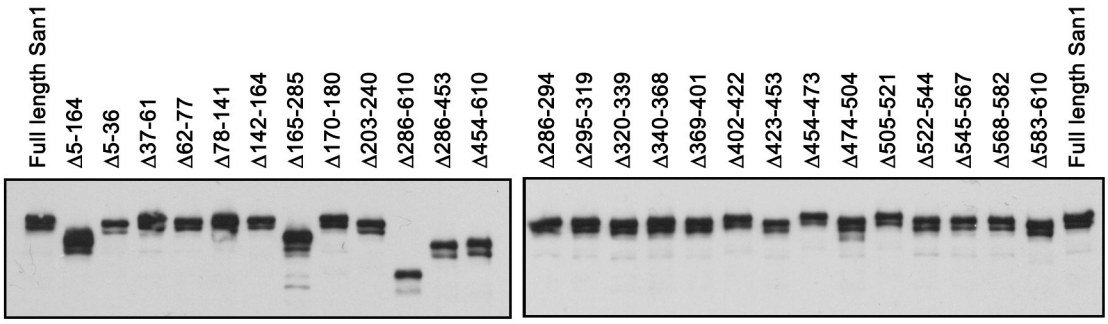


Figure 2.17: Steady state levels of GBD-San1 deletions. 2-hybrid strains expressing the indicated GBD-San1 fusions were lysed and separated by SDS-PAGE. Western blots with probed with anti-HSV, which each GBD fusion was tagged with.

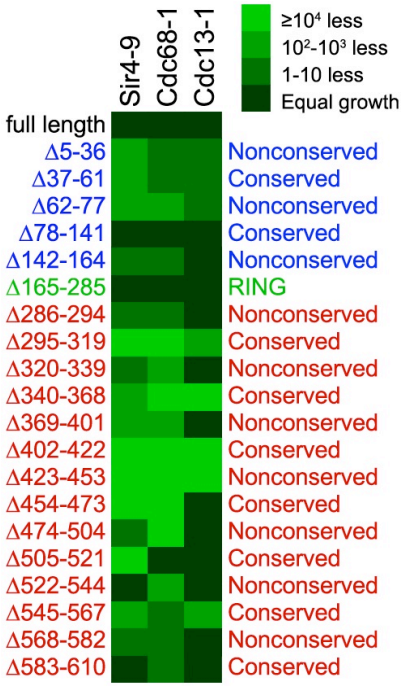


Figure 2.18: San1 binds missense substrates using its disordered regions. Cells expressing the indicated San1 deletion (top) and GAD fusion (left) were spotted media plus or minus histidine to measure spotting efficiency and the 2-hybrid interaction. Interaction dependency is indicated by green color intensity.

Identification of additional substrates of San1

The 2-hybrid deletion analysis provided compelling evidence that San1 interacts in distinct ways with each substrate through substrate-recognition modules located in the disordered regions. However, this is a small set of substrates, all of which are missense mutants, and a larger collection of substrates would allow us to test the hypothesis more comprehensively. Because San1 interacts with substrates in the 2-hybrid, we decided to use a 2-hybrid genetic selection to discover additional substrates and potential binding partners. Zara transformed a plasmid library made from yeast cDNA fused to the GAD into cells expressing GBD-San1^{C279S} with the *HIS3* gene under control of a Gal4-inducible promoter. From 4x10^{6} transformants, 98 grew into colonies on selective media. Of these, 32 different cDNAs were represented, 28 of which retested positive for an interaction with San1 (Figure 2.19).

None of the San1 interactors has a known function related to nuclear PQC degradation (Figure 2.19). Thus, we suspected they aren't partners of San1 but are substrates due to some structural abnormality. We analyzed the actual sequences of the proteins fused to the GAD to see if they could be considered abnormal and found that 22 are truncated, 5 are encoded by cDNAs in their reverse orientation thus generating small peptide sequences, and 1 is full length but with an added 5' non-coding sequence (Figure 2.19). Because most of the proteins identified were truncated, I retested the 2-hybrid interaction with full-length versions of these proteins and observed that San1 did not interact with the full-length versions (Figure 2.20), again demonstrating San1's specificity for abnormal proteins in the 2-hybrid assay.

To verify the interactors are San1 substrates, Rich analyzed the degradation of each GAD fusion in *SAN1* and *san1*Δ cells. He found 25 of 28 required San1 to be fully degraded: 14 underwent primarily San1-dependent degradation, 11 were subject to degradation that was partially dependent on San1, and 3 were stable when San1 was present (Figure 2.21). The three proteins that were stable were not required for the degradation of other substrates (data not shown), and thus the reason for their interaction remains unknown. Michelle also examined San1-dependent ubiquitination of select GAD-fusion substrates using the *in coli* assay and found they were ubiquitinated in a San1-dependent manner (Figure 2.22), indicating San1 directly targets these substrates as well.

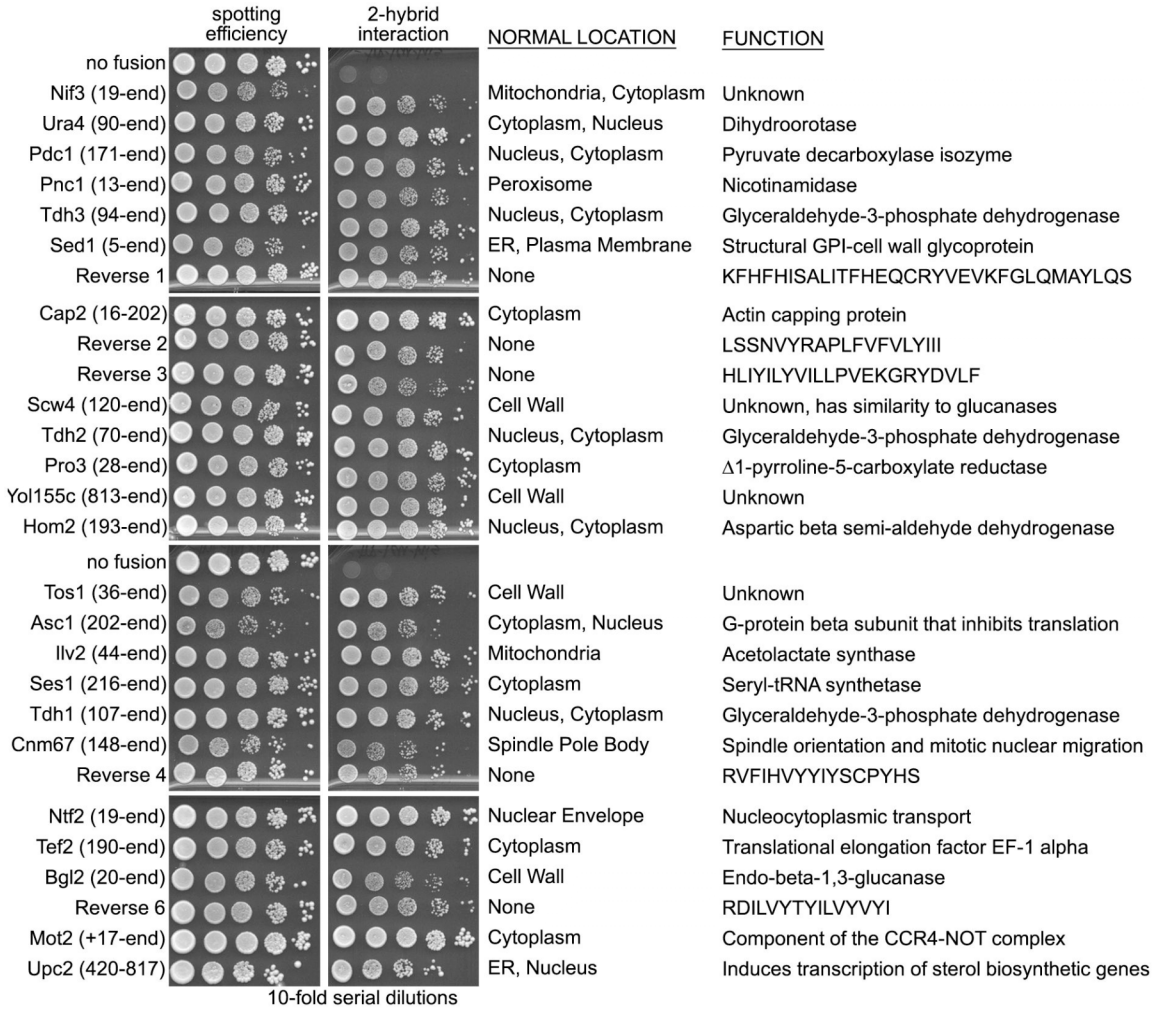


Figure 2.19: Interacting proteins identified in 2-hybrid selection with San1. Cells expressing GBD-San1^{C279S} and each of the listed proteins fused to GAD were spotted onto media plus or minus histidine to measure spotting efficiency and 2-hybrid interaction. The portion of each protein fused to GAD is in parentheses. The sequences of the reverse fusions are on the right.

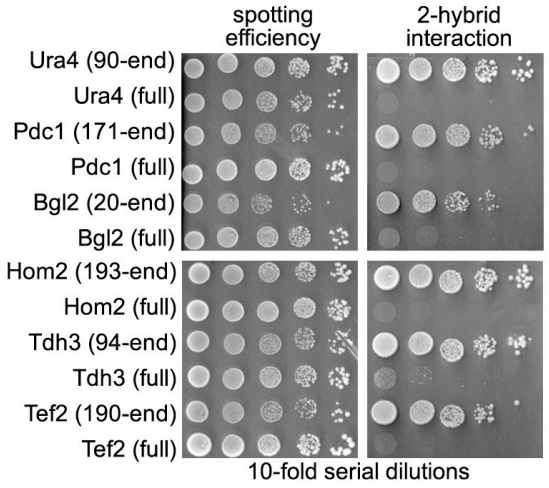


Figure 2.20: San1 does not interact with full-length proteins. Cells expressing GBD-San1^{C279S} and each of the listed proteins and their full-length counterpart fused to GAD were spotted on media to measure spotting efficiency and 2-hybrid interaction.

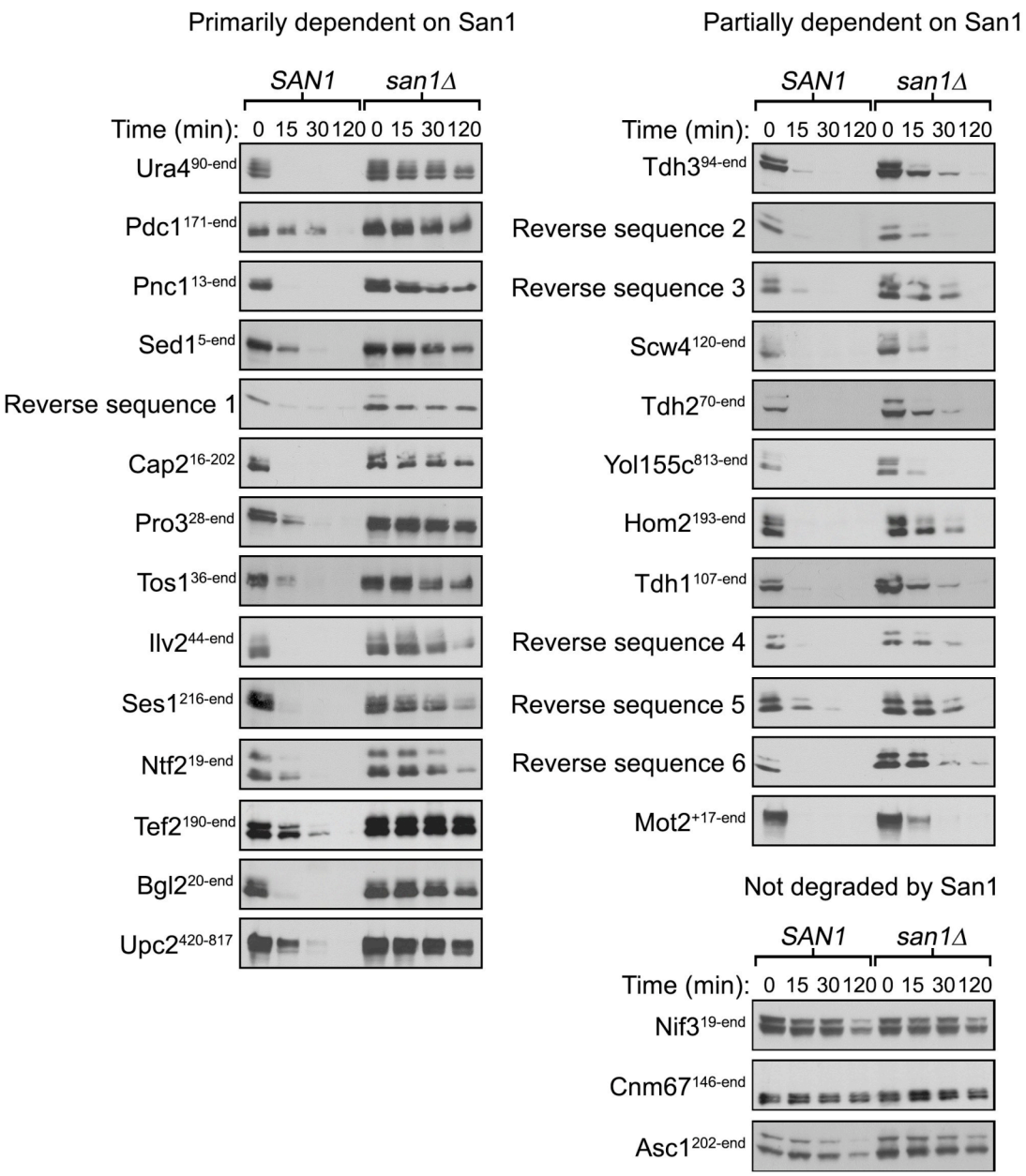


Figure 2.21: San1-dependent degradation of interacting proteins. Cycloheximide-chase assays of cells expressing the indicated GAD fusion were performed to assess stability in the presence or absence of *SAN1*. Time after cycloheximide addition is indicated above each lane. Anti-GAD antibodies were used to detect each GAD fusion. Those GAD fusions whose degradation was primarily dependent upon San1 are grouped in the left column. Those GAD fusions whose degradation was slightly dependent upon San1 are grouped in the right column. Those GAD fusions that were not degraded are grouped at the bottom of the right column.

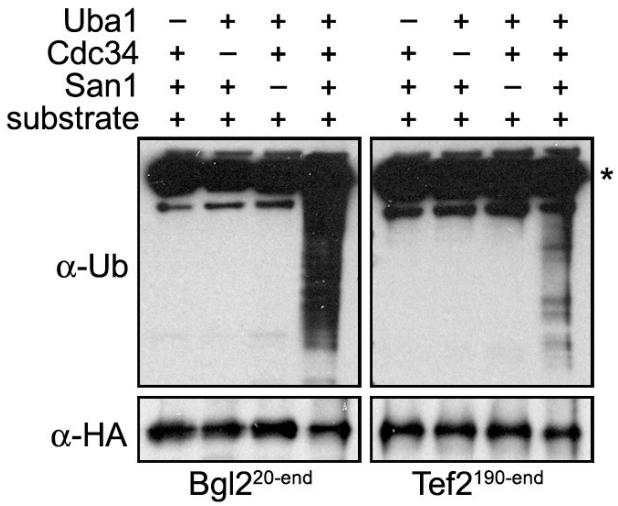


Figure 2.22: San1 ubiquitinates substrates in *E. coli*. 2xHA-tagged GAD substrates were immunoprecipitated from *E. coli* cells expressing the San1 ubiquitination cascade. Western blots were probed with anti-ubiquitin antibodies to assess substrate ubiquitination or anti-HA antibodies to assess substrate IP. Dashes above the lanes indicate which components were not expressed. Asterisk marks the antibody band.

San1 recognizes the 2-hybrid-derived substrates in distinct ways

Using the 2-hybrid assay and the same small San1 deletions described in Figure 2.18, Carrie and I assessed which portions of San1's N- and C-terminal disordered regions were involved in interactions with this larger substrate collection. An important point for this collection is that the San1 2-hybrid interaction could be assessed not only on media minus histidine, but on media minus adenine and media minus histidine plus 3-aminotriazole (+3AT), both of which are higher stringency media that require a stronger 2-hybrid interaction to register growth. This was not the case for Sir4-9, Cdc68-1, or Cdc13-1, likely because they were not selected based on their ability to strongly interact with San1 via two-hybrid.

Carrie and I assayed each interaction in duplicate on each different 2-hybrid media. As with Sir4-9, Cdc68-1, and Cdc13-1, the San1 deletions had variable effects for each substrate interaction. Although many regions throughout San1's N- and C-terminal regions were involved in substrate interactions, the conserved segments in the C-terminal region – particularly those spanning residues 295-319, 340-368, 402-422, 454-473, and 505-521 – appear to be the most important and widely involved in San1's interactions with these substrates. For most substrates, the deletions had little effect on interaction when cells were plated on the lowest stringency media (Figure 2.23, -his media), but showed greater effects as the stringency for the 2-hybrid interaction was increased (Figure 2.24, -ade and -his+3AT media). This result is consistent with San1 having multiple, variable contacts with each substrate such that loss of one contact affects substrate binding in a way that only becomes more apparent as the requirement for a stable 2-hybrid interaction increases.

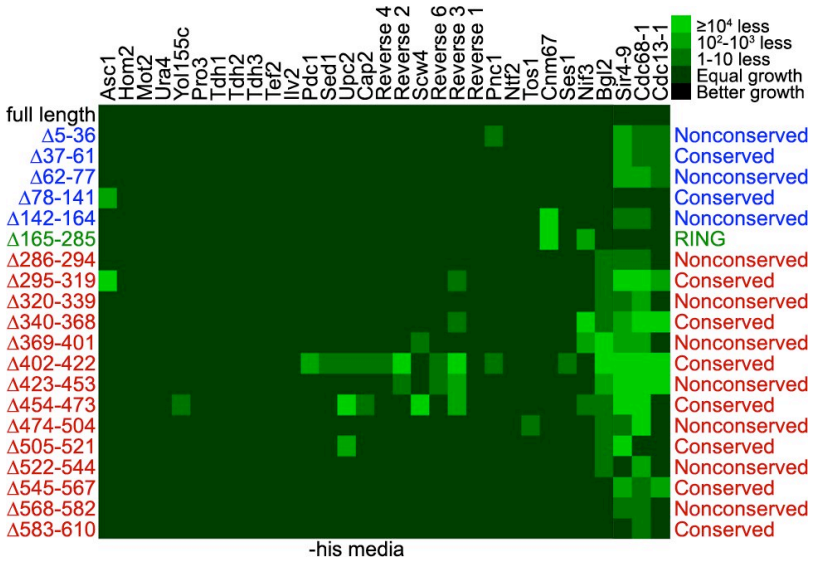


Figure 2.23: San1 binds 2-hybrid substrates distinctly. Cells expressing the indicated San1 deletion (top) and GAD fusion (left) were spotted onto media plus or minus histidine to measure spotting efficiency and the two-hybrid interaction. Interaction dependency is indicated by green color intensity.

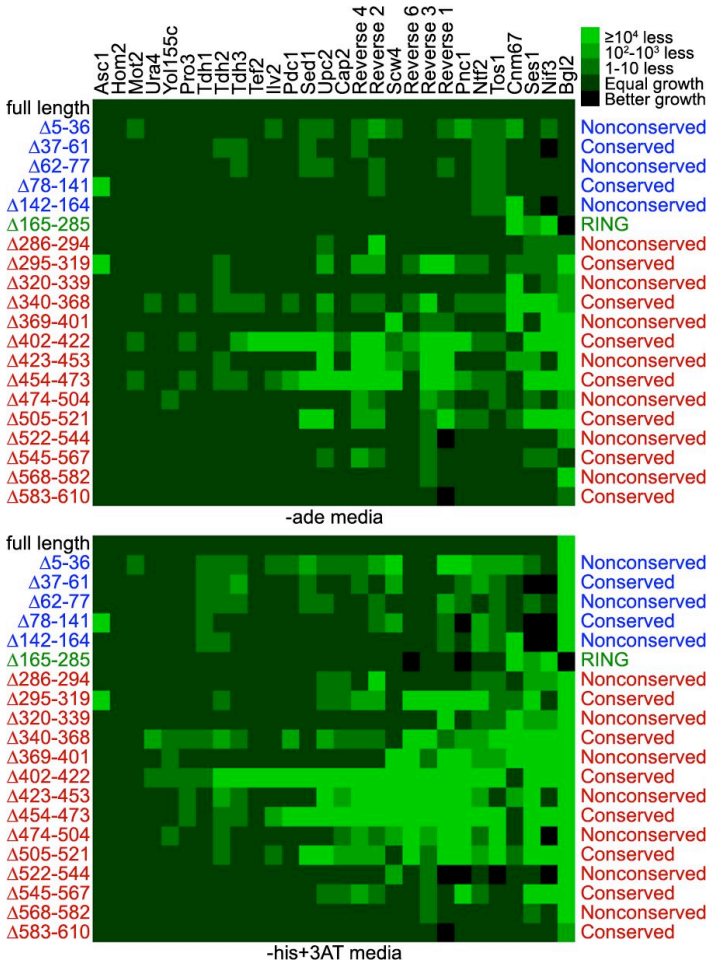


Figure 2.24: Stringent media uncovers more interaction requirements for substrates. Cells expressing the indicated San1 deletion and GAD fusion were spotted onto media minus histidine and adenine (top) or media minus histidine plus 3-AT (bottom) to measure the two-hybrid interaction under more stringent conditions. Interaction dependency is indicated by green color intensity.

Disorder in PQC

In exploring how San1 recognizes misfolded proteins, we discovered that San1 uses multiple binding sites within its highly disordered N- and C-terminal domains to engage its disparate substrates. The use of multiple substrate binding sites within disordered regions was previously uncharacterized in PQC degradation. In fact, other PQC ubiquitin-protein ligases CHIP, Hrd1, Doa10 and Ubr1 contain very little predicted disorder compared to San1 (our unpublished observation). However, sHSPs, which are involved in PQC repair, use disordered regions for interaction with client proteins (Haslbeck et al., 2004; Jaya et al., 2009; Reichmann et al., 2012; Stromer et al., 2004). sHSP 18.1 from *Pisum sativum* (pea) is especially interesting because it contains specific substrate binding sites within its disordered region (Stromer et al., 2004), which is analogous to San1. However, unlike San1, sHSP 18.1's binding sites are individual residues, not small ordered regions. It remains an open question how similar substrate binding actually is between sHSPs and San1.

Upon substrate or cofactor binding, intrinsically disordered proteins can undergo structural changes. One potential change is a disorder to order transition that leads to a single ordered conformation. Another possibility is a disorder to order transition where only local regions of disorder become structured. In the case of sHSP, substrate binding often does not lead to disorder to order transitions (Lindner et al., 1998; Tompa and Csermely, 2004). Future studies are needed to determine if San1 undergoes disorder to order transitions or remains disordered upon substrate binding.

A remaining question is why disorder is used in PQC degradation? One reason may be because IDPs have low affinity for their substrates, likely due to energy costs associated with disorder to order transitions, which means only a small pool of substrate is bound at any one time (Zhou, 2012). Low affinity for a San1-substrate interaction makes sense because ubiquitination occurs rapidly. Additionally, San1 potentially needs to query a large number of proteins and low affinity substrate binding should lead to enhanced interaction speed (Zhou, 2012). In fact, it has been postulated enhanced interaction speed in sHSP may help prevent aggregation of misfolded proteins (Tompa and Csermely, 2004). Another reason to use disorder in PQC is the potential for substrate specificity while using multiple conformations. IDPs can theoretically present a large, convoluted interaction surface due to their flexibility, which can lead to specific

interactions (Zhou, 2012). For San1, substrate specificity is critical because of the need to differentiate normal from abnormal proteins. Overall, the actual reason San1 uses disorder remains unknown.

How does San1 engage substrates?

While the 2-hybrid deletion analysis demonstrated San1 contains multiple substrate binding sites, it did not answer three key points about San1-substrate interaction. First, we could not distinguish the number of conformations San1 adopts upon engaging substrates. One possibility is that San1 possess a broad, but limited number of distinct substrate-binding modalities. Alternatively, certain substrates may share comparable abnormal features and shapes that are recognized by San1 in a similar way. Secondly, we could not determine how multiple binding sites are used to engage a single substrate. We predict the binding sites in San1 either engage substrates independently, or work in conjunction with each other. Lastly, the two-hybrid does not measure affinity, but only indicates a relative strength of interaction. This prevents us from determining if multiple binding sites increase the affinity for substrate binding. For example, the ubiquitin-protein ligase SCF^{cdc4} binds individual phosphorylation events in its substrate Sic1 weakly, but multiple sites strongly (Tang et al., 2012). Without a more quantitative San1-substrate interaction assay, these key questions will remain unanswered.

To better understand San1-substrate interaction, especially at the molecular level, we will need to develop methods for visualization of the interaction. For instance, we could determine disorder to order transition regions of San1 using hydrogen-deuterium exchange coupled with mass spectrometry. The regions that no longer exchange upon substrate binding would indicate a transition to a structured state. Additionally, NMR and SAXS experiments on isolated San1 substrate binding sites with and without substrate could elucidate disorder to order transitions and more precise substrate contact points. Another potential experiment to determine the exact residue contact points between San1 and substrates is to use amber codon crosslinking technology.

CHAPTER III: CHAPERONE ROLE IN PQC DEGRADATION

Introduction

An open question in the PQC field is how repair (chaperone) and degradation (protease) PQC systems in each cellular compartment coordinate and cooperate with each other to effect optimal protein homeostasis. In many cases, it is thought that repair and degradation PQC systems form a triage hierarchy, wherein the decision to refold or destroy a misfolded protein resides with the repair machinery (Wickner et al., 1999). From the studies of misfolded proteins targeted for PQC degradation in the ER and cytoplasm, it is clear chaperones play a key role in the degradation process (see chapter I and II for details). In most cases, however, it is not known if the chaperones' primary function in PQC degradation is to maintain the solubility of misfolded proteins so that they can be subsequently recognized by PQC ubiquitin-protein ligases for ubiquitination, or if they perform a more direct role in misfolded protein recognition by the PQC ubiquitin ligases.

In the nucleus, a compartment without *de novo* protein folding, we wanted to understand the use of chaperones in PQC degradation. We chose to study San1 because San1 is capable of ubiquitinating misfolded proteins directly *in vitro* (Figure 2.4), suggesting that chaperones are not required for the recognition of misfolded proteins by San1. However, two recent studies revealed Hsp70 chaperones (Ssa1/Ssa2) and a Hsp110 chaperone (Sse1) are required *in vivo* for the San1-mediated degradation of misfolded proteins that originate in the cytoplasm and are transported to the nucleus (Heck et al., 2010; Prasad et al., 2010). Due to this discrepancy, we decided to explore if there is a universal *in vivo* requirement for these chaperones or other nuclear chaperones in San1-mediated nuclear PQC degradation. Here, we used a diverse collection of substrates to reveal variability in the chaperone requirement for *in vivo* San1-mediated degradation.

The requirement for Hsp70 and Hsp110 chaperones differ between San1 substrates

Ssa1/Ssa2 were previously shown to be required for the degradation of San1 substrates that were initially presumed to be in the cytoplasm (Heck et al., 2010; Prasad et al., 2010). However, analysis of their localization revealed that these San1 substrates were present in the nucleus (Heck et al., 2010; Prasad et al., 2010), despite the fact that they are not normally nuclear proteins and thus lack canonical nuclear localization sequences (NLS). To explore if Ssa1/Ssa2 are generally required for San1-mediated degradation or solely for the degradation of misfolded proteins without an NLS, I examined the role of Ssa1/Ssa2 in the degradation of San1 substrates that are localized to the nucleus by virtue of a NLS. We previously characterized two different classes of San1 substrates with a NLS: missense mutant proteins (Gardner et al., 2005), and truncated proteins (chapter II). I chose representative examples of these classes of San1 substrates and compared the involvement of Ssa1/Ssa2 in their degradation to the involvement of Ssa1/Ssa2 in the degradation of two San1 substrates lacking an NLS, ΔssCpy* and Δ2GFP, that were previously shown to require Ssa1/Ssa2 for their degradation (Heck et al., 2010; Prasad et al., 2010).

Because we previously examined degradation out to three hours after protein synthesis inhibition, I did the same for ΔssCpy* and Δ2GFP. First, I verified the degradation of ΔssCpy* and Δ2GFP was dependent on Ssa1/Ssa2 (Figure 3.1). Because I examined degradation out to three hours, I observed degradation of Δ2GFP was not completely dependent on Ssa1/Ssa2. Next, I examined degradation of the missense mutant San1 substrates Cdc68-1 and Cdc13-1 fused to the Gal4 activation domain, previously used for two-hybrid analysis (Figure 2.2). With both Cdc68-1 and Cdc13-1, I did not observe any requirement for Ssa1/Ssa2 in their degradation (Figure 3.2). Lastly, I examined the degradation of truncated San1 substrates Ura4* and Pro3*, which we previously identified in the 2-hybrid selection (Figure 2.19). With both Ura4* and Pro3*, I observed a partial involvement of Ssa1/Ssa2 in their degradation (Figure 3.3). This partial stabilization matches the degradation of ΔssCpy* and Δ2GFP, but differs from the degradation of Cdc68-1 and Cdc13-1. From this data, I conclude that Ssa1/Ssa2 involvement in the degradation of San1 substrates is not universal and is substrate specific.

I next examined the universality of Sse1 involvement in the degradation of the different classes of San1 substrates. Sse1 was previously found to be required for the degradation of ΔssCpy* and Δ2GFP (Heck et al., 2010; Prasad et al., 2010). First, I verified Sse1 is involved in the degradation of ΔssCpy* and Δ2GFP (Figure 3.4). Interestingly, degradation was completely dependent on Sse1, which was different than Ssa1/2. Similar to Ssa1/Ssa2, Sse1 was partially involved in the degradation of Ura4* and Pro3* (Figure 3.5). Sse1 was partially required for Cdc13-1 degradation and not required for Cdc68-1 degradation (Figure 3.6). As with the case of Ssa1/Ssa2, I conclude that Sse1 involvement in the degradation of San1 substrates is also not universal and is substrate specific.

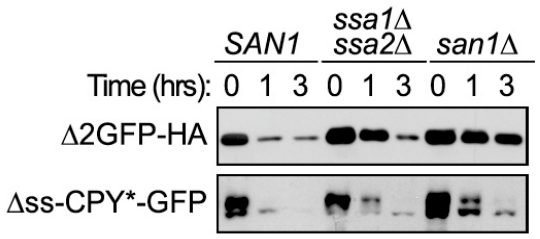


Figure 3.1: Ssa1/2 are partially required for no NLS substrate degradation. Cycloheximide-chase assays of cells expressing the indicated substrate that contains no NLS were performed to assess stability in *SAN1*, *san1*Δ, or *ssa1*Δ*ssa2*Δ cells. Time after cycloheximide addition is indicated above each lane. Anti-HA (Δ2GFP) or anti-GFP (Δss-CPY) antibodies were used to detect the appropriate fusion.

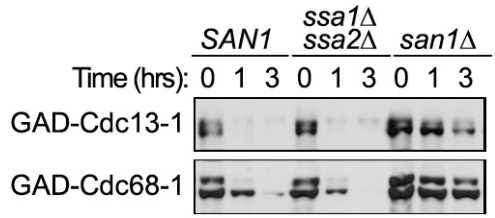


Figure 3.2: Ssa1/2 are not required for missense mutant degradation. Cycloheximide-chase assays of cells expressing the indicated missense mutant substrate were performed to assess stability in *SAN1*, *san1*Δ, or *ssa1*Δ*ssa2*Δ cells. Time after cycloheximide addition is indicated above each lane. Anti-GAD antibodies were used to detect each GAD fusion.

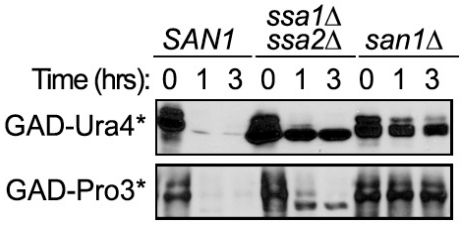


Figure 3.3: Ssa1/2 are partially required for truncated substrate degradation. Cycloheximide-chase assays of cells expressing the indicated truncated substrate were performed to assess stability in *SAN1*, *san1*Δ, or *ssa1*Δ*ssa2*Δ cells. Time after cycloheximide addition is indicated above each lane. Anti-GAD antibodies were used to detect each GAD fusion.

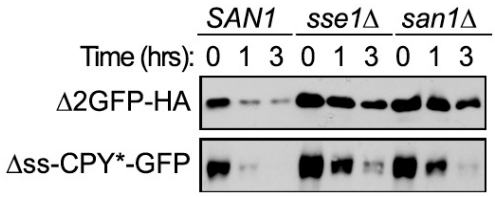


Figure 3.4: Sse1 is required for non-NLS substrate degradation. Cycloheximide-chase assays of cells expressing the indicated substrate that contains no NLS were performed to assess stability in *SAN1*, *san1*Δ*,* or *sse1*Δ cells. Time after cycloheximide addition is indicated above each lane. Anti-HA (Δ2GFP) or anti-GFP (Δss-CPY) antibodies were used to detect the appropriate fusion.

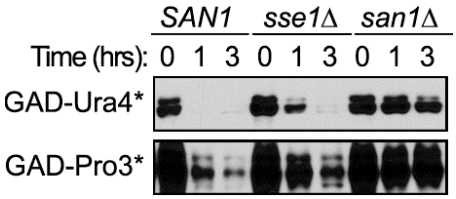


Figure 3.5: Sse1 is partially required for truncated substrate degradation. Cycloheximide-chase assays of cells expressing the indicated missense mutant substrate were performed to assess stability in *SAN1*, *san1*Δ*,* or *sse1*Δ cells. Time after cycloheximide addition is indicated above each lane. Anti-GAD antibodies were used to detect each GAD fusion.

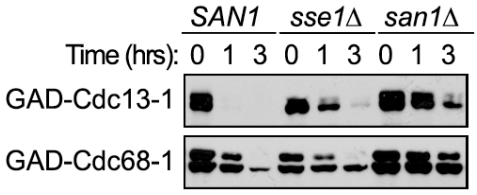


Figure 3.6: Sse1 is partially required for missense mutant degradation. Cycloheximide-chase assays of cells expressing the indicated truncated substrate were performed to assess stability in *SAN1*, *san1*Δ*,* or *ssa1*Δ*ssa2*Δ cells. Time after cycloheximide addition is indicated above each lane. Anti-GAD antibodies were used to detect each GAD fusion.

Other chaperones are not required for San1-mediated degradation

In addition to Ssa1/Ssa2 and Sse1, I also examined the participation of other chaperones in San1-mediated degradation including the Hsp40 chaperone Ydj1, Hsp90 chaperones Hsc82/Hsp82, the Hsp100 chaperone Hsp104, and the small heat shock proteins Hsp26 and Hsp42. I also intend to test the participation of the Hsp70 chaperones Ssb1/Ssb2 and the Hsp90 co-chaperone Sti1.Ydj1, Sti1, and Hsc82/Hsp82 have been previously shown to be involved in the degradation of some misfolded proteins (Huyer et al., 2004; McClellan et al., 2005). Ydj1, Hsc82/Hsp82, Hsp104, Hsp26, and Hsp42 are also nuclear localized (Gong et al., 2009), making them important chaperones to query for involvement in San1-mediated nuclear PQC degradation. So far, I have observed no involvement of these chaperones in the degradation of any of the San1 substrates tested, with the exception of Ydj1 being required for ∆2GFP degradation (our preliminary data). While it remains a formal possibility that an as yet unexamined chaperone is universally required for San1-mediated degradation, the evidence so far indicates that chaperone requirements for San1-mediated degradation are substrate specific. A differential requirement of chaperones indicates differences between San1 substrates.

The requirement for Hsp70 and Hsp110 chaperones may correlate with substrate solubility

Previously, we found that San1 is capable of ubiquitinating a denatured form of luciferase *in vitro* (Figure 2.4). However, we also noticed that San1 was incapable of ubiquitinating higher molecular weight forms of denatured luciferase that were likely aggregates. Thus, one possibility for the role of Ssa1/Ssa2 and Sse1 in the degradation of select San1 substrates is that these chaperones maintain the solubility of aggregation-prone misfolded proteins so that San1 can recognize them as substrates. Supporting this possibility is that ∆ssCpy* was shown to be insoluble in the absence of Ssa chaperones (Park et al., 2007). Therefore, in the future, we will determine if solubility correlates with the requirement of Ssa proteins for degradation. Overall, the lack of a universal chaperone requirement for San1 substrates suggests chaperones do not mediate substrate recognition, but form some other function that differs among substrates.

CHAPTER IV: THE ABNORMAL FEATURE RECOGNIZED IN NUCLEAR PQC

Introduction

A key feature of PQC degradation is the ability to distinguish an abnormal protein among the pool of its normal counterparts. This ability must transcend primary structure, since the abnormal protein can have the same sequence as its normal counterpart. Without this feature, PQC degradation systems would subject structurally intact normal proteins to random destruction. Additionally, because any protein could lose its structure and become misfolded, PQC degradation machinery must also be able to recognize a diverse assortment of unrelated substrates by targeting a common feature of abnormality. At present, we understand poorly what abnormal features PQC ubiquitin-protein ligases recognize in their misfolded substrates.

Because many PQC ubiquitin-protein ligases indirectly recognize misfolded proteins, the abnormal feature targeted can often be elucidated from required adapter proteins. For example, select ERAD-L substrates require Kar2 (Plemper et al., 1997), which is an Hsp70 chaperone. Hsp70 chaperones bind stretches of hydrophobicity surrounded by basic residues (Rudiger et al., 1997). Thus ERAD-L substrates might expose hydrophobic residues that are bound by Kar2 and delivered to Hrd1 for ubiquitination. However, exposed hydrophobicity in ERAD-L substrates has not yet been demonstrated as the abnormal feature recognized by the Hrd1 complex. Because Hsp70 chaperones are required for degradation of CHIP substrates (Murata et al., 2001) and Ubr1 substrates (Heck et al., 2010; Nillegoda et al., 2010; Prasad et al., 2010), the abnormal feature recognized in each pathway might also be exposed hydrophobicity. Certainly, exposed hydrophobicity is not the only abnormal feature targeted by PQC ubiquitin-protein ligases. For example, Yos9 recognizes specific N-linked glycans on ERAD-L substrates (Bhamidipati et al., 2005; Clerc et al., 2009; Kim et al., 2005; Quan et al., 2008). Because adaptor protein usage is variable in PQC degradation, the abnormal feature targeted is often hard to infer.

One theory in the PQC field is that exposed hydrophobicity is a key feature recognized in abnormal proteins. This theory exists because exposed hydrophobicity in an aqueous environment is energetically unfavorable and leads to protein aggregation. The best evidence for exposed hydrophobicity as the abnormal feature in PQC

substrates is with Doa10, because synthetic hydrophobic peptides attached to reporters require Doa10 for degradation (Gilon et al., 2000; Metzger et al., 2008; Ravid et al., 2006; Sadis et al., 1995). Furthermore, the transcription factor Matα2 is also targeted for Doa10-dependent degradation, with Doa10 recognizing the exposed hydrophobic face of an amphipathic helix normally buried when Matα2 forms a heterodimer with Mata1 (Johnson et al., 1998; Ravid et al., 2006). Lastly, degradation of the mutant nuclear protein Ndc10-2 requires the hydrophobic side of an amphipathic helix as well as a hydrophobic C-terminal tail (Furth et al., 2011). In all cases of PQC degradation where hydrophobicity is exposed, the minimal hydrophobic feature recognized has not been demonstrated. In the case of San1, which does not have a universal requirement of chaperones and directly binds its substrates, we wanted to examine the abnormal feature of San1 substrates.

Classes of PQC substrates San1 recognizes

Under normal physiological conditions, PQC degradation systems typically target only a small, random portion of an individual protein's total pool that has become misfolded through synthesis errors or post synthesis damage, or has failed to complex correctly with its partner proteins. Because of the difficulties in examining a small portion of any protein's pool, PQC degradation studies have routinely used three different types of substrates whose entire pool can be induced to misfold. The first type of PQC substrate comprises synthetic peptides, which when fused to reporter proteins can cause their PQC ubiquitin-protein ligase-dependent degradation (Gilon et al., 1998, 2000; Metzger et al., 2008). It is thought that synthetic short peptides mimic some feature present in misfolded proteins that is typically recognized by PQC ubiquitin-protein ligases. Synthetic short peptides represent "gibberish" sequences that can arise in a cell through a number of common mechanisms: termination codon read through, failure to suppress cryptic transcription, frameshift mutations, and mRNA splicing defects. With San1, we have identified 6 such peptides fused to the GAD from the 2-hybrid selection (Figure 2.19).

The second type of PQC substrate consists of truncated proteins (Fredrickson et al., 2011; Heck et al., 2010; Park et al., 2007; Prasad et al., 2010; Ward et al., 1995). Presumably, truncated proteins cannot fold into a proper structure and thus represent constitutively misfolded proteins targeted by PQC ubiquitin-protein ligases. Truncated

proteins can arise in a cell through a number of common mechanisms: frameshift mutations, nonsense mutations, failures in nonsense-mediated decay, premature transcription or translation termination, mRNA splicing defects, and the action of cellular proteases. In fact, proteolytic cleavage precedes aggregation and neurodegeneration in a number of tauopathies and polyQ-expansion disorders (Walsh et al., 2005). We have identified 20 truncated proteins that require San1 for their degradation (Figure 2.19).

The third type of PQC substrate encompasses missense mutant proteins (Betting and Seufert, 1996; Biederer et al.; Bordallo et al., 1998; Dasgupta et al., 2004; Estruch et al., 2009; Gardner et al., 2005; Kaganovich et al., 2008; Lewis and Pelham, 2009; Matsuo et al., 2011; Ravid et al., 2006; Wang et al., 2006; Wang and Prelich, 2009). In these cases, single missense mutations lead to misfolding of a protein and subsequent recognition by PQC ubiquitin-protein ligases. Often, the missense mutations cause temperature-dependent loss of protein function that can be recovered by eliminating their PQC degradation (Betting and Seufert, 1996; Bordallo et al., 1998; Dasgupta et al., 2004; Estruch et al., 2009; Gardner et al., 2005; Ravid et al., 2006; Wang et al., 2006), indicating that the misfolding caused by the mutations is not so severe that all protein function is lost. Missense mutants historically have been a main class of substrates used to study PQC degradation systems, due to the ability to genetically selection for temperature sensitive mutants in yeast. With San1, there are seven missense mutants that are dependent on San1 for their degradation (Estruch et al., 2009; Gardner et al., 2005; Lewis and Pelham, 2009).

Because San1 degrades misfolded proteins from each class, we hypothesized San1 is likely recognizing a similar abnormal feature in each substrate. For our studies on the abnormal feature San1 recognizes, we used examples from each type of substrate to see if we could identify a universal feature of abnormality targeted by San1.

Small hydrophobic peptides function as degrons for San1-mediated degradation

From the two-hybrid selection we performed with San1 and a yeast cDNA library (Figure 2.19), we noticed one interesting class of substrates were cDNAs oriented in the reverse direction to the GAD, causing the translation of random peptide sequences fused to the GAD rather than the desired protein encoded by the cDNA. Inspection of the peptide sequences revealed that 5 of the 6 peptides contain long tracts of hydrophobic

residues (Figure 4.1, peptides 1-6), suggesting that San1 targets exposed hydrophobicity. To explore this further, we obtained a quantitative measure of peptide hydrophobicity by averaging the Kyte-Doolittle hydrophobicity (Kyte and Doolittle, 1982) values for each residue in the peptide, thus yielding an average hydrophobicity score for each peptide (which we call a K-D score). We classified a peptide as hydrophobic if it has a K-D score >0 and as hydrophilic if it has a K-D score <0. All 6 peptides have K-D scores >0 (Figure 4.1).

To determine if San1 generally binds hydrophobic peptides, I subsequently performed a 2-hybrid selection using a library of ~10^7 random 16-mer peptides fused to the GAD (Yang 1995). I identified 10 unique peptides that interact with San1, and 9 of the 10 have K-D scores >0 (Figure 4.1, peptides A-J). The K-D scores again suggested that San1 targets hydrophobicity. To determine the likelihood that this result could occur by chance, Tom selected at random 10^6 simulated peptides from the library (based on the library's idealized composition) and found that the probability of selecting a peptide with a K-D score >0 is 0.336 (Figure 4.2). Therefore, it is unlikely that 9 of 10 peptides chosen at random from the library would have K-D scores >0 (*P* value = 3.8×10^{-4}). To verify the library itself is not biased towards hydrophobicity, Melissa and I sequenced 10 randomly chosen peptides and found 4 to have K-D scores >0 (Figure 4.3, random peptides K-T), indicating that the peptide composition of the library is as expected (*P* value = 0.448). None of the randomly chosen peptides interact with San1 by 2-hybrid (Figure 4.3), which we expected since they weren't identified as interactors in the 2-hybrid selection.

Because San1 is highly flexible (chapter II), we considered the possibility that the San1-interacting hydrophobic peptides might simply interact with *any* IDP non-specifically. To test this, Melissa measured the 2-hybrid interaction between all peptides and the deubiquitinating enzyme Ubp10, which is predicted to be as highly disordered outside of its catalytic domain as San1 (Figure 4.4). No peptide interacted with Ubp10 (Figure 4.5), demonstrating that the hydrophobic peptides' interactions are specific for San1 and ruling out the possibility that they simply interact non-specifically with any intrinsically disordered protein.

Melissa and I verified that the San1-interacting peptides target reporter fusions for San1-mediated degradation by attaching the peptides to green fluorescent protein (GFP) containing a nuclear localization signal (NLS), which allowed us to target the GFP fusions to the nucleus where San1 resides (Gardner 2005). Melissa and I analyzed the degradation of the GFP^{NLS}-peptide fusions in mutant *san1*Δ and wild-type *SAN1* cells using a cycloheximide-chase assay. Each GFP^{NLS}-peptide is degraded in wild-type *SAN1* cells and stabilized in mutant *san1*Δ cells (see representative examples in Figure 4.6), which is consistent with their interactions with San1 in the 2-hybrid assay (Figure 4.1). By contrast, the randomly selected peptides, which do not interact with San1 in the 2-hybrid assay (Figure 4.3), are not subject to San1-dependent degradation; the GFP^{NLS}-random peptide fusions are either stable or degraded in a San1-independent manner (see representative examples in Figure 4.7). Because the San1-interacting cDNA-derived peptides have stretches of hydrophobic residues and hydrophilic residues, Joel confirmed that the hydrophobic residues are solely responsible for San1-dependent degradation by splitting peptides 2 and 3 into their constituent hydrophobic and hydrophilic parts. He observed San1-dependent degradation only when the hydrophobic, but not the hydrophilic, part of these peptides is fused to GFP^{NLS} (Figure 4.8).

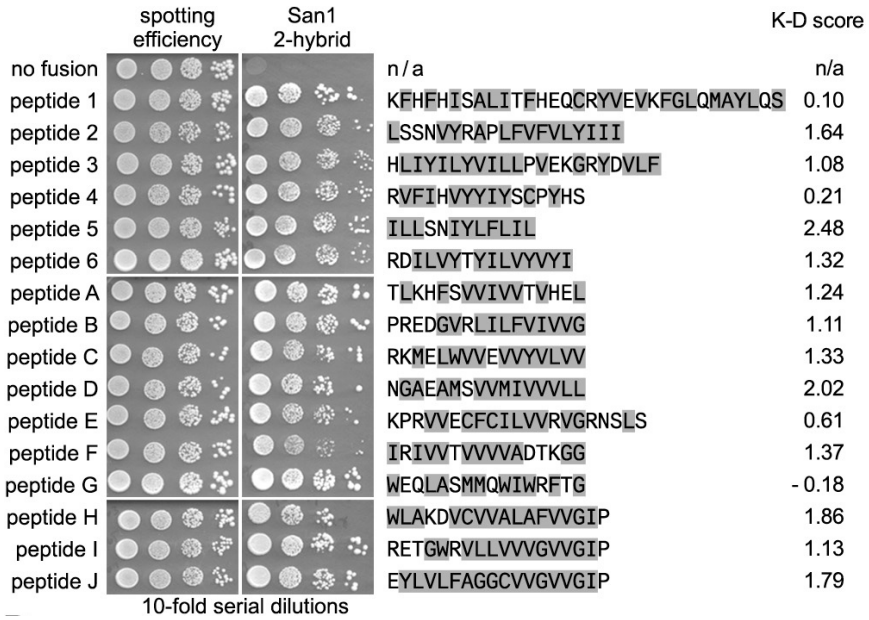


Figure 4.1: San1 interacts with hydrophobic peptides fused to the GAD. Cells expressing GBD-San1^{C279S} and each listed peptide fused to GAD were spotted onto media with or without histidine to measure spotting efficiency and 2-hybrid interaction. Hydrophobic residues are in gray. Average Kyte-Doolittle (K-D) scores for each peptide are listed on the right.

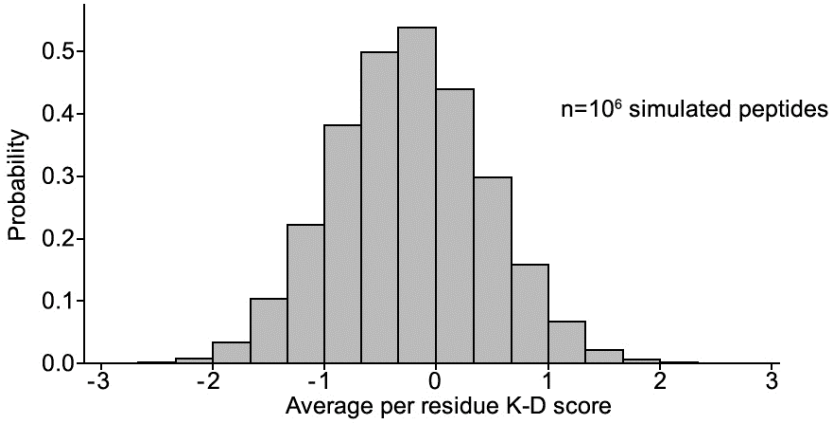


Figure 4.2: Hydrophobic content in simulation of peptide library. Histogram of Kyte-Doolittle scores generated from 10^6 simulated peptides selected randomly from the 16-mer peptide library.

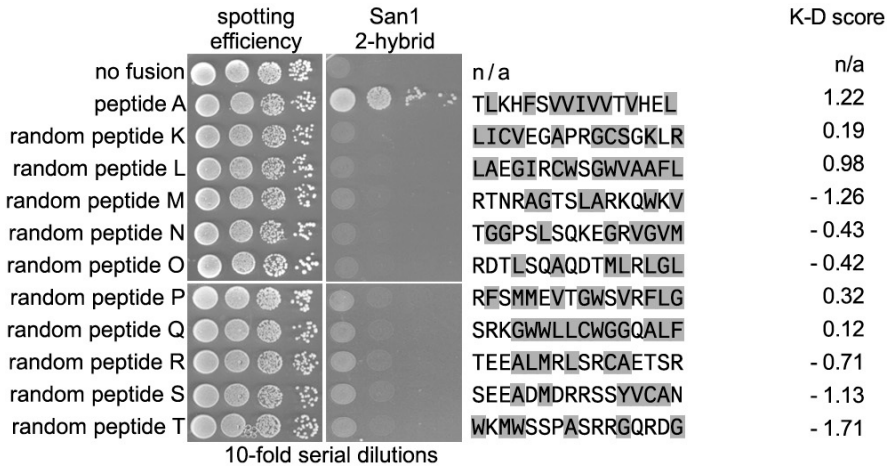


Figure 4.3: Random peptides do not interact with San1. Cells expressing GBD-San1^{C279S} and the indicated randomly selected peptide fused to GAD were spotted onto media with or without histidine to measure spotting efficiency and 2-hybrid interaction. Hydrophobic residues are in gray. Average Kyte-Doolittle (K-D) scores for each peptide are listed on the right.

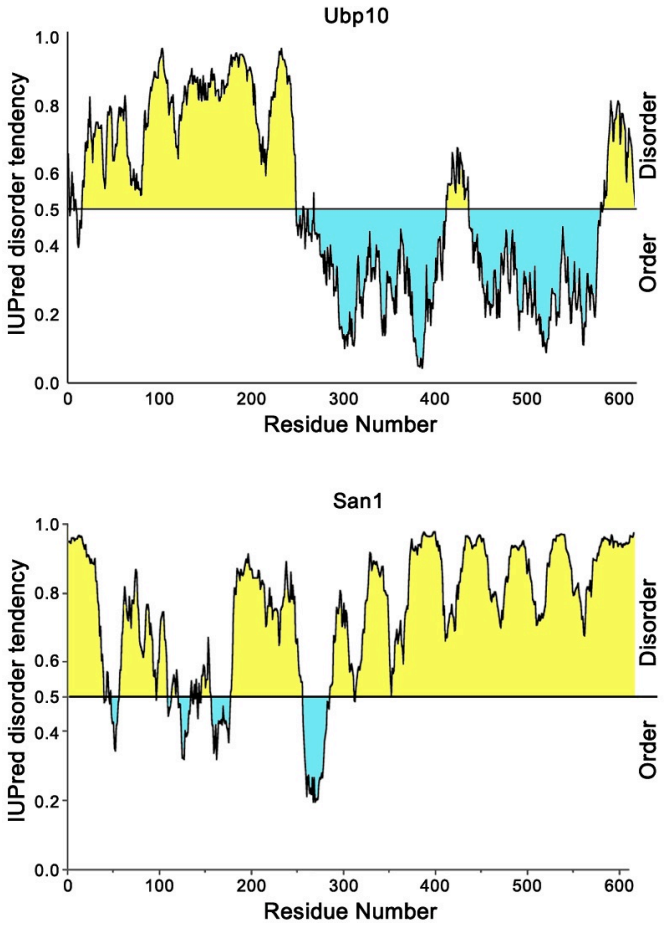


Figure 4.4: Disorder of Ubp10 and San1 is similar. Disorder predictions of Ubp10 and San1. Panels were generated from IUPred (http://iupred.enzim.hu/index.html). Predicted disordered regions are in yellow and predicted ordered regions are in blue.

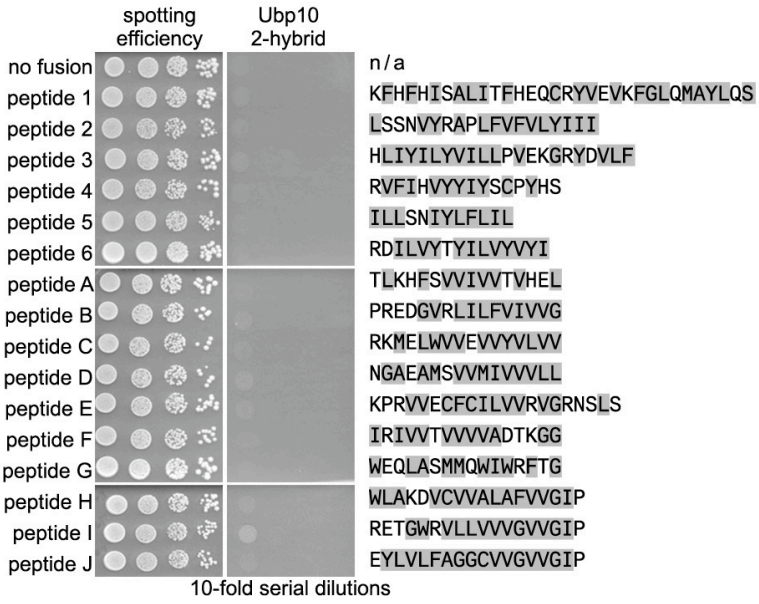


Figure 4.5: San1-interacting peptides do not interact with Ubp10. Cells expressing GBD-Ubp10 and each listed peptide fused to GAD were spotted onto media with or without histidine to measure spotting efficiency and 2-hybrid interaction. Hydrophobic residues are in gray.

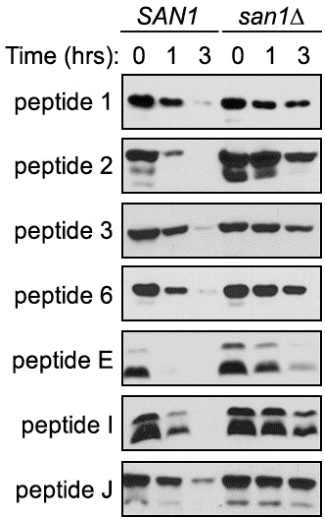


Figure 4.6: San1-dependent degradation of hydrophobic peptides. Cycloheximide-chase assays were performed to assess stability of the indicated GFP^{NLS}-peptide fusion in the presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots.

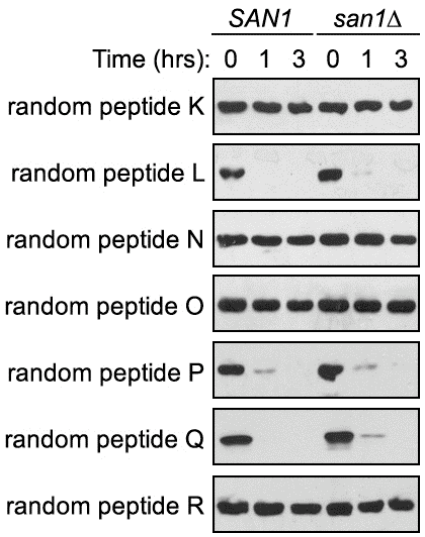


Figure 4.7: Random peptides are not San1 substrates. Cycloheximide-chase assays were performed to assess stability of the indicated GFP^{NLS}-peptide fusion in the presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots. None of the peptides evaluated exhibited *SAN1*-dependent stability, indicating that they are not San1 substrates.

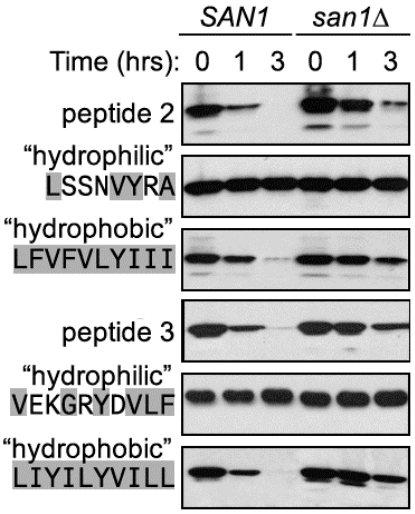


Figure 4.8: San1 targets hydrophobic parts of peptides. Cycloheximide-chase assays were performed to assess stability of GFP^{NLS}-Peptide 2 and 3 fusions and their isolated hydrophobic and hydrophilic components in the presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots.

A window of hydrophobicity defines peptide recognition via San1

A few of the randomly selected peptides do not interact with San1 but do have K-D scores >0 (Figure 4.3), so it is likely that San1 recognizes something other than the overall hydrophobicity of the peptide. Because each peptide contained long, uninterrupted stretches of hydrophobic residues, we hypothesized that San1 recognizes regions within each peptide where hydrophobic residues are concentrated. To test this, Joel determined the minimum number of contiguous hydrophobic residues required for San1-dependent degradation by substituting an increasing number of serine or threonine residues for the N-terminal hydrophobic residues in the peptide 2 and 3 hydrophobic variants. He found that San1 requires a minimum of 5 contiguous hydrophobic residues to target the GFP^{NLS}-peptides for degradation (Figure 4.9). Consistent with the idea that a window of hydrophobicity rather than overall hydrophobicity is recognized by San1, Joel found that the peptide 2 and 3 variants with only 4 contiguous hydrophobic residues are not degraded despite the fact that they have overall K-D scores >0 (TSTSTSYIII = 0.77, TSTSTSVILL = 1.18). Thus, we conclude that a local window of concentrated hydrophobicity rather than overall peptide hydrophobicity is the primary determinant of San1 recognition.

To examine the "window of hydrophobicity" hypothesis further, Joel evaluated the local K-D hydrophobicity score in a 5-residue sliding window across the length of all San1-interacting and non-interacting peptides. He recorded the maximum 5-residue window value found for each peptide and plotted the values of San1-interacting and non-interacting peptides separately for comparison (Figure 4.10). He observed a significantly higher mean K-D window value for the San1-interacting peptides compared with that for the non-interacting randomly selected peptides (*P* value = 2.9×10^{-5}). Kyte-Doolittle has particular biases in characterizing hydrophobicity (Kyte and Doolittle, 1982), so we also evaluated the maximum hydrophobicity window value in all peptides using AGGRESCAN – an algorithm trained to predict the aggregation propensity of protein regions based on their residues' hydrophobic content (Conchillo-Sole et al., 2007). Joel found that the San1-interacting peptides also have higher maximum AGGRESCAN window values than their non-interacting counterparts (Figure 4.10), with even greater statistical separation of the mean values (*P* value = 2.8×10^{-7}). Altogether, the experimental evidence coupled with the hydrophobicity analyses indicates that San1 recognizes a window of concentrated hydrophobicity in its target peptides.

Although there is a significant correlation between the window of hydrophobicity in a peptide and its interaction with San1 and San1-dependent degradation, there are some interesting outliers that can't be explained solely by this simple model. Both peptides 1 and G interact with San1 and are degraded in a San1-dependent manner, but each has significantly lower K-D and AGGRESCAN maximum window averages than the other interacting peptides. Conversely, random peptides L and Q have maximum window averages that resemble those of San1-interacting peptides. However, these peptides do not interact with San1 and are subject to San1-independent degradation. See the Discussion section for the potential reasons why we think these outliers are consistent with the "window of hydrophobicity" model of San1 recognition.

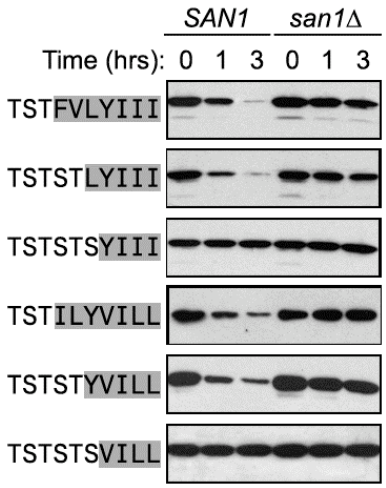
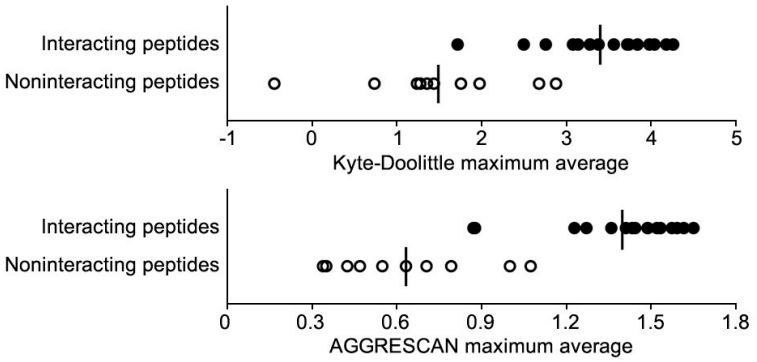


Figure 4.9: San1 can target five contiguous hydrophobic residues. Polar substitutions were made in the N-termini of Peptide 2 and 3 hydrophobic regions and were evaluated for *SAN1* dependency via cycloheximide-chase. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots. Peptides evaluated are *SAN1* substrates as long as they have 5 contiguous hydrophobic residues.



Peptide tested	Kyte-Doolittle maximum average	AGGRESCAN maximum average
peptide 1	2.76	0.887
peptide 2	3.56	1.615
peptide 3	3.14	1.555
peptide 4	2.50	1.379
peptide 5	3.74	1.543
peptide 6	3.08	1.510
peptide A	4.26	1.640
peptide B	3.98	1.674
peptide C	3.28	1.464
peptide D	4.18	1.597
peptide E	3.84	1.431
peptide F	3.72	1.289
peptide G	1.72	0.882
peptide H	3.38	1.246
peptide I	4.04	1.508
peptide J	3.28	1.453
sample mean	3.40	1.417
standard deviation	0.67	0.239

Peptide tested	Kyte-Doolittle maximum average	AGGRESCAN maximum average
random peptide K	2.68	1.013
random peptide L	2.88	0.802
random peptide M	0.74	0.340
random peptide N	1.98	0.713
random peptide O	1.44	0.473
random peptide P	1.24	0.639
random peptide Q	1.76	1.088
random peptide R	1.36	0.427
random peptide S	1.28	0.554
random peptide T	-0.44	0.352
sample mean	1.49	0.640
standard deviation	0.95	0.263

Figure 4.10: Peptides are defined by a local window of hydrophobicity. Kyte-Doolittle (top panel) and AGGRESCAN (second panel) sliding scales of 5 residues were used to determine the maximum hydrophobic window in San1-interacting (●) and non-interacting (○) peptides. A vertical line marks the mean of each group. Maximum window values for interacting and non-interacting peptides are listed below.

Type of hydrophobic residue is relevant in San1 recognition

Although our "window of hydrophobicity" model of San1 recognition correlated with nearly all peptides, we only tested the degradation of two isolated, contiguous 5-residue stretches of hydrophobic peptides (Figure 4.9). These two isolated 5-residue stretches did not cover the entire spectrum of hydrophobicity. Both 5-residue stretches contained Leu, Tyr, and Ile residues, but only one stretch contained a Val residue, and neither contained Phe, Ala, Met, or Trp residues. Because each of these hydrophobic residues was present in at least one of the peptides we identified in the 2-hybrid selection (Figure 4.1), we decided to test the ability for San1 to recognize each individual hydrophobic amino acid.

To determine San1's ability to recognize specific hydrophobic residues, I examined the degradation of five hydrophobic residues where each residue is identical, which we called "homopentamers". I predicted the window of hydrophobicity score for the homopentamer would correlate with San1-dependent degradation. Only five Leu, Val, or Ile residues have a K-D hydrophobic window value similar to the San1-interacting peptides (Figure 4.11-4.13). By using AGGRESCAN hydrophobic window values, Leu, Val, Ile, and Phe residues have values similar to the San1-interacting peptides (Figure 4.11-4.13). Thus I predicted only Leu, Val, Ile and Phe residues would be degraded as homopentamers. To experimentally test this prediction, I fused a TSTST-5mer sequence to GFP^{NLS} and examined degradation in *SAN1* and *san1*Δ cells. The degradation of each homopentamer fit into one of three distinct groups. The first group consisted of Phe (5F) and Ile (5I) residues, which were clearly degraded in a San1-dependent manner (Figure 4.11). The second group consisted of Tyr (5Y), Met (5M) and Trp (5W) residues, which were stable for the entirety of the assay (Figure 4.12). The final group consisted of Leu (5L), Val (5V) and Ala (5A) residues, which were degraded in a San1-independent manner (Figure 4.13). In each case, the degradation of "TSTSTYVILL" was similar to Figure 4.9. Overall, the homopentamers that are degraded are consistent with higher AGGRESCAN 5-residue window values, but surprisingly, San1 was not universally required for their degradation.

Because many peptides we identified in the 2-hybrid contained long stretches of only Val and Leu residues (Figure 4.1), I was surprised to see little to no dependence on San1 for the degradation of 5V or 5L. I hypothesized the lack of San1-dependence was

either because San1 prefers a non-homogenous stretch of Val and Leu residues or because other degradation pathways preferentially target these residues. To determine if San1 more readily targets a mix of Val and Leu residues, I examined the degradation of "LVLVL" fused to GFP^{NLS} in *SAN1* and *san1*Δ cells. The degradation of this peptide mirrored that of 5V (Figure 4.14), suggesting San1 recognition for Val and Leu residues does not change in a heterogeneous 5mer. To determine the minimal number of Ile residues in a stretch of Val residues that leads to San1-dependent degradation, I examined the degradation of "VIVIV" in *SAN1* and *san1*Δ cells. "VIVIV" was degraded in a San1-dependent manner, as was "VVVIV" (Figure 4.14), suggesting a single Ile residue imbedded in a stretch of Val residues is sufficient for San1-dependent degradation.

Next I wanted to determine what PQC degradation machinery targets the San1-independent class of homopentamers. I focused on 5V, as it was degraded the most rapidly. To find the other pathway(s) that target 5V for degradation, I first tested dependence on the proteasome for 5V degradation. To test dependence of the proteasome, I used the proteasome inhibitor MG-132, and this use led to a complete stabilization of 5V (Figure 4.15), demonstrating the other pathway involves proteasomal degradation. With proteasome degradation in mind, I selected Doa10 and Ubr1, both ubiquitin-protein ligases known to target abnormal proteins, as candidates for 5V degradation. To test the dependence of Doa10 and Ubr1 for 5V degradation, I examined degradation in *ubr1*Δ*doa10*Δ and *ubr1*Δ*san1*Δ*doa10*Δ cells and similar to *san1*Δ*,* it was almost completely degraded by one hour (Figure 4.16). This result indicates an unknown PQC ubiquitin-protein ligase likely exists in the nucleus.

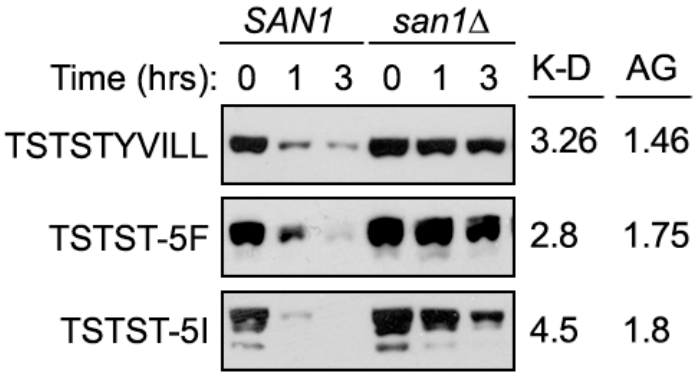


Figure 4.11: San1 targets five contiguous Ile or Phe residues. Cycloheximide-chase assays were performed to assess stability of GFP^{NLS}-homopentamer fusions in the presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots. Peptides here are degraded via San1. The Kyte-Doolittle (K-D) and AGGRESCAN (AG) scores for each 5-mer are listed on the right.

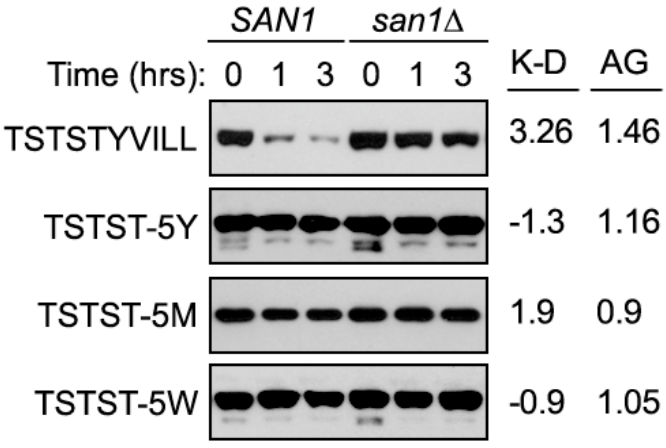


Figure 4.12: Five contiguous Met, Trp or Tyr residues are stable. See Figure 4.11. Peptides here are not degraded. The Kyte-Doolittle (K-D) and AGGRESCAN (AG) scores for each 5-mer are listed on the right.

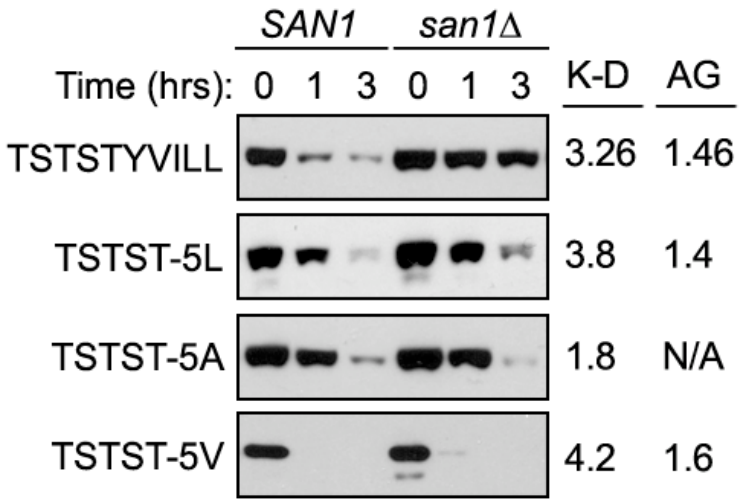


Figure 4.13: Five contiguous Val, Leu or Ala residue degradation is San1-independent. See Figure 4.11. Peptides here are degraded via a pathway that is not San1. The Kyte-Doolittle (K-D) and AGGRESCAN (AG) scores for each 5-mer are listed on the right.

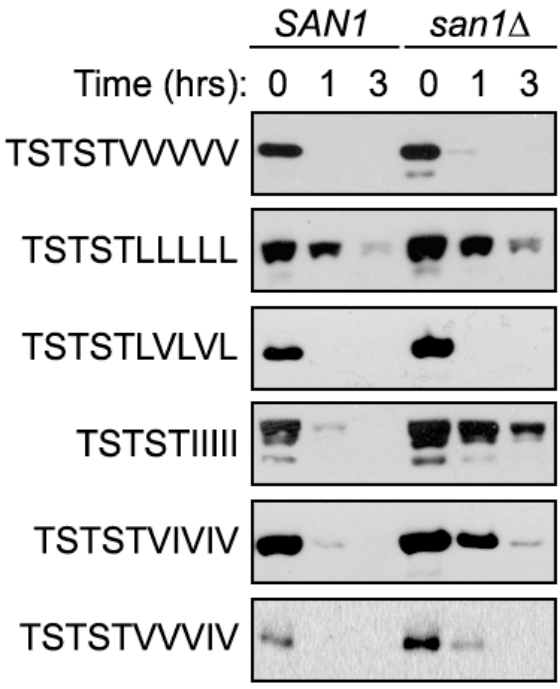


Figure 4.14: San1-dependent degradation of a mix of Val, Leu and Ile residues. Cycloheximide-chase assays were performed to assess stability of GFP^{NLS} peptide fusions in the presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots. Hydrophobic peptides with only one Ile residue are degraded via San1, while a mix of Leu and Val residues are not.

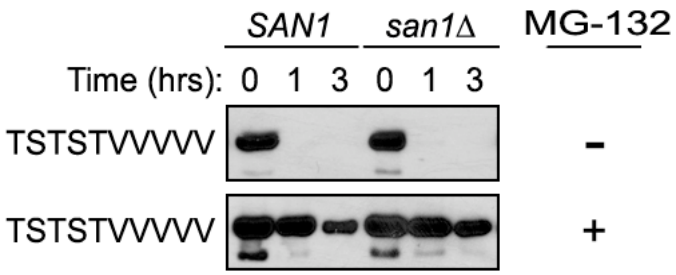


Figure 4.15: Five Val residue degradation is proteasome-dependent. Cycloheximide-chase assays were performed to assess stability of GFP^{NLS}-5V in the presence or absence of *SAN1* and MG-132, a proteasome inhibitor. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots.

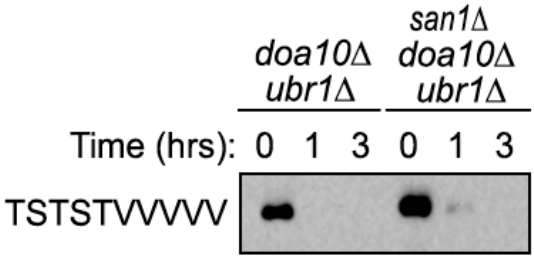


Figure 4.16: Five Val residue degradation is not via Doa10 or Ubr1. Cycloheximide-chase assays were performed to assess stability of GFP^{NLS}-5V in the presence or absence of *SAN1, UBR1* and *DOA10*. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots.

Hydrophobic residue degradation changes with length

Because my initial observations were made using a minimal stretch of hydrophobicity, I wanted to determine if the residue preferences of San1 change with longer stretches of hydrophobicity. I had two reasons to believe a longer stretch of hydrophobic residues might change San1 dependency. First, most peptides we identified in the 2-hybrid selection had stretches of hydrophobicity longer than five residues (Figure 4.1). Secondly, the two isolated 5 residue hydrophobic windows are degraded sluggishly compared to the peptides from which they are derived (Figure 4.8 and 4.9), suggesting length differences affect the rate of degradation. To determine if San1's ability to recognize specific hydrophobic residues changes in longer stretches of hydrophobicity, I fused TST-7mers (homosepatmer) to GFP^{NLS} and their examined degradation in *SAN1* and *san1*Δ cells. As with the homopentamers, the degradation of the homoseptamers comprised three distinct groups. The first group, consisting of 7I, 7F, 7L, 7V, and 7Y, were degraded in a San1-dependent manner (Figure 4.17). The second group, consisting of 7M and 7A, were degraded, but almost entirely in a San1-independent manner (Figure 4.17). Lastly, 7W was a stable protein (Figure 4.17, bottom). San1-dependent degradation of the homoseptamers correlated with higher AGGRESCAN 5-residue window values, except with 7Y (Figure 4.17). Interestingly, 5Y was stable, indicating a threshold effect for San1 recognition (See discussion for more). Overall, the San1-dependent degradation of contiguous stretches of hydrophobic residues is altered by length.

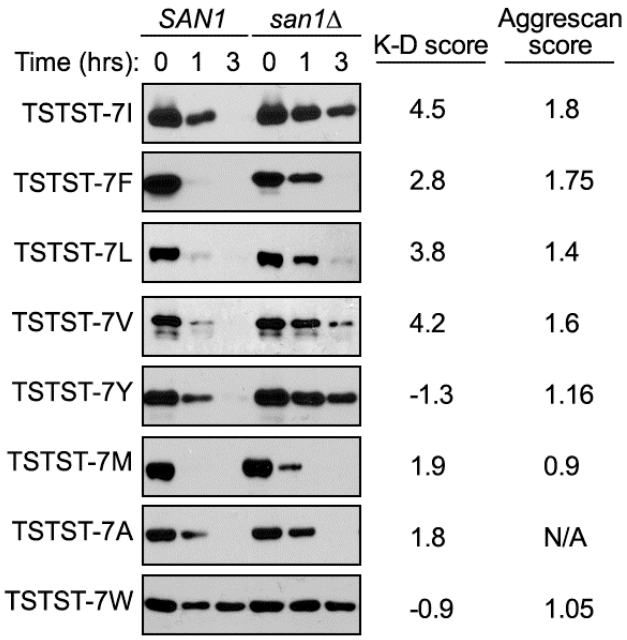


Figure 4.17: Seven contiguous hydrophobic residues changes San1-dependency. Cycloheximide-chase assays were performed to assess stability of GFP^{NLS}-homoseptamer fusions in the presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies. Time after cycloheximide addition is indicated above the blots. Residues targeted by San1 are different from homopentamers. On the right are the K-D and AGGRESCAN 5-residue scores.

San1 targets hydrophobicity in truncated substrates

The hydrophobicity that San1 recognizes in the peptides likely reflects a similar feature that San1 targets in the misfolded proteins it typically encounters in the nucleus. To see if this is the case, we turned to the other San1 substrates that we discovered in the 2-hybrid cDNA selection (Figure 2.19), which are predominantly truncated proteins fused to the NLS-containing GAD. For these substrates, we hypothesized that the truncations cause hydrophobic regions normally buried in the full-length protein to become surface exposed and accessible to San1 when they are localized to the nucleus by virtue of their fusion to the GAD. We identified candidate hydrophobic regions in each truncated substrate by plotting K-D hydrophobicity values along their sequence lengths. Fifteen truncated substrates have multiple hydrophobic stretches scattered throughout the protein; we categorized these as group I substrates (Figure 4.18, left plot). Six truncated substrates contain only a single pronounced stretch of hydrophobicity; we categorized these as group II substrates (Figure 4.18, right plot).

To determine if San1 recognizes hydrophobicity in truncated substrates, we focused on members of group II due to their single stretch of hydrophobicity. One group II substrate is Rad16*, a mutant form of the nuclear excision repair protein Rad16 in which residues 116-664 have been deleted. Rad16* possesses a 28 residue hydrophobic patch spanning residues 142-169 in the truncated Rad16* protein (Figure 4.19, left panel). This region corresponds to residues 691-718 in full-length Rad16 that normally would have been buried. Another group II substrate is Sed1*, which is a mutant form of the cell wall protein Sed1 in which the first 20 residues have been deleted. The truncated Sed1* protein contains a hydrophobic region spanning its final 44 residues (Figure 4.19, right panel). The hydrophobic patches in both Rad16* and Sed1* each contain a maximum 5-residue hydrophobic window values similar to the San1-interacting peptides (Figure 4.20), making them the likely regions recognized by San1 when localized to the nucleus. Both Rad16* and Sed1* interact with San1 in the 2-hybrid assay (Figure 4.21), and undergo San1-dependent degradation (Figure 4.22). Deletion of the hydrophobic patch in Rad16* and Sed1* disrupts their interaction with San1 (Figure 4.20, GAD-Rad16*$^{\Delta h}$ and -Sed1*$^{\Delta h}$), and eliminates their San1-dependent degradation (Figure 4.21). Conversely, the isolated hydrophobic patch from Rad16* (residues 691-718 of full-length Rad16) and Sed1* (residues 294-338 in full-length Sed1) is sufficient to cause interaction with San1 (Figure 4.21, GAD-Rad16*$^{\text{hydro only}}$ and -

Sed1*$^{hydro\ only}$), and induce San1-dependent degradation (Figure 4.22). Based on these results, we conclude that San1 also recognizes a window of exposed hydrophobicity in its truncated substrates that is both necessary and sufficient.

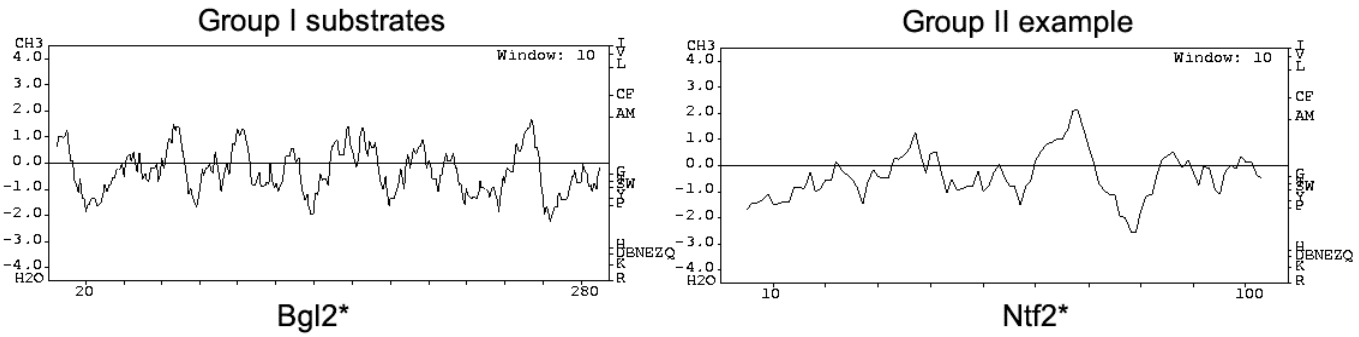


Figure 4.18: Examples of hydrophobicity profiles in truncated substrates. Kyte-Doolittle hydrophobicity evaluation is shown for a truncated San1 substrate with dispersed hydrophobicity (left) and a truncated San1 substrate with a distinct (right) hydrophobic region.

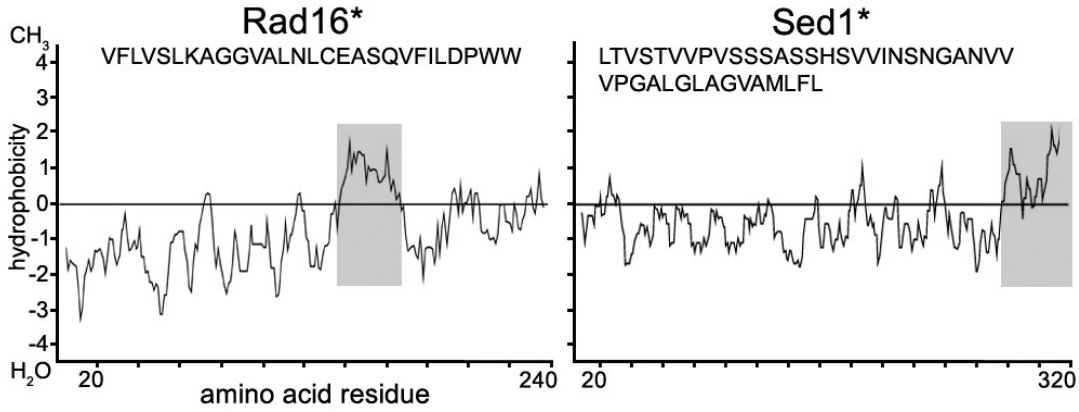


Figure 4.19: Rad16* and Sed1* are defined by a single hydrophobic patch. Kyte-Doolittle plots of Rad16* and Sed1* hydrophobicity. Gray indicates a stretch of hydrophobicity. The amino acid sequence of Rad16* and Sed1* in the gray box is indicated.

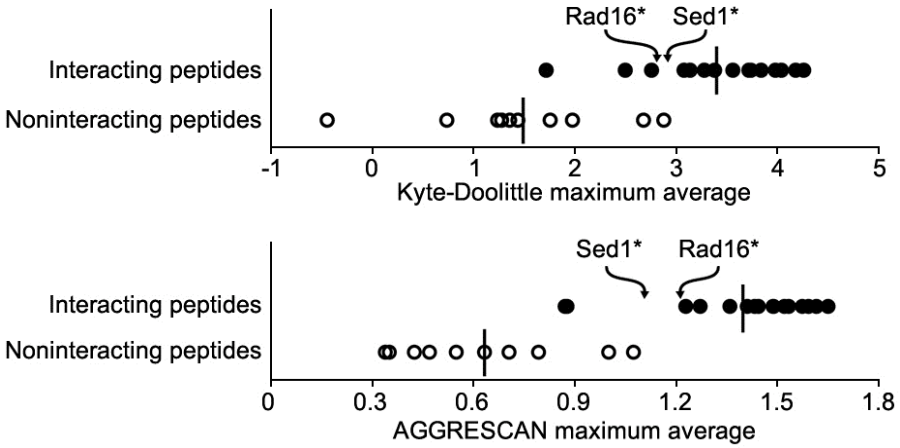


Figure 4.20: Rad16* and Sed1* in context of 5-residue hydrophobic window. Kyte-Doolittle and AGGRESCAN sliding-scale analyses were performed to determine the maximum average hydrophobic window in the hydrophobic stretches of Rad16* and Sed1*. Arrows indicate their maximum average hydrophobicity in the context of the interacting and non-interacting hydrophobic peptides shown in Figure 4.10.

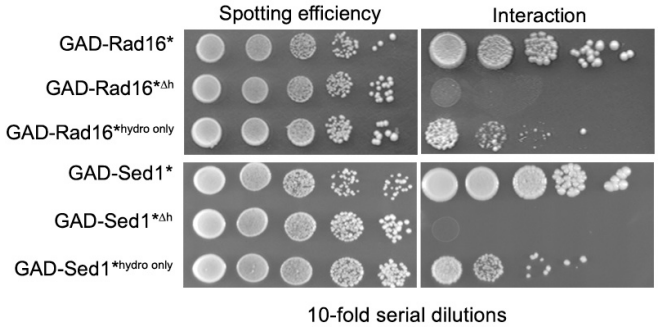


Figure 4.21: Two-hybrid between San1 and Rad16* or Sed1*. Cells expressing GBD-San1^{C279S} and the indicated GAD-Rad16*and GAD-Sed1* fusions were spotted onto media with or without histidine to measure spotting efficiency and 2-hybrid interaction. Middle rows indicate necessity (by deletion of the hydrophobic region) and bottom rows indicate sufficiency (hydrophobic region expressed alone) of substrate hydrophobic regions for San1 interaction.

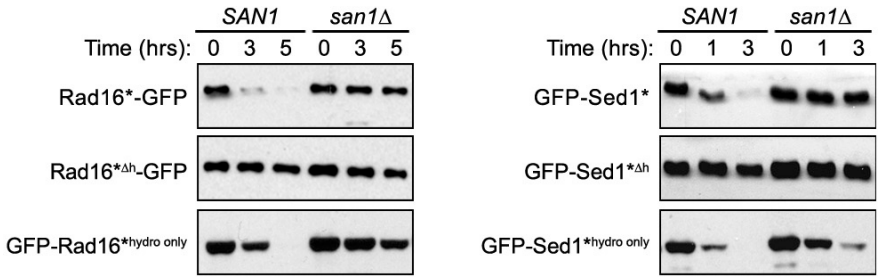


Figure 4.22: San1-dependent degradation of Rad16* and Sed1*. Cycloheximide-chase assays were performed to assess stability of the indicated GFP-Rad16* and GFP^{NLS}-Sed1*. Bottom panels indicate necessity (by deletion of the hydrophobic region) and sufficiency (hydrophobic region expressed alone) of substrate hydrophobic regions for San1 degradation.

Missense mutant San1 substrates have increased surface hydrophobicity

In addition to the peptides and truncated proteins, San1 also targets seven different proteins containing single missense mutations for degradation Sir4-9, Cdc68-1, Sir3-8, Cdc13-1, Mex67-5, Ura3-2 and Ura3-3 (Dasgupta et al., 2004; Estruch et al., 2009; Evans et al., 1998; Gardner et al., 2005; Lewis and Pelham, 2009). Most missense mutant proteins degraded via San1 retain a significant portion of wild-type function after elimination of their PQC degradation by deletion of San1 (Addinall et al., 2008; Estruch et al., 2009; Gardner et al., 2005; Lewis and Pelham, 2009; Schnell et al., 1989; Xu et al., 1993), suggesting that the functional structure of the mutant proteins is largely intact. In these cases, we hypothesized that the mutation causes a local misfolding that exposes normally buried hydrophobic residues in a part of the protein not primarily involved in the protein's function.

To test this hypothesis, we wanted to measure the surface hydrophobicity of missense mutant San1 substrates to see if the mutant versions had increased exposed hydrophobicity compared with their wild-type counterpart. For this analysis, we focused on missense mutant San1 substrates that have a crystal structure solved for the wild-type protein because we reasoned that these should be the best behaved during purification from bacteria and use *in vitro*. The yeast protein Ura3 has a known crystal structure (Chan et al., 2009), and two mutants Ura3-2 (P144L) and Ura3-3 (D243G) that are subject to San1-mediated degradation (Lewis and Pelham, 2009). Each wild-type Ura3 residue that is mutated is positioned near hydrophobic regions of at least 5 residues long buried in the structure (Figure 4.23): P144 is positioned near regions 84-89 (YNFLLF), 146-151 (GLLMLA), and 178-184 (FVIGFIA), whereas D243 is positioned near regions 197-206 (WLIMTPGVGL) and 230-234 (IIIVG). When we examined Ura3's sequence using the K-D or AGGRESCAN sliding window measures, we found that these local hydrophobic regions possess a maximum hydrophobicity window value comparable to the values in the San1-interacting peptides and truncated proteins (Figure 4.24).

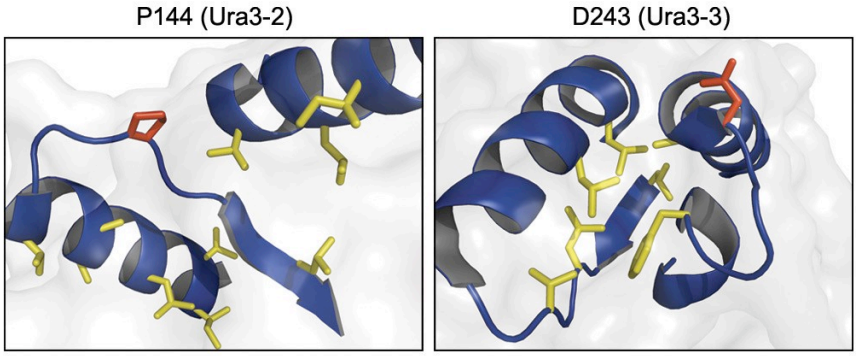
I purified Ura3, Ura3-2, and Ura3-3 from *E. coli* and Joel and I examined their surface hydrophobicity by measuring their ability to bind the dye 1-anilinonaphthalene-8-sulfonate (ANS) *in vitro*. ANS is negligibly fluorescent in an aqueous environment but binds hydrophobic residues and fluoresces in a hydrophobic environment (Hawe et al., 2008). ANS incubated with recombinant Ura3-2 and Ura3-3 produced a fluorescence

signal significantly brighter than seen with recombinant Ura3 (Figure 4.25), indicating greater surface exposed hydrophobicity in the mutants. This result correlates with the San1-dependent degradation of Ura3-2 and Ura3-3 *in vivo* (Figure 4.26), and is consistent with the hypothesis that San1 also targets exposed hydrophobicity in this type of substrate.

Although it is difficult to ascertain how a missense mutation causes misfolding by sequence and position alone (Gianni et al., 2010), we hypothesized that the misfolding might be local. Therefore, we focused our attention on the proximal hydrophobic regions located near each residue mutated in the Ura3 structure. I first examined if these mutation-proximal hydrophobic regions could function as San1 degrons *in vivo* by fusing 16mer peptides containing the five highest maximum Ura3 hydrophobic windows (Figure 4.24) to GFP^{NLS} and testing for San1-dependent degradation. We chose peptides with lengths of 16 residues so that the experiment is consistent with the peptide lengths used in Figure 4.1. I found only the GFP^{NLS} fusions with the Ura3 peptides spanning residues 140-155 and 226-241, the most proximal to each mutated residue (P144 and D243), are degraded in wild-type *SAN1* cells (Figure 4.27, left). Each is partially stabilized in the absence of *SAN1* (Figure 4.27, right), indicating that San1 can recognize these regions if they become exposed.

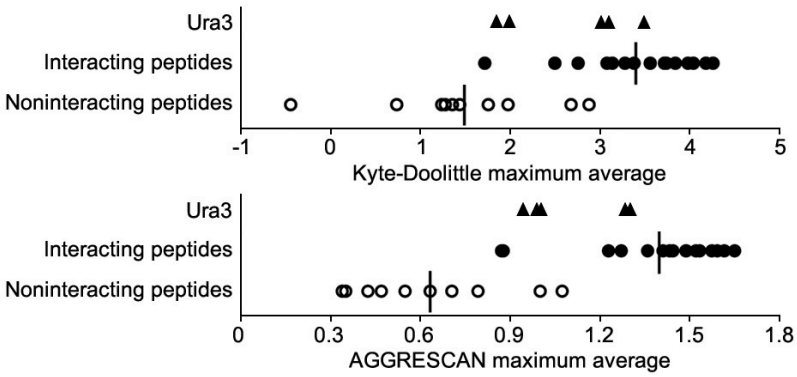
Because the peptide 226-241 was competent for San1-dependent degradation when fused to GFP^{NLS}, we decided to see if we could mutate this region near the D243G mutation to reduce the hydrophobicity in Ura3-3. Unfortunately, when I mutated the hydrophobic residues proximal to the Ura3-3 mutation (I231 and V233), I found the additionally mutated Ura3-$3^{I231A, V232A}$ protein was now subject to primarily San1-independent PQC degradation (Figure 4.28), thus hindering our ability to test the exposed hydrophobicity model in Ura3-3 through mutation. This was not unexpected, however, in that additional mutations to hydrophobic residues could cause further misfolding that now reveals an abnormality recognized by other PQC degradation systems. As an alternative to lowering hydrophobicity, I attempted to “lock” the region around the D243 mutation by introducing two cysteines that could form a disulfide bond (Clarke and Fersht, 1993). We predicted a disulfide bound would prevent exposure of hydrophobic residues near the D243G mutation in Ura3-3. To test this prediction I made two pairs of mutations based on the crystal structure, D221C/A258C and K54C/E249C,

and examined their degradation compared to Ura3-3. Both mutants Ura3-3$^{D221C,\ A258C}$ and Ura3-3$^{K54C,\ E249C}$ were subjected to San1-independent degradation (Figure 4.29), suggesting no disulfide bond formed. There are two potential reasons the locking mechanism did not work. First, the conditions used for the assay may not have been ideal for disulfide bond formation and second, the misfolding event might occur in a different region of the protein. Despite our inability to test the hypothesis in full-length Ura3-2 and Ura3-3 by mutational analyses, the rest of the data is consistent with San1 recognizing exposed hydrophobicity in missense mutant substrates.




```
  1 MSKATYKERA ATHPSPVAAK LFNIMHEKQT NLCASLDVRT TKELLELVEA
 51 LGPKICLLKT HVDILTDFSM EGTVKPLKAL SAKYNFLLFE DRKFADIGNT
101 VKLQYSAGVY RIAEWADITN AHGVVGPGIV SGLKQAAEEV TKEPRGLLML
151 AELSCKGSLA TGEYTKGTVD IAKSDKDFVI GFIAQRDMGG RDEGYDWLIM
201 TPGVGLDDKG DALGQQYRTV DDVVSTGSDI IIVGRGLFAK GRDAKVEGER
251 YRKAGWEAYL RRCGQQN
```

Figure 4.23: Hydrophobic residues near Ura3 mutations. The top panel illustrates local regions in the structure of Ura3 that surround residues P144 and D243. Red indicates the residue that is mutated in each version of Ura3 (P144 and D243), and yellow indicates surrounding hydrophobic residues. Images were produced from PDB 3GDK using macPyMOL. The bottom panel shows the amino acid sequence of Ura3 with relevant mutations highlighted in yellow and hydrophobic stretches underlined.



Ura3 region	Kyte-Doolittle maximum average	AGGRESCAN maximum average
84-89 YNFLLF	1.94	0.993
146-151 GLLMLA	3.02	0.938
178-184 FVIGFIA	3.12	1.291
197-206 WLIMTPGVGL	1.72	0.998
230-234 IIIVG	3.46	1.305

Figure 4.24: Ura3 in context of 5-residue hydrophobic window. Kyte-Doolittle and AGGRESCAN sliding-scale analyses were performed to identify hydrophobic windows in Ura3 (closed triangles, ▲). Triangles are placed in the context of the interacting and non-interacting hydrophobic peptides shown in Figure 4.10. Below, the Kyte-Dootlittle and AGGRESCAN hydrophobic window values are listed for the Ura3 region. Shown in both panels are the Ura3 regions with the top 5 maximum values.

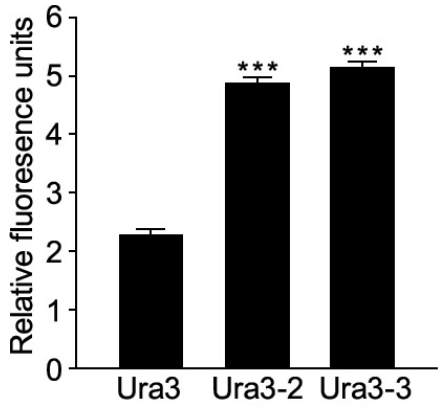


Figure 4.25: Ura3 missense mutants expose hydrophobicity. Purified recombinant Ura3 and missense variants Ura3-2 and Ura3-3 were incubated with 30 μM ANS for 1 hour and evaluated for fluorescence. Relative fluorescence was measured with excitation and emission wavelengths of 375nm and 485nm, respectively. Samples were measured in triplicate, with means and standard error shown. Statistical significance of < 0.001 was found when comparing means of Ura3 wild-type and mutant groups, and is indicated by asterisks (***).

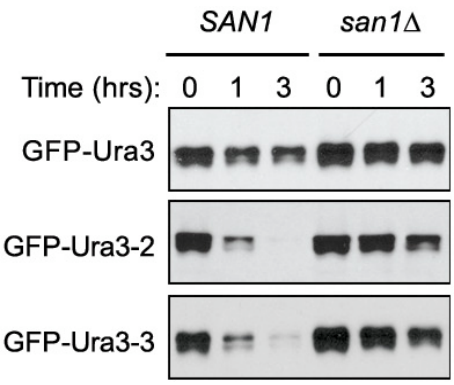


Figure 4.26: San1-dependent Ura3 missense mutant degradation. Cycloheximide-chase assays were performed to assess stability of the indicated GFP-Ura3 fusion in the presence presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies.

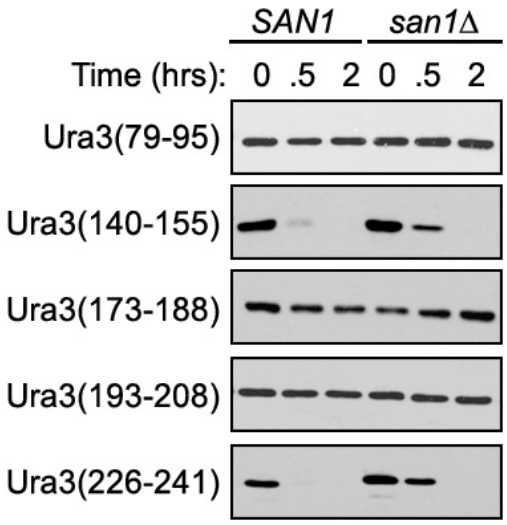


Figure 4.27: Peptides derived from Ura3 are degraded via San1. Cycloheximide-chase assays were performed to assess stability of the indicated GFP^{NLS}-Ura3 peptide fusion in the presence or absence of *SAN1*. Western blots were probed with anti-GFP antibodies. Peptides proximal to Ura3 mutations are San1 substrates.

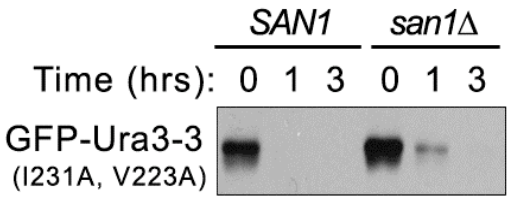


Figure 4.28: Mutation of Ura3 hydrophobicity leads to San1-independent degradation. Cycloheximide-chase assay was performed to assess stability of GFP^{NLS}-Ura3-$3^{I231A, V232A}$ fusion in the presence or absence of *SAN1*. Time after cycloheximide addition is indicated. Western blots were probed with anti-GFP antibodies.

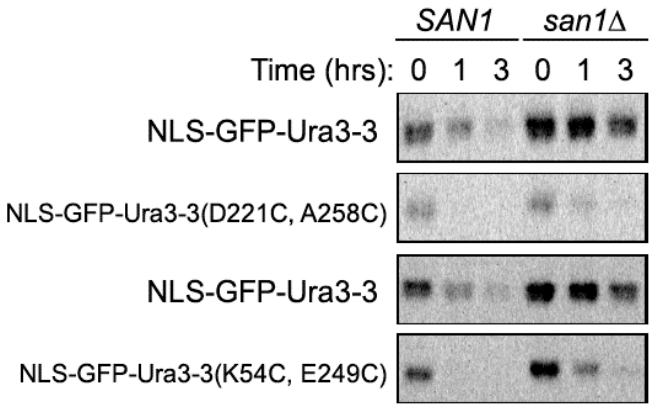


Figure 4.29: Cysteine mutations do not prevent Ura3-3 degradation. Cycloheximide-chase assays were performed to assess the stability of GFP^{NLS}-$Ura3\text{-}3^{D221C,\ A258C}$ and $Ura3\text{-}3^{K54C,\ E249C}$ fusions in the presence or absence of *SAN1*. Time after cycloheximide addition is indicated. Western blots were probed with anti-GFP antibodies.

San1 substrates are toxic in a hydrophobicity dependent manner

From an evolutionary perspective, San1 likely evolved to recognize an abnormal feature that has negative consequences if allowed to persist in the cell. Misfolded proteins tend to aggregate due to the interaction of inappropriately exposed hydrophobic surfaces (Chiti, 2006). Protein aggregation can have deleterious effects on cell viability as evidenced by the many known protein aggregation disorders (Skovronsky et al., 2006). When Rich went to verify the truncated proteins Zara found in the two-hybrid are San1 substrates and undergo San1-mediated degradation, he transformed the GAD-fusion plasmids into *san1*Δ cells. In many cases, He obtained few, if any transformants indicating that some of the GAD fusions might be toxic in San1's absence. Because the fusions are expressed from a constitutive promoter, Rich placed them behind a galactose-inducible promoter in order to regulate their expression. He examined toxicity by comparing the growth of *SAN1* and *san1*Δ cells containing each GAD fusion on media that repressed (glucose) or induced (galactose) expression. Of the 28 proteins, 11 conferred observable growth defects in *san1*Δ cells (Figures 4.30 for representative examples). As expected from San1's nuclear PQC function, the toxicity of all fusions depended upon their nuclear localization. When the NLS was removed from the GAD, the growth defects for all fusions was attenuated (Figure 4.30, bottom half).

Since San1 targets hydrophobicity, we reasoned that aggregation resulting from exposed hydrophobicity in these substrates is likely responsible for their toxic effects. We tested this hypothesis using Rad16*, which confers toxicity when expressed in *san1*Δ cells (Figure 4.31, Rad16*). Deletion of the hydrophobic region in Rad16* ameliorates its toxicity (Figure 4.31, Rad16*$^{\Delta h}$), linking exposed hydrophobicity to the detrimental effects of Rad16* expression. Using AGGRESCAN, we found that the Rad16* hydrophobic region is an aggregation "hot spot" (Figure 4.32). Concordant with this prediction, Rad16*-GFP forms an inclusion body in the cytoplasm of wild-type *SAN1* cells but is absent from the nucleus due to San1-mediated degradation (Figure 4.33, Rad16*-GFP). In *san1*Δ cells, Rad16*-GFP additionally accumulates throughout the nucleus and in a nuclear inclusion. Because Rad16* is toxic only in *san1*Δ cells (Figure 4.31), we concluded that it is the nuclear pool of Rad16* that confers toxicity, not the cytoplasmic pool. Deletion of the hydrophobic region in Rad16* results in uniform cytoplasmic and predominantly nuclear localization in both *SAN1* and *san1*Δ cells, with no observable inclusion formation (Figure 4.33, Rad16*$^{\Delta h}$). Thus, the hydrophobic region

in Rad16* targeted by San1 is responsible for both Rad16* aggregation and toxicity when present in the nucleus. In other words, San1 targets the very abnormal feature in a misfolded protein that can cause aggregation and toxicity.

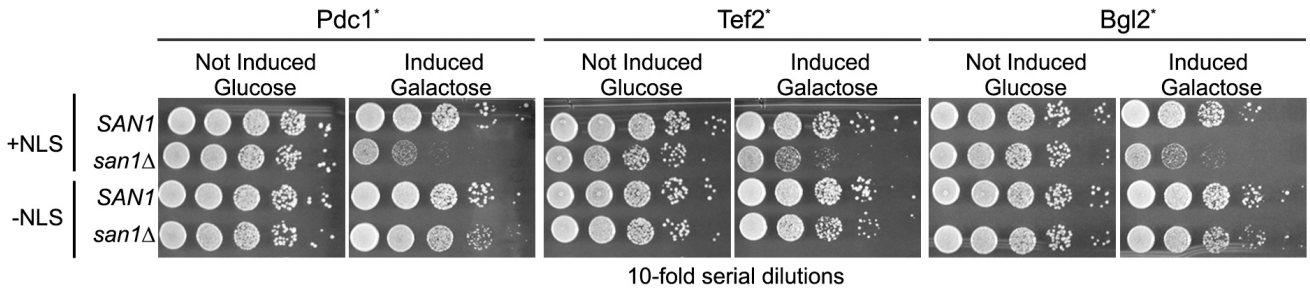


Figure 4.30: Truncated substrates are toxic in a San1-dependent manner. Wild-type *SAN1* or mutant *san1*Δ cells expressing the indicated GAD fusion, with or without an NLS, were spotted in 10-fold serial dilutions onto media with glucose to measure spotting efficiency or galactose to induce expression and toxicity.

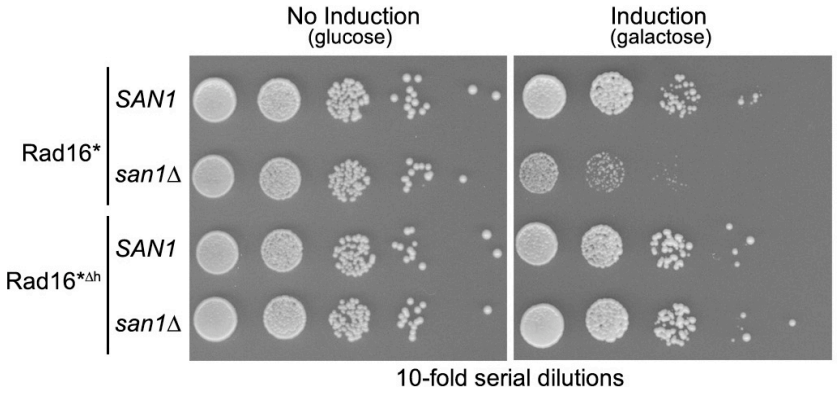


Figure 4.31: Rad16* toxicity is mediated by hydrophobic region. *SAN1* or *san1*Δ cells. expressing either Rad16* or Rad16*$^{\Delta h}$ were spotted via ten-fold dilutions onto media containing glucose to measure spotting efficiency or galactose to induce expression.

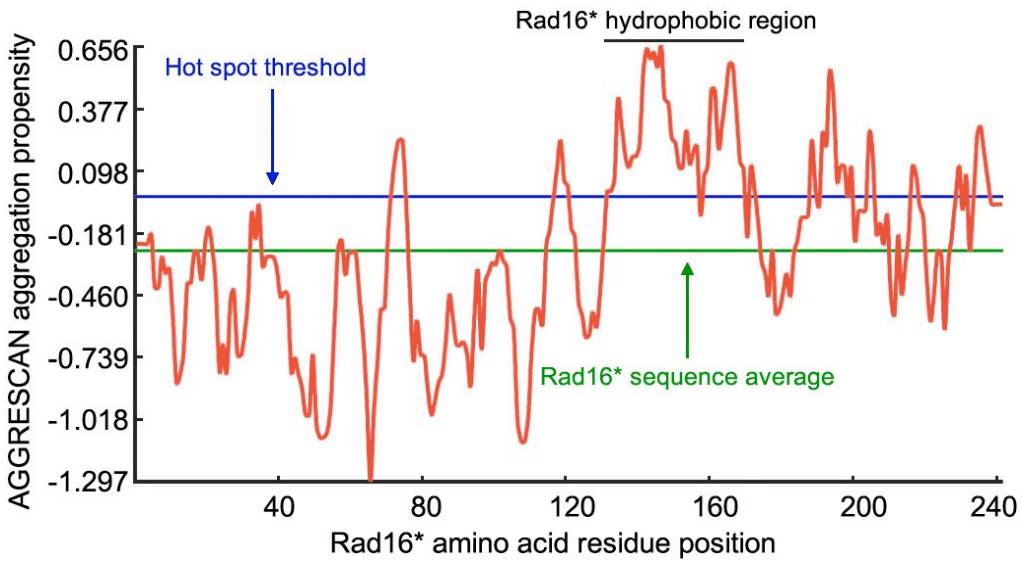


Figure 4.32: Rad16* hydrophobic region is predicted to aggregate. AGGRESCAN scores (y-axis) of Rad16* sequence (x-axis) are plotted from the N-terminus to the C-terminus. Regions with residue scores that fall above the blue threshold are considered aggregation prone.

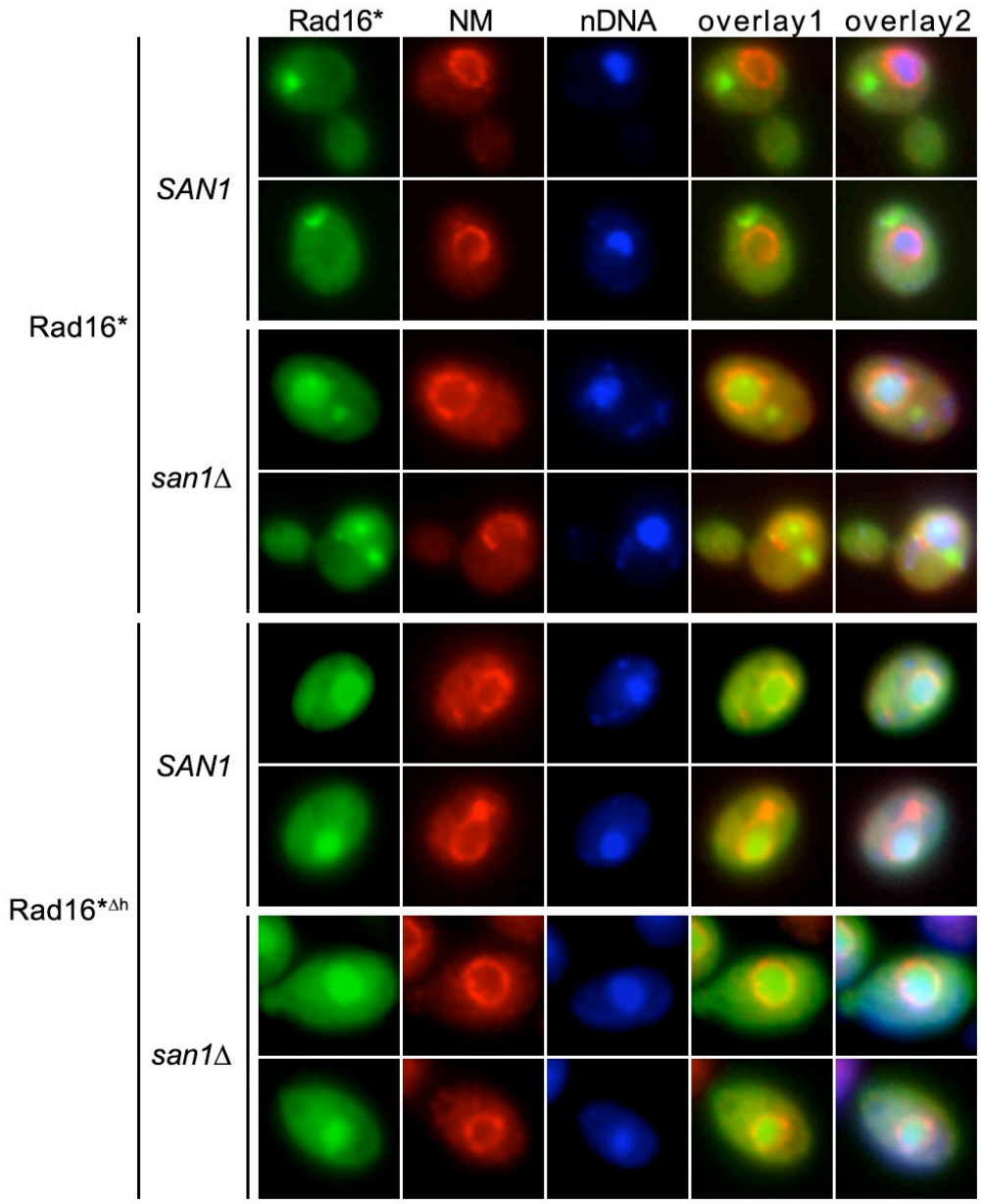


Figure 4.33: Rad16* nuclear aggregation is mediated by hydrophobic region. Expression of GFP-tagged Rad16* or Rad16*$^{\Delta h}$ was induced in *SAN1* and *san1*Δ cells by growth in galactose for 8 hours. Left panel shows GFP fluorescence, middle panels show nuclear membrane (NM) visualized by a Dbp5-DsRed fusion, and nuclear DNA (nDNA) visualized by staining with DAPI. Right two panels show overlays of the previous panels.

Discussion

For effective PQC degradation to occur in the cell, PQC ubiquitin ligases must target an abnormal feature that is commonly shared among many divergent misfolded proteins. Our analysis of San1 and its substrates indicates that exposed hydrophobicity is the feature San1 recognizes in its disparate substrates, with a window of 5 contiguous hydrophobic residues defining the minimum amount of exposed hydrophobicity required for San1-mediated degradation. To determine how common this feature is among yeast proteins, we used the PatMatch tool in the *S. cerevisiae* database (http://www.yeastgenome.org/cgi-bin/PATMATCH/nph-patmatch). The search identified 31583 instances of 5 contiguous hydrophobic residues in 2471 of 5885 proteins (42% of the queried proteins), demonstrating that this feature is very common in the yeast proteome. It is reasonable to assume that the vast majority of these hydrophobic stretches are normally buried in a protein's core, at a protein-protein interface, or within a membrane environment, and that their exposure would signify a defect in folding, assembly, or membrane insertion.

The hydrophobic preference of San1 was further defined by the degradation of contiguous stretches of identical hydrophobic residues. Homopentamer degradation demonstrated that San1 only degrades 5 Phe or Ile residues (Figure 4.11). Interestingly, the 5-residue AGGRESCAN values for 5I and 5F are higher than all other hydrophobic residues (Figure 4.11 to 4.13), suggesting San1 preferentially targets the residues most likely to aggregate. To determine if 5 contiguous Ile or Phe residues are present in the yeast genome, we again used PatMatch. The search identified five proteins with a stretch of five contiguous Ile residues and three proteins with a stretch of five contiguous Phe residues. Thus, while not common, homopentameric stretches of the hydrophobic residues San1 targets do exist in the yeast genome. Homoseptamer degradation demonstrated San1 targets Val, Leu and Tyr residues when present in longer contiguous stretches (Figure 4.17). The 5-residue AGGRESCAN hydrophobicity values of the Val and Leu residue septamers are predicted to be San1 substrates, but the Tyr residue septamer is not predicted to be a substrate. The most likely reason for this discrepancy is that the 5-residue sliding window of hydrophobicity value does not reflect the increase of local hydrophobicity in a septamer. Overall, the hydrophobic residue degradation results support the idea that there is a hydrophobicity threshold for San1.

The San1-interacting and non-interacting peptides (Figure 4.1 and 4.3) show statistically significant differences in the means of their 5-residue maximum sliding window Kyte-Doolittle and AGGRESCAN hydrophobicity values (Figure 4.10). There was some overlap, however, in the range of values for San1-interacting peptides and non-interacting peptides. In particular, San1-interacting peptides 1 and G have maximum window values that fall within the range observed for the non-interacting peptides (Figure 4.10). There are two potential reasons these peptides interact with San1. The first possibility is that these peptides contain hydrophobic residues that San1 preferentially targets. Supporting this reason is that Peptide 1 contains a stretch with Ile and Phe residues, which San1 prefers (Figure 4.1). Second, neither measure takes into account secondary structure. If a linear stretch of sequence forms an amphipathic α-helix or β-strand with hydrophobicity concentrated on one side of the secondary structure element, the peptide would not have a high maximum window value by either measure. Supporting this idea, using the JPRED 3 secondary structure predictor (Cole et al., 2008) we find that peptide 1 is predicted to form an amphipathic β-strand and peptide G is predicted to form an amphipathic α-helix (Figure 4.34).

Two of the random peptides, L and Q, possess maximum window values that are in the range of the San1-interacting peptides, but neither peptide interacts with San1 or is degraded in a San1-dependent manner. There are two possibilities why this might occur. The first is that these peptides are recognized by San1 but are also degraded by additional PQC degradation systems. In this scenario, the degradation of the GAD-peptide fusions would prevent acquisition of a sufficient steady-state level to trigger the 2-hybrid interaction with San1. Supporting this idea is that other PQC ubiquitin ligases are involved in the degradation of Peptide L (data not shown). However, the two-hybrid interaction was not retested in an *ubr1*Δ*doa10*Δ strain, so it remains unknown if the peptides would now interact with San1. The second possibility is that Peptide L and Q do not possess hydrophobic residues San1 can recognize. This possibility seems likely, as Peptide L contains Trp and Ala residues and Peptide Q contains multiple Trp residues, which are residues that San1 is unable to target (Figure 4.17).

In considering what misfolding means in the context of the Ura3 missense mutants, it is probable that overall secondary structures, including amphipathic ones, will remain relatively intact while their positioning in the Ura3 tertiary structure becomes

disrupted. In support of this, a recent mutational study was undertaken to probe the folding kinetics of the PDZ domain from the D1 C-terminal processing protease (D1p) of *Scenedesmus obliquus* (Gianni et al., 2010). Interestingly, 10 of 40 different point mutations were found to trap the D1pPDZ domain in a common misfolded intermediate, which possesses nearly the same overall secondary structure content as the native protein but the positions of its secondary structure elements are perturbed in its tertiary structure. By comparing the solvent-exposed surface area in the structure of the native protein with that of the misfolded intermediate, it was also found that the misfolded intermediate displays substantial surface exposure of hydrophobic regions that are normally buried in the native structure and distal to the mutations. This result mirrors that for the Ura3 missense mutant proteins that we used in this study, in which we observed increased surface hydrophobicity of the mutant proteins by ANS fluorescence *in vitro*. To determine what San1 actually recognizes in the Ura3 mutants we are collaborating with Jeremy England. Interestingly, through modeling, he has found the 150-180 residue range in Ura3-3 and the 50-80 and 100-125 residue ranges in Ura3-2 are likely to have increased surface exposure compared to Ura3 (data not shown). These ranges are not proximal to the mutations; demonstrating misfolding is not necessarily a local event. Definitive conclusions as to what San1 actually recognizes in the Ura3 mutants must await detailed structural studies of the mutants, but the results presented here are consistent with the hypothesis that San1 recognizes exposed hydrophobicity in misfolded missense mutant proteins.

It is generally assumed PQC degradation systems target misfolded proteins that are toxic to the cell but there are few published examples of PQC degradation substrates that are toxic. It surprised us then that toxic proteins constituted a large proportion of the San1 substrates identified by our 2-hybrid selection, particularly as this was performed without a selection for toxicity being imposed. We think the high number of toxic proteins identified in the 2-hybrid analysis reflects the fundamental reason for PQC degradation – to protect the cell from harm caused by the accumulation of misfolded proteins. Supporting this idea, the hydrophobic region Rad16* that San1 targets is also responsible for toxicity and aggregation. This underscores the functional purpose of PQC degradation systems like San1 – to recognize structural abnormalities within misfolded proteins that can harm the cell. It is well established that exposed hydrophobicity can lead to deleterious protein aggregation (Chiti, 2006). The nucleus in

particular is highly susceptible to aggregation as evidenced by the many human disorders associated with nuclear aggregation (Woulfe, 2007). From an evolutionary perspective, it is not surprising then that exposed hydrophobicity is the specific structural abnormality recognized by San1.

Peptide 1: KFHFHISALITFHEQCRYVEVKFGLQMAYLQS
Jnet: EEEEEEEEEEEE-----EEEEE--HHEEE---

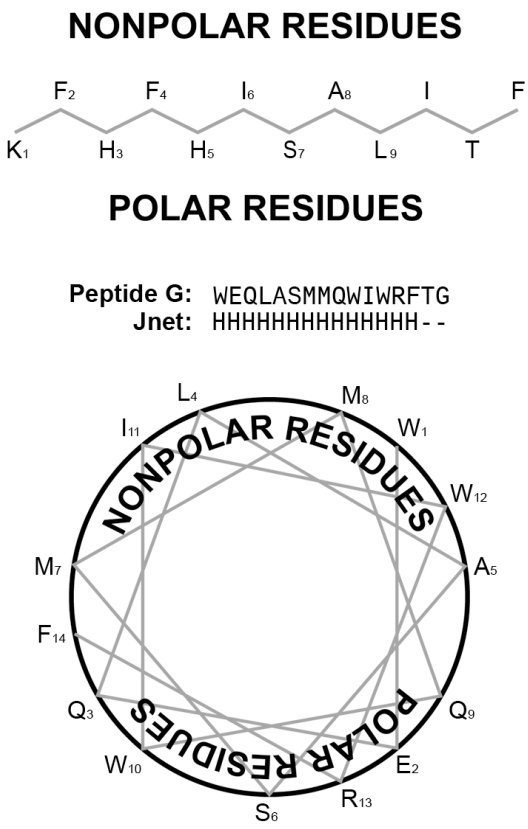


Figure 4.34: Peptide 1 and G are predicted to form secondary structures. JPRED3 predictions (http://www.compbio.dundee.ac.uk/www-jpred/) of peptide 1 and peptide G secondary structures are shown with cartoon representations of predicted amphipathic beta-strand and alpha-helical structures for each peptide.

CHAPTER V: SELF-CONTROL OF SAN1 STABILITY

Introduction

Because E3s interact with E2s charged with ubiquitin as well as substrates (Glickman and Ciechanover, 2002), there is the potential for an E3 to cause its own autoubiquitination, which can lead to the E3's proteasomal degradation. Unnecessary autoubiquitination and degradation would reduce the functional levels of active E3s, thus mechanisms must exist to prevent this possibility. Thus far, three mechanisms have been identified that minimize E3 autoubiquitination and maintain E3 stability. The first mechanism is through the interaction of the E3 with its complex partners. For example, interaction of the yeast E3 Hrd1 with its complex partner Hrd3 prevents autoubiquitination of Hrd1 and its degradation by the proteasome (Gardner et al., 2000). The second mechanism is through post-translational modification of the E3. For example, phosphorylation by protein kinase B or acetylation by CREB-binding protein decreases autoubiquitination of the mammalian E3 Mdm2 and its subsequent degradation (Feng et al., 2004; Wang et al., 2004). The third mechanism is removal of ubiquitin from the E3 by the activity of a deubiquitinating enzyme. For example, the stability of the mammalian E3 Nrd1 is dependent on the catalytic activity of USP8 (Wu et al., 2004). In each case of E3 autoubiquitination, it is not clear if the autoubiquitination is a functional means to regulate the E3, similar to the specific ubiquitination of a target substrate, or is simply due to the inherent activity of the E3.

Due to our previous findings, we decided to focus on potential self-stability mechanisms of San1. First, San1 possesses highly disordered N- and C-terminal regions that are used for misfolded substrate recognition (chapter II). In particular, San1 recognizes exposed hydrophobicity within its misfolded substrates (chapter IV). These observations lead to two questions concerning how San1 prevents its own ubiquitination and subsequent proteasome degradation. First, how does San1 prevent its highly disordered and conformationally flexible arms from *in cis* ubiquitination if/when they become positioned near the charged E2 bound by San1's RING domain? Second, if San1 is highly disordered and lacks structure, why is one San1 molecule not recognized for *in trans* ubiquitination by another San1 molecule if San1 is capable of targeting proteins that have lost their native structures? Here, we show that San1 has minimized the presence of key residues in its highly disordered regions to prevent its degradation through *in cis* and *in trans* autoubiquitination.

San1 lacks Lys residues in its disordered N- and C-terminal domains

An initial clue for how San1 might prevent *in cis* autoubiquitination came from examining the amino acid characteristics of its disordered N- and C- terminal domains. I noticed that the disordered N- and C-terminal domains of San1 were devoid of Lys residues. The 13 Lys residues that are present in San1 are clustered in or near the RING domain (Figure 5.1, top). Based on codon usage, a yeast protein the size of San1 is predicted to have 44 Lys residues (http://www.yeastgenome.org), which suggests San1 has a significant reduction of Lys residues. To determine if a lack of Lys residues is a common feature among PQC E3s, I examined the Lys residue content of the E3s CHIP, Ubr1, Hrd1 and Doa10. Each E3 contained a number of Lys residues more consistent with codon usage and also a more even distribution of Lys residues compared to San1. A key difference between San1 and these other PQC E3s is their disorder content based on disorder prediction algorithms. Thus I predicted intrinsic disorder is the reason for the reduced Lys residue content in San1.

To test my prediction, I plotted San1's disorder using PONDR-VLT (Dunker et al., 2008) and examined the distribution of Lys residues. Supporting my hypothesis, I found the regions in San1 without Lys residues are all highlighted by a high degree of intrinsic disorder (Figure 5.1, bottom). Interestingly, the Lys residues that exist in San1 are also in a disordered region. However, these Lys residues make up the nuclear localization signal within the RING domain and probably need to be in a region of disorder for proper nuclear import. Because ubiquitin is covalently attached to the free amino group on Lys residues, I predicted the lack of Lys residues in the N- and C- terminal disordered regions of San1 is to prevent *in cis* autoubiquitination.

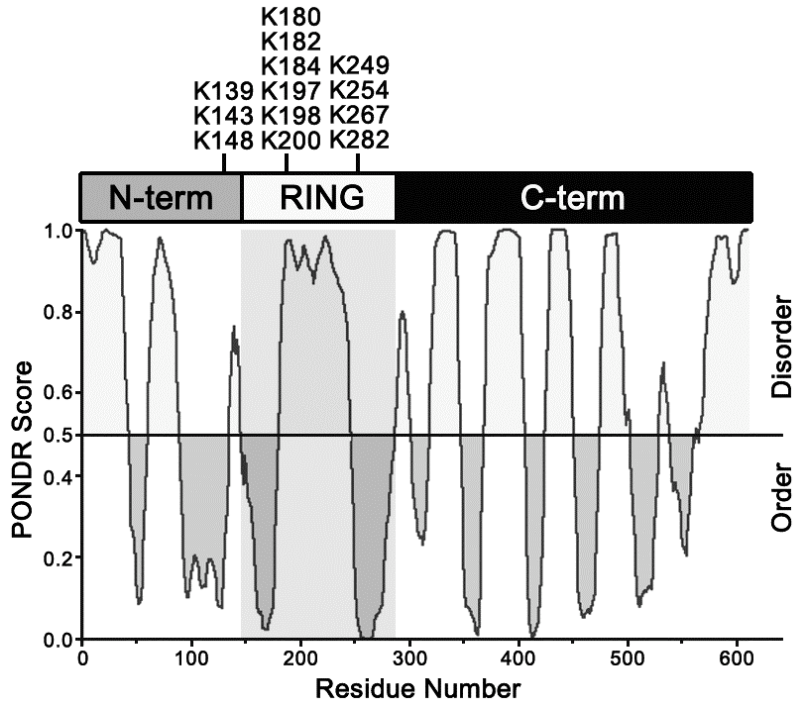


Figure 5.1: Lys residue layout in San1. San1 topology, with Lys residue positions marked (top), aligned with PONDR evaluation of San1's disorder (bottom). Light gray indicates disorder and dark gray indicates order. Gray box highlights the RING domain. The disordered N- and C- terminal substrate binding regions lack Lys residues.

Introduction of Lys residues in San1's disordered N- and C-terminal domains causes degradation via an *in cis* mechanism

I predicted that if the lack of Lys residues in San1's disordered substrate binding domains prevents autoubiquitination, then the addition of a single Lys residue in these regions would disrupt San1's stability. To test my prediction, I examined the stability of single Arg or Asn to Lys residue point mutants (+Lys) in San1's substrate binding domains (San1^{N13K}, San1^{N332K}, San1^{N444K}, and San1^{R556K}). To monitor their stability, each San1^{+Lys} mutant was tagged with 3HSV, which is functional in the degradation of Sir4-9 (Gardner et al., 2005). To test the stability of the San1^{+Lys} mutants, I examined their degradation in *SAN1* and *san1*Δ cells. Each San1^{+Lys} mutant was rapidly degraded in both the presence and absence of endogenous San1 (Figure 5.2). The instability was due to the additional Lys residue because wild-type San1-3HSV was stable (Figure 5.2, top). This result indicates San1 stability is regulated through a lack of Lys residues in its disordered N- and C- terminal domains.

Next, I wanted to determine the mechanism by which San1^{+Lys} mutants are degraded. I predicted degradation of the San1^{+Lys} mutants was either *in cis*, where a single San1 molecule ubiquitinates a Lys residue within its own disordered domains, or *in trans*, where one San1 molecule ubiquitinates a Lys residue in the disorder of another San1 molecule (Figure 5.3). If degradation was *in cis*, I predicted a catalytically inactive San1^{+Lys} mutant would be stable in presence of wild-type San1 (Figure 5.3, model). Thus I introduced a catalytically inactive mutation (R280A) to the San1+Lys mutants and examined their degradation in *SAN1* and *san1*Δ cells. Each mutant was not degraded in the presence of wild-type San1 and this stability matched that of San1^{R280A} (Figure 5.4), demonstrating an *in cis* degradation mechanism for the San1^{+Lys} mutants.

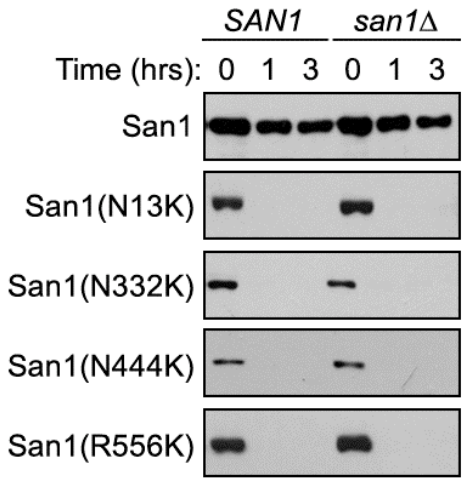


Figure 5.2: Introduction of Lys residues in San1 disorder alters stability. Cycloheximide-chase assays were performed to assess the stability of San1^{N13K}, San1^{N332K}, or San1^{R556K} in the presence or absence of *SAN1*. Time after cycloheximide addition is indicated. Western blots were probed with anti-HSV antibodies. Each San1 mutant is degraded in both *SAN1* and *san1*Δ cells.

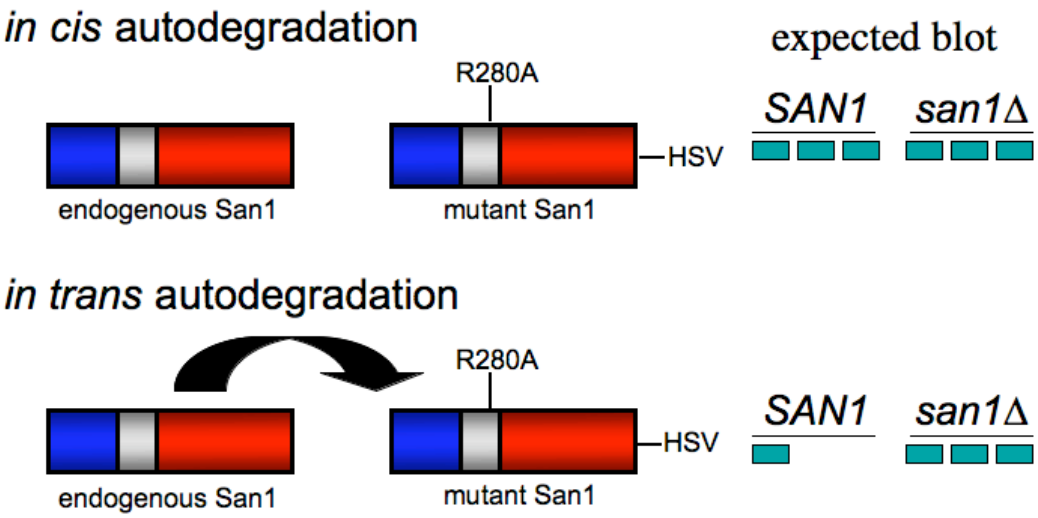


Figure 5.3: Model of *in cis* vs *in trans* San1 autodegradation. Top represents *in cis* autodegradation. If endogenous San1 cannot target mutant San1, catalytically inactive (R280A) mutant San1 will be stable. Bottom represents *in trans* autodegradation. If endogenous San1 can target mutant San1^{R280A}, mutant San1 will be degraded only in the presence of endogenous San1. The expected anti-HSV blot for each mechanism is indicated on the right.

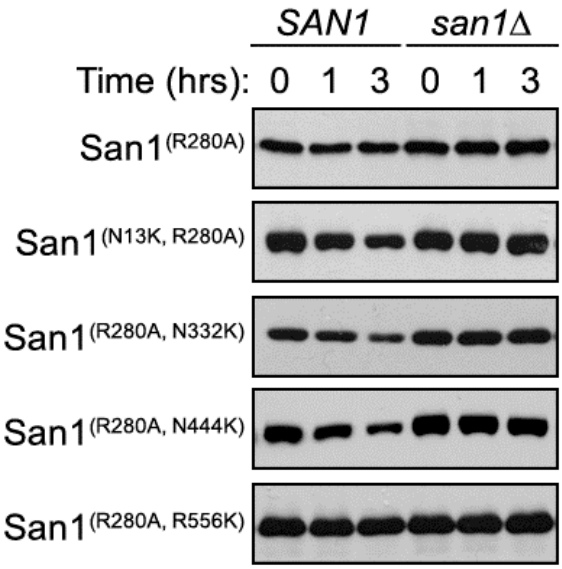


Figure 5.4: San1 with a Lys residue in a disordered region are degraded *in cis*. Cycloheximide-chase assays were performed to assess the stability of San1^{N13K}, San1^{N332K}, or San1^{R556K} coupled with R280A (catalytically inactive) in the presence or absence of *SAN1*. Time after cycloheximide addition is indicated. Western blots were probed with anti-HSV antibodies. Each San1^{+Lys} mutant is stable.

San1 has minimal hydrophobicity in its disordered N- and C-terminal domains

An initial clue for how San1 might prevent *in trans* autoubiquitination also came from examination of the residue content in San1's disordered N- and C- terminal domains. I noticed San1 had a low hydrophobic content in these regions, which was apparent when I plotted Kyte-Doolittle hydrophobicity values along San1's sequence length (Figure 5.5). A key feature of intrinsically disordered proteins is a low hydrophobic content (Tompa, 2002), which fits my San1 observation. I predicted low hydrophobic content in the disordered regions of San1 is important for autodegradation prevention, because San1 targets exposed hydrophobicity in its substrates (chapter IV). To examine local San1 hydrophobicity more quantitatively, I used Kyte-Doolittle hydrophobicity or AGGRESCAN sliding window of hydrophobicity measures (Figure 4.10). Surprisingly, the maximum 5-residue hydrophobicity window values within San1's disordered N- and C- terminal domains are only slightly below the average for San1-interacting peptides (Figure 5.6). However, these hydrophobic regions within San1 are not sufficient for degradation, which could due to a number of reasons. First, the sequences of the maximum 5-residue hydrophobic window values are "LVLSF" and "FFLVP". The lack of five contiguous hydrophobic residues in both windows is significant because San1 can target five contiguous hydrophobic residues, but not four (Figure 4.9). This lack of five contiguous hydrophobic residues within San1 may explain why San1 does not autoubiquitinate another San1 molecule *in trans*. Another possibility is that the hydrophobicity within San1 may not be exposed as it is in misfolded substrates, which would prevent recognition by another San1 molecule.

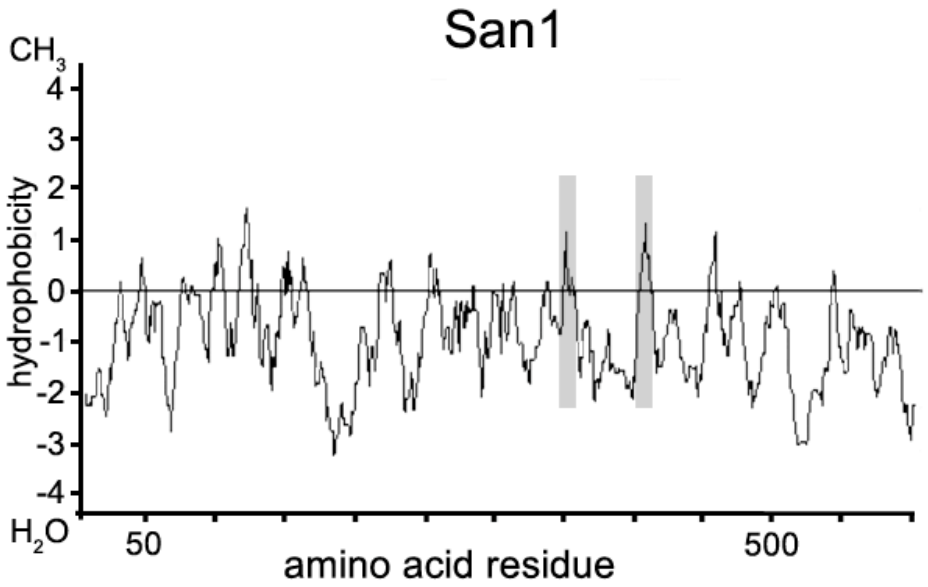


Figure 5.5: San1 contains little hydrophobicity. Kyte-Doolittle hydrophobicity evaluation of San1 is shown. Two of the most hydrophobic regions are highlighted in gray. Overall San1 contains very little hydrophobicity.

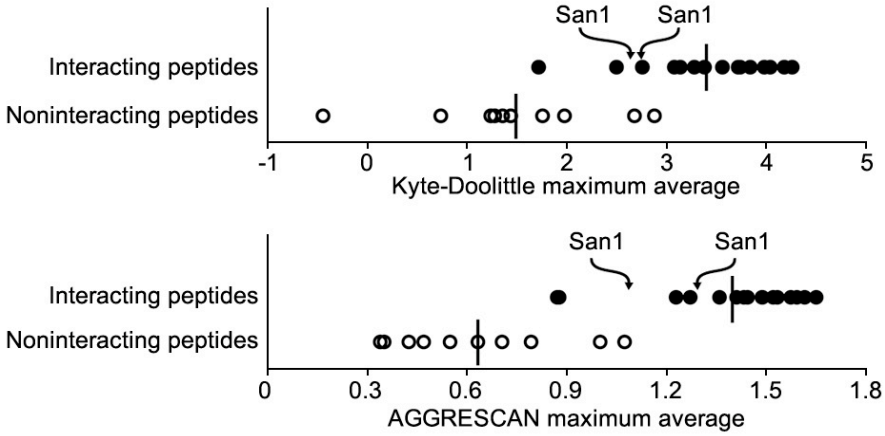


Figure 5.6: San1 in context of 5-residue hydrophobic window. Kyte-Doolittle and AGGRESCAN sliding-scale analyses were performed to identify hydrophobic windows in San1. Arrows are placed in the context of the interacting and non-interacting hydrophobic peptides shown in Figure 4.10.

Introduction of an appropriate hydrophobic region in San1's disordered N- and C-terminal domains leads to degradation via an *in trans* mechanism

If low hydrophobicity in San1's disordered N- and C- terminal domains prevents San1 *in trans* autodegradation, I predicted adding hydrophobicity to these regions would alter San1's stability. To test our prediction, I increased the number of contiguous hydrophobic residues in two regions of San1's C-terminal domain that have high window of hydrophobicity values (grey boxes, Figure 5.5). Mutation of two or three non-hydrophobic residues to hydrophobic residues in these regions resulted in the highly hydrophobic sequences of "FFLVLIV" and "WVLLILTLF". Similar to the San1^{+Lys} mutants, these "plus hydro" mutants contained a 3HSV tag to assess stability. To determine if the plus hydro San1 mutants are unstable via *in trans* autodegradation, I examined the degradation of catalytically inactive (R280A) San1$^{P411L,\ I412L,\ P413I}$ and San1$^{P361L,\ Q362I,\ N363V}$ in *SAN1* and *san1*Δ cells. Both plus hydro San1 mutants were slightly degraded in *SAN1* cells, but stable in *san1*Δ cells (Figure 5.7), suggesting an *in trans* mechanism. Overall, these results support the idea that low hydrophobicity in San1's C-terminal disordered domain prevents *in trans* autodegradation.

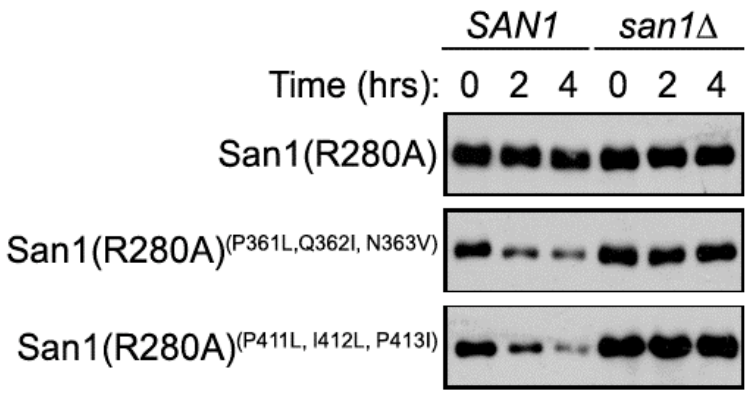


Figure 5.7: *in trans* degradation of San1 "plus hydro" mutants. Cycloheximide-chase assays were performed to assess the stability of San1$^{P411L,\ I412L,\ P413I}$ and San1$^{P361L,\ Q362I,\ N363V}$ coupled with R280A (catalytically inactive) in the presence or absence of *SAN1*. Time after cycloheximide addition is indicated. Western blots were probed with anti-HSV antibodies. Each mutant is slightly degraded in a San1-dependent manner.

Some Lys substitutions in San1's disordered N- and C-terminal domains cause degradation via an *in trans* mechanism

While testing the stability of San1^{+Lys} mutants, we noticed not all mutations behaved equally. The degradation of the four San1^{+Lys} mutants in Figure 5.3 was *in cis*, but other San1^{+Lys} mutants were different. The catalytically inactive (R280A) mutants San1^{N71K}, San1^{R104K}, San1^{R354K} were degraded rapidly in the presence of *SAN1*, but stable in *san1*Δ cells (Figure 5.8), suggesting *in trans* autoubiquitination. Different degradation mechanisms for San1^{+Lys} mutants suggest the location of the Lys residue within San1's disordered N- and C- terminal domains is important. Because the addition of hydrophobicity leads to *in trans* autodegradation (Figure 5.7), we hypothesized the *in trans* San1^{+Lys} mutants cause the exposure of hydrophobicity. Supporting our hypothesis, each *in trans* San1^{+Lys} mutant is located near more hydrophobic residues compared to the *in cis* San1^{+Lys} mutants (our observation). In fact, three of the highest 5-residue window of hydrophobicity values in San1 exist near the *in trans* San1^{+Lys} mutants (Figure 5.8, right). Future work will test our hypothesis that the *in trans* San1^{+Lys} mutants expose hydrophobicity.

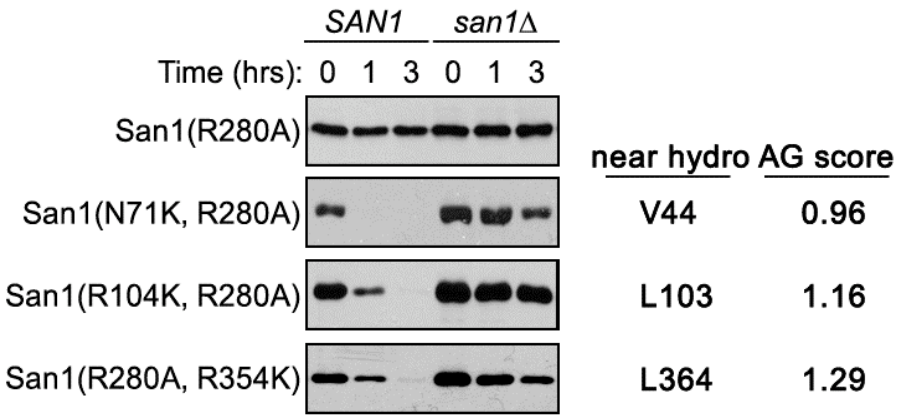


Figure 5.8: *in trans* degradation of San1 mutants with an additional Lys residue. Cycloheximide-chase assays were performed to assess the stability of San1^{N71K}, San1^{R104K}, or San1^{R354K} coupled with R280A (RING mutation) in the presence or absence of *SAN1*. Time after cycloheximide addition is indicated. Western blots were probed with anti-HSV antibodies. Each San1^{+Lys} mutant is degraded in a San1-dependent manner. Hydrophobic residues near the added Lys residue with the highest 5-residue hydrophobic window AGGRESCAN value are listed on the right.

Lack of Lys residues in San1's disordered N- and C-terminal domains is conserved

While my results demonstrated *S. cerevisiae* San1's lack of Lys residues in its N- and C- terminal disordered domains are critical for self-stability, it was unclear if this feature exists in other disordered PQC E3s. To date, no other disordered PQC E3s have been identified, so to determine the importance of a lack of Lys residues, I examined San1 homologs in other fungal species. We previously demonstrated *S. bayanus* and *S. mikatae* San1 homologs complement *S. cerevisiae* function (Figure 2.12 and 2.13). Additionally, we identified San1 genes in *S. castellii, Ashbya gossypii*, and *Candida albicans.* Lastly, a homolog of San1 was recently identified and characterized in *S. pombe* (Matsuo et al., 2011). This collection of homologs ranged from closely related (*S. bayanus*) to distantly related (*S. pombe*). I predicted if a lack of Lys residues in disordered regions is important for self-stability, it would be conserved in distant homologs. Before examining the Lys residue content, I plotted the order/disorder topology of each homolog using PONDR. Each homolog is predicted to contain large regions disorder interspersed within regions of order, similar to *S. cerevisiae* San1 (Figure 5.9). Next, I examined the Lys residue content in each homolog. As with *S. cerevisiae*, each homolog contained very few Lys residues and those that did exist were contained in or near the RING domain (Figure 5.9). The one outlier is *S. pombe*, which contains three Lys residues far from the RING domain. One explanation is that this region may not be important for substrate interaction. Supporting this explanation, the region with the three Lys residues contains the remnants of another RING domain (Cys and His residues). Such a region does not exist in any of the other homologs. Overall, the conservation of a lack of Lys residues within disordered regions among San1 homologs supports the idea that this feature is critical for the stability of disordered E3s.

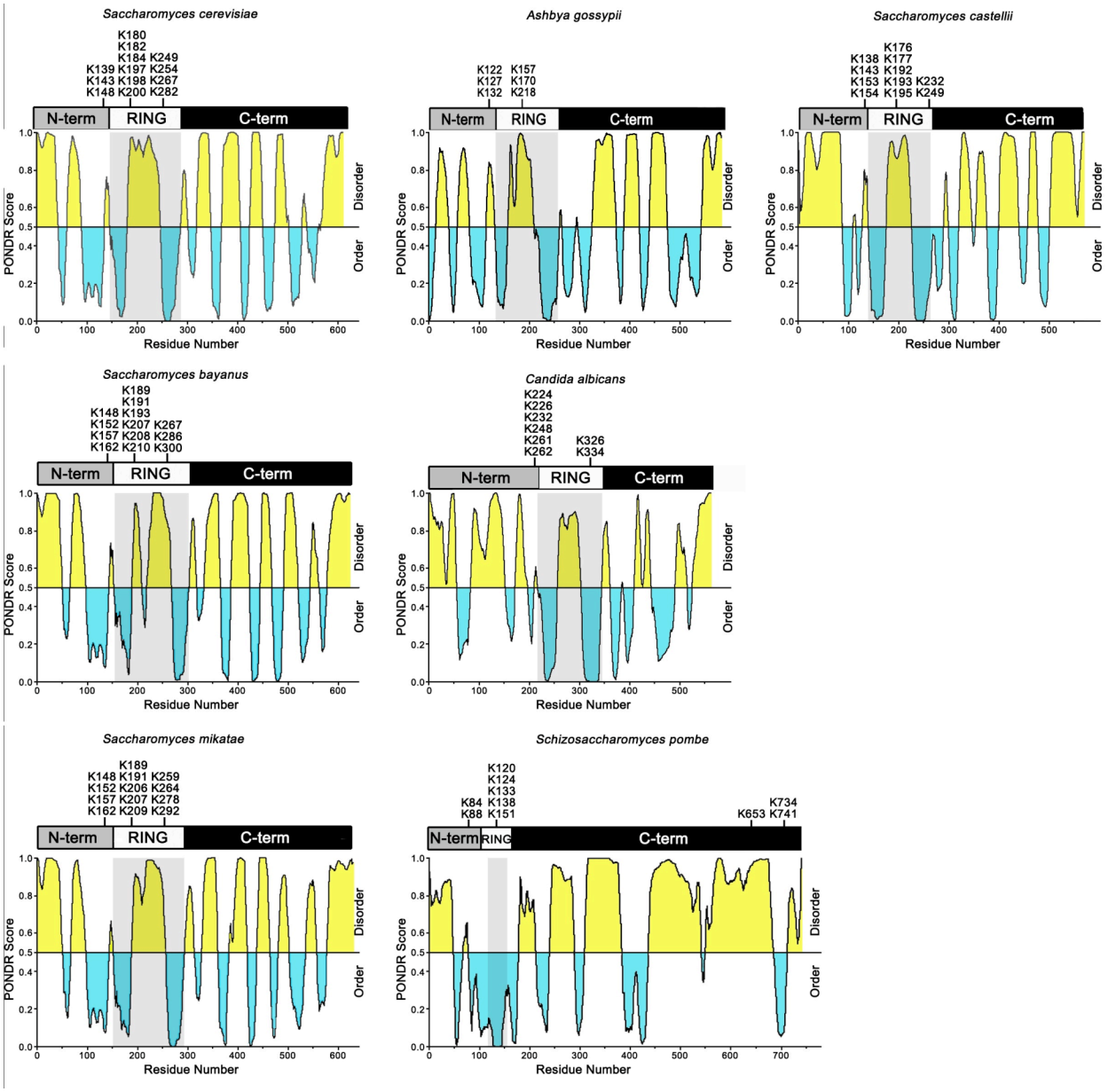


Figure 5.9: Lack of Lys residues and disorder is conserved among San1 homologs. As in Figure 5.1, San1 topology with Lys residues indicated by their position (top) aligned with PONDR evaluation of San1's disorder (bottom). Yellow indicates disorder and blue indicates order. *S. bayanus*, *S. mikatae*, *S. castellii,, Ashbya gossypii*, *Candida albicans,* and *Schizosaccharomyces pombe* were each evaluated in the same manner. All Lys residues fall in the RING domain or ordered region in the N-terminal domain, except *S. pombe*, which contains three Lys residues in the C-terminus.

Discussion

For effective ubiquitination to occur, an E3 needs to maintain its stability. These San1 findings identified two previously unknown mechanisms by which an E3 maintains its stability. First, San1 lacks Lys residues in its disordered N- and C- terminal domains. Each Lys residue that does exist in San1 resides in or near the RING domain (Figure 5.1). This lack of Lys residues in disordered regions prevents *in cis* autodegradation (Figure 5.2 and 5.4). I think the lack of Lys residues in San1's substrate binding regions suggests other E3s might lack Lys residues in their substrate binding domains. Any region of an E3 that could become close to the catalytic site upon substrate binding might also lack Lys residues. Testing this is difficult, as the full structures of most E3's are unknown. Additionally, I think San1's lack of Lys residues in its disordered domains suggests other disordered proteins might be depleted of amino acids that could be improperly modified. In disordered proteins, there is no folded core and thus all residues are surface exposed and likely susceptible to improper modification. This differs from structured proteins, where a modifiable residue could be inaccessible in a structure and thus not improperly modified. Thus a lack of certain modifiable residues in intrinsically disordered proteins might be essential, especially if the modification improperly alters the function of the protein. My finding that the addition of a Lys residue in San1's disordered regions leads to rapid degradation supports this idea. It remains unclear if other disordered proteins are devoid of certain amino acids, like San1.

The second mechanism of E3 stability identified using San1 is that an E3 lacks the very feature it targets in substrates. I found San1 lacks five contiguous hydrophobic residues in its N- and C- terminal disordered domains, which is the feature San1 targets in substrates. From this observation, I predicted that if an E3 contains the feature that it targets in its substrates, it would see itself as a substrate. Supporting this idea, placement of longer stretches of hydrophobicity in San1 led to *in trans* autodegradation (Figure 5.7). It is unclear if other E3s lack the feature or sequence they recognize in substrates, but most E3s are structured, where the feature or sequence might be inaccessible and thus not targeted as a substrate. An open question in E3 autoubiquitination is if autoubiquitination is similar to substrate recognition, where a specific feature or sequence is targeted. These results with San1 suggest autoubiquitination can be similar to substrate recognition (*in trans*) and also very different (*in cis*).

Both a lack of Lys residues and a pattern of disorder/order are conserved among fungal homologs of San1 (Figure 5.9). To date, no E3 in higher organisms has been identified that functions like San1. One issue with identifying a homolog to San1 is that highly disordered proteins have low conservation of linear sequence, which is apparent in the fungal homologs of San1 (our unpublished observation, Figure 2.15). A lack of linear sequence conservation would confound current homology analyses (Brown, 2002), thus a different approach is necessary. Because a lack of Lys residues and an order/disorder topology are conserved across fungal homologs, I predict higher organism homologs may also have these features. To determine if these features exist in higher organisms, I examined the Lys residue content of all human proteins predicted to have a RING domain. I identified multiple proteins with a reduced number of Lys residues that have large regions of disorder. It remains unclear if any candidate identified functions in PQC degradation, but searching for these features appears to be a feasible strategy for homolog identification. Overall, the conserved sequence features in San1 demonstrate sequence characteristics in disordered proteins are probably crucial for homolog identification.

CHAPTER VI: CONCLUSIONS

Overall, the research in this dissertation elucidated much about how nuclear PQC degradation functions, which differs from much of the previous ubiquitin-mediated PQC degradation work. First, we demonstrated the nuclear PQC ubiquitin-protein ligase San1 directly binds misfolded proteins using disordered N- and C- terminal domains. Next, we demonstrated there is not a universal chaperone requirement for San1-mediated degradation. Then we showed San1 targets exposed hydrophobicity in multiple classes of misfolded proteins. San1 can target as few as five contiguous hydrophobic residues and recognizes Phe and Ile residues most readily, which are the two residues most likely to aggregate. Additionally, we found San1 protects the nucleus from toxic proteins and that San1 recognizes the very feature that confers toxicity. Lastly, we discovered two features of San1 that maintain San1's stability, a lack of Lys residues and reduced hydrophobicity in the N- and C- terminal disordered domains. While each aspect of this research furthered our understanding of how San1 functions in nuclear PQC degradation, many questions remain, such as: Are there other functions of San1? What does San1-substrate interaction look like at the molecular level? Are there normal proteins that are substrates of San1? Future studies will address these questions among others.

One concern in the PQC degradation field is that much of our knowledge is based on one or two substrates that are grossly misfolded and overexpressed. With San1, many of the substrates we identified were grossly misfolded and overexpressed, but the missense mutant substrates were not. However, we still do not know the actual substrates of San1 *in vivo*. One idea is that a small portion of most nuclear proteins is ubiquitinated via San1, solely for PQC degradation. The ability to query the folded state of individual members of a protein pool over time would help answer this question. Another finding that supports the idea of San1 targeting portions of proteins pools is that San1 expression is repressed during heat shock and increased upon a return to room temperature (Gasch et al., 2000). Higher temperatures increase protein misfolding and thus San1 would degrade more proteins. Perhaps San1 is repressed to give chaperones the chance to refold proteins and then increased to degrade any misfolded proteins that were unable to be refolded. Another substrate possibility is that San1 degrades proteins for non-PQC functions within the nucleus. This seems somewhat unlikely, but there were

nucleolar proteins identified in the IP-Mass spec experiment that could be San1 substrates or interactors (Figure 2.3).

One observation from these studies is that not all San1 substrates are completely stable in the absence of San1, which suggests other E3s exist in the nucleus of *S. cerevisiae*. I previously described experiments with the E3s Ubr1 and Doa10 (chapter V), which are able to degrade portions of certain San1 substrates. It is difficult to determine where the degradation of these substrates occurs. The issue is that San1 substrates are translated in the cytoplasm and then transported into the nucleus via importins. It's probable that Ubr1, a cytoplasmic PQC E3, targets a pool of San1 substrates before they reach the nucleus. Even in the absence of Ubr1, Doa10 and San1, we noticed some substrates were still not completely stable. Future studies are needed to identify other potential nuclear PQC E3s and determine if Ubr1 and Doa10 are involved in nuclear PQC degradation.

An important next step in nuclear PQC research is determining if an E3 like San1 exists in higher organisms. The mammalian E3s PML-IV and URHF-2 have been implicated in mammalian nuclear PQC degradation San1 (Fu et al., 2005; Iwata et al., 2009; Janer al., 2006), but neither possesses the disordered topology of San1. Thus, a true San1 functional homolog has yet to be discovered in mammals. As described in chapter V, perhaps the use of sequence features will be crucial in finding a homolog of San1. It's likely other disordered proteins have features that are conserved, even if primary sequence is not, and these features may help identify homologs to these IDPs.

APPENDIX 1: MATERIALS AND METHODS

2-hybrid Assays

All interaction tests were performed in duplicate using two independent isolates. Growth of cells expressing a GAD fusion and a GBD-San1 deletion was scored relative to growth of cells expressing the identical GAD fusion and full-length GBD-San1 on each different stringency plate. Growth of cells expressing a San1 deletion that resulted in colony sizes larger than cells expressing full-length San1 was scored 5. Growth similar to cells expressing full-length was scored 4. Growth resulting in smaller colony sizes or reduced 10 fold in spotting was scored 3. Growth reduced 10^2-10^3 fold was scored 2. Growth reduced $\geq 10^4$ fold was scored 1. For the dendogram in Figure 2.24, an agglomerative hierarchical clustering was computed for the interactions on -his+3AT media using Euclidean distance to quantify the dissimilarity of substrates and Ward linkage to quantify the dissimilarity of clusters. The hierarchical clustering package of Mathematica (Version 7.0, Wolfram Research Inc) was used to generate the dendrogram.

Mass spectrometry

Independent 3HSV-San1^{C279S} or untagged San1^{C279S} cultures were grown to ~1x10^7 cells/ml. Formaldehyde was added to 1% (v/v) and cells incubated for 2 minutes. Glycine was added to 125mM to quench crosslinking. Cells were lysed in SUME (8M urea, 1% SDS, 10mM MOPS, pH6.8, 10mM EDTA). Lysates were diluted 1:5 in IP buffer (15mM Na_2HPO_4, 150mM NaCl, 2% Triton X-100, 0.1% SDS, 0.5% deoxycholate, 10mM EDTA) and incubated for 16 hours at 4˚C with 1:1000 mouse anti-HSV antibody (Novagen) bound to 1.25mg/ml Protein A Dynabeads (Invitrogen). Beads were washed 3x in IP buffer and 2x in IP wash buffer (50mM NaCl, 10mM Tris-Cl, pH7.5). Proteins were eluted by incubation at 65˚C for 10 minutes in SUMEB (SUME + 0.01% bromophenol blue). Samples were run 1cm into an 8-16% SDS-PAGE gel and gel slices excised. Proteins in the slices were trypsin digested and peptides identified by tandem MS using an Orbitrap mass spectrometer (Fred Hutchinson Cancer Research Center Proteomics facility). We used Peptide Prophet (Keller et al., 2002) to evaluate the validity of peptide identifications eliminating peptides with a score <0.85 (<2.5% false discovery rate).

***In vitro* ubiquitination**

0.27μM luciferase in 50mM Tris-HCl (pH7.5) was denatured by incubating at 23˚C for 10 minutes in 4M urea. 15μl reactions containing 0.08μg ubiquitin activase (Uba1), 0.1μg ubiquitin conjugase (UbcH5a), 0.76μg San1, 2.5μg ubiquitin, 2mM Mg-ATP, 50mM Tris-HCl (pH7.5), 2.5mM $MgCl_2$, and 0.9μg denatured or native luciferase were incubated at 30˚C for 60 minutes. For heat denaturation, the reaction was incubated at 42˚C. 15μl of SUMEB was added to each sample and incubated at 65˚C for 10 minutes to stop the reaction.

***In coli* ubiquitination**

Components of the San1 ubiquitination cascade were cloned behind T7*lac* promoters in Duet expression vectors (Novagen) that were maintained in T7 Express competent *E. coli* (New England Biolabs) with appropriate antibiotic selection. Cells were grown at 37˚C to OD_{600} ~1.0 and induced with 800μM IPTG for 6 hours at 30˚C. Cells were lysed in Bugbuster (Novagen) with 10mM PMSF. To immunoprecipitate the 2xHA-tagged substrates, soluble extract was incubated for 16 hours at 4˚C with 1:2000 mouse anti-HA antibodies (Sigma) conjugated to 0.75mg/ml Protein A Dynabeads (Invitrogen) in IP buffer. Beads were washed 3x in IP buffer then incubated at 65˚C for 10 minutes in SUMEB.

Expression and purification of San1

StrepII-San1-3HSV-8His was cloned into pACYCDuet and transformed into BL-21 cells. Cells were grown in LB + chloramphenicol to an OD_{600} ~1.0 and induced with 600μM IPTG overnight at 16˚C. Cells were harvested and lysed in BugBuster with PMSF, aprotinin, leupeptin, and pepstatin A. Soluble extract was applied to TALON resin (Clontech) equilibrated in a 300mM NaCl, 50mM Na_2HPO_4 pH8.0 buffer and eluted with 1M imidazole added to the buffer. The elution fraction was applied to Strep-Tactin resin (Novagen) equilibrated with a 150mM NaCl, 100mM Tris-HCl, 1mM EDTA, pH8.0 buffer. San1 was eluted with 2.5mM desthiobiotin and concentrated using an Amicon Ultracel concentrator (cut-off 30 kDa). Purified San1 was centrifuged at 100,000xg for 45 minutes to eliminate aggregates prior to gel filtration.

Limited Proteolysis

Proteolysis of San1 and BSA was carried out by trypsin and thermolysin digestion at 25˚C in digestion buffer (50mM NaCl and 15mM Na_2HPO_4, pH7.3). The enzyme to substrate ratio was 1:50. Proteolysis was stopped by addition of SUMEB and boiling for 10 minutes.

Circular dichroism spectroscopy

Far-UV CD spectra of purified San1 or BSA were obtained on an Aviv 62DS CD spectrometer. Spectra were recorded at 25˚C with a 1mm optical cuvette from 260 to 190 nm. An average of three runs was obtained by sampling every 2nm with a 30 second averaging time. Each sample was in 50mM NaCl and 15mM Na_2HPO_4, pH7.3 and the protein concentrations were between 5 and 10μM. We subtracted the background buffer spectrum from each protein spectrum.

Degradation Assays

Cycloheximide-chase degradation assays were performed similar to previously described (Gardner et al., 2005). Cells were grown at 23˚C or 30˚C in media with 3% raffinose to ~$1x10^7$ cells/ml. Galactose was added to 3% and the cells incubated 3-5 hours at 23˚C or 30˚C. Cycloheximide was added to 50μg/ml and the cells further incubated at 30˚C or 37˚C for 0-3 hours. Cells were lysed at the appropriate time point in 200μl SUMEB. Cycloheximide-chase assays in chapter 3 and 5 were performed using media with 2% glucose because the substrates were constitutively expressed.

Sliding window analysis

Kyte-Doolittle and AGGRESCAN hydrophobicity values for each peptide tested were measured using a 5-residue moving window. The maximum values recorded for each peptide were grouped according to that peptide's ability to interact with San1 by 2-hybrid. Interacting and non-interacting groups were evaluated by normal quantile plot and both conformed to normal distributions. However, the sample groups had unequal sample sizes and variances. Therefore, we compared the means for interacting and non-interacting groups using Welch's unpaired t-Test. One-tailed *P*-values are reported because we are testing the hypothesis that San1-interacting peptides are more hydrophobic than non-interacting peptides.

***In vitro* hydrophobic binding assay**

Plasmid DNA coding for Ura3, Ura3-2 and Ura3-3 6His fusions (see plasmid list) was transformed into T7 express cells (New England Bio Labs). Cells were grown in lysogeny broth (LB) with chloramphenicol at 37˚C until an OD_{600} ~ 1.0 was reached. Protein expression was induced by adding IPTG (isopropyl β-D-1-thiogalactopyranoside) (Gold Biotechnology Inc.) to a final concentration of 300μM followed by incubation at 16˚C overnight. Harvested cells were lysed in 10 ml of BugBuster (Novagen) with 10mM of phenylmethylsulfonyl fluoride (PMSF). Soluble extract was applied to TALON resin (Novagen) equilibrated in a 300mM NaCl, 50mM Na_2HPO_4, pH7.0 buffer. The column was washed with 7.5mM imidazole (Sigma) added to the equilibration buffer and protein was eluted with 150mM imidazole. 40μL of 0.6mg/ml recombinant protein and 10μL of varying concentrations of ANS (8-anilinonaphthalene-1-sulfonate) (Acros Organics) were mixed in a 96-well plate and incubated at 25˚C for 1 hour. ANS fluorescence was excited at 375nm and absorbance measurements taken at 485nm. All ANS assays were conducted in triplicate.

Microscopy

Fixation and DAPI-staining of cells was performed as previously described (Gardner 2005). Cells expressing GFP(S65T) or DsRED fusion proteins were grown in synthetic media with 3% raffinose to ~$1x10^7$ cells/ml. Galactose was added to 3% and the cells incubated 8 hours. Cells were fixed for 15 minutes in 4% paraformaldehyde. Fixed cells were washed and resuspended in potassium phosphate buffer (1.2M sorbitol, 100mM potassium phosphate, pH7.5). Cells were then permeabilized by addition of 2% (v/v) Triton X-100 and stained for 1 minute with the blue nucleic acid-binding dye DAPI (4',6-diamidino-2-phenylindole) (Molecular Probes). Following staining, cells were washed and resuspended in potassium phosphate buffer. Fixed and stained cells were visualized at room temperature on a Nikon Eclipse E600 microscope (Nikon Inc.) equipped with a 60X oil immersion objective, a 100W mercury lamp, and filter sets to visualize GFP (FITC-HYQ), DsRED (Texas Red-HYQ), and DAPI (UV-2A). Images were captured with a Photometrics CoolSNAP fx, 12-bit, cooled CCD camera (Roper Scientific) and the accompanying RS Image software. All images were cropped and processed for publication using the Mac version of Photoshop CS.

REFERENCES

Addinall, S.G., Downey, M., Yu, M., Zubko, M.K., Dewar, J., Leake, A., Hallinan, J., Shaw, O., James, K., Wilkinson, D.J.*, et al.* (2008). A genomewide suppressor and enhancer analysis of cdc13-1 reveals varied cellular processes influencing telomere capping in Saccharomyces cerevisiae. Genetics *180*, 2251-2266.

Akerfelt, M., Morimoto, R.I., and Sistonen, L. (2010). Heat shock factors: integrators of cell stress, development and lifespan. Nat Rev Mol Cell Biol *11*, 545-555.

Balch, W.E., Morimoto, R.I., Dillin, A., and Kelly, J.W. (2008). Adapting proteostasis for disease intervention. Science *319*, 916-919.

Ballinger, C.A., Connell, P., Wu, Y., Hu, Z., Thompson, L.J., Yin, L.Y., and Patterson, C. (1999). Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions. Mol Cell Biol *19*, 4535-4545.

Benitez, E.M., Stolz, A., and Wolf, D.H. (2011). Yos9, a control protein for misfolded glycosylated and non-glycosylated proteins in ERAD. FEBS Lett *585*, 3015-3019.

Betting, J., and Seufert, W. (1996). A yeast Ubc9 mutant protein with temperature-sensitive in vivo function is subject to conditional proteolysis by a ubiquitin- and proteasome-dependent pathway. J Biol Chem *271*, 25790-25796.

Bhamidipati, A., Denic, V., Quan, E.M., and Weissman, J.S. (2005). Exploration of the topological requirements of ERAD identifies Yos9p as a lectin sensor of misfolded glycoproteins in the ER lumen. Mol Cell *19*, 741-751.

Biederer, T., Volkwein, C., and Sommer, T. (1996). Degradation of subunits of the Sec61p complex, an integral component of the ER membrane, by the ubiquitin-proteasome pathway. Embo J *15*, 2069-2076.

Bordallo, J., Plemper, R.K., Finger, A., and Wolf, D.H. (1998). Der3p/Hrd1p is required for endoplasmic reticulum-associated degradation of misfolded lumenal and integral membrane proteins. Mol Biol Cell *9*, 209-222.

Buschhorn, B.A., Kostova, Z., Medicherla, B., and Wolf, D.H. (2004). A genome-wide screen identifies Yos9p as essential for ER-associated degradation of glycoproteins. FEBS Lett *577*, 422-426.

Carvalho, P., Goder, V., and Rapoport, T.A. (2006). Distinct ubiquitin-ligase complexes define convergent pathways for the degradation of ER proteins. Cell *126*, 361-373.

Chan, K.K., Wood, B.M., Fedorov, A.A., Fedorov, E.V., Imker, H.J., Amyes, T.L., Richard, J.P., Almo, S.C., and Gerlt, J.A. (2009). Mechanism of the orotidine 5'-monophosphate decarboxylase-catalyzed reaction: evidence for substrate destabilization. Biochemistry *48*, 5518-5531.

Chiti, F. (2006). Relative importance of hydrophobicity, net charge, and secondary structure propensities in protein aggregation. In Protein misfolding, aggregation, and conformational diseases Part A: protein aggregation and conformational diseases, V.N. Uversky, and A.L. Fink, eds.

Christianson, J.C., Shaler, T.A., Tyler, R.E., and Kopito, R.R. (2008). OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. Nat Cell Biol *10*, 272-282.

Clarke, J., and Fersht, A.R. (1993). Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. Biochemistry *32*, 4322-4329.

Clerc, S., Hirsch, C., Oggier, D.M., Deprez, P., Jakob, C., Sommer, T., and Aebi, M. (2009). Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum. J Cell Biol *184*, 159-172.

Cole, C., Barber, J.D., and Barton, G.J. (2008). The Jpred 3 secondary structure prediction server. Nucleic Acids Res *36*, W197-201.

Conchillo-Sole, O., de Groot, N.S., Aviles, F.X., Vendrell, J., Daura, X., and Ventura, S. (2007). AGGRESCAN: a server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. BMC Bioinformatics *8*, 65.

Connell, P., Ballinger, C.A., Jiang, J., Wu, Y., Thompson, L.J., Hohfeld, J., and Patterson, C. (2001). The co-chaperone CHIP regulates protein triage decisions mediated by heat-shock proteins. Nat Cell Biol *3*, 93-96.

Dasgupta, A., Ramsey, K.L., Smith, J.S., and Auble, D.T. (2004). Sir antagonist 1 (San1) is a ubiquitin ligase. J Biol Chem *279*, 26830-26838.

de la Torre-Ruiz, M.A., Mozo-Villarias, A., Pujol, N., and Petkova, M.I. (2010). How budding yeast sense and transduce the oxidative stress signal and the impact in cell growth and morphogenesis. Curr Protein Pept Sci *11*, 669-679.

Deak, P.M., and Wolf, D.H. (2001). Membrane topology and function of Der3/Hrd1p as a ubiquitin-protein ligase (E3) involved in endoplasmic reticulum degradation. J Biol Chem *276*, 10663-10669.

Denic, V., Quan, E.M., and Weissman, J.S. (2006). A luminal surveillance complex that selects misfolded glycoproteins for ER-associated degradation. Cell *126*, 349-359.

Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics *21*, 3433-3434.

Dosztanyi, Z., Meszaros, B., and Simon, I. (2009). ANCHOR: web server for predicting protein binding regions in disordered proteins. Bioinformatics *25*, 2745-2746.

Dunker, A.K., Oldfield, C.J., Meng, J., Romero, P., Yang, J.Y., Chen, J.W., Vacic, V., Obradovic, Z., and Uversky, V.N. (2008). The unfoldomics decade: an update on intrinsically disordered proteins. BMC Genomics *9 Suppl 2*, S1.

Eisele, F., and Wolf, D.H. (2008). Degradation of misfolded protein in the cytoplasm is mediated by the ubiquitin ligase Ubr1. FEBS Lett *582*, 4143-4146.

Estruch, F., Peiro-Chova, L., Gomez-Navarro, N., Durban, J., Hodge, C., Del Olmo, M., and Cole, C.N. (2009). A genetic screen in Saccharomyces cerevisiae identifies new genes that interact with mex67-5, a temperature-sensitive allele of the gene encoding the mRNA export receptor. Mol Genet Genomics *281*, 125-134.

Evans, D.R., Brewster, N.K., Xu, Q., Rowley, A., Altheim, B.A., Johnston, G.C., and Singer, R.A. (1998). The yeast protein complex containing cdc68 and pob3 mediates core-promoter repression through the cdc68 N-terminal domain. Genetics *150*, 1393-1405.

Fang, S., Ferrone, M., Yang, C., Jensen, J.P., Tiwari, S., and Weissman, A.M. (2001). The tumor autocrine motility factor receptor, gp78, is a ubiquitin protein ligase implicated in degradation from the endoplasmic reticulum. Proc Natl Acad Sci U S A *98*, 14422-14427.

Feng, J., Tamaskovic, R., Yang, Z., Brazil, D.P., Merlo, A., Hess, D., and Hemmings, B.A. (2004). Stabilization of Mdm2 via decreased ubiquitination is mediated by protein kinase B/Akt-dependent phosphorylation. J Biol Chem *279*, 35510-35517.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature *340*, 245-246.

Fontana, A., de Laureto, P.P., Spolaore, B., Frare, E., Picotti, P., and Zambonin, M. (2004). Probing protein structure by limited proteolysis. Acta Biochim Pol *51*, 299-321.

Fredrickson, E.K., Rosenbaum, J.C., Locke, M.N., Milac, T.I., and Gardner, R.G. (2011). Exposed hydrophobicity is a key determinant of nuclear quality control degradation. Mol Biol Cell *22*, 2384-2395.

Furth, N., Gertman, O., Shiber, A., Alfassy, O.S., Cohen, I., Rosenberg, M., Kleinberger-Doron, N., Friedler, A., and Ravid, T. (2011). Exposure of Bipartite Hydrophobic Signal Triggers Nuclear Quality Control of Ndc10 at the Endoplasmic Reticulum/Nuclear Envelope. Mol Biol Cell.

Gardner, R.G., Nelson, Z.W., and Gottschling, D.E. (2005). Degradation-mediated protein quality control in the nucleus. Cell *120*, 803-815.

Gardner, R.G., Swarbrick, G.M., Bays, N.W., Cronin, S.R., Wilhovsky, S., Seelig, L., Kim, C., and Hampton, R.Y. (2000). Endoplasmic reticulum degradation requires lumen to cytosol signaling. Transmembrane control of Hrd1p by Hrd3p. J Cell Biol *151*, 69-82.

Gasch, A.P., Spellman, P.T., Kao, C.M., Carmel-Harel, O., Eisen, M.B., Storz, G., Botstein, D., and Brown, P.O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell *11*, 4241-4257.

Gauss, R., Jarosch, E., Sommer, T., and Hirsch, C. (2006). A complex of Yos9p and the HRD ligase integrates endoplasmic reticulum quality control into the degradation machinery. Nat Cell Biol *8*, 849-854.

Gianni, S., Ivarsson, Y., De Simone, A., Travaglini-Allocatelli, C., Brunori, M., and Vendruscolo, M. (2010). Structural characterization of a misfolded intermediate populated during the folding process of a PDZ domain. Nat Struct Mol Biol *17*, 1431-1437.

Gilon, T., Chomsky, O., and Kulka, R.G. (1998). Degradation signals for ubiquitin system proteolysis in Saccharomyces cerevisiae. EMBO J *17*, 2759-2766.

Gilon, T., Chomsky, O., and Kulka, R.G. (2000). Degradation signals recognized by the Ubc6p-Ubc7p ubiquitin-conjugating enzyme pair. Mol Cell Biol *20*, 7214-7219.

Glickman, M.H., and Ciechanover, A. (2002). The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. Physiol Rev *82*, 373-428.

Gong, Y., Kakihara, Y., Krogan, N., Greenblatt, J., Emili, A., Zhang, Z., and Houry, W.A. (2009). An atlas of chaperone-protein interactions in Saccharomyces cerevisiae: implications to protein folding pathways in the cell. Mol Syst Biol *5*, 275.

Greenfield, N.J. (2006). Using circular dichroism spectra to estimate protein secondary structure. Nat Protoc *1*, 2876-2890.

Hampton, R.Y., Gardner, R.G., and Rine, J. (1996). Role of 26S proteasome and HRD genes in the degradation of 3-hydroxy-3-methylglutaryl-CoA reductase, an integral endoplasmic reticulum membrane protein. Mol Biol Cell *7*, 2029-2044.

Hartl, F.U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. Nature *475*, 324-332.

Haslbeck, M., Ignatiou, A., Saibil, H., Helmich, S., Frenzl, E., Stromer, T., and Buchner, J. (2004). A domain in the N-terminal part of Hsp26 is essential for chaperone function and oligomerization. J Mol Biol *343*, 445-455.

Hawe, A., Sutter, M., and Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. Pharm Res *25*, 1487-1499.

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

Hosokawa, N., Kamiya, Y., Kamiya, D., Kato, K., and Nagata, K. (2009). Human OS-9, a lectin required for glycoprotein endoplasmic reticulum-associated degradation, recognizes mannose-trimmed N-glycans. J Biol Chem *284*, 17061-17068.

Huyer, G., Piluek, W.F., Fansler, Z., Kreft, S.G., Hochstrasser, M., Brodsky, J.L., and Michaelis, S. (2004). Distinct machinery is required in Saccharomyces cerevisiae for the endoplasmic reticulum-associated degradation of a multispanning membrane protein and a soluble luminal protein. J Biol Chem *279*, 38369-38378.

Jaenicke, L.A., Brendebach, H., Selbach, M., and Hirsch, C. (2011). Yos9p assists in the degradation of certain non-glycosylated proteins from the endoplasmic reticulum. Mol Biol Cell.

Jaya, N., Garcia, V., and Vierling, E. (2009). Substrate binding site flexibility of the small heat shock protein molecular chaperones. Proc Natl Acad Sci U S A *106*, 15604-15609.

Jiang, J., Ballinger, C.A., Wu, Y., Dai, Q., Cyr, D.M., Hohfeld, J., and Patterson, C. (2001). CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation. J Biol Chem *276*, 42938-42944.

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

Kaganovich, D., Kopito, R., and Frydman, J. (2008). Misfolded proteins partition between two distinct quality control compartments. Nature *454*, 1088-1095.

Kaneko, M., Ishiguro, M., Niinuma, Y., Uesugi, M., and Nomura, Y. (2002). Human HRD1 protects against ER stress-induced apoptosis through ER-associated degradation. FEBS Lett *532*, 147-152.

Kaser, M., and Langer, T. (2000). Protein degradation in mitochondria. Semin Cell Dev Biol *11*, 181-190.

Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem *74*, 5383-5392.

Kikkert, M., Doolman, R., Dai, M., Avner, R., Hassink, G., van Voorden, S., Thanedar, S., Roitelman, J., Chau, V., and Wiertz, E. (2004). Human HRD1 is an E3 ubiquitin ligase involved in degradation of proteins from the endoplasmic reticulum. J Biol Chem *279*, 3525-3534.

Kim, W., Spear, E.D., and Ng, D.T. (2005). Yos9p detects and targets misfolded glycoproteins for ER-associated degradation. Mol Cell *19*, 753-764.

Kitada, T., Asakawa, S., Hattori, N., Matsumine, H., Yamamura, Y., Minoshima, S., Yokochi, M., Mizuno, Y., and Shimizu, N. (1998). Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. Nature *392*, 605-608.

Kreft, S.G., Wang, L., and Hochstrasser, M. (2006). Membrane topology of the yeast endoplasmic reticulum-localized ubiquitin ligase Doa10 and comparison with its human ortholog TEB4 (MARCH-VI). J Biol Chem *281*, 4646-4653.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol *157*, 105-132.

Lerner, M., Corcoran, M., Cepeda, D., Nielsen, M.L., Zubarev, R., Ponten, F., Uhlen, M., Hober, S., Grander, D., and Sangfelt, O. (2007). The RBCC gene RFP2 (Leu5) encodes a novel transmembrane E3 ubiquitin ligase involved in ERAD. Mol Biol Cell *18*, 1670-1682.

Lewis, M.J., and Pelham, H.R. (2009). Inefficient quality control of thermosensitive proteins on the plasma membrane. PLoS One *4*, e5038.

Lilley, B.N., and Ploegh, H.L. (2005). Multiprotein complexes that link dislocation, ubiquitination, and extraction of misfolded proteins from the endoplasmic reticulum membrane. Proc Natl Acad Sci U S A *102*, 14296-14301.

Lindner, R.A., Kapur, A., Mariani, M., Titmuss, S.J., and Carver, J.A. (1998). Structural alterations of alpha-crystallin during its chaperone action. Eur J Biochem *258*, 170-183.

Matsuo, Y., Kishimoto, H., Tanae, K., Kitamura, K., Katayama, S., and Kawamukai, M. (2011). Nuclear protein quality is regulated by the ubiquitin-proteasome system through the activity of Ubc4 and San1 in fission yeast. J Biol Chem.

Mayer, M.P., and Bukau, B. (2005). Hsp70 chaperones: cellular functions and molecular mechanism. Cell Mol Life Sci *62*, 670-684.

McClellan, A.J., Scott, M.D., and Frydman, J. (2005). Folding and quality control of the VHL tumor suppressor proceed through distinct chaperone pathways. Cell *121*, 739-748.

Metzger, M.B., Maurer, M.J., Dancy, B.M., and Michaelis, S. (2008). Degradation of a cytosolic protein requires endoplasmic reticulum-associated degradation machinery. J Biol Chem *283*, 32302-32316.

Mueller, B., Lilley, B.N., and Ploegh, H.L. (2006). SEL1L, the homologue of yeast Hrd3p, is involved in protein dislocation from the mammalian ER. J Cell Biol *175*, 261-270.

Murata, S., Minami, Y., Minami, M., Chiba, T., and Tanaka, K. (2001). CHIP is a chaperone-dependent E3 ligase that ubiquitylates unfolded protein. EMBO Rep *2*, 1133-1138.

Nadav, E., Shmueli, A., Barr, H., Gonen, H., Ciechanover, A., and Reiss, Y. (2003). A novel mammalian endoplasmic reticulum ubiquitin ligase homologous to the yeast Hrd1. Biochem Biophys Res Commun *303*, 91-97.

Nillegoda, N.B., Theodoraki, M.A., Mandal, A.K., Mayo, K.J., Ren, H.Y., Sultana, R., Wu, K., Johnson, J., Cyr, D.M., and Caplan, A.J. (2010). Ubr1 and Ubr2 Function in a Quality Control Pathway for Degradation of Unfolded Cytosolic Proteins. Mol Biol Cell *21*, 2102-2116.

Park, S.H., Bolender, N., Eisele, F., Kostova, Z., Takeuchi, J., Coffino, P., and Wolf, D.H. (2007). The cytoplasmic Hsp70 chaperone machinery subjects misfolded and endoplasmic reticulum import-incompetent proteins to degradation via the ubiquitin-proteasome system. Mol Biol Cell *18*, 153-165.

Petroski, M.D., and Deshaies, R.J. (2005). Function and regulation of cullin-RING ubiquitin ligases. Nat Rev Mol Cell Biol *6*, 9-20.

Plemper, R.K., Bohmler, S., Bordallo, J., Sommer, T., and Wolf, D.H. (1997). Mutant analysis links the translocon and BiP to retrograde protein transport for ER degradation. Nature *388*, 891-895.

Plemper, R.K., Bordallo, J., Deak, P.M., Taxis, C., Hitt, R., and Wolf, D.H. (1999). Genetic interactions of Hrd3p and Der3p/Hrd1p with Sec61p suggest a retro-translocation complex mediating protein transport for ER degradation. J Cell Sci *112 (Pt 22)*, 4123-4134.

Prasad, R., Kawaguchi, S., and Ng, D.T. (2010). A nucleus-based quality control mechanism for cytosolic proteins. Mol Biol Cell *21*, 2117-2127.

Prilusky, J., Felder, C.E., Zeev-Ben-Mordehai, T., Rydberg, E.H., Man, O., Beckmann, J.S., Silman, I., and Sussman, J.L. (2005). FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. Bioinformatics *21*, 3435-3438.

Quan, E.M., Kamiya, Y., Kamiya, D., Denic, V., Weibezahn, J., Kato, K., and Weissman, J.S. (2008). Defining the glycan destruction signal for endoplasmic reticulum-associated degradation. Mol Cell *32*, 870-877.

Ravid, T., and Hochstrasser, M. (2008). Diversity of degradation signals in the ubiquitin-proteasome system. Nat Rev Mol Cell Biol *9*, 679-690.

Ravid, T., Kreft, S.G., and Hochstrasser, M. (2006). Membrane and soluble substrates of the Doa10 ubiquitin ligase are degraded by distinct pathways. EMBO J *25*, 533-543.

Receveur-Brechot, V., Bourhis, J.M., Uversky, V.N., Canard, B., and Longhi, S. (2006). Assessing protein disorder and induced folding. Proteins *62*, 24-45.

Reichmann, D., Xu, Y., Cremers, C.M., Ilbert, M., Mittelman, R., Fitzgerald, M.C., and Jakob, U. (2012). Order out of disorder: working cycle of an intrinsically unfolded chaperone. Cell *148*, 947-957.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. Proteins *42*, 38-48.

Ron, D., and Walter, P. (2007). Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol *8*, 519-529.

Rosser, M.F., Washburn, E., Muchowski, P.J., Patterson, C., and Cyr, D.M. (2007). Chaperone functions of the E3 ubiquitin ligase CHIP. J Biol Chem *282*, 22267-22277.

Rudiger, S., Germeroth, L., Schneider-Mergener, J., and Bukau, B. (1997). Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. Embo J *16*, 1501-1507.

Sadis, S., Atienza, C., Jr., and Finley, D. (1995). Synthetic signals for ubiquitin-dependent proteolysis. Mol Cell Biol *15*, 4086-4094.

Sato, B.K., Schulz, D., Do, P.H., and Hampton, R.Y. (2009). Misfolded membrane proteins are specifically recognized by the transmembrane domain of the Hrd1p ubiquitin ligase. Mol Cell *34*, 212-222.

Schnell, R., D'Ari, L., Foss, M., Goodman, D., and Rine, J. (1989). Genetic and molecular characterization of suppressors of SIR4 mutations in Saccharomyces cerevisiae. Genetics *122*, 29-46.

Shimura, H., Hattori, N., Kubo, S., Mizuno, Y., Asakawa, S., Minoshima, S., Shimizu, N., Iwai, K., Chiba, T., Tanaka, K.*, et al.* (2000). Familial Parkinson disease gene product, parkin, is a ubiquitin-protein ligase. Nat Genet *25*, 302-305.

Skovronsky, D.M., Lee, V.M., and Trojanowski, J.Q. (2006). Neurodegenerative diseases: new concepts of pathogenesis and their therapeutic implications. Annu Rev Pathol *1*, 151-170.

Stankiewicz, M., Nikolay, R., Rybin, V., and Mayer, M.P. (2010). CHIP participates in protein triage decisions by preferentially ubiquitinating Hsp70-bound substrates. FEBS J *277*, 3353-3367.

Stromer, T., Fischer, E., Richter, K., Haslbeck, M., and Buchner, J. (2004). Analysis of the regulation of the molecular chaperone Hsp26 by temperature-induced dissociation: the N-terminal domail is important for oligomer assembly and the binding of unfolding proteins. J Biol Chem *279*, 11222-11228.

Sultana, R., Theodoraki, M.A., and Caplan, A.J. (2011). UBR1 promotes protein kinase quality control and sensitizes cells to Hsp90 inhibition. Exp Cell Res *318*, 53-60.

Swanson, R., Locher, M., and Hochstrasser, M. (2001). A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and Matalpha2 repressor degradation. Genes Dev *15*, 2660-2674.

Szathmary, R., Bielmann, R., Nita-Lazar, M., Burda, P., and Jakob, C.A. (2005). Yos9 protein is essential for degradation of misfolded glycoproteins and may function as lectin in ERAD. Mol Cell *19*, 765-775.

Tang, X., Orlicky, S., Mittag, T., Csizmok, V., Pawson, T., Forman-Kay, J.D., Sicheri, F., and Tyers, M. (2012). Composite low affinity interactions dictate recognition of the cyclin-dependent kinase inhibitor Sic1 by the SCFCdc4 ubiquitin ligase. Proc Natl Acad Sci U S A *109*, 3287-3292.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

Tompa, P. (2002). Intrinsically unstructured proteins. Trends Biochem Sci *27*, 527-533.

Tompa, P., and Csermely, P. (2004). The role of structural disorder in the function of RNA and protein chaperones. FASEB J *18*, 1169-1175.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. Protein Sci *20*, 1298-1345.

Voisine, C., Pedersen, J.S., and Morimoto, R.I. (2010). Chaperone networks: tipping the balance in protein folding diseases. Neurobiol Dis *40*, 12-20.

Walsh, R., Storey, E., Stefani, D., Kelly, L., and Turnbull, V. (2005). The roles of proteolysis and nuclear localisation in the toxicity of the polyglutamine diseases. A review. Neurotox Res *7*, 43-57.

Wang, X., Taplick, J., Geva, N., and Oren, M. (2004). Inhibition of p53 degradation by Mdm2 acetylation. FEBS Lett *561*, 195-201.

Wang, Z., Jones, G.M., and Prelich, G. (2006). Genetic analysis connects SLX5 and SLX8 to the SUMO pathway in Saccharomyces cerevisiae. Genetics *172*, 1499-1509.

Wang, Z., and Prelich, G. (2009). Quality control of a transcriptional regulator by SUMO-targeted degradation. Mol Cell Biol *29*, 1694-1706.

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

Wickner, S., Maurizi, M.R., and Gottesman, S. (1999). Posttranslational quality control: folding, refolding, and degrading proteins. Science *286*, 1888-1893.

Woulfe, J.M. (2007). Abnormalities of the nucleus and nuclear inclusions in neurodegenerative disease: a work in progress. Neuropathol Appl Neurobiol *33*, 2-42.

Wu, X., Yen, L., Irwin, L., Sweeney, C., and Carraway, K.L., 3rd (2004). Stabilization of the E3 ubiquitin ligase Nrdp1 by the deubiquitinating enzyme USP8. Mol Cell Biol *24*, 7748-7757.

Xu, Q., Johnston, G.C., and Singer, R.A. (1993). The Saccharomyces cerevisiae Cdc68 transcription activator is antagonized by San1, a protein implicated in transcriptional silencing. Mol Cell Biol *13*, 7553-7565.

Yonashiro, R., Ishido, S., Kyo, S., Fukuda, T., Goto, E., Matsuki, Y., Ohmura-Hoshino, M., Sada, K., Hotta, H., Yamamura, H.*, et al.* (2006). A novel mitochondrial ubiquitin ligase plays a critical role in mitochondrial dynamics. Embo J *25*, 3618-3626.

Zavacki, A.M., Arrojo, E.D.R., Freitas, B.C., Chung, M., Harney, J.W., Egri, P., Wittmann, G., Fekete, C., Gereben, B., and Bianco, A.C. (2009). The E3 ubiquitin ligase

TEB4 mediates degradation of type 2 iodothyronine deiodinase. Mol Cell Biol *29*, 5339-5347.

Zhang, Y., Gao, J., Chung, K.K., Huang, H., Dawson, V.L., and Dawson, T.M. (2000). Parkin functions as an E2-dependent ubiquitin- protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1. Proc Natl Acad Sci U S A *97*, 13354-13359.

Zhou, H.X. (2012). Intrinsic disorder: signaling via highly specific but short-lived association. Trends Biochem Sci *37*, 43-48.

CURRICULUM VITAE

EDUCATION

PhD Pharmacology (2012); University of Washington, Seattle, WA
Advisor: Richard G. Gardner, Ph.D.

BS Biology, **BA** Chemistry (2007); *magna cum laude*, Whitworth University, Spokane, WA

AWARDS & SCHOLARSHIPS

July 2010 – June 2012	**NIGMS in Pharmacological Sciences** Institutional National Research Service Award
May 2007	**Outstanding Senior** Department of Biology, Whitworth University

TEACHING AND RESEARCH TRAINING

Teaching Assistant: University of Washington, Seattle, WA September 2008-March 2009.

- Led a weekly discussion group for General Pharmacology 401 and 402
- Wrote quizzes and graded exams

Research Mentor: Gardner Lab, University of Washington September 2009-Present

- Mentored four rotation students and two undergraduate students
- Aided in developing their research projects, technical skills and ability to interpret results

PUBLICATIONS

1. **Fredrickson, E. F**. and R. G. Gardner. **2012**. Selective degradation of abnormal proteins by ubiquitin-mediated protein quality control. *Seminars in Cell and Developmental Biology* **22**: in press.
2. **Fredrickson, E. K.***, Rosenbaum, J. C.*, Locke, M. N., Milac, T. I., and R. G. Gardner. **2011**. Exposed hydrophobicity is a key determinant for nuclear protein quality control degradation. *Molecular Biology of the Cell* **22**: 2384-2395.
3. Rosenbaum, J. C.*, **Fredrickson, E. K.***, Oeser, M. L., Garrett-Engele, C. M., Locke, M. N., Richardson, L. A., Nelson, Z. W., Hetrick, E. D., Milac, T. I., Gottschling, D. E., and R. G. Gardner. **2011**. Disorder targets misorder in nuclear quality control degradation: an intrinsically disordered ubiquitin ligase directly recognizes misfolded substrates. *Molecular Cell* **41**: 93-106.

   Featured on the Cover and in a Preview in the same issue of *Molecular Cell* 41: 2-3.

IN PREPARATION

1. **Fredrickson, E. F**. and R. G. Gardner. Regulation of a flexible protein quality control ubiquitin- protein ligase. Anticipate **June 2012** submission.
2. **Fredrickson, E. F**., Gallagher, P.S. and R. G. Gardner. Chaperone requirements for nuclear protein quality control degradation. Anticipate **Summer 2012** submission.
3. **Fredrickson, E.F** and R.G. Gardner. Not all hydrophobicity is the same: specific hydrophobicity recognized in nuclear protein quality control degradation. Anticipate **Summer 2012** submission.

POSTER PRESENTATIONS

1. **Fredrickson, E. K.**, Rosenbaum, J. C. and R. G. Gardner. The interplay between degradation and repair machinery in nuclear protein quality control. Page 48. Molecular Chaperones and Stress Responses. May 4^{th}-8^{th} 2010. Cold Spring Harbor. New York.
2. **Fredrickson, E. K.**, Rosenbaum, J. C. and R. G. Gardner. The interplay between degradation and repair machinery in nuclear protein quality control. Abstract #113C page 98. Abstracts of the Yeast Genetics and Molecular Biology Meeting, July 27^{th}-August 1^{st}, 2010, University of British Columbia, Vancouver, Canada.