DISSERTATION

MOLECULAR BASIS OF YEAST PRION FORMATION

Submitted by

James A. Toombs

Department of Biochemistry and Molecular Biology

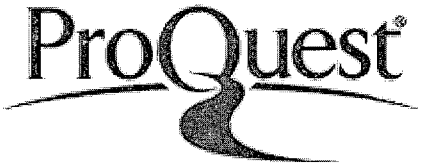
In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2009



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

July 9, 2009

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY JAMES A. TOOMBS ENTITLED "MOLECULAR DETERMINANTS OF YEAST PRION FORMATION" BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate work

Dr. Laurie Stargell

Dr. Olve Peersen

Dr. Edward Hoover

Dr. Asa Ben-Hur

Advisor Dr. Eric Ross

Department Chair Dr. P. Shing Ho

ABSTRACT OF DISSERTATION

MOLECULAR BASIS OF YEAST PRION FORMATION

Amyloid fibers are highly organized protein aggregates that are associated with many fatal diseases. Prions represent a unique class of amyloid fibers that are distinguished by their infectivity and inheritability. In the yeast *S. cerevisiae*, there are several known prion forming proteins. Since the discovery of the first yeast prions in the early 1990s, they have provided a useful model system for studying the biology of prion proteins. While it has been determined that amino acid composition is important to prion formation, there has not yet been any quantitative study aimed at determining how composition promotes or inhibits prion formation. Without this knowledge, our understanding of the events that drive prion formation and our ability to identify new prion-forming proteins is severely limited. In this dissertation, we describe our experiments with the yeast prion protein Sup35p that have illuminated the sequence requirements for yeast prion formation. From these results, we conclude that: (i) amino acid composition, not primary sequence, is the major driving force behind yeast prion propagation, and (ii) prion formation occurs in domains characterized by relatively few prion promoting residues dispersed throughout an intrinsically disordered region.

James A. Toombs
Department of Biochemistry and Molecular Biology
Colorado State University
Fort Collins, CO 80523

Summer 2009

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES vi

CHAPTER 1: Introduction 1
- I. Perspective 1
- II. Yeast Prions 6
  - a. Prion-Forming Domains 7
  - b. Critical Processes of Prion Biology 9
  - c. Prion Strains 12
- III. The Sup35p/[*PSI*$^+$] Model System 12
  - a. Molecular Organization of Sup35 13
  - b. The Sup35p PFD 15
  - c. Monitoring the Presence of [*PSI*$^+$] 17
- IV. Composition: Primary Driving Force of Prion Formation 18
  - a. Influence of Composition on Amyloid Formation 19
  - b. Q/N and Hydrophobic Residues 19
  - c. Bioinformatics 20
- V. Dissertation Overview 22
- VI. References 24

CHAPTER 2: Hsp104-Dependet [*PSI*$^+$] Maintenance Independent of Oligopeptide Repeats 29
- Summary 29
- Introduction 30
- Results 36
- Discussion 47
- Materials and Methods 51
- Acknowledgements 53
- References 53

CHAPTER 3: Compositional Determinants of Prion Formation In Yeast 57
- Summary 57
- Introduction 58
- Results 61
- Discussion 85

TABLE OF CONTENTS (Continued)

Materials and Methods 92
Acknowledgements 96
References 96

CHAPTER 4: Conclusion and Future Perspectives 100
I. Primary sequence of ORD not involved in [*PSI*$^+$] maintenance 101
II. Compositional Determinants of Yeast Prion Formation 103
III. Summary/Future Directions 105
IV. References 107

APPENDIX: Reporter assay systems for [URE3] detection and analysis 108
Abstract 108
Introduction 109
Description of Methods 111
Concluding Remarks 127
Acknowledgments 128
References 128

LIST OF TABLES AND FIGURES

CHAPTER 1:

Figure 1.1:	Molecular Structure of Amyloid Fibrils	2
Figure 1.2:	Kinetics of In Vitro Prion Formation	10
Figure 1.3:	Structural/Functional domains of Sup35p	14

CHAPTER 2:

Figure 2.1:	Structural/Functional domains of Sup35p	32
Figure 2.2:	Fusion proteins maintain wild type [*PSI*$^+$]	38
Figure 2.3:	Scrambling the ORD does not prevent prion maintenance	41
Figure 2.4:	ScrPrP mutants maintain wild type [*PSI*$^+$]	43
Figure 2.5:	ScrNuc mutants maintain wild type [*PSI*$^+$]	45
Figure 2.6:	Hsp104p overexpression cures Sup35 ORD mutants	46

CHAPTER 3:

Figure 3.1:	Truncation mapping of the Sup35-27p PFD	64
Table 3.1:	Mutated sequences from selected and unselected libraries	66
Table 3.2:	Library 1 Amino Acid Representation	68
Table 3.3:	Physical properties of prion forming isolates from Library 1	70
Figure 3.2:	Relationship between the properties of an amino acid and its prevalence among Library 1 prion-forming isolates	72
Figure 3.3:	95% confidence intervals for Libraries 1 and 2	73
Table 3.4:	Library 2 Amino Acid Representation	75
Figure 3.4:	Library 2 shows weaker biases than Library 1	76
Figure 3.5:	Predicted prion-prone regions	78
Figure 3.6:	Predicting prion propensity for Q/N-rich domains	80
Figure 3.7:	Maps of potential prion-forming Q/N-rich proteins	82
Figure 3.8:	Maps of potential non-prion-forming Q/N-rich proteins	83
Figure 3.9:	Summing overlapping windows to calculate total prion propensity of a region	84

Figure 3.10: The Ure2p, Sup35p, and Rnq1p PFDs are biased towards amyloidogenic disorder-promoting residues 87

APPENDIX

Figure A.1: Reporter assay systems for the detection of [URE3] 113

Table A.1: Representative reporter strains 114

Figure A.2: Examples of media component influence on reporter activity 116

Table A.2: Media for [URE3] detection and analysis 117

Figure A.3: Induction of [URE3] by overexpression 119

Figure A.4: Transformation of [URE3] filaments into P_{DAL5}:*ADE2* reporter strains 120

Table A.3: Protocol for transformation of yeast cells with Ure2p filaments 123

Figure A.5: Repressor activity of a Sup35p$^{1-251}$-Ure2p$^{66-354}$ fusion protein. 126

Chapter 1: Introduction

I. Perspective

Many fatal and incurable human diseases are associated with the existence of long filamentous protein aggregates known as amyloid fibers. These fibers are formed by the structural conversion of normal, soluble proteins into an alternate conformation rich in β-strands that promotes well-ordered aggregation of the proteins. Over twenty biochemically distinct types of amyloid fibers have been discovered in the animal kingdom, with each being specifically associated with a unique clinical syndrome (Sipe and Cohen 2000). In addition to being the causative agent of numerous amyloidosis diseases, amyloid fibers are also associated with many fatal neurological diseases including Alzheimer's, Parkinson's and Huntington's, making them the target of intense research over the past several decades.

Amyloid fibers are characterized by a cross-beta sheet quaternary structure in which the beta-strands from individual protein monomers stack and align perpendicular to the axis of the fibril (Figure 1) (Sunde, Serpell et al. 1997). The amyloid core of these fibers, which consists of tightly aligned stacks of beta sheets, is protease resistant due to the limited solvent exposure within this densely packed structure. Detection of amyloid fibers is accomplished through taking advantage of their biophysical properties, which include enhancing the fluorescence of the benzothiazole dye Thioflavin T and displaying apple-green birefringence when stained with Congo Red in the presence of polarized light (Nilsson 2004).

Figure 1.1

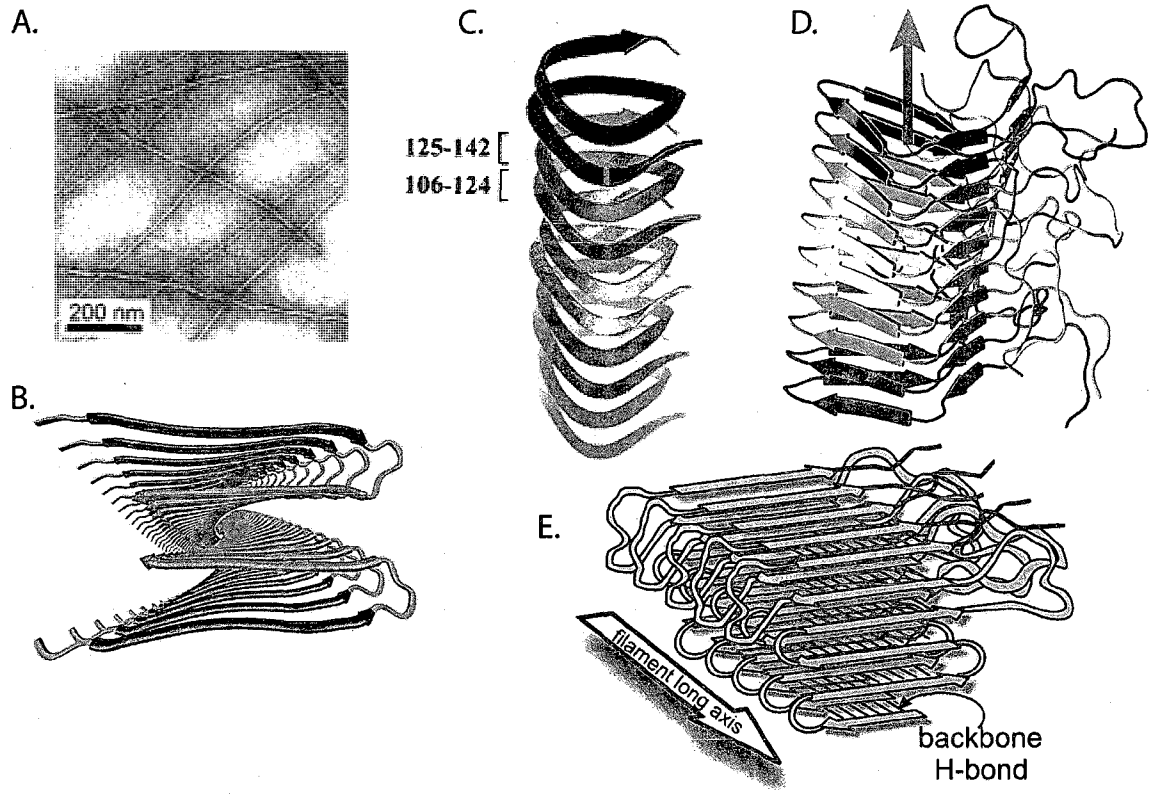


Figure 1.1: Molecular structure of various amyloid forming proteins. The Aβ (1-40) peptide was determined by A) transmission electron microscopy and B) solid-state NMR. Each monomer contributes two beta strands (one blue and one red). Images were adapted from Balbach, Petkova et al. 2002. C) Hypothesized model of PrP prion domain. Figure adapted from Boshuizen, Schulz et al. 2009. D) Solid state NMR structure of the HET-s prion. Figure adapted from Wasmer, Lange et al. 2008. E) Solid state NMR structure of the Sup35 N-domain (blue), which contributes four beta strands, and M-domain (yellow), which contributes 2 beta strands. Figure adapted from Shewmaker, Wickner et al. 2006.

Visualization of amyloid fibers in vitro using transmission electron microscopy (Figure 1A) and atomic force microscopy has revealed that each individual amyloid fiber consists of several (typically 2-6) protofilaments, each about 2-5 nm in diameter (Serpell, Sunde et al. 2000). These protofilaments twist together to form rope-like fibers that are on average 7-13 nm in diameter (Sunde, Serpell et al. 1997; Serpell, Sunde et al. 2000), or they associate laterally to form long ribbons that are 2-5 nm thick and up to 30 nm wide (Bauer, Aebi et al. 1995; Saiki, Honda et al. 2005; Pedersen, Dikov et al. 2006). X-ray crystallographic examinations have provided high-resolution structures that reveal a parallel, in register conformation for the cross β spine of amyloid fibers (Makin, Atkins et al. 2005; Nelson, Sawaya et al. 2005). Although these X-ray structures were based on synthetic amyloids made from short peptides of known amyloid forming domains, they posses the key characteristics of amyloid fibers and have provided detailed molecular resolution unattainable by microscopy techniques. More recently there has been success elucidating amyloid fiber structures using solid-state nuclear magnetic resonance (SSNMR) techniques (Jaroniec, MacPhee et al. 2002; Petkova, Ishii et al. 2002; Ritter, Maddelein et al. 2005). While these structures determined by SSNMR are not as detailed as X-ray crystal structures, they still provide high-resolution molecular detail and can be used to examine entire amyloid domains and not just peptide fragments.

Using SSNMR, Tycko and colleagues found that within the core of amyloid fibers, β-sheets from individual monomers are stacked onto one another in a parallel, in register arrangement (Antzutkin, Balbach et al. 2000; Balbach, Petkova et

al. 2002; Petkova, Ishii et al. 2002; Shewmaker, Wickner et al. 2006; Shewmaker, Ross et al. 2008) (Figure 1B and E). Although this arrangement has been found for many amyloids including Aβ and [*PSI*$^+$], other amyloids including PrP (Boshuizen, Schulz et al. 2009) (Figure 1C) and HET-s (Wasmer, Lange et al. 2008) (Figure 1D) form left handed β-solenoids. While these structures are similar to the parallel, in register, intermolecular β-sheet stacking of cross-β spines, they also contain intramolecular β-sheet stacking in addition to the intermolecular β-sheet stacking.

While most amyloid fibers are not infectious, prions, a special class of amyloid, are capable of spreading from one organism to another (Prusiner 1982; Prusiner 1998). These self-propagating, infectious/inheritable amyloid fibers represent a unique form of protein-only genetic inheritance in which a particular trait is transmitted not through nucleic acid, but through alternative structural conformational forms of specific cellular proteins (Chien, Weissman et al. 2004). This is fundamentally different from epigenetic inheritance, which is based on protein post-translational modifications, because prions are transmissible and not limited to direct inheritance through the cell lineage.

In mammals, prions are the cause of the Transmissible Spongiform Encephalopathies (TSE's), which include Bovine Spongiform Encephalopathy in cattle (a.k.a. Mad Cow Disease), Chronic Wasting Disease in deer and elk, Scrapie in sheep and a slew of diseases in humans including Creutzfelt-Jacob Disease, Gerstmann-Sträussler-Scheinker disease, Fatal Familial Insomnia and Kuru. Affecting about one individual out of a million people per year, prion diseases can be sporadic (spontaneous), familial (genetic/inherited) or acquired (transmitted

through infection) (Prusiner 1998). Due to the rapid progression, ultimate fatality and lack of a cure for these prion diseases, there is currently intense research into the processes and mechanisms of prion biology.

Currently, the only known prion forming protein in mammals and the causative agent of the diseases listed above is PrP. Despite being present in all mammals examined thus far (Prusiner 1998), the function of this cell-surface glycoprotein (Westergard, Christensen et al. 2007) has not yet been discovered; transgenic mice devoid of the PrP gene (*Prnp*$^{0/0}$) display normal development although altered sleep wake cycles have been reported (Tobler, Gaus et al. 1996). While the prion form of PrP, termed PrP^{SC}, efficiently converts the normally folded version, PrP^{C}, into the prion form, a species barrier exists between homologous proteins from different mammals. For example, mice infected with hamster PrP^{SC} prions rarely develop the disease, but transgenic mice expressing the hamster PrP^{C} protein become ill within about two months after infection with hamster PrP^{SC} (Prusiner 1998). Despite this species barrier, there is overwhelming evidence that humans can acquire prion diseases from beef infected with Mad Cow disease (Roma and Prayson 2005) even though the PrP molecules between humans and cattle are divergent at 30 different positions. The reality of this danger is highlighted by the appearance of variant Creutzfeld-Jacob Disease in humans coinciding with the high incidence of Mad Cow disease occurrences of the late 1990's.

II. Yeast Prions

In the yeast *Saccromyces cerevisiae*, there are at least seven known prion forming proteins: Sup35p forms the [*PSI*$^{+}$] prion (Cox 1965; Ter-Avanesyan, Dagkesamanskaya et al. 1994); Ure2p forms [*URE3*] (Lacroute 1971; Wickner 1994); Rnq1p forms [*PIN*$^{+}$] (Sondheimer and Lindquist 2000; Derkatch, Bradley et al. 2001), Swi1p forms [*SWI*$^{+}$] (Du, Park et al. 2008), Cyc8p forms [*OCT*$^{+}$] (Patel, Gavin-Smyth et al. 2009), Mot3p forms [*MOT3*] (Alberti, Halfmann et al. 2009) and Mca1p forms [*MCA*] (Nemecek, Nakayashiki et al. 2009). In addition to these known prions, the proposed prion protein candidate New1p, found by a genomic compositional search, forms [*NU*$^{+}$] (Michelitsch and Weissman 2000; Santoso, Chien et al. 2000). More recently, an in depth systematic survey of the yeast proteome has provided strong evidence for the prion forming potential of several other yeast proteins (Alberti, Halfmann et al. 2009).

Yeast prions, and in particular [*PSI*$^{+}$] and [*URE3*], have provided powerful model systems for investigating the structural, kinetic and biological properties of prions due to their short incubation times, ease of genetic manipulation and general biosafety. However, there is increasing evidence that some yeast prions, unlike PrP^{SC}, do not cause diseases and may actually be advantageous (Tuite and Cox 2009), at least in certain environmental conditions (True and Lindquist 2000). Although the function of Rnq1p and New1p have yet to be discovered, all other yeast prion proteins are either transcriptional regulators (Ure2p, Swi1p, Cyc8p, Mca1p and Mot3p) or translational regulators (Sup35p). This revelation along with the absence of disease in prion containing yeast, point to the hypothesis that prions,

at least in *S. cerevisiae*, can serve as a unique method for transmitting genetic information by remodeling gene expression (Tuite and Cox 2009) or altering protein translation (True, Berlin et al. 2004). This may also be true in the fungi *Podospora anserine*, where another known prion, [Het-s] alters mating compatibility, but does not cause disease (Coustou, Deleu et al. 1997; Maddelein, Dos Reis et al. 2002).

Further evidence that yeast prions are not diseases and are actually unique methods for transmitting non-Mendalian genetic traits is the fact that yeast are capable of converting back and forth between the two protein structural conformations; the soluble non-prion form and the aggregated prion form. The rare natural occurrence of prion containing yeast strains suggests that the prion phenotype is not generally beneficial; however, under certain environmental conditions the prion phenotype is advantageous and allows prion cells to outcompete non-prion cells (True and Lindquist 2000). Spontaneous prion conformational switches in yeast are rare events with frequencies typically between 10^{-5} to 10^{-7}, a rate that is not greatly different from the occurrence of nuclear gene mutations.

A. Prion-forming domains

A common structural feature of prions is that the prion forming ability of a protein resides within a specific domain that is necessary and sufficient for prion formation. These prion-forming domains (PFDs) are structurally and functionally separate from the functional domain(s) of the protein. When the prion protein is in

its soluble/functional conformation, the PFD is often intrinsically disordered, however, upon prion formation, these domains adopt a β-sheet secondary structure and stack to form the cross β-sheet quaternary structure common to all amyloid fibers. In addition to the PFDs being structurally and functionally distinct from the ordered domain(s) of the protein, they are also "convertible," meaning that if they are spliced onto other proteins such as green fluorescent protein, they will confer prion-forming capabilities onto that protein.

Aside from being transmissible and inheritable, one of the major differences between yeast prions and other amyloids is the amino acid composition of the PFD. Yeast PFDs are characterized by high glutamine and asparagine (Q/N) content, low hydrophobic content, and an overall low compositional complexity. Whether this compositional bias separates all yeast proteins from other amyloid forming proteins, or whether just one class of yeast prion proteins has so far been discovered, has not yet been determined. Additionally, the functional/structural basis for this compositional bias is also currently unknown.

The fact that the PFDs of yeast prions are disordered in the native state of the protein distinguishes them from many other amyloid forming proteins wherein the amyloid-forming region is folded in the native conformation of the protein. Due to the lack of order, yeast prion domains do not have to compete between a natively folded state and an amyloid state. Therefore, strong amyloid propensities are not required for prion aggregation by yeast PFDs, and may even be selected against.

B. Critical processes of prion biology

Prion fibers form through a nucleated growth mechanism involving the two distinguishable processes of fiber nucleation and fiber elongation, followed by the third process of fiber cleavage. Monitoring the conversion of soluble proteins into its aggregated fibular form by measuring the increase in ThT fluorescence typically shows a lag phase followed by a rapid exponential growth phase (Figure 2A). The lag phase is assumed to be the time required for nuclei to form (Mukhopadhyay, Krishnan et al. 2007), and once this occurs, the rapid association of soluble proteins with the fiber nucleus followed by conversion to amyloid (Scheibel, Bloom et al. 2004) causes the exponential growth phase and represents the elongation process. Furthermore, if preformed fibers or "seeds" are added to the mixture, the lag phase is completely abolished (Serio, Cashikar et al. 2000; Scheibel and Lindquist 2001), indicating that the nucleation process is the rate-limiting step.

While the details of the nucleation process, also referred to as de novo aggregation, have not been fully elucidated, there is evidence that this occurs through a two step process in which several monomers first come together as an oligomeric species wherein the intrinsically disordered PFDs sample an ensemble of rapidly fluctuating structures until the proper alignment is found. Once this occurs, the oligomer then reorganizes into a nucleating prion structure (Mukhopadhyay, Krishnan et al. 2007). Although there may be control mechanisms yet to be discovered, prion nucleation seems to occur by random chance; by increasing the number of prion forming molecules through over-expression of either the full-length protein or just the PFD, prion formation can be induced in vivo. Additionally,

Figure 1.2

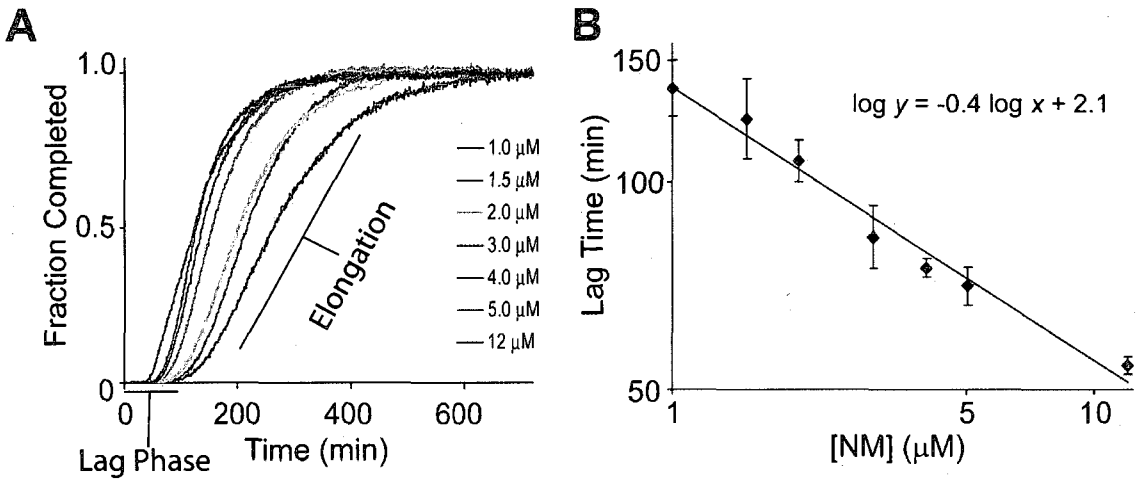


Figure 1.2: Kinetics of in vitro prion formation. A) De novo polymerization of the N and M domains of Sup35p (1.0–12 uM) followed by a continuous thioflavin T binding assay. B) Lag time (measured as time to 5% completion of polymerization) versus NM concentration for the polymeriza- tions shown in (A). Figure adapted from (Collins, Douglass et al. 2004)

increasing the protein concentration in vitro decreases the lag time required for fiber nucleation (Figure 2B).

Once prion fibers have been established in a cell, whether by de novo aggregation, inheritance or infection, fiber growth occurs through the elongation process, which involves at least two separate steps (Scheibel, Bloom et al. 2004). Soluble protein first associates with a preformed prion fiber and forms an assembly intermediate followed by conversion of the intrinsically disordered PFD to the β-sheet structural conformation of the fiber core. While elongation of the growing fiber occurs at both ends, it is not known whether this occurs by the addition of monomers (Collins, Douglass et al. 2004) or oligomers (Narayanan, Walter et al. 2006), or whether elongation occurs preferentially at one end vs. the other (Scheibel, Kowal et al. 2001; DePace and Weissman 2002).

Finally, a third process is required for the stable propagation of prion fibers during cell division. This process of fiber cleavage is dependent on Hsp104p (Shorter and Lindquist 2004), a member of the AAA^+ (Adenosine triphosphotases Associated with diverse Activities) family of chaperone proteins, and is necessary for the production of new inheritable/infectious prion fibers by breaking single fibers into several smaller fibers. Without this process, cells would maintain just a single large prion aggregate that would be incapable of infecting other cells or being inherited by daughter cells during mitosis. Since this process is not required for non-infectious amyloids, it has been hypothesized to be the difference between prions and other amyloids (Osherovich, Cox et al. 2004).

C. Prion Strains

Prion fibers made from the same protein can adopt several different cross β-sheet quaternary structures within the fiber core. These structural variations lead to different fiber characteristics - for example weak vs. strong prions and stable vs. unstable prions. A good example of this is highlighted in an experiment where Sup35p fibers were grown in vitro at either 4°C (Sc4) or 37°C (Sc37) (Toyama, Kelly et al. 2007). Sc4 fibers had a solvent free core that extended through the first 40 amino acids while the solvent free core of Sc37 fibers extended through the first 70 amino acids. Previous examination showed that the Sc4 fibers had a stronger strain phenotype than Sc37 fibers and that this was directly related to the fact that Sc4 fibers had a higher fragmentation rate than Sc37 fibers (Tanaka, Collins et al. 2006), likely due to the smaller and therefore weaker amyloid core.

III. The Sup35p/[*PSI*$^+$] model system

Of the yeast prions, *Sup35p* and *Ure2p* along with their respective prions [PSI$^+$] and [URE3] are the most thoroughly studied since they were the first to be discovered. Both prions provide well-characterized model systems that allows for the study of all three critical processes of prion biology. Throughout this dissertation, all experiments were performed using the Sup35p model system for prion formation and propagation. In this system, cells containing prions are designated [*PSI*$^+$] while cells lacking the prion are designated [*psi*$^-$].

A. Molecular organization of Sup35p

In yeast, Sup35p is a homologue of the translational release factor eRF3. In concert with Sup45p, the homologue of eRF1, this dimer complex recognizes the nonsense codons UAA, UAG and UGA, and stimulates the release of the nascent polypeptide from the ribosome in an ATP consuming process (Inge-Vechtomov, Zhouravleva et al. 2003). The translation termination function of this protein is relegated to the globular C-terminal domain, which extends from amino acid 254-685 (see Figure 3). This domain, often referred to as the C-domain or the EF-domain, is all that is required for efficient translation termination by Sup35p.

The prion-forming domain (PFD), also referred to as the N-domain, is located at the N-terminal region of the protein, spanning amino acids 1-114 (Ter-Avanesyan, Dagkesamanskaya et al. 1994). This intrinsically disordered domain is necessary and sufficient for prion fiber formation (Chernoff, Derkach et al. 1993; Ter-Avanesyan, Dagkesamanskaya et al. 1994; Derkatch, Chernoff et al. 1996; King, Tittmann et al. 1997).

The middle (M) domain of Sup35p, which spans from amino acid 115-253, provides a linker between the globular C domain and the intrinsically disordered PFD. This highly charged domain has no known function other than its ability to stabilize [*PSI*$^+$] fibers (Ter-Avanesyan, Kushnirov et al. 1993; Ter-Avanesyan, Dagkesamanskaya et al. 1994).

Figure 1.3

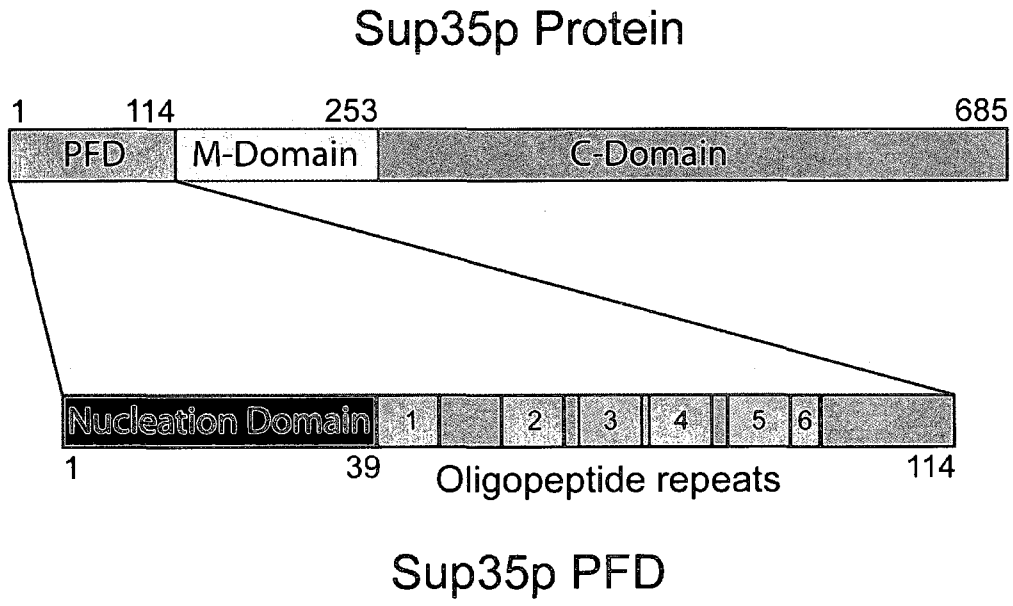


Figure 1.1: Schematic of Sup35p. The PFD, the highly charged middle domain (M-domain) and the C-terminal translation termination domain (C-domain) are shown. The PFD is enlarged below, showing the Q/N rich nucleation domain and the oligopeptide repeat domain (ORD), which consists of 5½ repeats of the consensus sequence ($^{P}/_{Q}$)QGGYQ($^{Q}/_{S}$)YN.

B. Sup35p PFD

The PFD of Sup35p is composed of two separate sub-domains (see Figure 3) that are believed to have distinct functions. The nucleation domain, spanning amino acids 1-39, is rich with glutamine (Q) and asparagine (N) residues and is believed to be responsible for [*PSI*$^+$] nucleation and elongation of [*PSI*$^+$] fibers (DePace, Santoso et al. 1998; Osherovich, Cox et al. 2004). Evidence for this hypothesis is provided by the fact that single point mutations that abolished the [*PSI*$^+$] phenotype in vivo were located with in the first 39 amino acids and mostly altered Q or N residues, highlighting their importance (DePace, Santoso et al. 1998). Subsequent in vitro kinetic studies monitoring de novo aggregation and fiber growth found that these mutants displayed decreased rates compared to wild-type proteins. A separate experiment found that the first 49 amino acids are all that is necessary for addition onto existing prion fibers, yet not sufficient to maintain the prion (Osherovich, Cox et al. 2004). It has been hypothesized that Q/N residues are critical to the fiber core due to their unique hydrogen bonding capabilities in forming polar zippers (Perutz, Johnson et al. 1994; Perutz, Pope et al. 2002; Tsai, Reches et al. 2005); when β-sheets are stacked in a parallel, in register fashion, the R-group amide of Q/N residues interacts with the backbone amide of its neighboring Q/N residue.

The second domain within the Sup35p PFD is the oligopeptide repeat domain (ORD), which spans amino acids 40-96 and contains 5½ degenerate repeats of the consensus sequence ($^{P}/_{Q}$)QGGYQ($^{Q}/_{S}$)YN (Liu and Lindquist 1999; Parham, Resende et al. 2001; Crist, Nakayashiki et al. 2003). This domain is hypothesized to be necessary for Hsp104p dependent fiber cleavage as truncation of more than one

repeat (Liu and Lindquist 1999; Parham, Resende et al. 2001; Osherovich, Cox et al. 2004; Shkundina, Kushnirov et al. 2006) or replacement of this domain with a random sequence (Crist, Nakayashiki et al. 2003) eliminates [*PSI*$^+$]. Furthermore, chimeric proteins in which the prion domain of Sup35p has been replaced with a poly Q tract (Q62) can form amyloids, but these amyloids are not stably propagated; however, addition of the ORD from Sup35p allows for stable propagation (Osherovich, Cox et al. 2004). While this region is clearly required for [*PSI*$^+$] maintenance, the actual repeat sequence most likely plays no role in this process since mutant versions of Sup35p with scrambled PFDs and no ORD are capable of forming and maintaining prions (Ross, Edskes et al. 2005).

Interestingly, the 5½ oligopeptide repeats in the Sup35p ORD are reminiscent of the five oligopeptide repeats found in the mammalian prion protein PrP, which have the sequence PHGGGWGQ. Similarly to the Sup35p ORD, the PrP repeats are critical to the prion phenotype; expansion of the oligopeptide repeat domain is associated with dominant inherited prion diseases (Prusiner, Scott et al. 1998; Wadsworth, Hill et al. 2003), while PrP devoid of the ORD has increased incubation periods and reduced prion titers in terminally ill mice (Flechsig, Shmerling et al. 2000). The similarity between the Sup35p and the PrP repeats has prompted researchers to use the Sup35p model system to study the PrP repeats (Parham, Resende et al. 2001; Dong, Bloom et al. 2007; Tank, Harris et al. 2007; Kalastavadi and True 2008). These studies have found that oligopeptide repeats from PrP can functionally replace the Sup35p ORD in supporting [*PSI*$^+$] maintenance and that increasing the number of PrP repeats inserted in place of the Sup35p ORD

shortens the lag time for in vitro fiber formation assays. However, the interpretation of these studies must be reexamined if the actual repeat sequence is not essential for [*PSI*$^+$] formation and maintenance, and since scrambled Sup35p mutants are capable of forming and maintaining prions (Ross, Edskes et al. 2005), a closer examination into the role of the Sup35p ORD is necessary.

C. Monitoring the presence of [*PSI*$^+$]

The common system for monitoring the presence of [*PSI*$^+$] takes advantage of the translation termination function of this protein. When in the [*PSI*$^+$] state, the activity of the protein is reduced due to its sequestration into prion fibers, which in turn leads to an increased rate of nonsense suppression, meaning that stop codons in the mRNA are not recognized with the same efficiency as they are in [*psi*$^-$] cells. In our yeast strains, a premature stop codon has been inserted into the *ADE2* allele (Cox 1965). In [*psi*$^-$] cells, the stop codon is efficiently recognized leading to the production of a truncated, inactive gene product. When this occurs cells are dependent on supplemental adenine for survival since they are no longer able to produce their own; furthermore, cells grown on media containing limiting adenine turn red due to the build up of a pigment derived from the Ade2p substrate. Alternatively, since [*PSI*$^+$] cells have a deficiency in recognizing stop codons, they are able to express full-length Ade2p, allowing them to grow without supplemental adenine and grow white in the presence of limiting adenine.

The appearance of the Ade$^+$ phenotype, meaning that the yeast are able to grow without supplemented adenine and remain white when grown on limiting

adenine, is not definitive proof of the presence of [*PSI*$^+$] since chromosomal mutations can also lead to this phenotype. One of the properties of [*PSI*$^+$] is that it is dependent on the chaperone protein Hsp104p for stable propagation. Inhibition of Hsp104p by low concentrations of guanidine HCl (GdHCl) (Ferreira, Ness et al. 2001; Jung and Masison 2001; Jung, Jones et al. 2002) eliminates [*PSI*$^+$] propagation and effectively cures the prion phenotype (Tuite, Mundy et al. 1981). Therefore, potential Ade$^+$ prion strains are grown on 4mM GdHCl to test for curability before they are concluded to be prions.

IV. Composition: the primary driving force behind prion formation

One of the most important questions in the prion field has been what allows a protein to become a prion. The biggest clue to this question was answered several years ago when it was found that for both Ure2p and Sup35p, prion formation and propagation was not inhibited by scrambling the primary sequence while maintaining the overall amino acid composition of the PFD (Ross, Baxa et al. 2004; Ross, Edskes et al. 2005). Therefore composition, and not primary sequence, is the major determinant for prion formation. Unfortunately however, we currently do not have a good understanding of what compositional elements drive or inhibit prion formation in yeast. Although there have been numerous studies correlating amino acid composition to amyloid-forming propensities (Chiti, Stefani et al. 2003; Fernandez-Escamilla, Rousseau et al. 2004; Pawar, Dubay et al. 2005), there has been very little attention paid to yeast prions, which have unique amino acid compositions within their PFDs not shared by other amyloid forming proteins.

A. Influence of Composition on Amyloid Formation

The physiochemical properties of the amino acid side chains in critical regions greatly influences the aggregation potential of amyloid nucleating domains (Chiti, Taddei et al. 2002; Wurth, Guimard et al. 2002). The aggregation rates of mutants of non-Q/N rich peptides were positively correlated with hydrophobicity and β-sheet propensity and negatively correlated with α-helix propensity and charge (Chiti, Stefani et al. 2003). Remarkably, an equation that accounts for these parameters was able to accurately predict the aggregation rate of 27 different amyloid-forming polypeptides. Additionally, if the hydrophobic patterning and the total length of the polypeptide is considered in addition to the previous parameters, then the aggregation propensities for the individual amino acids can also be calculated (Pawar, Dubay et al. 2005). While these studies have provided useful information for predicting amyloid formation, they are unable to predict prion formation in yeast. This shortfall is primarily due to the high concentration of glutamine and asparagine residues, which are predicted to have low amyloid-forming propensities, and the lack of hydrophobic residues in yeast PFDs, which are predicted to have high amyloid-forming propensities (Pawar, Dubay et al. 2005).

B. Hydrophobic and Q/N Residues

At first glance, the most obvious characteristic of yeast prions is their unusually high glutamine and asparagine (Q/N) content, which ranges from 30 to 50% within the known yeast PFDs. Not only is this representation high in comparison to the yeast genome, it is also high when compared to other known

amyloid-forming proteins. However, this is not a trait of all prions since neither the mammalian prion protein PrP nor the other fungal prion from *Podospora anserine*, [Het-s], has an unusually high representation of Q/N residues. Additionally, yeast prions all have a remarkably low representation of hydrophobic residues even though they are predicted to promote amyloid formation and are often found in other amyloid forming domains. This begs the question, why do all known yeast prions have such a strikingly unique composition from other amyloid and prion forming proteins?

It is possible that the skewed composition, consisting of unusually high Q/N content and remarkably low hydrophobic content, present in the known yeast PFDs is actually an artifact of how most of these prions have been discovered. Using the compositions of the Sup35p, Ure2p and Rnq1p PFDs as a training set, most of the other known prions were first identified by their compositional homology in genomic searches. Therefore, the compositional bias observed in the yeast PFDs is really a result of homology searches based on only three sequences. While there has been some evidence indicating the importance of Q/N residues (DePace, Santoso et al. 1998), there unfortunately have not yet been any detailed examinations into the compositional requirements for yeast prions.

C. Bioinformatics

Bioinformatic programs have been used to identify new potential PFDs within the yeast genome (Michelitsch and Weissman 2000; Sondheimer and Lindquist 2000; Alberti, Halfmann et al. 2009). While two bona fide prion proteins

(Rnq1p and Mot3p) and one prion candidate (New1p) have been identified through bioinformatic searches, the strength and accuracy of these programs are quite limited. These programs are designed to scan genomes in search of open reading frames that have domains with similar compositions to known PFDs. The limitation to this method is two-fold: first, the data training set with which these programs base their compositional searches on are small, and second, these programs search only for what is there, but not why it is there.

The bioinformatic searches that identified New1p (Michelitsch and Weissman 2000) and Rnq1p (Sondheimer and Lindquist 2000) built their training set on the compositions of the Sup35p and Ure2p PFDs while the bioinformatic search that identified Mot3p (Alberti, Halfmann et al. 2009) was based on the compositions of the Sup35p, Ure2p, Rnq1p and New1p PFDs. Since Rnq1p and New1p were identified by their compositional similarities to Sup35p and Ure2p, essentially all bioinformatic searches have been performed using the compositions of the Sup35p and Ure2p PFDs as a training set. Both of these PFDs have extremely high Q/N content and very low representation of hydrophobic residues; therefore, all new prion proteins discovered thus far by bioinformatics share these same compositional features, which may or may not be critical for prion formation.

Current prion prediction algorithms such as betascan (Bryan, Menke et al. 2009) assign penalties for the presence of charged residues and hydrophobic residues simply because they are under-represented in the known yeast prion domains. However, the reason for these biases are unknown, and so assigning the same penalty may not reflect the true nature of the biases. For example, certain

residues may be excluded from prion domains because they inhibit prion formation yet other residues may be excluded from prion domains because they promote prion formation too strongly and prevent the protein from existing in its native soluble state. Defining the basis for the biases seen in yeast prions will improve the strength and accuracy of bioinformatic search programs. Currently, the most sophisticated search algorithm (Alberti, Halfmann et al. 2009) can only predict new potential prion domains with 20% accuracy, making large scale genomic searches for prion proteins in other organisms unfeasible.

V. Dissertation overview

The overriding goals of my research are to discern the molecular basis of prion formation by the yeast prion protein Sup35p, and to apply this information to improving the ability to identify new PFDs through genomic searches. The experiments described herein address the questions of whether the primary sequence of the Sup35p ORD contributes to the [*PSI*$^+$] phenotype and how amino acid composition within the PFD affects prion formation.

Chapter 1 focuses on a series of mutants that are designed to elucidate the role of the Sup35p ORD. Analysis of these mutants regarding prion nucleation and maintenance indicates that the primary sequence of oligopeptide repeats 1 and 2 are important for avoiding a molecular incompatibility that prevents mutant Sup35p from interacting with wild-type [*PSI*$^+$] fibers. However, in contrast to the current thought in the yeast prion field, we have found that the primary sequence of the entire ORD plays no role in Hsp104p-dependent [*PSI*$^+$] maintenance.

Additionally, the composition of this domain is not necessary as either scrambled PrP repeats or scrambled Sup35p nucleation domains can replace the Sup35p oligopeptide repeats with no discernable difference in prion maintenance from wild-type [*PSI*$^+$] fibers.

Chapter 2 describes two mutant libraries that have been designed to characterize the affects of amino acid composition on prion formation. By randomly mutating specific regions of a particular PFD and screening for mutants capable of forming prions, we have determined what amino acid compositions promote and inhibit prion formation in yeast. Specifically, we find hydrophobic and aromatic residues promote prion formation while proline and charged residues inhibit prion formation. Surprisingly, we found no bias for Q/N residues. Using these prion propensities identified by our screen in conjunction with the Fold Index, which provides a prediction for the degree of intrinsic disorder, we are now able to accurately predict prion formation by a given domain with greater than 95% accuracy.

My work demonstrates that yeast prion formation is determined solely by the amino acid composition within the PFD. Despite the appearance of oligopeptide repeats within some of the known PFDs which hint at a primary sequence component to aggregation, amyloid propensity combined with the intrinsic disorder of a given domain are what promote prion formation in yeast. Using these parameters we have developed a novel prion prediction algorithm that identifies PFDs with unprecedented accuracy.

VI. References

Alberti, S., R. Halfmann, et al. (2009). "A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins." Cell **137**(1): 146-158.

Antzutkin, O. N., J. J. Balbach, et al. (2000). "Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of beta-sheets in Alzheimer's beta-amyloid fibrils." Proc Natl Acad Sci U S A **97**(24): 13045-50.

Balbach, J. J., A. T. Petkova, et al. (2002). "Supramolecular structure in full-length Alzheimer's beta-amyloid fibrils: evidence for a parallel beta-sheet organization from solid-state nuclear magnetic resonance." Biophys J **83**(2): 1205-16.

Bauer, H. H., U. Aebi, et al. (1995). "Architecture and polymorphism of fibrillar supramolecular assemblies produced by in vitro aggregation of human calcitonin." J. Struct. Biol. **115**(1): 1-15.

Boshuizen, R. S., V. Schulz, et al. (2009). "Heterologous Stacking of Prion Protein Peptides Reveals Structural Details of Fibrils and Facilitates Complete Inhibition of Fibril Growth." J Biol Chem **284**(19): 12809-12820.

Bryan, A. W., Jr., M. Menke, et al. (2009). "BETASCAN: probable beta-amyloids identified by pairwise probabilistic analysis." PLoS Comput Biol **5**(3): e1000333.

Chernoff, Y. O., I. L. Derkach, et al. (1993). "Multicopy SUP35 gene induces de-novo appearance of psi-like factors in the yeast Saccharomyces cerevisiae." Curr Genet **24**(3): 268-70.

Chien, P., J. S. Weissman, et al. (2004). "Emerging principles of conformation-based prion inheritance." Annu. Rev. Biochem. **73**: 617-56.

Chiti, F., M. Stefani, et al. (2003). "Rationalization of the effects of mutations on peptide and protein aggregation rates." Nature **424**(6950): 805-8.

Chiti, F., N. Taddei, et al. (2002). "Kinetic partitioning of protein folding and aggregation." Nat Struct Biol **9**(2): 137-43.

Collins, S. R., A. Douglass, et al. (2004). "Mechanism of Prion Propagation: Amyloid Growth Occurs by Monomer Addition." PLoS Biology **2**(10): e321.

Coustou, V., C. Deleu, et al. (1997). "The protein product of the het-s heterokaryon incompatibility gene of the fungus Podospora anserina behaves as a prion analog." Proc. Natl. Acad. Sci. USA **94**(18): 9773-8.

Cox, B. S. (1965). "PSI, a cytoplasmic suppressor of super-suppressor in yeast." Heredity **26**: 211-232.

Crist, C. G., T. Nakayashiki, et al. (2003). "[PHI+], a novel Sup35-prion variant propagated with non-Gln/Asn oligopeptide repeats in the absence of the chaperone protein Hsp104." Genes Cells **8**(7): 603-18.

DePace, A. H., A. Santoso, et al. (1998). "A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion." Cell **93**(7): 1241-52.

DePace, A. H. and J. S. Weissman (2002). "Origins and kinetic consequences of diversity in Sup35 yeast prion fibers." Nat. Struct. Biol. **9**(5): 389-96.

Derkatch, I. L., M. E. Bradley, et al. (2001). "Prions affect the appearance of other prions: the story of [PIN(+)]." Cell **106**(2): 171-82.

Derkatch, I. L., Y. O. Chernoff, et al. (1996). "Genesis and variability of [PSI] prion factors in Saccharomyces cerevisiae." Genetics **144**(4): 1375-86.

Dong, J., J. D. Bloom, et al. (2007). "Probing the role of PrP repeats in conformational conversion and amyloid assembly of chimeric yeast prions." J Biol Chem **282**(47): 34204-12. Epub 2007 Sep 24.

Du, Z., K. W. Park, et al. (2008). "Newly identified prion linked to the chromatin-remodeling factor Swi1 in Saccharomyces cerevisiae." Nat Genet **40**(4): 460-5. Epub 2008 Mar 23.

Fernandez-Escamilla, A. M., F. Rousseau, et al. (2004). "Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins." Nat. Biotechnol. **22**(10): 1302-6.

Ferreira, P. C., F. Ness, et al. (2001). "The elimination of the yeast [PSI+] prion by guanidine hydrochloride is the result of Hsp104 inactivation." Mol. Microbiol. **40**(6): 1357-69.

Flechsig, E., D. Shmerling, et al. (2000). "Prion protein devoid of the octapeptide repeat region restores susceptibility to scrapie in PrP knockout mice." Neuron **27**(2): 399-408.

Inge-Vechtomov, S., G. Zhouravleva, et al. (2003). "Eukaryotic release factors (eRFs) history." Biol Cell **95**(3-4): 195-209.

Jaroniec, C. P., C. E. MacPhee, et al. (2002). "Molecular conformation of a peptide fragment of transthyretin in an amyloid fibril." Proc. Natl. Acad. Sci. USA **99**(26): 16748-53.

Jung, G., G. Jones, et al. (2002). "Amino acid residue 184 of yeast Hsp104 chaperone is critical for prion- curing by guanidine, prion propagation, and thermotolerance." Proc. Natl. Acad. Sci. USA **99**(15): 9936-41.

Jung, G. and D. C. Masison (2001). "Guanidine hydrochloride inhibits Hsp104 activity in vivo: a possible explanation for its effect in curing yeast prions." Curr. Microbiol. **43**(1): 7-10.

Kalastavadi, T. and H. L. True (2008). "Prion protein insertional mutations increase aggregation propensity but not fiber stability." BMC Biochem **9**: 7.

King, C. Y., P. Tittmann, et al. (1997). "Prion-inducing domain 2-114 of yeast Sup35 protein transforms in vitro into amyloid-like filaments." Proc. Natl. Acad. Sci. USA **94**(13): 6618-22.

Lacroute, F. (1971). "Non-Mendelian mutation allowing ureidosuccinic acid uptake in yeast." J. Bacteriol. **106**(2): 519-22.

Liu, J. J. and S. Lindquist (1999). "Oligopeptide-repeat expansions modulate 'protein-only' inheritance in yeast." Nature **400**(6744): 573-6.

Maddelein, M. L., S. Dos Reis, et al. (2002). "Amyloid aggregates of the HET-s prion protein are infectious." Proc Natl Acad Sci U S A **99**(11): 7402-7.

Makin, O. S., E. Atkins, et al. (2005). "Molecular basis for amyloid fibril formation and stability." Proc. Natl. Acad. Sci. USA **102**(2): 315-20.

Michelitsch, M. D. and J. S. Weissman (2000). "A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions." Proc. Natl. Acad. Sci. USA **97**(22): 11910-5.

Mukhopadhyay, S., R. Krishnan, et al. (2007). "A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures." Proc Natl Acad Sci U S A **104**(8): 2649-54.

Narayanan, S., S. Walter, et al. (2006). "Yeast prion-protein, sup35, fibril formation proceeds by addition and substraction of oligomers." Chembiochem **7**(5): 757-65.

Nelson, R., M. R. Sawaya, et al. (2005). "Structure of the cross-beta spine of amyloid-like fibrils." Nature **435**(7043): 773-8.

Nemecek, J., T. Nakayashiki, et al. (2009). "A prion of yeast metacaspase homolog (Mca1p) detected by a genetic screen." Proc Natl Acad Sci U S A **106**(6): 1892-6. Epub 2009 Jan 27.

Nilsson, M. R. (2004). "Techniques to study amyloid fibril formation in vitro." Methods **34**(1): 151-60.

Osherovich, L. Z., B. S. Cox, et al. (2004). "Dissection and design of yeast prions." PLoS Biol. **2**(4): E86.

Parham, S. N., C. G. Resende, et al. (2001). "Oligopeptide repeats in the yeast protein Sup35p stabilize intermolecular prion interactions." EMBO J. **20**(9): 2111-9.

Patel, B. K., J. Gavin-Smyth, et al. (2009). "The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion." Nat Cell Biol **11**(3): 344-9. Epub 2009 Feb 15.

Pawar, A. P., K. F. Dubay, et al. (2005). "Prediction of "Aggregation-prone" and "Aggregation-susceptible" Regions in Proteins Associated with Neurodegenerative Diseases." J. Mol. Biol. **350**(2): 379-92.

Pedersen, J. S., D. Dikov, et al. (2006). "The changing face of glucagon fibrillation: structural polymorphism and conformational imprinting." J Mol Biol **355**(3): 501-23.

Perutz, M. F., T. Johnson, et al. (1994). "Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases." Proc Natl Acad Sci U S A **91**(12): 5355-8.

Perutz, M. F., B. J. Pope, et al. (2002). "Aggregation of proteins with expanded glutamine and alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid beta -peptide of amyloid plaques." Proc. Natl. Acad. Sci. USA **99**(8): 5596-5600.

Petkova, A. T., Y. Ishii, et al. (2002). "A structural model for Alzheimer's beta -amyloid fibrils based on experimental constraints from solid state NMR." Proc Natl Acad Sci U S A **99**(26): 16742-7.

Prusiner, S. B. (1982). "Novel proteinaceous infectious particles cause scrapie." Science **216**(4542): 136-44.

Prusiner, S. B. (1998). "Prions." Proc Natl Acad Sci U S A **95**(23): 13363-83.

Prusiner, S. B. (1998). "The prion diseases." Brain Pathol **8**(3): 499-513.

Prusiner, S. B., M. R. Scott, et al. (1998). "Prion protein biology." Cell **93**(3): 337-48.

Ritter, C., M. L. Maddelein, et al. (2005). "Correlation of structural elements and infectivity of the HET-s prion." Nature **435**(7043): 844-8.

Roma, A. A. and R. A. Prayson (2005). "Bovine spongiform encephalopathy and variant Creutzfeldt-Jakob disease: how safe is eating beef?" Cleve Clin J Med **72**(3): 185-6, 189-90, 192-4.

Ross, E. D., U. Baxa, et al. (2004). "Scrambled Prion Domains Form Prions and Amyloid." Mol. Cell. Biol. **24**(16): 7206-7213.

Ross, E. D., H. K. Edskes, et al. (2005). "Primary sequence independence for prion formation." Proc. Natl. Acad. Sci. USA **102**(36): 12825-12830.

Saiki, M., S. Honda, et al. (2005). "Higher-order molecular packing in amyloid-like fibrils constructed with linear arrangements of hydrophobic and hydrogen-bonding side-chains." J Mol Biol **348**(4): 983-98.

Santoso, A., P. Chien, et al. (2000). "Molecular basis of a yeast prion species barrier." Cell **100**(2): 277-88.

Scheibel, T., J. Bloom, et al. (2004). "The elongation of yeast prion fibers involves separable steps of association and conversion." Proc Natl Acad Sci U S A **101**(8): 2287-92.

Scheibel, T., A. S. Kowal, et al. (2001). "Bidirectional amyloid fiber growth for a yeast prion determinant." Curr Biol **11**(5): 366-9.

Scheibel, T. and S. L. Lindquist (2001). "The role of conformational flexibility in prion propagation and maintenance for Sup35p." Nat. Struct. Biol. **8**(11): 958-62.

Serio, T. R., A. G. Cashikar, et al. (2000). "Nucleated conformational conversion and the replication of conformational information by a prion determinant." Science **289**(5483): 1317-21.

Serpell, L. C., M. Sunde, et al. (2000). "The protofilament substructure of amyloid fibrils." J Mol Biol **300**(5): 1033-9.

Shewmaker, F., E. D. Ross, et al. (2008). "Amyloids of shuffled prion domains that form prions have a parallel in-register beta-sheet structure." Biochemistry **47**(13): 4000-7. Epub 2008 Mar 7.

Shewmaker, F., R. B. Wickner, et al. (2006). "Amyloid of the prion domain of Sup35p has an in-register parallel beta-sheet structure." Proc Natl Acad Sci U S A **103**(52): 19754-9. Epub 2006 Dec 14.

Shkundina, I. S., V. V. Kushnirov, et al. (2006). "The role of the N-terminal oligopeptide repeats of the yeast Sup35 prion protein in propagation and transmission of prion variants." Genetics **172**(2): 827-35.

Shorter, J. and S. Lindquist (2004). "Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers." Science **304**(5678): 1793-7.

Sipe, J. D. and A. S. Cohen (2000). "Review: history of the amyloid fibril." J. Struct. Biol. **130**(2-3): 88-98.

Sondheimer, N. and S. Lindquist (2000). "Rnq1: an epigenetic modifier of protein function in yeast." Mol. Cell **5**(1): 163-72.

Sunde, M., L. C. Serpell, et al. (1997). "Common core structure of amyloid fibrils by synchrotron X-ray diffraction." J. Mol. Biol. **273**(3): 729-39.

Tanaka, M., S. R. Collins, et al. (2006). "The physical basis of how prion conformations determine strain phenotypes." Nature **442**(7102): 585-589.

Tank, E. M., D. A. Harris, et al. (2007). "Prion protein repeat expansion results in increased aggregation and reveals phenotypic variability." Mol Cell Biol **27**(15): 5445-55. Epub 2007 Jun 4.

Ter-Avanesyan, M. D., A. R. Dagkesamanskaya, et al. (1994). "The SUP35 omnipotent suppressor gene is involved in the maintenance of the non-Mendelian determinant [psi+] in the yeast Saccharomyces cerevisiae." Genetics **137**(3): 671-6.

Ter-Avanesyan, M. D., V. V. Kushnirov, et al. (1993). "Deletion analysis of the SUP35 gene of the yeast Saccharomyces cerevisiae reveals two non-overlapping functional regions in the encoded protein." Mol. Microbiol. **7**(5): 683-92.

Tobler, I., S. E. Gaus, et al. (1996). "Altered circadian activity rhythms and sleep in mice devoid of prion protein." Nature **380**(6575): 639-42.

Toyama, B. H., M. J. Kelly, et al. (2007). "The structural basis of yeast prion strain variants." Nature **449**(7159): 233-7. Epub 2007 Sep 2.

True, H. L., I. Berlin, et al. (2004). "Epigenetic regulation of translation reveals hidden genetic variation to produce complex traits." Nature **431**(7005): 184-187.

True, H. L. and S. L. Lindquist (2000). "A yeast prion provides a mechanism for genetic variation and phenotypic diversity." Nature **407**(6803): 477-83.

Tsai, H. H., M. Reches, et al. (2005). "Energy landscape of amyloidogenic peptide oligomerization by parallel-tempering molecular dynamics simulation: significant role of Asn ladder." Proc Natl Acad Sci U S A **102**(23): 8174-9.

Tuite, M. F. and B. S. Cox (2009). "Prions remodel gene expression in yeast." Nat Cell Biol **11**(3): 241-3.

Tuite, M. F., C. R. Mundy, et al. (1981). "Agents that cause a high frequency of genetic change from [psi+] to [psi-] in Saccharomyces cerevisiae." Genetics **98**(4): 691-711.

Wadsworth, J. D., A. F. Hill, et al. (2003). "Molecular and clinical classification of human prion disease." Br Med Bull **66**: 241-54.

Wasmer, C., A. Lange, et al. (2008). "Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core." Science **319**(5869): 1523-6.

Westergard, L., H. M. Christensen, et al. (2007). "The cellular prion protein (PrP(C)): its physiological function and role in disease." Biochim Biophys Acta **1772**(6): 629-44.

Wickner, R. B. (1994). "[URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae." Science **264**(5158): 566-9.

Wurth, C., N. K. Guimard, et al. (2002). "Mutations that reduce aggregation of the Alzheimer's Abeta42 peptide: an unbiased search for the sequence determinants of Abeta amyloidogenesis." J Mol Biol **319**(5): 1279-90.

Chapter 2: Hsp104p-dependent [*PSI*$^+$] maintenance independent of oligopeptide repeats

James A. Toombs, Nathan Liss and Eric D. Ross

Nathan Liss was responsible for the work contributing to Figure 2.2D

A modified version of this work has been submitted for publication (2009).

Summary

Prions represent a unique class of amyloid fibers that are distinguished by their infectious, self-propagating nature. The infectivity of prions is thought to result from chaperone-dependent fiber cleavage that breaks large prion fibers into smaller, inheritable propagons. [*PSI*$^+$] is the prion form of the yeast Sup35p protein. Randomizing the order of the amino acids in the Sup35p prion domain does not block prion formation or propagation, suggesting that amino acid composition is the primary determinant of Sup35p's prion propensity. However, like the mammalian prion protein PrP, Sup35p contains an oligopeptide repeat domain (ORD). Deletion and mutational analysis indicate that the ORD is critical for [*PSI*$^+$] propagation. The PrP ORD can substitute for the Sup35p ORD in supporting [*PSI*$^+$] propagation, suggesting a common role for repeats in supporting prion maintenance. Here, we carefully examine the sequence features of the ORD that allow for [*PSI*$^+$] propagation. We find that the ability of the Sup35p and PrP ORDs to support [*PSI*$^+$] propagation is primary sequence independent, and that the compositional requirements for the ORD region are highly flexible.

Introduction

Many human diseases are associated with the formation of amyloid fibers including Alzheimer's disease, Huntington's disease, type-II diabetes and the transmissible spongiform encephalopathies. Amyloid fibers are comprised of well-ordered protein aggregates characterized by filamentous morphology, cross-ß-sheet structure, protease resistance, and yellow-green birefringence upon staining with Congo red (Kisilevsky and Fraser 1997; Sipe and Cohen 2000). A subset of amyloid fibers, termed prions, is infectious (Prusiner 1998; Prusiner, Scott et al. 1998). PrP is the only known prion-forming protein in mammals. First discovered as the infectious agent of scrapie (Prusiner 1982), PrP is responsible for mad cow disease, chronic wasting disease in deer and elk, and several diseases in humans including Kuru, Creutzfeldt-Jakob disease and fatal familial insomnia (Prusiner 1998).

In the yeast *S. cerevisiae*, there are several known prion-forming proteins. The non-chromosomal genetic elements [*PSI*$^+$](Cox 1965) and [*URE3*] (Lacroute 1971) which are prion forms of Sup35p and Ure2p, respectively (Wickner 1994), are the two most widely studied. Due to their short incubation times, ease of genetic manipulation and relative biosafety, yeast prions have provided useful model systems to study the mechanisms by which prions are nucleated and propagated in vivo.

Sup35p is an essential component of the translation termination machinery in yeast (Stansfield, Jones et al. 1995; Zhouravleva, Frolova et al. 1995). When in the [*PSI*$^+$] state, the translation termination activity of Sup35p is compromised due to its sequestration into amyloid fibers. The reduced concentration of soluble, functional

Sup35p results in increased stop codon read-through (Patino, Liu et al. 1996; Paushkin, Kushnirov et al. 1996). Therefore, [*PSI*$^+$] can be monitored by suppression of nonsense mutations.

The Sup35p protein is divided into three functionally and structurally distinct domains (Fig. 1): the C-terminal domain (a.a. 254-685) that is essential for the translation termination activity of Sup35p; the prion-forming domain (PFD; a.a. 1-114) that drives prion formation; and the highly charged middle domain that has no known function other than its ability to stabilize [PSI$^+$] fibers (Ter-Avanesyan, Kushnirov et al. 1993; Ter-Avanesyan, Dagkesamanskaya et al. 1994; DePace, Santoso et al. 1998; Liu, Sondheimer et al. 2002; Inge-Vechtomov, Zhouravleva et al. 2003). CD analysis has determined that the PFD is natively disordered in solution (Serio, Cashikar et al. 2000), while solid-state NMR data indicates that upon amyloid formation, the PFDs stack to form in-register parallel β-sheets (Shewmaker, Wickner et al. 2006; Shewmaker, Ross et al. 2008), although an alternative model has been proposed in which intermolecular contacts within the amyloid fibril occur primarily at the ends of the PFD (Krishnan and Lindquist 2005).

The Sup35p PFD is composed of two separate sub-domains thought to have distinct functions. The glutamine/asparagine-rich (Q/N-rich) tract (amino acids 1-39) is responsible for prion nucleation (DePace, Santoso et al. 1998; Osherovich, Cox et al. 2004), while the oligopeptide repeat domain (ORD), which spans amino acid 40-96, allows for efficient Hsp104p-dependent prion propagation (Liu and Lindquist 1999; Parham, Resende et al. 2001; Osherovich, Cox et al. 2004; Shkundina, Kushnirov et al. 2006). The ORD consists of 5½ degenerate repeats of

Figure 2.1

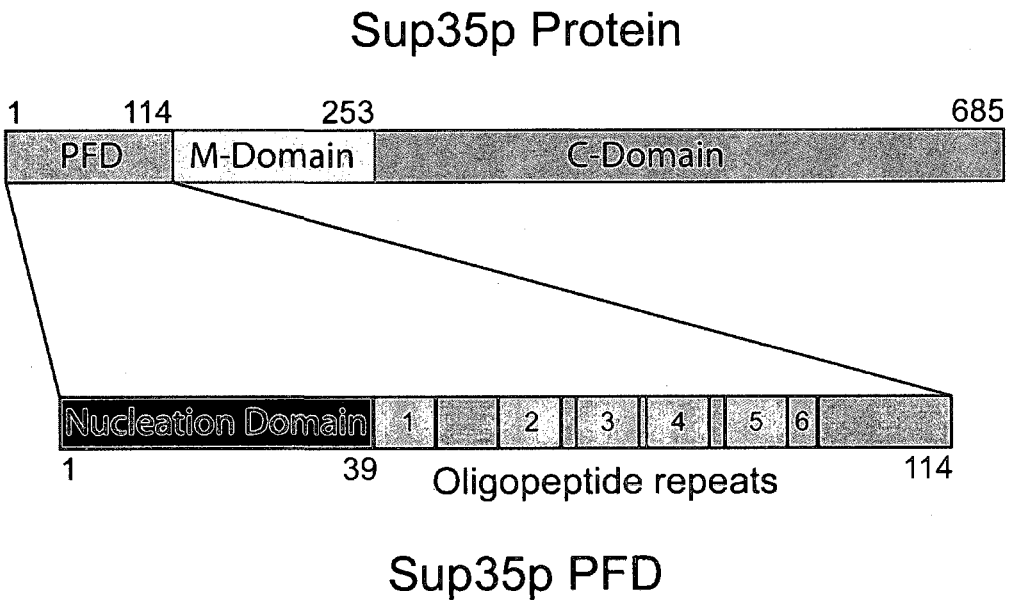


Figure 2.1: Schematic of Sup35p. The PFD, the highly charged middle domain (M-domain) and the C-terminal translation termination domain (C-domain) are shown. The PFD is enlarged below, showing the Q/N rich nucleation domain and the oligopeptide repeat domain (ORD), which consists of 5½ repeats of the consensus sequence $(^{P}/_{Q})$QGGYQ$(^{Q}/_{S})$YN.

the consensus sequence ($^{P}/_{Q}$)QGGYQ($^{Q}/_{S}$)YN (Liu and Lindquist 1999; Parham, Resende et al. 2001; Crist, Nakayashiki et al. 2003). Truncation of all or part of the ORD (Ter-Avanesyan, Kushnirov et al. 1993; Liu and Lindquist 1999; Parham, Resende et al. 2001; Osherovich, Cox et al. 2004) or replacement of the repeats with a random sequence (Crist, Nakayashiki et al. 2003) destabilizes or eliminates [*PSI*$^{+}$]. Deletion of one or more repeats also increases the average [*PSI*$^{+}$] aggregate size (Shkundina, Kushnirov et al. 2006), suggesting that the repeats facilitate Hsp104p-dependent fragmentation, either by acting as a direct binding site for Hsp104p or by changing the conformation of the amyloid core to allow for Hsp104p access. Additionally, replacement of the Sup35p ORD in *S. cerevisiae* with the oligopeptide repeat motifs from the Sup35p of *Y. lipolytica,* which has previously been shown to be potent [*PSI*$^{+}$] forming protein (Nakayashiki, Ebihara et al. 2001), significantly affects prion propagation and allows for prion maintenance in the absence of Hsp104p (Crist, Nakayashiki et al. 2003). Finally, chimeric proteins in which the PFD of Sup35p has been replaced with a poly Q tract (Q62) can form amyloids, but these amyloids are not stably propagated; however, addition of the ORD from Sup35p allows for stable propagation (Osherovich, Cox et al. 2004). Therefore, it has been proposed that efficient chaperone-dependent aggregate cleavage may represent the difference between infectious and non-infectious amyloids, and that repeat sequences may play a critical role in allowing for chaperone-dependent cleavage (Osherovich, Cox et al. 2004).

Although numerous studies have demonstrated that the ORD is important for [*PSI*$^{+}$] maintenance, none of these studies has examined whether repeats *per se* are

required for this function, or whether some other feature of the ORD allows for prion maintenance. Ure2p does not contain repeats, demonstrating that repeats are not a necessary feature for prion maintenance. Additionally, four of five scrambled mutants of Sup35p, in which the order of the amino acids in the PFD was randomized while maintaining amino acid composition, were able to stably maintain [*PSI*$^+$] (Ross, Edskes et al. 2005), suggesting that composition, not primary sequence, is the major determinant of the [*PSI*$^+$] phenotype.

However, these scrambling results have not diminished the attention focused on the ORD. Recently, attention has focused on the striking similarity between the ORDs of Sup35p and PrP. PrP has five octa-peptide repeats of PHGGGWGQ, vs. 5½ ($^{P}/_{Q}$)QGGYQ($^{Q}/_{S}$)YN repeats for Sup35p. Mutations in the PrP gene, *PRNP*, resulting in the expansion of the oligopeptide repeat domain are associated with dominant inherited prion diseases (Prusiner, Scott et al. 1998; Wadsworth, Hill et al. 2003), while PrP devoid of the ORD has increased incubation periods and reduced prion titers in terminally ill mice (Flechsig, Shmerling et al. 2000). Due to the similarities of the Sup35p and PrP oligopeptide repeats, Sup35p has recently been used as a model for examining the role of the PrP repeats in prion formation and propagation (Parham, Resende et al. 2001; Dong, Bloom et al. 2007; Tank, Harris et al. 2007; Kalastavadi and True 2008). These studies have established that oligopeptide repeats from PrP can functionally replace the Sup35p ORD in supporting [*PSI*$^+$] maintenance, and that increasing the number of PrP repeats inserted in place of the Sup35p ORD shortens the lag time for in vitro fiber formation assays.

The continued focus on the ORD in spite of the insensitivity of Sup35p to

scrambling might seem surprising. However, a primary sequence element was recently identified within another prion protein, Rnq1p (Sondheimer and Lindquist 2000; Derkatch, Bradley et al. 2001), that is required for interaction with the chaperone Sis1p (Sondheimer, Lopez et al. 2001; Douglas, Treusch et al. 2008), bolstering the idea that primary sequence elements such as the ORD can affect chaperone interactions. Given the similarity between the Sup35p and PrP ORDs, it seems likely that the primary sequence of the repeats is similarly important for some aspect of [*PSI*$^+$] formation or propagation.

Various hypotheses have been proposed to explain how the primary sequence of the ORD could be critical for [*PSI*$^+$] propagation in spite of the lack of an ORD in Ure2p and in spite of the insensitivity of Sup35p to scrambling. Shkundina et al. proposed that the repeats are not necessary in the context of artificial prions such as those formed by scrambled Sup35p, but that the repeats may serve a specific function within the context of naturally occurring [*PSI*$^+$] variants (Shkundina, Kushnirov et al. 2006). Alternatively, the role of the repeats could be subtler. [*PSI*$^+$] is more sensitive to Hsp104p levels than [*URE3*]; although both [*URE3*] and [*PSI*$^+$] are efficiently eliminated by Hsp104p deletion, only [*PSI*$^+$] is eliminated by Hsp104p over-expression (Shorter and Lindquist 2004). Similarly, high concentrations of Hsp104p completely dissolve Sup35p amyloid aggregates in vitro, but not Ure2p aggregates (Shorter and Lindquist 2006). Additionally, many spontaneous [*URE3*] isolates are unstable (Schlumpberger, Prusiner et al. 2001), which may be a due to the lack of repeats. (Osherovich, Cox et al. 2004) Therefore, the repeats may not be an absolute requirement for prions, but instead explain some of Sup35p's unique

properties (Osherovich, Cox et al. 2004). Alternatively, we hypothesized that the ORD is important for its composition, not its primary sequence.

To further investigate this conflicting data and to explore the sequence requirements for Hsp104p-dependent prion maintenance, we have made a broad range of mutant Sup35p proteins in which the ORD has been disrupted. Upon analysis of these mutants, we find that the primary sequence of the oligopeptide repeats is nonessential for Hsp104p-dependent [*PSI*$^+$] maintenance, disproving the widely held hypothesis that the oligopeptide repeats provide a specific recognition sequence for chaperone-mediated fiber cleavage.

Results

Replacing the ORD with segments from scrambled Sup35p mutant proteins

Because four of five scrambled versions of Sup35p were able to efficiently propagate prions, we hypothesized that each of these proteins must have propagation domains analogous to the ORD. By identifying these regions, we hoped to identify the common features that allow for Hsp104p-dependent prion maintenance. Therefore, we replaced the 75 amino acid maintenance domain (a.a. 40-114) of wild-type Sup35p, which includes the entire ORD, with either the first 75 or the last 75 amino acids of the scrambled Sup35p mutants (Sup35-21p, -24p, -25p, -26p and -27p) (Ross, Edskes et al. 2005). The fusion constructs were named FP21N (fusion protein Sup35-21 N-terminus, indicating that the ORD of wild-type Sup35p was replaced with the N-terminal 75 amino acids from the scrambled prion protein Sup35-21p), FP21C, FP24N, FP24C, FP26N, FP26C, FP27N and FP27C.

Using plasmid shuffling, we introduced each of the constructs into a [*PSI*$^+$] strain in the place of the endogenous *SUP35* and assessed whether these proteins could propagate [*PSI*$^+$]. Propagation of [*PSI*$^+$] was assayed by monitoring nonsense suppression of the *ade2-1* allele (Cox 1965). In the absence of [*PSI*$^+$], *ade2-1* mutants are unable to grow without adenine and form red colonies in the presence of limiting adenine due to accumulation of a pigment derived from the substrate of Ade2p. [*PSI*$^+$] causes stop-codon read-through, allowing for growth without adenine and white or pink colony formation in the presence of limiting adenine. Surprisingly, seven out of the eight fusion mutants were capable of maintaining the Ade$^+$ phenotype (Fig. 2A and 2B). To confirm that the maintenance of the Ade$^+$ phenotype was a result of [*PSI*$^+$] maintenance, we tested whether the Ade$^+$ phenotype could be cured by treatment with low concentrations of guanidine HCl. Guanidine HCl cures both [URE3] and [*PSI*$^+$] (Aigle and Lacroute 1975; Tuite, Mundy et al. 1981) by inhibiting Hsp104p (Ferreira, Ness et al. 2001; Jung and Masison 2001; Jung, Jones et al. 2002). In all cases, the Ade$^+$ phenotype was efficiently cured by treatment with 4 mM guanidine HCl, demonstrating that these mutants were indeed maintaining [*PSI*$^+$] in an Hsp104p-dependent manner (Fig. 2B).

In the plasmid shuffle assay, successful prion propagation requires two steps. First, the mutant must be compatible with, and therefore able to add onto, the pre-existing wild-type [*PSI*$^+$] aggregates. This molecular compatibility is analogous to the species barrier seen for the mammalian prion proteins, in which prion transmission is inefficient between different species due to primary sequence differences between the proteins (Santoso, Chien et al. 2000). Second, the cellular

Figure 2.2

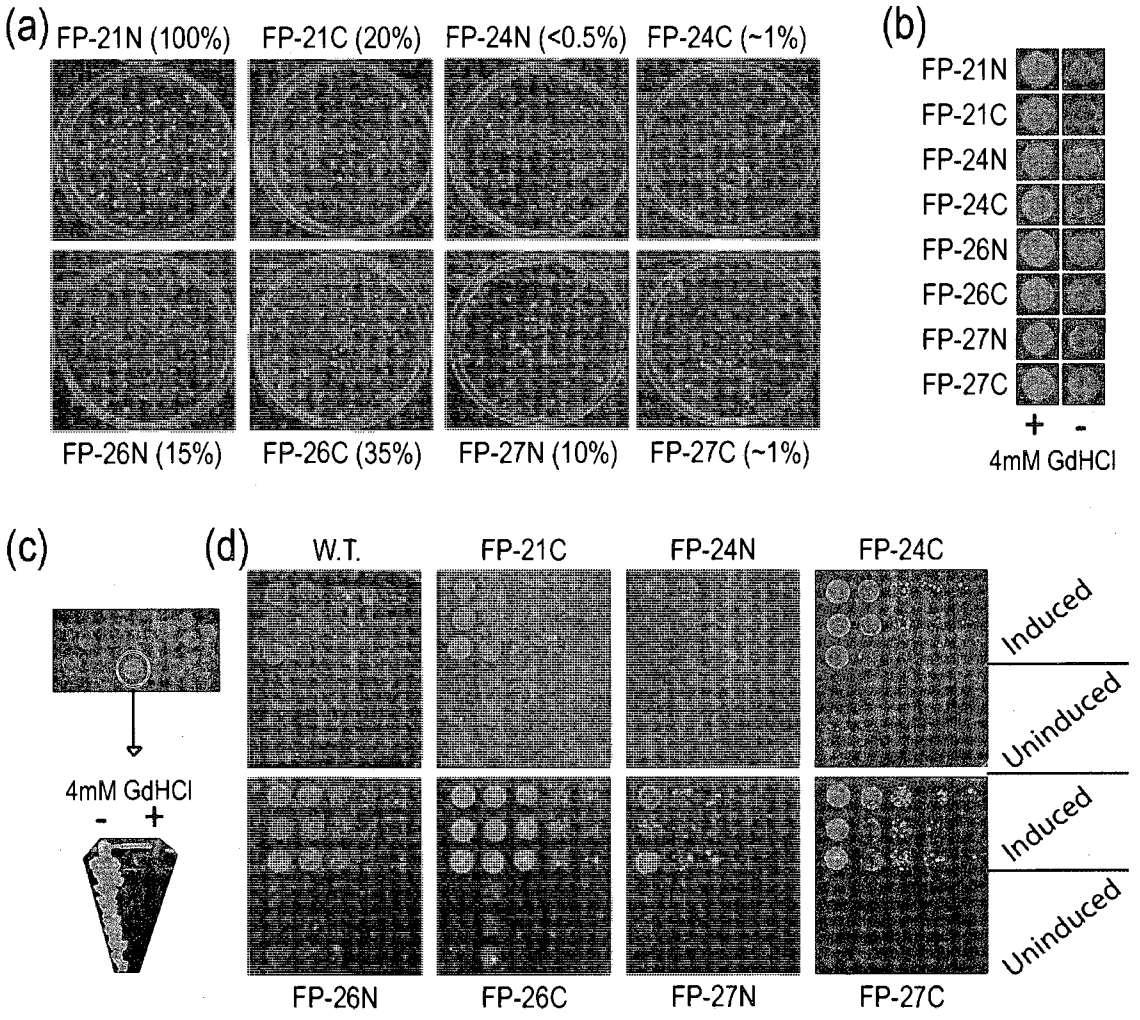


Figure 2.2: Fusion proteins maintain wild type [*PSI*$^+$]. a) Molecular compatibility between wild-type [*PSI*$^+$] aggregates and fusion proteins in which the ORD was replaced with fragments from the scrambled PFDs. Plasmids expressing the fusion proteins were transformed into a *SUP35*-deleted [*PSI*$^+$] strain that expressed wild-type *SUP35* from a *URA3* plasmid. After counterselection to remove the wild-type *SUP35* plasmid, strains were plated for single colonies on YPD to test for [*PSI*$^+$] by color phenotype. Approximate percentages of [*PSI*$^+$] cells for each strain are indicated in parentheses. b) [*PSI*$^+$] stability and curing. [*PSI*$^+$] colonies expressing the fusion proteins were streaked for single colonies on YPAD or YPAD plus 4mM guandine HCl and then spotted onto YPD to test for [*PSI*$^+$]. c) A rare white colony from FP-21 (Panel A) was streaked onto YPD to test for stability of the Ade$^+$ phenotype. To confirm that the colony was [*PSI*$^+$], it was also streaked onto YPAD plus 4 mM guanidine and then streaked onto YPD to test for [*PSI*$^+$]. d) Strains expressing wild-type *SUP35* (W.T.) or the fusion proteins were transformed with either plasmid pKT24 containing the *GAL1* promoter (uninduced) or with a derivative pKT24 expressing the NM domain of the same variant of *SUP35* from the *GAL1* promoter (induced). Strains were grown in galactose/raffinose dropout medium and serial dilutions plated onto medium lacking adenine to select for [*PSI*$^+$].

machinery must efficiently cleave fibers composed of the mutant in order to allow for efficient inheritance to daughter cells.

Interestingly, all of the mutants except FP21N had some degree of a molecular incompatibility that had to be overcome for maintenance to occur (Fig. 2A). While the degree of incompatibility varied, all mutants except FP24N efficiently propagated [*PSI*$^+$] with no detectable prion loss after overcoming this barrier (Fig. 2C and data not shown).

For all mutants that displayed a molecular incompatibility with wild-type [*PSI*$^+$] fibers, we further tested their ability to form prions de novo. Since prion formation occurs by a spontaneous molecular conversion event, increasing the number of molecules increases the likelihood of prion formation. Therefore, we transiently overexpressed the N and M domains of each mutant in the corresponding strain to induce prion formation and tested for the appearance of Ade$^+$ colonies (Fig. 2D). In all mutants tested, Ade$^+$ colony formation was detectable, and increased with overexpression. Additionally, all Ade$^+$ colonies were curable by guanidine HCl, confirming that the Ade$^+$ phenotype was the result of [*PSI*$^+$] formation (data not shown).

Scrambling the ORD of Sup35p

While the first two repeats are necessary for efficient addition of monomers onto pre-existing fibers, deletion of any repeat inhibits Hsp104p dependent prion propagation (Parham, Resende et al. 2001; Osherovich, Cox et al. 2004; Shkundina, Kushnirov et al. 2006). To separately examine the primary sequence requirements

of the ORD region for [*PSI*$^+$] propagation and molecular compatibility, we randomized the order of the amino acids in either all of the repeats or just the last 3½ repeats while keeping amino acid composition constant. In each case, three scrambled constructs were generated.

Deletion of the last 3½ repeats blocks [*PSI*$^+$] propagation. However, all three of the mutants in which the last 3½ repeats were scrambled (Scr½ORD1, 2 and 3) showed no molecular incompatibility with wild-type [*PSI*$^+$] aggregates and efficiently maintained [*PSI*$^+$] without any detectable prion loss (Fig. 3A and data not shown) demonstrating that the primary sequence of the last 3½ repeats plays little or no role in molecular compatibility or efficient prion propagation. By contrast, scrambling all 5½ repeats resulted in significant molecular incompatibility between the scrambled mutants and wild-type [*PSI*$^+$] aggregates. Of the three constructs in which all 5½ repeats were scrambled (ScrORD1, 2 and 3), only one mutant, ScrORD-1, was capable of overcoming this molecular incompatibility (Fig. 3A). Although this was a rare event, once overcome, maintenance of the mutant prion was indistinguishable from wild type [*PSI*$^+$] propagation (data not shown). Although the ScrORD 2 and ScrORD 3 mutant proteins were completely incompatible with wild-type [*PSI*$^+$] aggregates, they could be induced to form stable, curable prions (Fig. 3B and 3C), demonstrating that their failure to propagate wild type [*PSI*$^+$] was a result of molecular incompatibility, not an intrinsic inability to propagate prions.

These results show that the primary sequence of the first two repeats is an important determinant of molecular compatibility, consistent with results from in vitro experiments indicating that the first two repeats are involved in critical

Figure 2.3

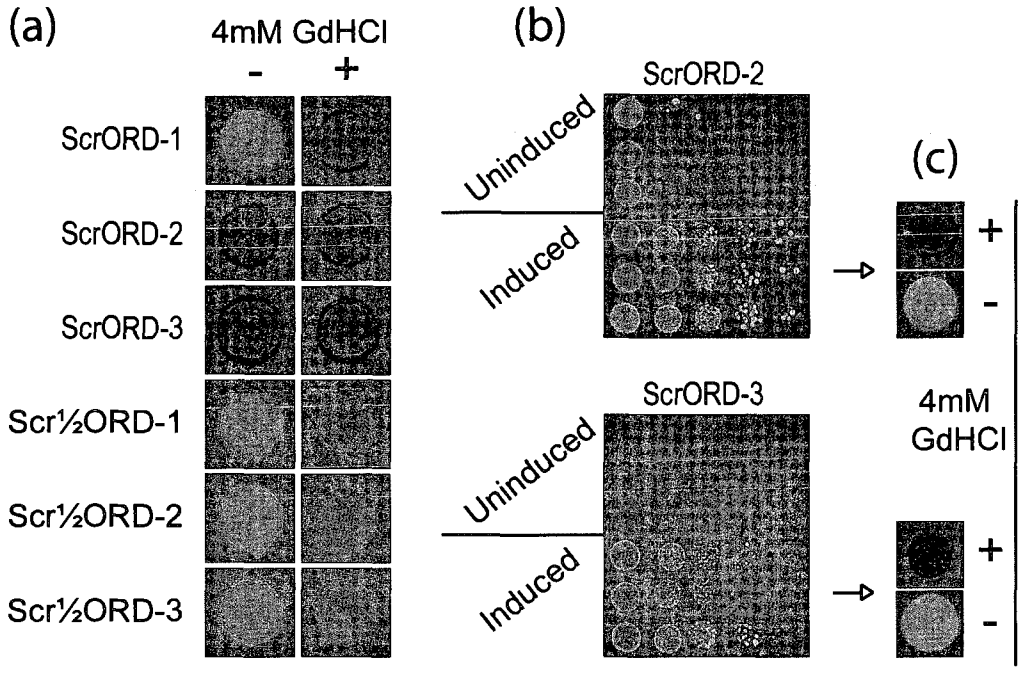


Figure 2.3: Scrambling the ORD does not prevent prion maintenance. a) Maintenance of [*PSI*$^+$] by *SUP35* derivatives in which the primary sequence of all of the ORD or the last 3½ repeats of the ORD was scrambled. Plasmids expressing mutant versions of *SUP35* were transformed into a *SUP35*-deleted [*PSI*$^+$] strain that expressed wild-type *SUP35* from a *URA3* plasmid. After counterselection to remove the wild-type *SUP35* plasmid, strains were plated for single colonies on YPD. Individual white colonies were streaked onto YPD or YPD plus 4 mM guanidine HCl and then spotted onto YPD to test for [*PSI*$^+$]. Because no white colonies were observed for ScrORD-2 and -3 after counterselection, pink colonies were tested. b) Scrambled ORD Sup35p mutants tested for de novo prion formation with (induced) and without (uninduced) overexpression of the matching NM domain. c) Ade$^+$ colonies induced by overexpression of ScrORD-2 and ScrORD-3 were grown on YPAD with and without 4mM GdHCl and tested for [*PSI*$^+$].

contacts within the core of Sup35p amyloid fibers (Toyama, Kelly et al. 2007). However, once this incompatibility barrier is overcome, the repeats per se are not required for [*PSI*$^+$] maintenance.

Replacing the ORD with scrambled PrP repeats

We hypothesized that the ability of the repeats from the mammalian prion protein PrP to substitute for Sup35p repeats in promoting prion maintenance would likewise be primary sequence independent. To test this hypothesis, we generated three constructs (ScrPrP-1, 2 and 3) in which we replaced the last 3½ repeats of Sup35p with a scrambled PrP ORD. Two of the three constructs were capable of maintaining [*PSI*$^+$] with no molecular incompatibility (Fig. 4A and data not shown). ScrPrP-3 was incompatible with the wild-type [*PSI*$^+$] prion, but was able to form prions de novo (Fig. 4B). Once formed, these prions were maintained efficiently in an Hsp104p dependent manner, with no detectable prion loss (Fig. 4C and data not shown). While these experiments do not reveal the role of the PrP ORD in the context of the PrP protein, these results do indicate that the ability of the PrP repeats to substitute for the Sup35p repeats in promoting [*PSI*$^+$] propagation is not dependent on the primary sequence of the repeats.

Replacing the ORD with a scrambled Sup35p nucleation domain.

Because the nucleation domain (amino acids 1-39) and the ORD have significantly different amino acid compositions, we hypothesized that these compositional differences could be responsible for their differing functions:

Figure 2.4

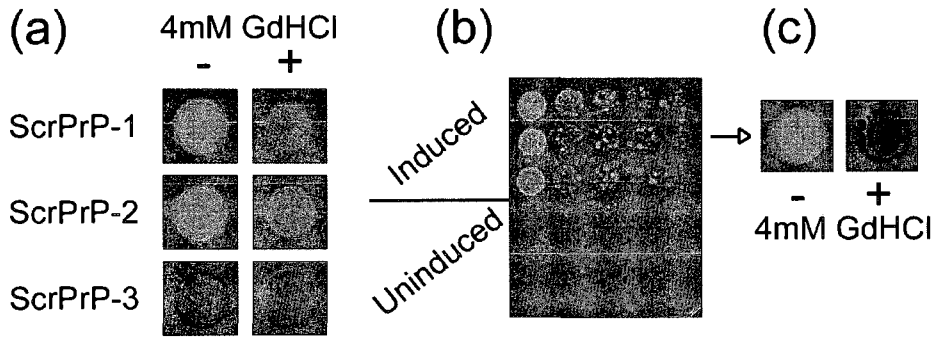


Figure 2.4: ScrPrP mutants maintain wild type [*PSI*$^+$]. a) Maintenance of [*PSI*$^+$] by fusion proteins in the last 3½ repeats of the ORD was replaced with scrambled PrP ORDs. b) De novo prion formation for scrambled PrP mutants. c) Curability of the Ade$^+$ phenotype in the scrambled PrP mutants.

supporting prion nucleation and chaperone-dependent propagation, respectively. To test this hypothesis, we replaced the last 3½ repeats of the ORD with a scrambled Sup35p PrD nucleation domain. Surprisingly, the resulting mutant proteins (ScrNuc1, 2 and 3) were all capable of efficiently maintaining and propagating the [*PSI*$^+$] prion without any detectable prion loss (Fig. 5A). ScrNuc-1 and ScrNuc-3 showed no molecular incompatibility with wild-type [*PSI*$^+$] aggregates; ScrNuc2 show modest molecular incompatibility, but once this barrier was overcome; it stably propagated [*PSI*$^+$] in an Hsp104p dependent manner (Fig. 5B). These results indicate that the compositional differences between the nucleation domain and the ORD do not explain their distinct functions.

HSP104p over-expression cures scrambled prions.

To test whether the ORD could be responsible for the unique sensitivity of [*PSI*$^+$] to Hsp104p overexpression, we tested whether scrambled mutants were similarly cured by Hsp104p overexpression. We transformed the Scrambled PFD Mutants (Ross, Edskes et al. 2005) and the ScrORD mutants with a plasmid that constitutively over-expresses *HSP104* from an *ADH1* promoter. After three days of growth, the transformed colonies were spotted onto YPD to detect prion loss (Fig. 6). Although prion curing is accomplished with varying efficiencies for each mutant, they are all clearly sensitive to Hsp104p overexpression despite lacking oligopeptide repeats, indicating that the primary sequence of the ORD is not responsible for the sensitivity of [*PSI*$^+$] aggregates to Hsp104p overexpression.

Figure 2.5

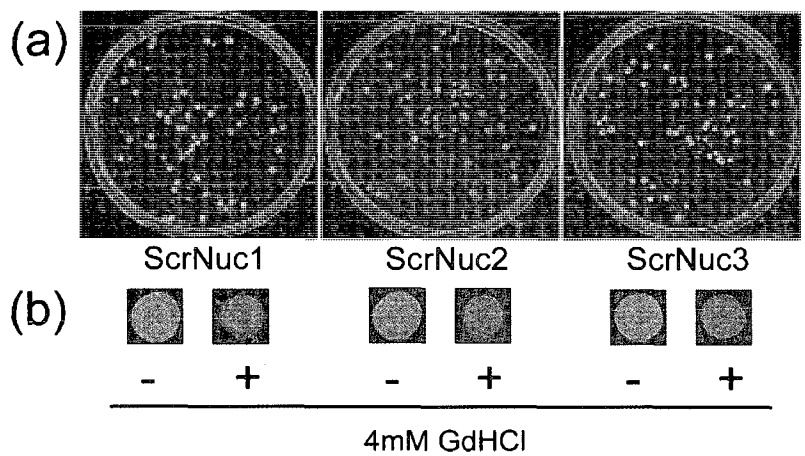


Figure 2.5: ScrNuc mutants maintain wild type [*PSI*$^+$]. a) Molecular compatibility between wild-type [*PSI*$^+$] aggregates and *SUP35* mutants in which the last 3½ repeats of the ORD was replaced with a scrambled Sup35p. b) Curability of Ade$^+$ phenotype in the scrambled nucleation domain mutants.

Figure 2.6

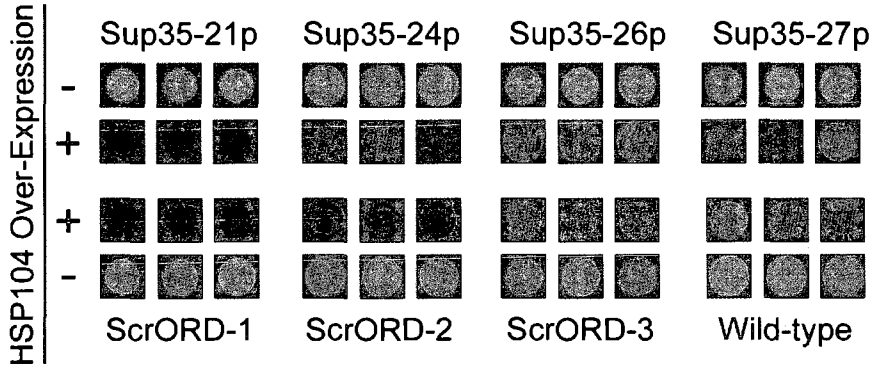


Figure 2.6: Hsp104p overexpression cures Sup35 ORD mutants. Strains expressing scrambled versions of *SUP35*, scrambled ORD mutants or wild-type *SUP35* were transformed either with a plasmid expressing *HSP104* from the *ADH1* promoter (+) or an empty vector. After 3 days growth, transformed strains were spotted onto YPD to test for [*PSI*$^+$].

Discussion

In contrast to the evidence that scrambling the Sup35p PFD does not inhibit [*PSI*$^+$] propagation (Ross, Edskes et al. 2005), there exists a large body of literature that touts the necessity of the Sup35p ORD for [*PSI*$^+$] maintenance (Liu and Lindquist 1999; Parham, Resende et al. 2001; Osherovich, Cox et al. 2004; Shkundina, Kushnirov et al. 2006). Two other yeast prion proteins, Rnq1p (Vitrenko, Pavon et al. 2007) and New1p (Osherovich, Cox et al. 2004), along with the mammalian prion protein PrP, have repeat-like domains; however, other yeast prion forming proteins such as wild-type Ure2p and the scrambled versions of Ure2p and Sup35p are capable of forming and propagating prions without the presence of oligopeptide repeats, suggesting that repeats are not essential to prion propagation. Three explanations for this disconnect have been proposed. Shkudina *et. al.* proposed that although repeats may not be required for artificial prions such as those formed by scrambled versions of Sup35p, they may be required within the context of naturally-formed [*PSI*$^+$] variants (Shkundina, Kushnirov et al. 2006). Alternatively, the repeats could explain the unique properties of wild-type Sup35p. Finally, we hypothesized that the distinct activities of the ORD and Q/N-rich nucleation domain of Sup35p were a result of their distinct compositions.

Our experiments, designed to resolve the conflicting data regarding the role of the ORD in [*PSI*$^+$] maintenance, suggests that none of these three explanations are correct. While this region of the PFD is important for efficient [*PSI*$^+$] propagation, neither the primary sequence nor the composition of the oligopeptide repeats is essential for Hsp104p-dependent [*PSI*$^+$] maintenance. Although there are subtle

differences in prion formation and Hsp104p sensitivity among our mutants (and there may be other subtle differences that we have not detected), disruption of the repeats did not consistently reduce the efficiency of prion formation, propagation or sensitivity to Hsp104p overexpression.

In scrambling the primary sequence of the ORD, we found that repeats 1 and 2 are important determinants for molecular compatibility between the wild-type prion and our mutant proteins, while repeats 3 through 5½ have only a weak affect on molecular compatibility. This observation is consistent with in vitro data that suggest repeats 1 and 2 are within the core of the amyloid fiber (Toyama, Kelly et al. 2007), and in vivo data that implicates these repeats being important to fiber formation (Liu and Lindquist 1999). However, these repeats are not critical to the integrity of the fiber; once mutants with repeats 1 and 2 disrupted overcame the molecular incompatibility with wild-type [*PSI*$^+$], [*PSI*$^+$] was propagated in a manner indistinguishable from wild-type.

Our data indicating that the ORD plays no role in [*PSI*$^+$] maintenance seems in contrast to the wide body of literature supporting the importance of the ORD. This new finding may be due to the novel approach employed; past experiments concluding the ORD is important to [*PSI*$^+$] propagation were based on data from Sup35p mutants containing either truncated or expanded ORDs. The problem with such studies is that in addition to changing the number of repeats, truncating or expanding the ORD also changes the total length of the PFD, as well as the spacing between the nucleation domain and the Sup35p C-terminus, making it easy to misinterpret the basis for the observed effects.

Since neither the primary sequence nor the composition of the ORD is required for efficient [*PSI*$^+$] formation and propagation, it remains puzzling why repeat deletions prevent prion propagation. Since deletions of repeats reduce Hsp104p-dependent fragmentation of [*PSI*$^+$] aggregates (Shkundina, Kushnirov et al. 2006), this region may act as an Hsp104p binding site as previously proposed (Osherovich, Cox et al. 2004; Shkundina, Kushnirov et al. 2006). If so, our data argue that the sequence requirements for Hsp104p binding and recruitment must be quite broad. Because Hsp104p needs to recognize a wide variety of targets, such sequence flexibility for chaperone recognition should not be too surprising.

Alternatively, the region of the ORD may act as a spacer separating the PFD from the M and C domains of Sup35p. For example, as discussed by Shkundina et al (Shkundina, Kushnirov et al. 2006) the MC domains of Sup35p may shield the nucleation domain from Hsp104p binding, such that without sufficient separation between the nucleation and MC domains, Hsp104p binding is prevented. This hypothesis is supported by truncation studies, which show that deleting one or all of the repeats within the Sup35p ORD prevents [*PSI*$^+$] propagation. Since neither the primary sequence nor the composition of the ORD is necessary for propagation, yet any random sequence can not replace this domain (Crist, Nakayashiki et al. 2003), we hypothesize that the only requirements for this region is that it provide enough length for efficient Hsp104p binding and be able to form a β-sheet structure upon prion formation. A bias towards disorder promoting residues may also be important in keeping the prion domain in an open conformation that is readily accessible for prion formation.

This hypothesis provides a simple rationalization for why repeats are present in several prion-forming proteins. Repeat expansions are a simple way to genetically expand a domain. Amyloid proteins that form fibers unable to undergo chaperone-mediated fiber cleavage to produce new heritable propagons may become prions once a repeat expanding genetic mutation occurs in the amyloid forming region. Such expansions would allow for efficient chaperone binding and cleavage. By analogy, expansion of DNA segments encoding for glutamine are seen in a variety of diseases (Shao and Diamond 2007), yet in all cases, the glutamines are encoded for by the codon CAG. This could be interpreted as suggesting that expanded CAG repeats are toxic while CAA repeats are not. However, a more likely explanation is that both would be equally toxic, but that CAG repeats are observed in disease because CAG repeats are more prone to expansion. Similarly, it is possible that any expansion within a PFD would increase amyloid-forming propensity, but that repeats (or repeat expansions) are specifically seen because they are the simplest mechanism for domain expansion.

In light of our data, the validity of using the Sup35p model system to assess the role of the PrP repeats must be reexamined. Our data shows that the ORD of Sup35p is not sequence or composition specific. Thus, the ability of the PrP repeats to substitute for the Sup35p ORD may simply reflect the fact that a diverse array of sequences can serve this function. If so, the repeats may serve a completely different function in mammalian prion diseases. Alternatively, the PrP and Sup35p oligopeptide repeats may play the same roles in prion formation; expansions within the PrP ORD are associated with disease (Collinge, Brown et al. 1992; Poulter, Baker

et al. 1992), but this may not be due to the actual repeats per se, but due to the fact that expansion of a disordered, aggregation prone region is likely to increase amyloid formation propensity.

Materials and Methods

Strains and media: Standard yeast media and methods were used as previously described (Sherman 1991), except the YPD contained 0.5% yeast extract instead of the standard 1%. In all experiments, yeast were grown at 30°C. All experiments were performed in *Saccharomyces cerevisiae* strain 780-1D/pJ533 (Song, Wu et al. 2005). This strain's genotype is *α kar1-1 SWQ5 ade2-1 his3 leu2 trp1 ura3 sup35*::KanMx [PSI^+] [PIN^+]; pJ533 expresses *SUP35* from a *URA3* plasmid as the sole copy of *SUP35* in the cell.

Replacing the Sup35p: ScrORD constructs were designed and constructed as previously described for *URE2*.(Ross, Baxa et al. 2004) Briefly, four overlapping oligonucleotides were used to construct each scrambled ORD (see Supplementary Table 1 for oligonucleotide sequences). Oligonucleotides were combined and amplified by PCR, generating the scrambled ORDs. The N-terminal (non-scrambled) portion of the PFD was amplified in a separate reaction. N-terminal and ORD fragments were combined and re-amplified with EDR869 and EDR871. PCR products were co-transformed with AatII/HindIII-cut pJ526 (from Dan Masison, National Institutes of Health) into yeast strain 780-1D/pJ533. Transformants were

selected on SD-leu and then stamped onto 5-fluoroorotic acid (5-FOA) containing medium to select for loss of pJ533. Plasmids expressing mutant *SUP35*s were confirmed by DNA sequencing.

FP21N, FP21C, FP24N, FP24C, FP26N, FP26C, FP27N and FP27C were similarly constructed in two steps. Fragments from the scrambled Sup35s were amplified by PCR, while N-terminal portion of the PFD was amplified in a separate reaction. N-terminal and scrambled fragments were combined, reamplified and cloned into pJ526.

For each of the ScrNuc, ScrPrP and Scr½ORD constructs except for ScrPrP2 and ScrNuc1, N-terminal and C-terminal fragments of *SUP35* were PCR amplified in separate reaction with primers EDR348 and EDR243, respectively, each paired with the mutagenic primer indicated in Supplementary Table 1. N-terminal and scrambled fragments were combined, reamplified and cloned into pJ526. ScrPrP2 and ScrNuc1 were constructed as the ScrORD constructs.

Testing for prion maintenance and curing: Transformed colonies were re-suspended in water in a 96-well microtiter plate, spotted onto minimal media plates containing 5-FOA and grown for 2-3 days at 30°C to select for loss of pJ533. Cells from the 5-FOA plates were streaked onto YPD plates to test for [*PSI*$^+$].

To test for curability, white Ade$^+$ colonies were grown on YPAD or YPAD plus 4mM guanidine HCl (GdHCl). Single colonies were spotted onto YPD to test for loss of [*PSI*$^+$].

Induction experiments: To generate induction plasmids, the N and M domains of the mutant *SUP35*s were amplified by PCR with primers EDR1008 and EDR969, installing a *Bam*HI site before the start codon, and a stop codon and *Pst*I site after the M domain. PCR products were digested with *Bam*HI and *Pst*I and inserted into *Bam*HI/*Pst*I-cut pKT24 (from Kim Taylor, NABI, Rockville, MD). Ligation products were transformed into *Escherichia coli* and analyzed by DNA sequencing. Induction experiments were performed as previously described (Ross, Edskes et al. 2005).

Prion curing by Hsp104p over-expression: Strains were transformed with pER10, which expresses *HSP104* from the *ADH1* promoter or a vector control (pER41). Transformed colonies were re-suspended in water in a 96-well microtiter plate, spotted onto YPD to test for loss of [PSI^+].

Acknowledgements

This work was supported in part by Basil O'Connor Starter Scholar Research Award Grant No. 5-FY07-104 from the March of Dimes to E.D.R. We would like to thank members of the Ross lab for helpful discussion.

References

Aigle, M. and F. Lacroute (1975). "Genetical aspects of [URE3], a non-mitochondrial, cytoplasmically inherited mutation in yeast." Mol Gen Genet **136**(4): 327-35.

Collinge, J., J. Brown, et al. (1992). "Inherited prion disease with 144 base pair gene insertion. 2. Clinical and pathological features." Brain **115 (Pt 3)**: 687-710.

Cox, B. S. (1965). "PSI, a cytoplasmic suppressor of super-suppressor in yeast." Heredity **26**: 211-232.

Crist, C. G., T. Nakayashiki, et al. (2003). "[PHI+], a novel Sup35-prion variant propagated with non-Gln/Asn oligopeptide repeats in the absence of the chaperone protein Hsp104." Genes Cells **8**(7): 603-18.

DePace, A. H., A. Santoso, et al. (1998). "A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion." Cell **93**(7): 1241-52.

Derkatch, I. L., M. E. Bradley, et al. (2001). "Prions affect the appearance of other prions: the story of [PIN(+)]." Cell **106**(2): 171-82.

Dong, J., J. D. Bloom, et al. (2007). "Probing the role of PrP repeats in conformational conversion and amyloid assembly of chimeric yeast prions." J Biol Chem **282**(47): 34204-12. Epub 2007 Sep 24.

Douglas, P. M., S. Treusch, et al. (2008). "Chaperone-dependent amyloid assembly protects cells from prion toxicity." Proc Natl Acad Sci U S A **105**(20): 7206-11.

Ferreira, P. C., F. Ness, et al. (2001). "The elimination of the yeast [PSI+] prion by guanidine hydrochloride is the result of Hsp104 inactivation." Mol. Microbiol. **40**(6): 1357-69.

Flechsig, E., D. Shmerling, et al. (2000). "Prion protein devoid of the octapeptide repeat region restores susceptibility to scrapie in PrP knockout mice." Neuron **27**(2): 399-408.

Inge-Vechtomov, S., G. Zhouravleva, et al. (2003). "Eukaryotic release factors (eRFs) history." Biol Cell **95**(3-4): 195-209.

Jung, G., G. Jones, et al. (2002). "Amino acid residue 184 of yeast Hsp104 chaperone is critical for prion- curing by guanidine, prion propagation, and thermotolerance." Proc. Natl. Acad. Sci. USA **99**(15): 9936-41.

Jung, G. and D. C. Masison (2001). "Guanidine hydrochloride inhibits Hsp104 activity in vivo: a possible explanation for its effect in curing yeast prions." Curr. Microbiol. **43**(1): 7-10.

Kalastavadi, T. and H. L. True (2008). "Prion protein insertional mutations increase aggregation propensity but not fiber stability." BMC Biochem **9**: 7.

Kisilevsky, R. and P. E. Fraser (1997). "A beta amyloidogenesis: unique, or variation on a systemic theme?" Crit. Rev. Biochem. Mol. Biol. **32**(5): 361-404.

Krishnan, R. and S. L. Lindquist (2005). "Structural insights into a yeast prion illuminate nucleation and strain diversity." Nature **435**(7043): 765-72.

Lacroute, F. (1971). "Non-Mendelian mutation allowing ureidosuccinic acid uptake in yeast." J. Bacteriol. **106**(2): 519-22.

Liu, J. J. and S. Lindquist (1999). "Oligopeptide-repeat expansions modulate 'protein-only' inheritance in yeast." Nature **400**(6744): 573-6.

Liu, J. J., N. Sondheimer, et al. (2002). "Changes in the middle region of Sup35 profoundly alter the nature of epigenetic inheritance for the yeast prion [PSI+]." Proc. Natl. Acad. Sci. USA **99 Suppl 4**: 16446-53.

Nakayashiki, T., K. Ebihara, et al. (2001). "Yeast [PSI+] "prions" that are crosstransmissible and susceptible beyond a species barrier through a quasi-prion state." Mol Cell **7**(6): 1121-30.

Osherovich, L. Z., B. S. Cox, et al. (2004). "Dissection and design of yeast prions." PLoS Biol. **2**(4): E86.

Parham, S. N., C. G. Resende, et al. (2001). "Oligopeptide repeats in the yeast protein Sup35p stabilize intermolecular prion interactions." EMBO J. **20**(9): 2111-9.

Patino, M. M., J. J. Liu, et al. (1996). "Support for the prion hypothesis for inheritance of a phenotypic trait in yeast." Science **273**(5275): 622-6.
Paushkin, S. V., V. V. Kushnirov, et al. (1996). "Propagation of the yeast prion-like [psi+] determinant is mediated by oligomerization of the SUP35-encoded polypeptide chain release factor." EMBO J. **15**(12): 3127-34.
Poulter, M., H. F. Baker, et al. (1992). "Inherited prion disease with 144 base pair gene insertion. 1. Genealogical and molecular studies." Brain **115 (Pt 3)**: 675-85.
Prusiner, S. B. (1982). "Novel proteinaceous infectious particles cause scrapie." Science **216**(4542): 136-44.
Prusiner, S. B. (1998). "Prions." Proc Natl Acad Sci U S A **95**(23): 13363-83.
Prusiner, S. B. (1998). "The prion diseases." Brain Pathol **8**(3): 499-513.
Prusiner, S. B., M. R. Scott, et al. (1998). "Prion protein biology." Cell **93**(3): 337-48.
Ross, E. D., U. Baxa, et al. (2004). "Scrambled Prion Domains Form Prions and Amyloid." Mol. Cell. Biol. **24**(16): 7206-7213.
Ross, E. D., H. K. Edskes, et al. (2005). "Primary sequence independence for prion formation." Proc. Natl. Acad. Sci. USA **102**(36): 12825-12830.
Santoso, A., P. Chien, et al. (2000). "Molecular basis of a yeast prion species barrier." Cell **100**(2): 277-88.
Schlumpberger, M., S. B. Prusiner, et al. (2001). "Induction of distinct [URE3] yeast prion strains." Mol. Cell. Biol. **21**(20): 7035-46.
Serio, T. R., A. G. Cashikar, et al. (2000). "Nucleated conformational conversion and the replication of conformational information by a prion determinant." Science **289**(5483): 1317-21.
Shao, J. and M. I. Diamond (2007). "Polyglutamine diseases: emerging concepts in pathogenesis and therapy." Hum Mol Genet **16 Spec No. 2**: R115-23.
Sherman, F. (1991). "Getting started with yeast." Methods Enzymol. **194**: 3-21.
Shewmaker, F., E. D. Ross, et al. (2008). "Amyloids of shuffled prion domains that form prions have a parallel in-register beta-sheet structure." Biochemistry **47**(13): 4000-7. Epub 2008 Mar 7.
Shewmaker, F., R. B. Wickner, et al. (2006). "Amyloid of the prion domain of Sup35p has an in-register parallel beta-sheet structure." Proc Natl Acad Sci U S A **103**(52): 19754-9. Epub 2006 Dec 14.
Shkundina, I. S., V. V. Kushnirov, et al. (2006). "The role of the N-terminal oligopeptide repeats of the yeast sup35 prion protein in propagation and transmission of prion variants." Genetics **172**(2): 827-35.
Shorter, J. and S. Lindquist (2004). "Hsp104 catalyzes formation and elimination of self-replicating Sup35 prion conformers." Science **304**(5678): 1793-7.
Shorter, J. and S. Lindquist (2006). "Destruction or potentiation of different prions catalyzed by similar Hsp104 remodeling activities." Mol Cell **23**(3): 425-38.
Sipe, J. D. and A. S. Cohen (2000). "Review: history of the amyloid fibril." J. Struct. Biol. **130**(2-3): 88-98.
Sondheimer, N. and S. Lindquist (2000). "Rnq1: an epigenetic modifier of protein function in yeast." Mol. Cell **5**(1): 163-72.
Sondheimer, N., N. Lopez, et al. (2001). "The role of Sis1 in the maintenance of the [RNQ+] prion." EMBO J. **20**(10): 2435-42.

Song, Y., Y. X. Wu, et al. (2005). "Role for Hsp70 chaperone in Saccharomyces cerevisiae prion seed replication." Eukaryot Cell **4**(2): 289-97.

Stansfield, I., K. M. Jones, et al. (1995). "The products of the SUP45 (eRF1) and SUP35 genes interact to mediate translation termination in Saccharomyces cerevisiae." EMBO J. **14**(17): 4365-73.

Tank, E. M., D. A. Harris, et al. (2007). "Prion protein repeat expansion results in increased aggregation and reveals phenotypic variability." Mol Cell Biol **27**(15): 5445-55. Epub 2007 Jun 4.

Ter-Avanesyan, M. D., A. R. Dagkesamanskaya, et al. (1994). "The SUP35 omnipotent suppressor gene is involved in the maintenance of the non-Mendelian determinant [psi+] in the yeast Saccharomyces cerevisiae." Genetics **137**(3): 671-6.

Ter-Avanesyan, M. D., V. V. Kushnirov, et al. (1993). "Deletion analysis of the SUP35 gene of the yeast Saccharomyces cerevisiae reveals two non-overlapping functional regions in the encoded protein." Mol. Microbiol. **7**(5): 683-92.

Toyama, B. H., M. J. Kelly, et al. (2007). "The structural basis of yeast prion strain variants." Nature **449**(7159): 233-7. Epub 2007 Sep 2.

Tuite, M. F., C. R. Mundy, et al. (1981). "Agents that cause a high frequency of genetic change from [psi+] to [psi-] in Saccharomyces cerevisiae." Genetics **98**(4): 691-711.

Vitrenko, Y. A., M. E. Pavon, et al. (2007). "Propagation of the [PIN+] prion by fragments of Rnq1 fused to GFP." Curr Genet **51**(5): 309-19.

Wadsworth, J. D., A. F. Hill, et al. (2003). "Molecular and clinical classification of human prion disease." Br Med Bull **66**: 241-54.

Wickner, R. B. (1994). "[URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae." Science **264**(5158): 566-9.

Zhouravleva, G., L. Frolova, et al. (1995). "Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3." EMBO J. **14**(16): 4065-72.

Chapter 3: Compositional determinants of prion formation in yeast

James A. Toombs, Blake R. McCarty and Eric D. Ross

Blake R. McCarty contributed to the work described in Figure 3.1

A modified version of this work was submitted for publication (2009)

Summary

Numerous prions (infectious proteins) have been identified in yeast that result from conversion of soluble proteins into β-sheet-rich amyloid-like protein aggregates. Yeast prion formation is driven primarily by amino acid composition. However, yeast prion domains are generally lacking in the bulky hydrophobic residues most strongly associated with amyloid formation, and are instead enriched in glutamines (Q) and asparagines (N). Q/N-rich domains are thought to be involved in both disease-related and beneficial amyloid formation. Q/N-rich domains are over-represented in eukaryotic genomes, but predictive methods have not yet been developed to efficiently distinguish between prion and non-prion Q/N-rich domains. We have developed a novel in vivo assay to quantitatively assess how composition affects prion formation. Using our results, we have defined the compositional features that promote prion formation, allowing us to accurately distinguish between Q/N-rich domains that can form prion-like aggregates and those that can not.

Introduction

Amyloid fibers are associated with a large number of neurodegenerative diseases and systemic amyloidoses. Amyloid fibrils are rich in cross-beta quaternary structure in which β-strands are perpendicular to the long axis of the fibril (Chiti and Dobson 2006).

[URE3] and [*PSI*$^+$] are the prion (infectious protein) forms of the *Saccharomyces cerevisiae* proteins Ure2p and Sup35p, respectively (Wickner 1994); formation of both prions involves conversion of the native proteins into an infectious, amyloid form. Ure2p and Sup35p have served as powerful model systems for examining the basis for amyloid formation and propagation. Both proteins possess a well-ordered functional domain responsible for executing the normal function of the protein, while a functionally and structurally separate glutamine/asparagine (Q/N) rich intrinsically disordered domain is necessary and sufficient for prion aggregation and propagation (Ter-Avanesyan, Dagkesamanskaya et al. 1994; Masison and Wickner 1995; Masison, Maddelein et al. 1997; Bradley and Liebman 2004).

Six other prions have also been identified in yeast: [*PIN*$^+$] (Sondheimer and Lindquist 2000; Derkatch, Bradley et al. 2001), [*NU*$^+$] (Santoso, Chien et al. 2000), [*SWI*$^+$] (Du, Park et al. 2008), [*OCT*$^+$] (Patel, Gavin-Smyth et al. 2009), [*MCA*] (Nemecek, Nakayashiki et al. 2009) and [*MOT3*] (Alberti, Halfmann et al. 2009), with evidence suggesting the existence of several others (Alberti, Halfmann et al. 2009). Like Ure2p and Sup35p, each of these prion proteins contains a Q/N-rich prion-forming domain (PFD). Mutational studies of the PFDs of Ure2p and Sup35p have

shown that amino acid composition is the major driving force behind prion formation (Ross, Baxa et al. 2004; Ross, Edskes et al. 2005). However, we currently have little ability to predict how changes in amino acid composition will affect prion propensity.

A variety of algorithms have been developed to predict a peptide's propensity to form amyloid fibrils based on its amino acid sequence, including BETASCAN (Bryan, Menke et al. 2009), TANGO (Fernandez-Escamilla, Rousseau et al. 2004), Zyggregator (Tartaglia, Pawar et al. 2008), SALSA (Zibaee, Makin et al. 2007) and PASTA (Trovato, Seno et al. 2007). These algorithms have been quite successful at identifying regions prone to amyloid aggregation and predicting the effects of mutations on aggregation propensity for many amyloid-forming proteins. However, they have generally been quite ineffective for Q/N-rich amyloid proteins such as the yeast prion proteins. For example, using the statistical mechanics-based algorithm TANGO (Fernandez-Escamilla, Rousseau et al. 2004), which predicts aggregation propensity based on a peptide's physico-chemical properties, Linding et al. found that the Sup35p and Ure2p PFDs both completely lack predicted β-aggregation nuclei (Linding, Schymkowitz et al. 2004). Similarly, Ure2p and Sup35p are lacking in the bulky hydrophobic residues predicted by Zyggregator to nucleate prion formation.

Why are these algorithms so effective for many amyloid-forming proteins, yet so ineffective for the yeast prion proteins? Most amyloid-forming domains are highly hydrophobic, and increased hydrophobicity is correlated with increased amyloid aggregation propensity (Chiti and Dobson 2006). By contrast, despite

having few charged residues, the yeast PFDs are all highly polar, largely due to the high concentration of Q/N residues and under-representation of hydrophobic residues.

Similar Q/N-rich domains are highly over-represented in eukaryotic genomes (Michelitsch and Weissman 2000), but not all Q/N-rich domains appear able to form prions. Recently, the 100 Q/N-rich domains from yeast that are most compositionally similar to known PFDs were analyzed in a series of in vivo and in vitro assays (Alberti, Halfmann et al. 2009). Only eighteen showed prion-like activity in all assays, and many failed to show prion-like activity in any assay. While the discovery of prion-like activity in many yeast proteins was very exciting, this low success rate highlights our current inability to distinguish Q/N-rich proteins that are likely to form prions from those that are not. Understanding the sequence features that promote prion formation by Q/N-rich domains is critical for accurately predicting prion potential.

The significant compositional differences between the yeast prions and most other amyloid proteins suggest that there may be two fundamentally distinct classes of amyloid-forming proteins driven by different types of interactions. Specifically, Q/N-residues, which have been predicted to have relatively low amyloid propensity in the context of hydrophobic amyloid domains (Pawar, Dubay et al. 2005), may promote amyloid formation when present at sufficiently high density. Stacking of Q/N residues to form polar zippers has been proposed to stabilize amyloid fibrils (Perutz, Pope et al. 2002). Consistent with this hypothesis, mutational studies of Sup35p indicate that Q/N residues are critical for driving [*PSI*$^{+}$] formation (DePace,

Santoso et al. 1998), and expanded poly-Q or poly-N tracts are sufficient to drive amyloid aggregation (Zoghbi and Orr 2000; Peters and Huang 2007). Alternatively, the compositional differences between the yeast prions and other amyloid proteins may simply reflect a function of these domains other than driving amyloid formation. There is no reason to think that the yeast PFDs are optimized for maximum aggregation potential.

We have developed an in vivo method to determine the amino acid requirements for prion formation by Q/N-rich proteins. As expected, we found proline and charged residues to be strongly inhibitory to prion formation; but surprisingly, despite being largely under-represented in yeast PFDs, hydrophobic residues strongly promoted prion formation. Using these data, we were able to distinguish with greater than 90% accuracy between Q/N-rich domains that can form prion-like aggregates and those that can not. These experiments provide detailed insight into the compositional requirements for yeast prion formation through a quantitative, in vivo approach.

Results:

Mapping the Sup35-27p prion domain

Randomizing the order of the amino acids in the Sup35p PFD while maintaining amino acid composition does not prevent [*PSI*$^+$] formation (Ross, Edskes et al. 2005). We used one of these scrambled versions of Sup35p, Sup35-27p, as a template for mutagenesis. Sup35-27p was chosen for three reasons. First, we hypothesized that most mutations would reduce prion propensity. Wild-type

Sup35p very rarely forms prions without over-expression, making it difficult to isolate prion-forming clones upon mutagenesis. Sup35-27p forms prions de novo with greater efficiency than wild-type Sup35p, allowing for isolation of a broader range of prion-forming clones. Second, any specific prion-promoting primary sequence elements or any binding sites within the PFD for interacting proteins were likely disrupted by randomization, simplifying interpretation of the results of our library screen. Finally, solid state NMR suggests that Sup35-27p forms fibrils that are structurally similar to those formed by wild-type Sup35p (Shewmaker, Ross et al. 2008).

To identify ideal regions of the Sup35-27p PFD to target for random mutagenesis, we mapped the prion-promoting regions of the PFD through deletion analysis. Prion formation was detected by monitoring nonsense suppression of the *ade2-1* allele (Cox 1965). *ade2-1* mutants are unable to grow without adenine and form red colonies when grown in the presence of limiting adenine due to accumulation of a pigment derived from the substrate of Ade2p. Sup35p is a translation termination factor. [*PSI*$^+$] formation inactivates Sup35p, resulting in increased read-through of stop codons (Kushnirov and Ter-Avanesyan 1998), allowing *ade2-1* [*PSI*$^+$] cells to grow without adenine and form white colonies on limiting adenine.

SUP35-27 carrying various deletions was expressed as the sole copy of *SUP35* in the cell. We monitored Ade$^+$ colony formation with and without transient overexpression of the matching PFD. To confirm that Ade$^+$ colony formation was a result of [*PSI*$^+$] formation, we tested individual Ade$^+$ colonies to determine whether

the Ade+ phenotype was curable by guanidine. Growth on medium containing low concentrations of guanidine cures [*PSI*+] (Tuite, Mundy et al. 1981) by inhibiting Hsp104p (Ferreira, Ness et al. 2001; Jung and Masison 2001).

We found that 20 amino acids could be removed from either end of the PFD while retaining [*PSI*+] formation, both with and without prion domain overexpression (Figure 1 and data not shown). Deletion analysis of the PFD showed varying levels of importance for different regions within the prion core (Figure 1B), with the region of amino acids 31-50 particularly sensitive to deletion.

Random mutagenesis of *SUP35-27*

Based on our deletion data, we targeted amino acids 31-50 Sup35-27p for random mutagenesis. In preliminary experiments, we tested the optimal size of the mutated region. Mutagenesis of 12 amino acids almost entirely eliminated prion formation, while mutagenesis of either four or eight amino acids still allowed for an easily detectable level of prion formation (data not shown). Because a larger region of mutagenesis would provide more data and increase the stringency of selection, eight amino acids were mutated in all subsequent experiments. Amino acids 39-46 of the PFD were initially targeted for random mutagenesis because these residues lie within the region that seems critical for prion formation, and because the composition of this region is fairly representative of the wild-type Sup35 PFD.

We used an oligonucleotide-based mutagenesis approach. An oligonucletide was designed that annealed to the regions flanking codons 39-46, but in which the codons 39-46 were replaced with the sequence $(NNB)_8$, where N represents any of

Figure 3.1

A. Sup35-27p PFD (amino acids 1-114)

MSYQGYQNGQ YYQNKFYDQN GSYQGSGYNN NQYGQQQQ***YN***
YQQAGGYNYA NQYG***NPNQRP*** ***GQ***GNGQQSQQ QQYQRPQYNQ
YYQAQNNQPQ GFQNDGYNQG GGNANGAPNY GLYQ

B.

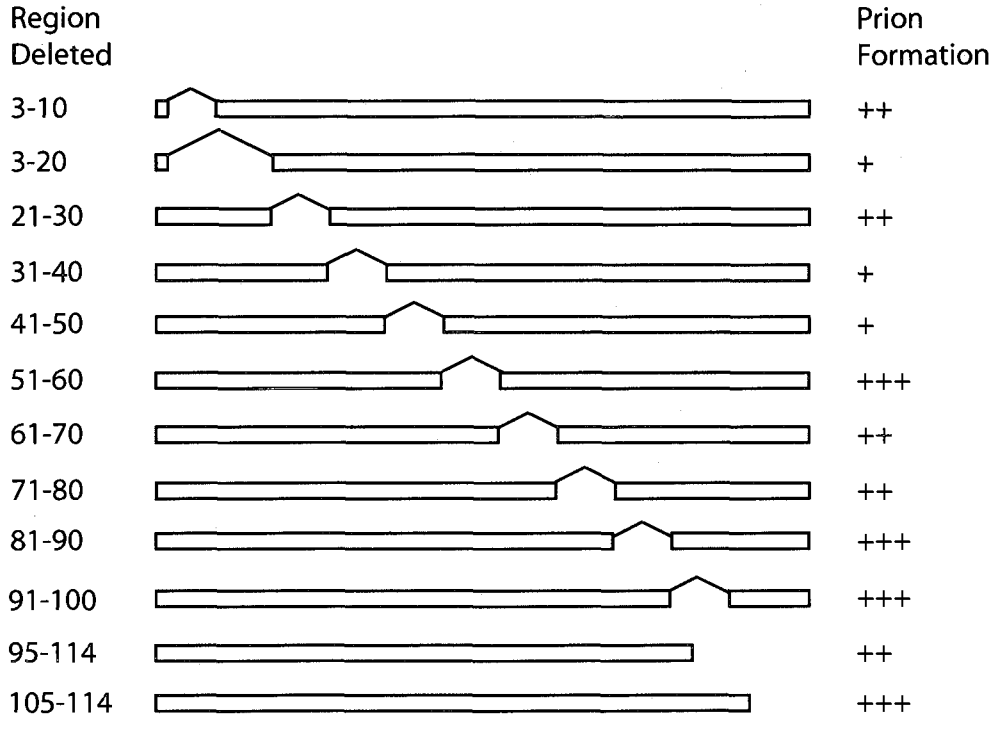


Figure 3.1: Truncation mapping of the Sup35-27p PFD. (A) Sequence of Sup35-27, with regions mutated in Library 1 and 2 in bold italics. (B) Ten amino acid segments were deleted from the Sup35-27p PFD. Prion formation was induced by over-expression. All mutants produced stable curable prions. The relative frequency of prion formation was recorded as low (+), medium (++) or high (+++).

the four nucleotides and B represents any nucleotide except adenine. Excluding adenine from the final position prevents insertion of two of the three stop codons without excluding any amino acids. This oligonucleotide was used to build a library of randomly mutated versions of *SUP35-27.* The library was transformed into yeast cells in which the sole copy of *SUP35* was expressed from a plasmid. Using plasmid shuffling, the wild-type *SUP35* was replaced with the random library.

Each clone was screened for Sup35p activity using the *ade2-1* allele. All library clones were spotted onto medium containing limiting adenine and onto medium lacking adenine. Colonies that grew red on limiting adenine and did not grow without adenine were pooled. To prevent rare strong prion-forming clones from dominating selection, clones were pooled into mini-libraries consisting of approximately 50 clones. With this size mini-library, we were able to isolate a single prion-forming clone from about half of the mini-libraries. *SUP35* was sequenced from individual library clones prior to prion selection to generate a naïve library data set.

Mini-libraries were plated onto medium lacking adenine to select for [*PSI*$^+$] formation. To distinguish Ade$^+$ cells resulting from [*PSI*$^+$] formation from those resulting from DNA mutation, we tested individual Ade$^+$ colonies to determine whether the Ade$^+$ phenotype was curable by guanidine. Cells were grown on YPD with and without guanidine, and then tested for loss of the Ade$^+$ phenotype (data not shown). *SUP35* from cells that stably maintained the Ade$^+$ phenotype on YPD but lost it on YPD plus guanidine were sequenced. We successfully isolated 27 such

Table 3.1: Mutated sequences from selected and unselected libraries.

Libaray 1[a]		Library 2[b]	
[PSI+] Isolates	Naïve library	[PSI+] Isolates	Naïve library
VNIFPYYN	TDPWVPHP AQDSHPDI	FANHAHWV SFSYVTFP	SVSDHTNP QTTTAIHA
VTSGSYNT	NPEVPNAN NNPQYLFK	GTTYAPLF CQINWRTA	KGRVSGPE PHEAVSSC
ASNIVMNC	THHSHTLP DERPWCPE	WNAFSTYS GPPFPGQN	ATSPVPRH RRHYAPSI
AHTTNMIV	YLPFMDTP GPTMNNRD	HTVHHIYP VASWASVG	YEYSPLQH KYMYHANM
YNCSVNML	PPIVKPRT THRHNKHR	LNTFPHSY YREGDNLW	TMTDLPYL LADSNTPR
FSIYMPYK	VDDRHMFS KGSPSTPT	DIMTNNAE HTLVFNDR	ESILWASQ SLAAPRDN
LLVHSNAI	CKSVCNFD EAPSKSAQ	SQDYSSYD	FTRAKSRT FWIDGSAD
WGARQFNI	GISTRSQE RPERRSNP	CINTGLWL	TTSYHPEL DRHYFAGS
VTTDILAM	VSLSKNRL ICWHTEPY	HLHMSMLS	VAHCRHPL IRTHMSSK
RRDYLTRF	LRDPDTCS CIKHINSI	DRHYFAGS	SSTLLDPK ARNMTRYL
STVICGVI	RKATDLFP PVPSSSQP	GGPIFNTK	IETHFTLS RAYDILPV
IHFWPRAP	TAYVRHID GANSAITN	SFMAVETR	APHGLGPT NEDPGTDT
HSNVSVIH	DRYKGKPH SHLWRRNR	TWDGIGYR	RCSDSQGV SRSIRYDN
TWAPIMVY	DPNAALVF DSHTGTPR	SPPFETSP	VHHDPVST SQDYSSYD
MFQHGIGV	HIHPLFIH STVPPPHH	GVNTHTSY	HPIMSSLS
TRIWNFSG	TLARRDPP VNCARGTA	SIHMRVSS	LGPVHYRN
YHSVEFRI	PNASGIHY QVASQNGR	HNDRTAFM	SMHNGTHR
TTVNHHFN	ADSASNAS SSNKFMHT	PQNQTWAD	DGPTYDWT
GSLSLQYF	NGPAYPLA GFTKALPG	PDYFFHPT	PYKAATRN
IFDIANHS	SVNPALYR ALSSRQWS	HVPSPAHQ	PTYNDPST
LQPCYCSR	SGVSTAVR IDKNLMSH	DSDHHFWP	LSQSYVQE
MLSSNFIH	LNRITLRN CFLRSYMG	TSNTIIRA	YDSGTPPK
SSGPLNFI	IVPRNVNC VALIPKTA	DCLGYPGL	SQQRFNPT
CLSPAECR	NISPFSKD HNLANHSH	SMHNGTHR	HRDNCRTR
QFVARVFR	MTQNPHIF KMTTNTKH	ESILWASQ	PPQAVYPP
LKSVITWN	LSARPLGH	PRLTNHSS	QHASGRDG
SVHVNSTS	LGNPTFHY	FWMQRNSC	QTRFYGIH

[a] Library 1 mutated amino acids 39-46 of the Sup35-27p PFD

[b] Library 2 mutated amino acids 55-62 of the Sup35-27p PFD

stable prion isolates from an initial library of 3016 clones. Sequences of the mutated regions of the naïve and prion-forming clones can be found in Table 1.

Compositional biases among the prion-forming isolates

For each amino acid, the observed odds ratio (OR_{obs}), representing the degree of over/under-representation of the amino acid within the prion-forming isolates, was determined (Table 2). OR_{obs} was defined as:

$$OR_{obs} = [p_p/(1-p_p)]/ [p_n/(1-p_n)] \quad (1)$$

where p_p is the per residue frequency of the amino acid among the prion-forming isolates, and p_n is the per residue frequency of the amino acid among the naïve library.

A statistically significant ($P<0.05$) over-representation for the amino acids Phe, Ile and Val was seen among the prion-forming isolates, and statistically significant bias was seen against the amino acids Asp, Lys and Pro (Table 2). Other more subtle biases were seen, but these were not statistically significant due to the limits of sample size. Grouping similar amino acids allows for detection of more subtle biases by effectively increasing the sample size. We observed a strong bias in favor of non-polar amino acids [as defined in (Broome and Hecht 2000)] and aromatic amino acids among the prion-forming library, while both positively and negatively charged residues were under-represented (Table 2). Surprisingly, Q/N

Table 3.2: Library 1 Amino Acid Representation

Amino Acid(s)	Selected [*PSI*+] Library [a]	Unselected Naïve Library [b]	Odds Ratio [c]	P-value [d]
Phenylalanine (F)	0.075	0.032	2.31	0.040
Isoleucine (I)	0.102	0.045	2.26	0.015
Valine (V)	0.102	0.045	2.26	0.015
Tyrosine (Y)	0.054	0.025	2.18	0.099
Methionine (M)	0.038	0.020	1.96	0.19
Tryptophan (W)	0.024	0.012	1.95	0.32
Cysteine (C)	0.033	0.022	1.52	0.43
Serine (S)	0.125	0.109	1.14	0.68
Asparagine (N)	0.096	0.089	1.08	0.88
Glutamine (Q)	0.024	0.022	1.07	1.00
Glycine (G)	0.038	0.040	0.96	1.00
Leucine (L)	0.059	0.061	0.96	1.00
Threonine (T)	0.069	0.078	0.89	0.75
Histidine (H)	0.059	0.078	0.76	0.50
Alanine (A)	0.042	0.072	0.67	0.38
Arganine (R)	0.054	0.081	0.67	0.31
Glutamic Acid (E)	0.009	0.017	0.55	0.51
Proline (P)	0.038	0.127	0.30	0.002
Aspartic Acid (D)	0.014	0.051	0.28	0.041
Lysine (K)	0.009	0.045	0.21	0.028
Groups				
Aromatic (FWY)	0.144	0.067	2.32	0.002
Hydrophobic (FILMV)	0.347	0.195	2.20	3.0×10^{-05}
Charged (DEKR)	0.083	0.183	0.41	8.8×10^{-04}
Positive (KR)	0.060	0.118	0.48	0.024
Negative (DE)	0.023	0.065	0.34	0.034
Polar (NQHST)	0.343	0.346	0.98	0.92
Q/N	0.111	0.103	1.08	0.79

[a] [*PSI*+] value represents the frequency of occurrence of the amino acid among the prion-forming isolates.

[b] Naïve value represents the frequency of occurrence of the amino acid among the unselected clones.

[c] Odds ratios were calculated using equation (1)

[d] P-value is based on the 2-tailed Fisher Exact Probability Test

residues, despite being strongly over-represented among all known yeast prions, did not display a statistically significant over-represention within the mutated region of the prion-forming isolates.

Characteristics that promote prion formation

The aggregation propensity for mutants of various non-Q/N-rich amyloidogenic polypeptides is positively correlated with hydrophobicity and β-sheet propensity and negatively correlated with charge and α-helix propensity (Pawar, Dubay et al. 2005). Consistent with these predictions, we found that hydrophobicity (Roseman 1988) and β-sheet propensity (Street and Mayo 1999) were significantly greater among the prion-forming clones than among the naïve library (for the β-sheet propensity scale, lower values represent increased β-sheet propensity), and that the absolute value of the net charge was significantly lower (Table 3). Surprisingly, α-helix propensity (Koehl and Levitt 1999) was modestly greater among the prion-forming clones. However, this can almost entirely be attributed to a bias against prolines among the prion-forming clones; when prolines are excluded from the calculation, the average per-residue α-helix propensity is statistically indistinguishable between the naïve and prion-forming libraries.

We examined the degree to which over/under-representation of each amino acid among the prion-forming isolates could be explained based on the physical properties of the amino acid. For each amino acid, the natural log of the amino

Table 3.3: Physical properties of prion forming isolates from Library 1

Property	Mean for naïve library ± SEM	Mean for prion-forming isolates ± SEM	P-value
Hydrophobicity[a]	-4.30 ± 0.72	0.02 ± 0.75	0.00018
β-sheet propensity (C)[b]	3.18 ± 0.10	2.34 ± 0.09	3.5×10^{-7}
α-helix propensity (K)[c]	1.35 ± 0.17	0.73 ± 0.18	0.028
Charge[d]	0.88 ± 0.11	0.44 ± 0.12	0.016

[a] (Roseman 1988). Higher values represent greater hydrophobicity. For each naïve or prion-forming isolate, the sum of the hydrophobicities of each residue within the mutagenized was calculated. Data are the average sum per construct for the respective libraries. Standard errors are indicated.
[b] (Street and Mayo 1999). Higher values represent lower β-sheet propensity.
[c] (Koehl and Levitt 1999). Higher values represent lower α-helix propensity.
[d] The absolute value of the net charge of the mutagenized region.

acid's odds ratio was plotted as a function of various physical properties. Hydrophobicity and β-sheet propensity were both positively correlated with the $\ln(OR_{obs})$, with R^2 values of 0.46 and 0.50, respectively (Figure 2A, B). By contrast, no significant correlation was seen between an amino acid's α-helix propensity and its odds ratio (Figure 2C).

Assuming that the effects of hydrophobicity and β-sheet propensity are additive, we combined these properties for each amino acid to predict the odds ratio:

$$\ln(OR_{pred}) = A(H) + B(P_\beta) \qquad (2)$$

where H and P_β are the hydrophobicity and β-sheet propensity of the amino acid, respectively. As a starting estimate for *A* and *B*, we used the slopes of Figures 2A and 2B (0.25 and -2.26, respectively). This function was used to calculate OR_{pred} for each amino acid. The predicted odds ratio showed a strong correlation with the observed odds ratio (R^2=0.74; Figure 2D). The observed slope was 0.71; it is not surprising that it would be less than one, as hydrophobicity and β-sheet propensity are not truly independent. Based on the curve fit, A and B in equation (3) were optimized, yielding the following equation which fits all data points within experimental error (see figure 3 for 95% confidence intervals):

$$\ln(OR_{pred}) = 0.18(H) + 1.61(P_\beta) + 0.66 \qquad (3)$$

Figure 3.2

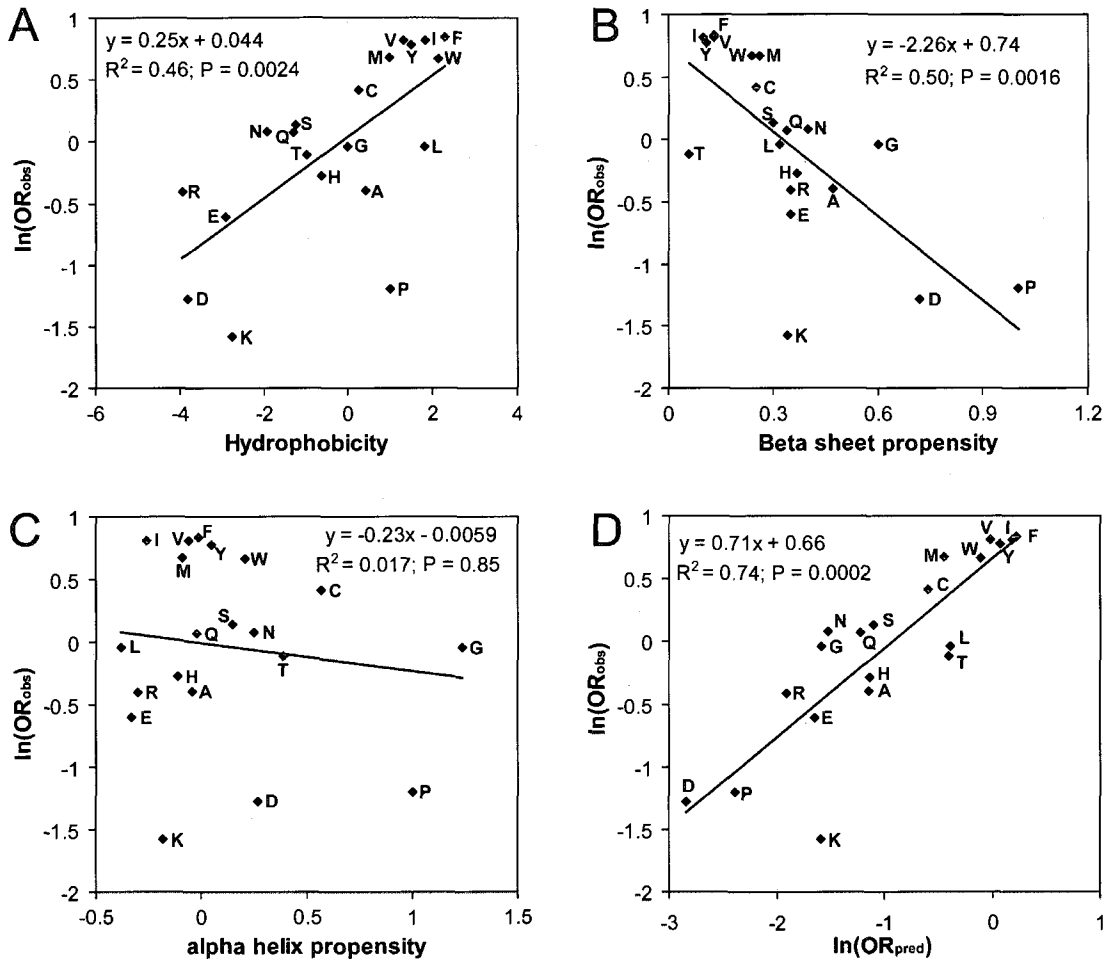


Figure 3.2: Relationship between the properties of an amino acid and its prevalence among Library 1 prion-forming isolates. $\ln(OR_{obs})$ for each amino acid in Library 1 plotted versus: hydrophobicity (A), β-sheet propensity (B; lower values represent greater β-sheet propensity), α-helix propensity (C; lower values represent greater α-helix propensity), and $\ln(OR_{pred})$ (D), as calculated using equation (3). Odds ratios were determine as in equation (1). P values were calculated by Spearman's rank correlation.

Figure 3.3

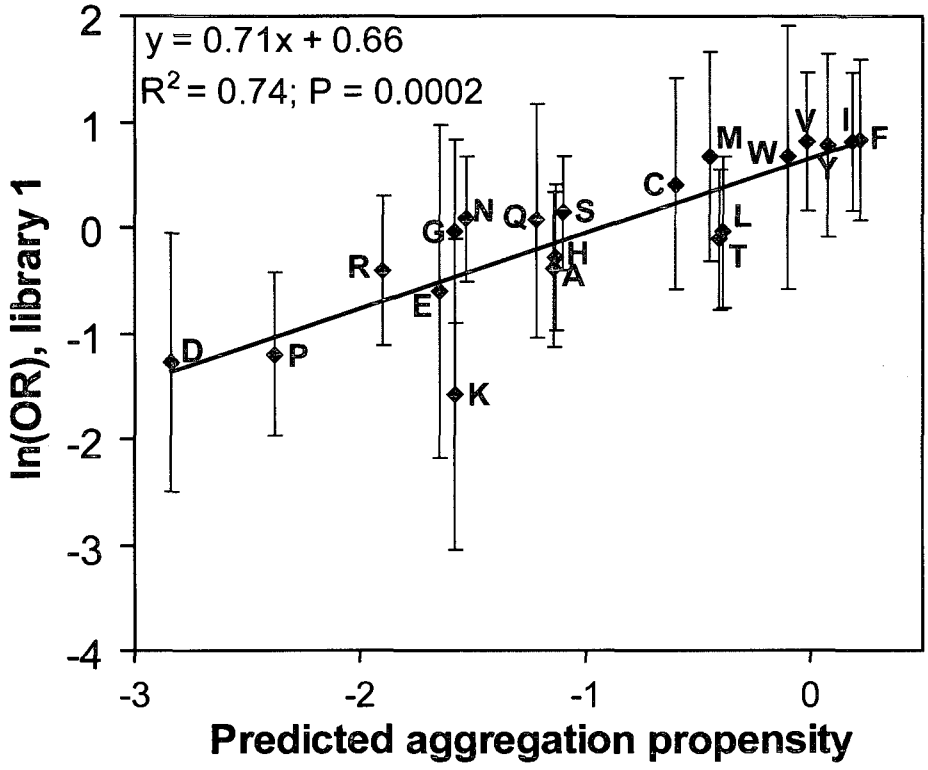


Figure 3.3: 95% confidence intervals for Libraries 1 and 2. Plot of $\ln(OR_{obs})$ for each amino acid in Library 1 plotted versus $\ln(OR_{pred})$, as in Figure 2D, with 95% confidence intervals added.

Similar biases are seen at a second position

To determine whether different regions of the PFD show different amino acid biases, we generated a second library (Library 2) targeted to amino acids 55-62 – a region that appears to be less critical for [*PSI*$^+$] formation based on our deletion mapping. Using the same methods as for Library 1, we screened 1033 clones and found 33 capable of forming prions – a success rate of 3.2% versus 0.90% for Library 1. This nearly 4-fold increase in prion-formation rate highlights the lesser importance of amino acids 55-62 to prion formation.

Analysis of individual amino acid representation (Table 4) shows weaker biases in Library 2 than Library 1. Phenylalanine and tryptophan were statistically significantly over-represented and lysine was under-represented at a level that approached statistical significance. When analyzing the amino acids by groups, the same biases were observed as in Library 1, but with decreased selection strength (Table 4). There was a strong correlation between the odds ratios for each amino acid for Library 1 and Library 2, with an R^2 of 0.48 (Figure 4A), confirming that the general biases are similar for the two libraries. This is even more apparent when amino acids are considered in groups (hydrophobic, charged, polar and aromatic), where the correlation plot of the odds ratios for the two libraries has an R^2 of 0.98 (Figure 4B). For both plots, the slope is approximately 0.65, indicating that although similar amino acids are selected for among the prion-forming clones in the two libraries, the strength of selection is stronger for Library 1 than for Library 2.

Table 3.4: Library 2 Amino Acid Representation

Amino Acid(s)	Selected [*PSI*+] Library [a]	Unselected Naïve Library [b]	Odds Ratio [c]	P-value [d]
Tryptophan (W)	0.042	0.009	4.71	0.013
Phenylalanine (F)	0.064	0.021	3.16	0.011
Asparagine (N)	0.068	0.040	1.77	0.14
Methionine (M)	0.030	0.021	1.43	0.60
Glycine (G)	0.057	0.046	1.26	0.58
Isoleucine (I)	0.038	0.030	1.25	0.65
Cysteine (C)	0.015	0.012	1.25	1.00
Histidine (H)	0.080	0.070	1.15	0.75
Valine (V)	0.038	0.037	1.04	1.00
Glutamic Acid (E)	0.023	0.024	0.93	1.00
Serine (S)	0.110	0.119	0.91	0.80
Threonine (T)	0.087	0.095	0.91	0.78
Alanine (A)	0.057	0.064	0.88	0.73
Leucine (L)	0.045	0.052	0.87	0.85
Glutamine (Q)	0.030	0.037	0.82	0.82
Tyrosine (Y)	0.049	0.064	0.76	0.48
Proline (P)	0.072	0.095	0.74	0.37
Aspartic Acid (D)	0.045	0.067	0.66	0.29
Arganine (R)	0.045	0.076	0.58	0.17
Lysine (K)	0.004	0.021	0.17	0.081
Groups				
Aromatic (FWY)	0.155	0.095	1.76	0.031
Hydrophobic (FILMV)	0.216	0.162	1.43	0.11
Charged (DEKR)	0.117	0.189	0.57	0.023
Positive (KR)	0.049	0.098	0.48	0.029
Negative (DE)	0.068	0.091	0.73	0.36
Polar (NQHST)	0.375	0.360	1.07	0.73
Q/N	0.098	0.076	1.32	0.38

[a] [*PSI*+] value represents the frequency of occurrence of the amino acid among the prion-forming isolates.

[b] Naïve value represents the frequency of occurrence of the amino acid among the unselected clones.

[c] Odds ratios were calculated using equation (1)

[d] P-value is based on the 2-tailed Fisher Exact Probability Test

Figure 3.4

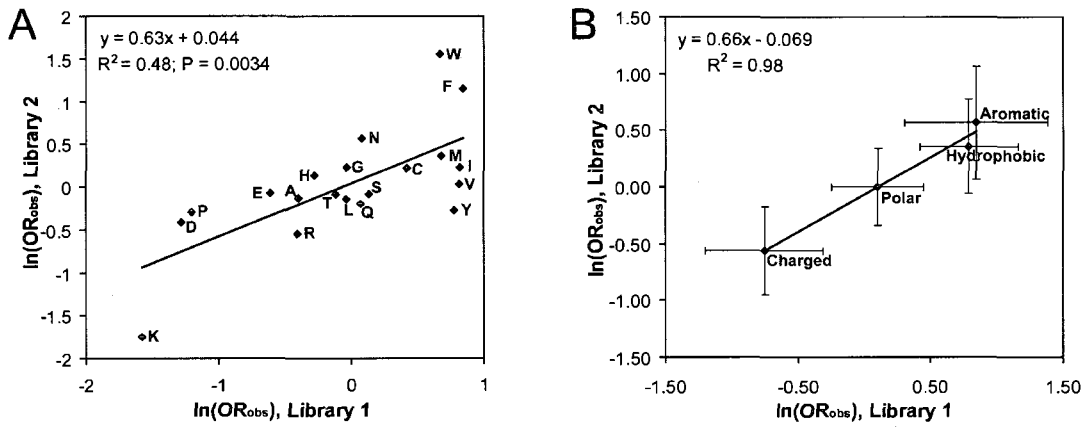


Figure 3.4. Library 2 shows weaker biases than Library 1. $\ln(OR_{obs})$ from Library 1 was plotted versus $\ln(OR_{obs})$ from Library 2 for each amino acid (A) and for groups of amino acids (B). Hydrophobic residues were defined as Phe, Ile, Leu, Met and Val. Polar amino acids were defined as Ser, Thr, His, Gln and Asn. Charged amino acids are Asp, Glu, Lys and Arg. Aromatic amino acids were defined as Trp, Tyr and Phe. Error bars represent 95% confidence intervals.

Identification of regions sensitive to mutagenesis

We hypothesized that the differential sensitivity to mutation seen in the regions targeted for mutagenesis in Library 1 and 2 could be explained by differences in prion propensity. We scanned the Sup25-27 PFD using a window size of eleven amino acids. For each window, we calculated the predicted prion potential as the sum of the $\ln(OR_{obs})$ (Figure 5A), based on the experimentally obtained values from Library 1. These plots predict the region targeted in Library 1 to have relatively high prion propensity and the region targeted in Library 2 to have low prion propensity. These results are consistent with the greater sensitivity to deletion (Figure 1) and stringency of selection (Figure 4) seen for Library 1 versus Library 2.

Similar analysis of the wild-type Sup35 PFD reveals two peaks in prion potential spanning amino acids 8-35 and 44-61 (Figure 5B). This nicely coincides with analysis of Sup35 showing that: (i) mutations that block [*PSI*$^+$] propagation specifically localize to amino acids 8-34 of the PFD (DePace, Santoso et al. 1998), (ii) the amyloid core, as defined by hydrogen-deuterium exchange, spans either the first 40 or first 70 amino acids of the PFD, depending on the structural variant analyzed (Toyama, Kelly et al. 2007), (iii) the minimal fragment required to efficiently induce [*PSI*$^+$] formation is amino acids 1-64 (Osherovich, Cox et al. 2004), and (iv) [*PSI*$^+$] propagation requires, depending on the prion variant, amino acids 7-21, 9-37 or 5-52 (Chang, Lin et al. 2008).

Figure 3.5

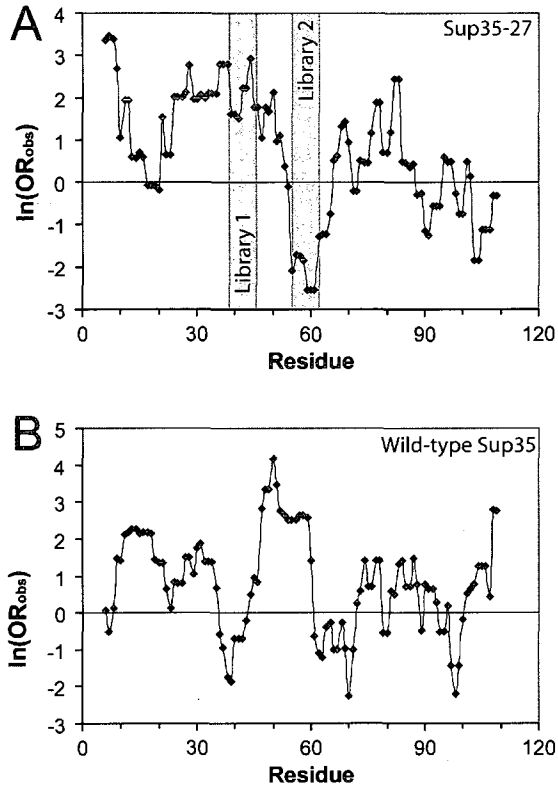


Figure 3.5. Predicted prion-prone regions. The PFDs of Sup35-27 (A) and wild-type Sup35 (B) were scanned using an eleven amino acid window size. At each position within the prion domains, the sum of $\ln(OR_{obs})$ for the indicated amino acid and the five amino acids on either side were calculated to determine the prion propensity of the window. Regions mutated in Library 1 and 2 are indicated.

Predicting prion propensity based on composition

A key question is whether results from mutagenesis of small regions can be extrapolated to predict prion formation by larger PFDs. To reduce noise in scanning entire proteins, we expanded our window size to 41 amino acids, which roughly correlates with the minimal fragment required to induce yeast prion formation (Ross, Edskes et al. 2005). Scanning of twenty randomly selected non-Q/N-rich proteins (Figure 6A; red circles) shows that regions with prion propensity equal to or greater than that of the yeast PFDs (Figure 6A; blue circles) are common. However, when the randomly selected proteins are scanned for order propensity using FoldIndex (Prilusky, Felder et al. 2005), regions of high prion propensity within the randomly selected proteins are consistently predicted to have high order propensity; by contrast, the yeast PFDs are intrinsically disordered (Figure 6A). Thus, the yeast PFDs are unique in being both prion-prone and intrinsically disordered. When full-length Ure2p and Sup35p are scanned using a 41 amino acid window size, the PFDs are the only regions of the proteins that have both negative values for order propensity and positive values for prion propensity (Figure 6B, C).

To determine whether a combination of prion propensity and disorder could be used to distinguish between prion-forming and non-prion forming Q/N-rich domains, we utilized the massive data set generated by Alberti et al (Alberti, Halfmann et al. 2009). Using a Hidden Markov Model, they scanned the yeast genome for proteins with domains that were most compositionally similar to the PFDs of Sup35p, Ure2p, Rnq1p and New1p. Four different assays were used to test

Figure 3.6

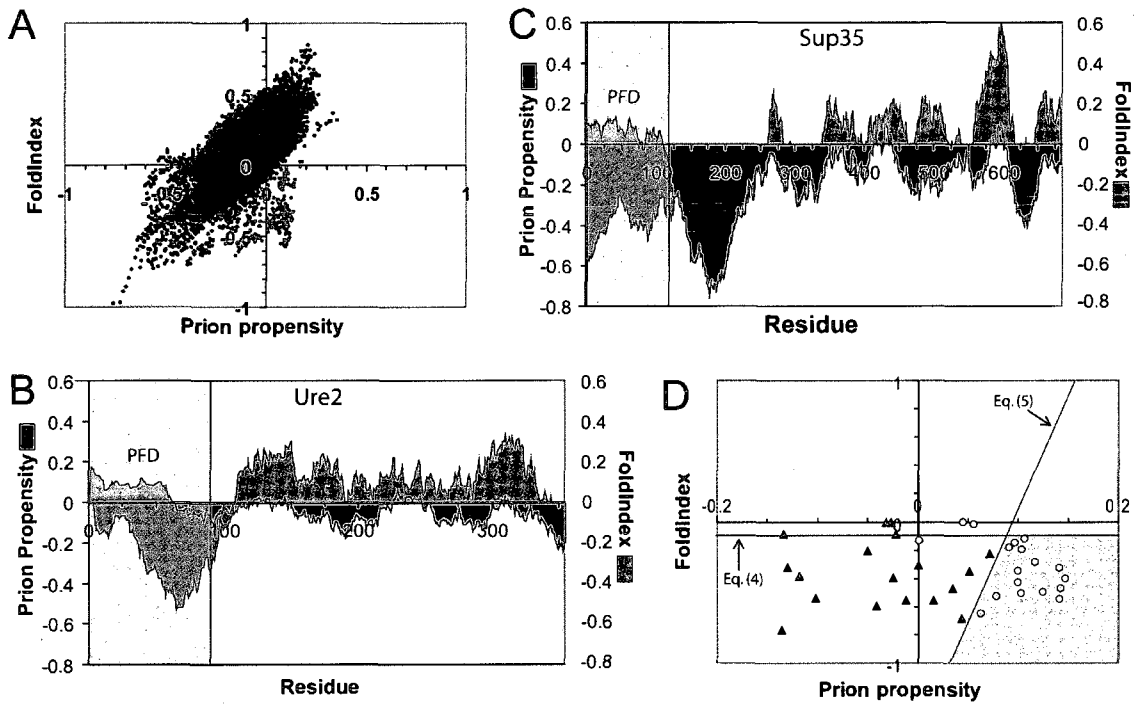


Figure 3.6: Predicting prion propensity for Q/N-rich domains. (A) Prion propensity and order predictions for twenty randomly selected open reading frames (red) and the PFDs of Sup35p and Ure2p (blue). For every 41 amino acid window throughout the proteins, predicted prion propensity (calculated as the average $\ln(OR_{obs})$ across the window, using the values from Table 1) versus predicted FoldIndex order propensity (where negative values are associated with disorder) was plotted. Randomly selected open reading frames are RPS17A, YGR235C, YPR1, NIP100, ERG12, ARP10, YAR003W, ECM7, YNL083W, YLR247C, YDR275W, YOR087W, SEC8, ALG6, YBR226C, ROM1, MAL33, MYO3, SFI1 and YJL039C. (B) Identification of prion-prone regions within Ure2p. Ure2p was scanned using a window size of 41 amino acids, calculating for each window the average order propensity (blue) and prion propensity (red). The PFD is shaded. (C) Identification of prion-prone regions within Sup35p using the same method. (D) Prion propensity and order prediction for Q/N-rich proteins that show prion-like activity (open circles) and those that do not (shaded triangles). For each potential prion domain, average prion propensity and average disorder are plotted for the 41 consecutive 41-amino acid windows with maximum average predicted prion propensity. Also plotted are Equations (4) and (5), and the region defined by Equations (4) and (5) as prion-prone is shaded.

the top 100 domains. All of these domains were highly enriched in Q/N residues. Eighteen proteins showed prion-like activity in all assays, while 18 did not show activity in any of the assays. We scanned each of these 36 potential PFDs using a 41 amino acid window size, calculating average disorder propensity and prion propensity. The profiles of the domains found to form prion-like aggregates were strikingly different from those that did not show prion activity. All of the domains that showed prion-like activity had multiple consecutive windows that were both disordered and prion-prone (Figure 7); by contrast, each of peptides that lack prion activity had either no regions that were both disordered and prion-prone, or only very short regions (Figure 8).

To determine whether the presence of consecutive prion-prone windows could be used to distinguish between prion-forming and non-prion Q/N domains, for each peptide we identified the 41 consecutive 41 amino acid windows that had maximum average predicted prion propensity. By averaging 41 consecutive windows, we are effectively calculating prion propensity for 81 consecutive amino acids (40 on each side of a central residue), but weighting each residue inversely proportional to its distance from the central residues (Figure 9). Therefore, this method incorporates the idea that yeast PFDs are often quite large, but that the sequence requirements for prion formation are more flexible further from the core of the PFD. When the prion propensity for the optimal region of each peptide was plotted versus average disorder, a clear difference is seen between those peptides that showed prion-like activity and those that did not (Figure 6D). In fact, if the criteria for a PFD are based on two equations:

Figure 3.7

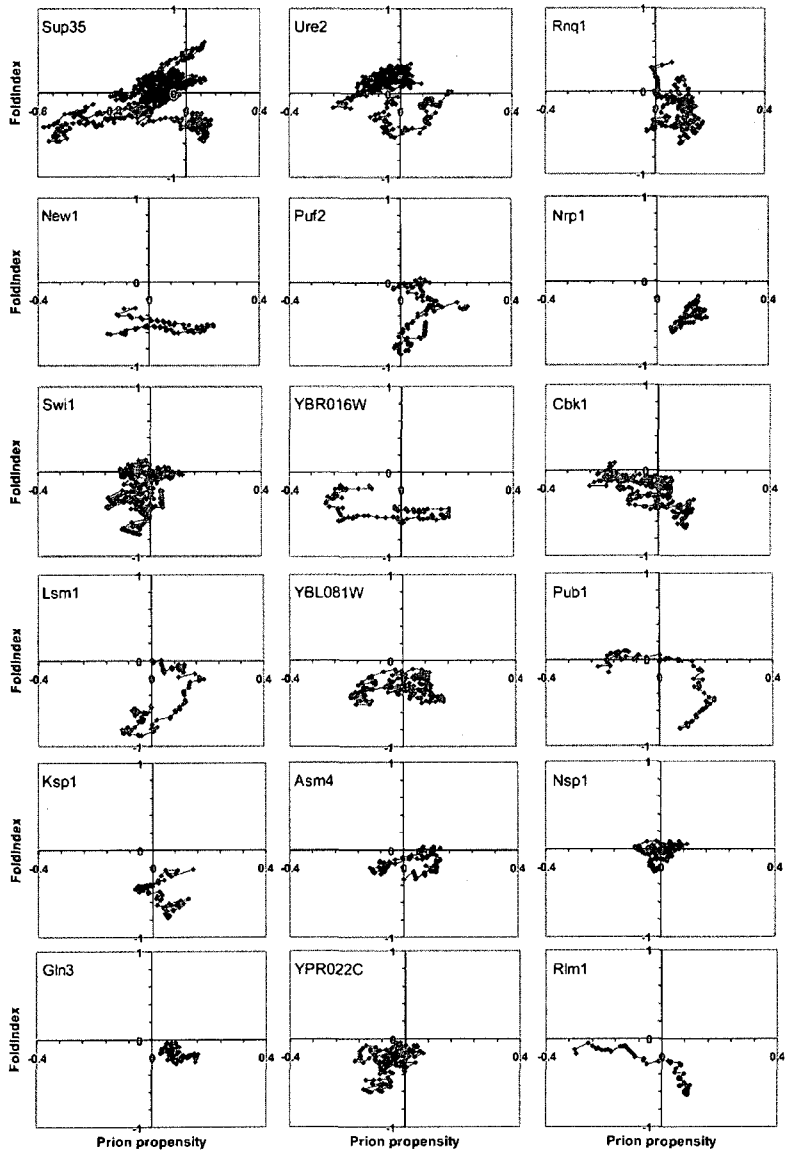


Figure 3.7: Maps of potential prion-forming Q/N-rich proteins. Each of the Q/N-rich domains identified and tested by Alberti et al. that showed prion-like activity in all assays was scanned using a 41 amino acid window size, calculating for each window the average FoldIndex order propensity and prion propensity. Prion propensity was calculated based on the average $\ln(OR_{obs})$ for each amino acid in the window, using the values from Table 1. Prion propensity versus order propensity was plotted for each 41 amino acid window (blue). For Sup35p and Ure2p, prion propensity and order propensity were also calculated for the non-prion domains (red).

Figure 3.8

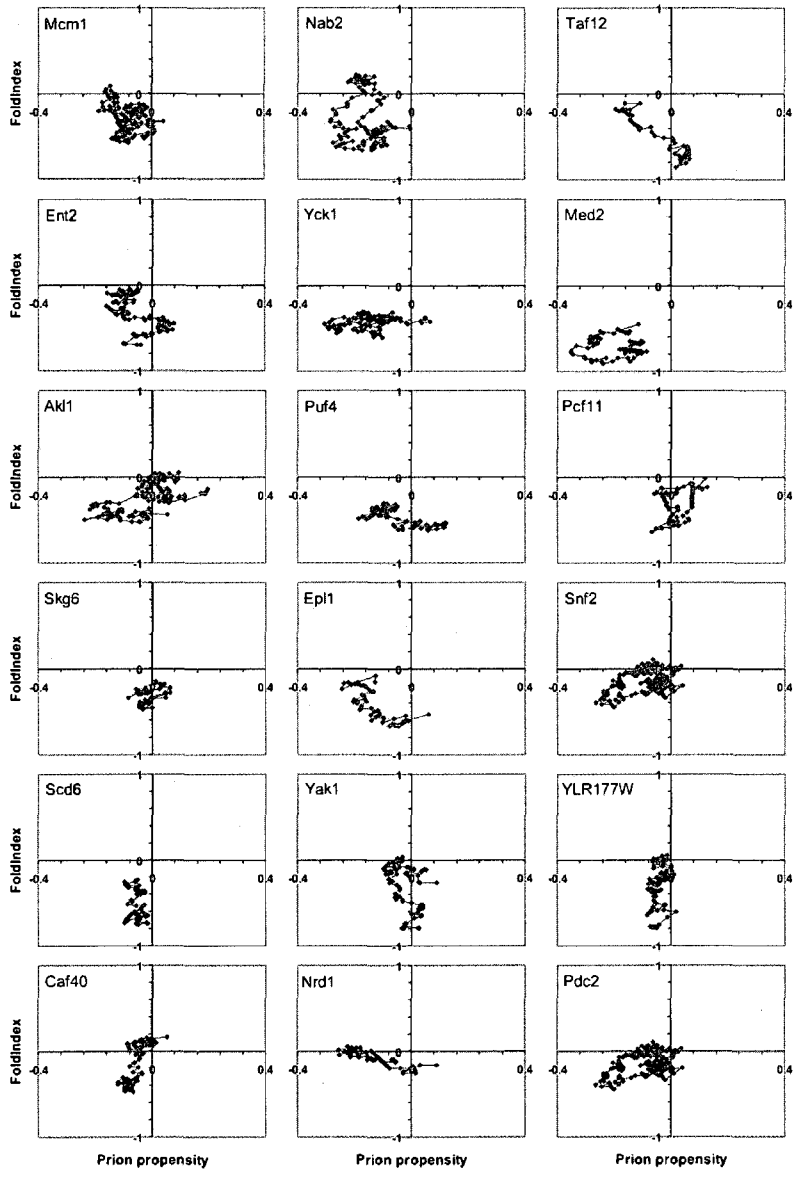


Figure 3.8: Maps of potential non-prion-forming Q/N-rich proteins. Each of the Q/N-rich domains identified and tested by Alberti et al. that failed to show prion-like activity in any assay was scanned using a 41 amino acid window size, calculating for each window the average FoldIndex order propensity and prion propensity. Prion propensity versus order propensity was plotted for each 41 amino acid window.

Figure 3.9

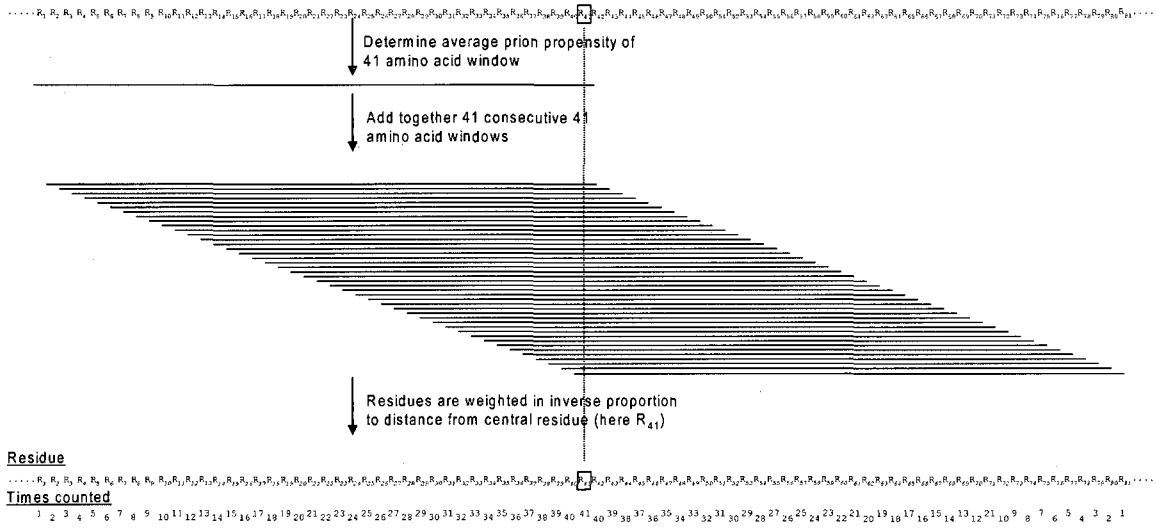


Figure 3.9. Summing overlapping windows to calculate total prion propensity of a region. (A) First, for each 41 amino acid window, prion propensity is calculated as the average $\ln(OR_{obs})$ for each amino acid in the window, using the experimentally obtained values from Table 1. (B) For each position in the protein, the forty-one windows that overlap with that position are averaged. (C) Prion propensity for the forty-one consecutive windows is therefore dependent on $\ln(OR_{obs})$ for 81 consecutive amino acids, but with each residue weighted into the calculation in inverse proportion to its distance from the central residue.

$$P > 0.05 \quad (4)$$

$$F < 16P - 1.5 \quad (5)$$

where P is predicted prion propensity and F is the FoldIndex order prediction, 33 of 36 peptides (91.7%) are correctly scored. 100% of the peptides lacking prion-like activity are properly scored as non-prion, and fifteen of eighteen domains with prion-like activity are correctly scored as prion-positive.

Discussion

We have developed the first method to quantitatively measure the prion-forming propensities of individual amino acids in vivo. This method has a number of advantages over the more commonly used error-prone PCR. First, mutations can be carefully controlled and targeted, allowing for quantitative analysis of prion propensities for each amino acid. Second, because multiple mutations within a given codon are unlikely in error-prone PCR, only a subset of codons is accessible from a given starting codon. By contrast, every amino acid should be represented at every position within our libraries, providing a broader data set.

Although the amino acid biases in Libraries 1 and 2 showed similar trends, the strength of the biases was much greater for Library 1. This link between compositional flexibility and position highlights one of the challenges of bioinformatic searches – PFDs likely contain regions with different levels of importance to prion formation. Although the region targeted in Library 2 was less sensitive to mutation, amino acid biases were still seen, indicating that this region

affects prion formation. Similar differential sensitivity is seen for wild-type Sup35p, where regions outside of the amyloid core are required for prion propagation, but are less sensitive to mutation than the amyloid core (DePace, Santoso et al. 1998; Osherovich, Cox et al. 2004). Therefore, an ideal search algorithm to identify new PFDs would look for a small nucleation domain of optimal composition, but also consider the flanking sequences, presumably weighting amino acid composition progressively less farther away from the nucleation domain. Our method of summing consecutive windows accomplishes this goal.

The fact that no amino acids were excluded from either library shows the highly flexible sequence requirements for prion formation. This is surprising since the known yeast PFDs have very limited amino acid representation. The effects of prion-inhibiting residues in our libraries may be offset by nearby prion-promoting residues within the mutated region; for example, in Library 1, sequences containing charged residues or prolines showed an over-representation of hydrophobic aromatic residues (Phe, Tyr and Trp; P=0.031 by Fisher's Exact Test). Additionally, this sequence flexibility may reflect the ability of yeast prion proteins to adopt multiple amyloid conformations (prion variants); non-ideal residues may be accommodated by alternate structures.

One surprising result from these experiments is that there is almost no correlation between the amino acids that most strongly promote prion formation and the compositions of the yeast PFDs (Figure 10A-C). We observed a bias against

Figure 3.10

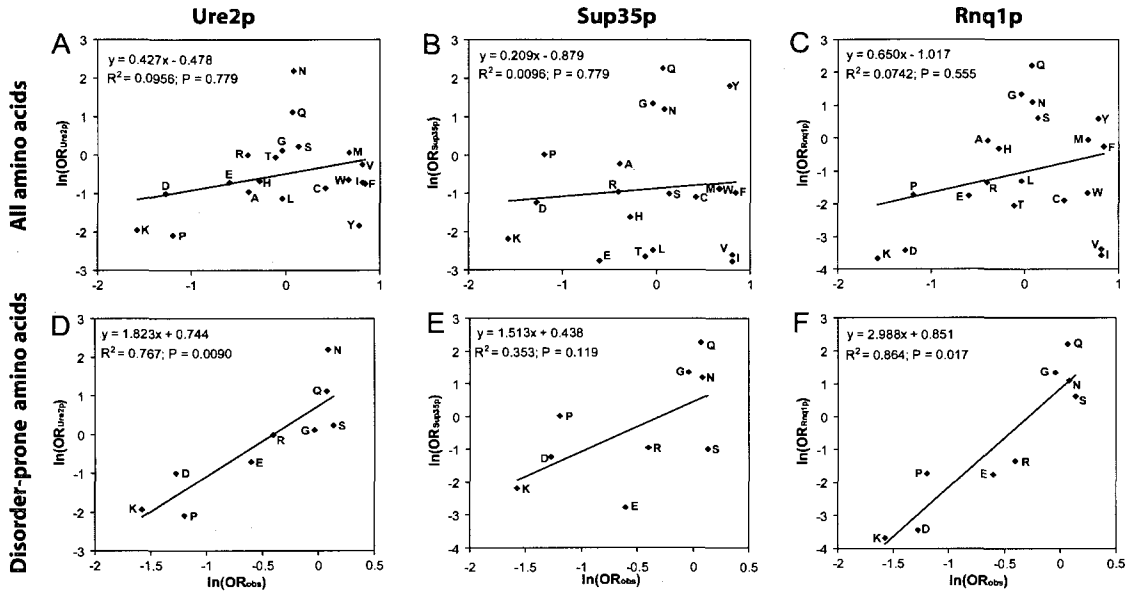


Figure 3.10: The Ure2p, Sup35p, and Rnq1p PFDs are biased towards amyloidogenic disorder-promoting residues. (A-C) Relationship between the degree to which an amino acid promotes prion formation and the amino acid's prevalence within yeast PFDs. $\ln(OR_{obs})$ (from Table 1) was plotted versus $\ln(OR_{PFD})$ (as calculated in Equation (6)) for the prion domains from Ure2p (A), Sup35p (B) and Rnq1p (C). (D-F) Analysis only of the prion promoting amino acids Lys, Pro, Gly, Arg, Asn, Gln, Ser, Glu, and Asp.

charged amino acids, a strong bias for hydrophobic residues and no apparent bias for polar or Q/N residues. By contrast, the yeast PFDs are enriched in Q/N residues and lacking in hydrophobic residues. The Sup35p PFD has only three non-polar residues [as defined by (Broome and Hecht 2000)], which constitutes 2.6% of the domain, while the average [*PSI*$^+$]-forming clone from Library 1 contained 2.8 non-polar residues out of eight amino acids, constituting 35%. This raises the question of why hydrophobic residues, which so clearly promote prion formation, are almost completely absent from yeast PFDs. Intrinsic disorder likely partially explains this disconnect. A key feature of the yeast PFDs is that they are intrinsically disordered (Serio, Cashikar et al. 2000; Pierce, Baxa et al. 2005). When we consider only those residues most strongly associated with intrinsic disorder – Lys, Pro, Gly, Arg, Asn, Gln, Ser, Glu, and Asp (Weathers, Paulaitis et al. 2004) – there is excellent correlation between prion propensity and the compositions of the yeast PFDs (Figure 10D-F). These residues are highly over-represented in each of the yeast PFDs ($P<0.0001$ each for the Sup35p, Rnq1p and Ure2p PFDs), accounting for 75%, 76% and 79% of the amino acids in the Sup35p, Ure2p and Rnq1p PFDs, respectively.

Thus, we hypothesize that the yeast PFDs are not optimized for maximum intrinsic amyloid propensity, but instead are biased towards those residues that have the highest prion propensity while still maintaining intrinsic disorder. While a small number of additional hydrophobic residues within the Sup35-27p PFD promotes prion formation, larger numbers would likely lead to hydrophobic

collapse, potentially inhibiting prion formation or leading to non-specific aggregation. Therefore, although the effects of individual mutations within a Q/N-rich domain can be accurately predicted based on hydrophobicity and β-sheet propensity [Equation (3)], it is necessary to consider order propensity when examining the composition of entire domains. This hypothesis provides a possible explanation for the prevalence of Qs and Ns in the yeast PFDs. Qs and Ns have relatively high prion propensity compared to other disorder-promoting residues (Figure 6). Alternatively, there may be a threshold number and/or density of Q/N residues required for prion formation, and above this threshold small changes in the number of Q/N residues may exert only a subtle effect.

Based on the apparent importance of intrinsic disorder to yeast prion formation, we propose that amyloid proteins can be divided into three broad classes. First are the Q/N-rich amyloid proteins. For many proteins, native state stability prevents amyloid formation by keeping amyloid-promoting residues buried. Our data suggest that rather than having a high concentration of strongly amyloidogenic (but also order-promoting) residues, the yeast PFDs form prions by excluding order-promoting residues and having large stretches with a high concentration of the most amyloidogenic disorder-promoting residues. Thus, yeast PFDs do not have to overcome native state stability to form prions, explaining their efficient prion formation despite their relatively low prion propensity. A second class of amyloid domains, which includes peptides such as Aβ, is also largely disordered. However, rather than amyloid formation being driven by many weak interactions of modestly amyloid-promoting residues (such as Qs and Ns), amyloid

formation is driven by short, highly amyloidogenic nucleation domains (Balbach, Ishii et al. 2000; Pawar, Dubay et al. 2005). Finally, proteins such as transthyretin (Kelly 1998) have highly amyloid-prone regions that are usually buried within the folded structure of the protein. As seen in Figure 6A, regions of high amyloid propensity (often greater than in the yeast PFDs) are common, but are generally found in ordered regions. For these proteins, native state stability will largely determine amyloid propensity (Kelly 1998).

Differences between classes of amyloid proteins may explain why α-helix propensity is strongly biased against in the amyloid prediction algorithm developed by Pawar et al., but no such effect is seen in our experiments. In the context of amyloid domains that are buried within a structured region of a protein, or in short hydrophobic stretches that are prone to amyloid formation, there is a competition between different structures. Therefore, α-helix formation will inhibit β-sheet formation and subsequent amyloid formation. By contrast, our proteins are disordered, so the insertion of a small number of α-helix prone amino acids is unlikely to have a strong effect.

Although intrinsic disorder likely explains why the yeast PFDs do not contain large numbers of hydrophobic residues, it is less clear why they do not contain a few more hydrophobic residues scattered throughout. Although insertion of large numbers of hydrophobic residues might inhibit prion formation by leading to hydrophobic collapse, our results clearly demonstrate that insertion of a small number of hydrophobic residues into a Q/N-rich domain promotes prion formation.

One possible explanation for the exclusion of hydrophobic residues in yeast PFDs may be to curb amyloid propensity. It is unclear whether yeast prion formation is a beneficial phenomenon providing a mechanism to regulate protein activity, or a detrimental phenomenon analogous to human amyloid disease. [*PSI*$^{+}$] can increase resistance to certain stress conditions (True and Lindquist 2000), but the failure to observe [*PSI*$^{+}$] in wild yeast strains (Nakayashiki, Kurtzman et al. 2005) argues that beneficial [*PSI*$^{+}$] formation is at most a rare event. If yeast prions are diseases, it is not surprising that the PFDs are not be optimized for maximum prion potential. In this case, the amino acid composition of a PFD presumably would reflect the normal function of the PFD. By contrast, if prion formation is a beneficial event allowing for rapid conversion between active and inactive states, it would only be beneficial if both the prion and non-prion states are stable. Thus, maximum prion efficiency would not be ideal. Instead, the prion potential of the PFD would be optimized such that the frequencies of prion formation and loss would yield a balance of prion and non-prion cells that would lead to optimal survival (Masel and Griswold 2009).

This highlights a risk of using known yeast prions as a training set for bioinformatic searches for new prions. Bioinformatic programs that have identified new prions have built their training programs based on the compositions of known PFDs. Our data suggest that the yeast prions have evolved not for maximum prion forming efficiency, but instead to be at the threshold of efficient prion formation. Bioinformatic programs that search for new prions based on known prion sequences penalize both charged and hydrophobic residues similarly, because they are both under-represented in known yeast PFDs. However, our data suggest very

different reasons for these biases; charged residues inhibit prion formation, while hydrophobic residues too strongly promote prion formation and/or order. Accurate prediction of prion propensity requires understanding which deviations from known prion-forming compositions will promote prion formation and which will inhibit. This may explain why even using the most sophisticated bioinformatic method available to identify new PFDs, less than 20% of identified domains showed prion-like activity in all assays (Alberti, Halfmann et al. 2009). Furthermore, of the top ten sequences predicted by Alberti's Hidden Markov Model to form prions, three sequences failed all four tests for prion behavior while only three passed all four. By contrast, our experimental based prediction algorithm accurately distinguishes among these proteins.

Overall, this data provides detailed insight into how amino acid composition affects prion formation by Q/N-rich domains. By highlighting the critical role for intrinsic disorder in yeast prion formation, our data explain the discrepancy between the compositions of the yeast PFDs and the amino acids thought to promote amyloid formation. Our data will allow for genome scanning to identify novel PFDs, and by using a smaller window size (as in Figure 5) to identify critical nucleating domains within known PFDs.

Methods:

Strains and media

Standard yeast media and methods were used as previously described (Sherman 1991), except that YPD contained 0.5% yeast extract instead of the standard 1%. In all experiments, yeast were grown at 30°C. All experiments were performed in *Saccharomyces cerevisiae* strain 780-1D/pJ533 (Song, Wu et al. 2005). This strain's genotype is *α kar1-1 SUQ5 ade2-1 his3 leu2 trp1 ura3 sup35*::KanMx [PIN$^+$]; pJ533 expresses *SUP35* from a *URA3* plasmid as the sole copy of *SUP35* in the cell.

PFD Truncation Mapping

SUP35-27 deletions were generated by a two-step PCR procedure in which the regions N-terminal and C-terminal to the site of deletion were amplified in separate reactions. Products of these reactions were combined and reamplified with outer primers. Final PCR products were co-transformed with *Bam*HI/*Hin*dIII-cut pJ526 (cen *LEU2;* from Dan Masison, National Institutes of Health) into yeast strain 780-1D/pJ533 and selected on SD-leu. Transformants were spotted onto 5-fluoroorotic acid (5-FOA) containing medium to select for loss of pJ533.

To generate plasmids for prion domain overexpression, each NM-domain was amplified with primers EDR654 and EDR969, which installs a stop codon at the end of the M domain. PCR products were digested with *Bam*HI and *Xho*I and ligated into *Bam*HI/*Xho*I-cut pKT24 (Ross, Edskes et al. 2005).

Creating mutant libraries

Degenerate oligonucleotides were used to randomly mutate regions of the *SUP35-27* PFD. Nucleotides 115-138 were mutated in Library 1 and nucleotides 163-186 in

Library 2. Primers (EDR1003 and EDR1121, respectively), made by Invitrogen, were antisense, containing degenerate segments such that the reverse complement encoded a 25% mix of each nucleotide at positions 1 and 2 of each mutated codon and a 33.3% mix of C, G and T at the third position. The 5' and 3' ends of EDR1003 and EDR1121 contained regions of homology to *SUP35-27*. These primers were paired with EDR259 to amplify the N-terminal region of *SUP35-27*. In a second PCR reaction, a primer complementary to the non-degenerate 5' region of EDR1003 or EDR1121 (EDR1007 or EDR672, respectively) was paired with EDR262 to amplify the C-terminal side of the *SUP35-27*. Products of these reactions were combined and reamplified with the outer primers. The final PCR products were co-transformed with *Bam*HI/*Hin*dIII-cut pJ526 into yeast strain 780-1D/pJ533 and selected on SD-leu. Transformants were spotted onto 5-FOA-containing medium to select for loss of pJ533.

Screening for [*PSI*$^+$] clones

Library mutants that grew on 5-FOA were then stamped onto SC-Ade, YPD and YPAD and grown for 3-5 days at 30°C. Only isolates that were red when grown on YPD and did not grow on SC-ADE were pooled into mini-libraries (~50 colonies each). Mini-libraries were plated onto SC-Ade at concentrations of 10^6 and 10^5 cells per plate and grown for 5 days at 30°C. To test curability, Ade$^+$ colonies were grown on YPAD plus 4mM GdHCl. Clones in which the Ade$^+$ phenotype was stable and curable were sequenced.

Analysis of the prion-forming libraries

Hydrophobicities (Roseman 1988), α-helix propensity (Koehl and Levitt 1999) and β-sheet propensity (Street and Mayo 1999) were calculated using previously reported scales. Because neither the α-helix propensity nor β-sheet propensity scales contained values for proline, we set the proline α-helix and β-sheet propensities equal to 1 to account for the known ability of proline to disrupt α-helices and β-sheets.

Compositions of yeast prion domains

For each yeast PFD, the odds ratio for each amino acid (OR_{PFD}) was calculated as:

$$OR_{PFD} = [p_{pfd}/(1-p_{pfd})]/ [p_{gen}/(1-p_{gen})] \qquad (6)$$

where p_{pfd} is fraction of residues in the PFD that are the indicated amino acid, and p_{gen} is the fraction of codons within all predicted open reading frames in the yeast genome that code for the amino acid. Codon frequencies are from the Saccharomyces Genome Database (http://www.yeastgenome.org/). For the plotting of OR_{PFD} versus OR_{obs}, for amino acids that were completely absent from the PFD, p_{pfd} was set as 0.5 divided by the length of the PFD to avoid zero values in the logarithm.

Calculating prion propensity and disorder

Proteins were scanned using a 41 amino acid window size. For plotting, the position of the window was defined based on the central amino acid. For windows near the termini of proteins (such that there are fewer than 20 amino acids on the C- or N-

terminal side of the central amino acid), the denominator was adjusted accordingly in calculating the average values. Randomly selected proteins were chosen from all annotated open reading frames using Excel's random number generating function.

Acknowledgements

This work was funded in part by an American Heart Association Grant (0735321N) to E.D.R. We would like to thank Olve Peersen, P. Shing Ho, Frank Shewmaker and members of the Ross lab for helpful discussion, and Nathan Liss, Holly Hamilton and Daniel Lee for their tireless help in screening thousands of colonies.

References

Alberti, S., R. Halfmann, et al. (2009). "A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins." Cell **137**(1): 146-158.

Balbach, J. J., Y. Ishii, et al. (2000). "Amyloid fibril formation by A beta 16-22, a seven-residue fragment of the Alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR." Biochemistry **39**(45): 13748-59.

Bradley, M. E. and S. W. Liebman (2004). "The Sup35 domains required for maintenance of weak, strong or undifferentiated yeast [PSI+] prions." Mol. Microbiol. **51**(6): 1649-1659.

Broome, B. M. and M. H. Hecht (2000). "Nature disfavors sequences of alternating polar and non-polar amino acids: implications for amyloidogenesis." J Mol Biol **296**(4): 961-8.

Bryan, A. W., Jr., M. Menke, et al. (2009). "BETASCAN: probable beta-amyloids identified by pairwise probabilistic analysis." PLoS Comput Biol **5**(3): e1000333.

Chang, H. Y., J. Y. Lin, et al. (2008). "Strain-specific sequences required for yeast [PSI+] prion propagation." Proc Natl Acad Sci U S A **105**(36): 13345-50. Epub 2008 Aug 29.

Chiti, F. and C. M. Dobson (2006). "Protein misfolding, functional amyloid, and human disease." Annu Rev Biochem **75**: 333-66.

Cox, B. S. (1965). "PSI, a cytoplasmic suppressor of super-suppressor in yeast." Heredity **26**: 211-232.

DePace, A. H., A. Santoso, et al. (1998). "A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion." Cell **93**(7): 1241-52.

Derkatch, I. L., M. E. Bradley, et al. (2001). "Prions affect the appearance of other prions: the story of [PIN(+)]." Cell **106**(2): 171-82.

Du, Z., K. W. Park, et al. (2008). "Newly identified prion linked to the chromatin-remodeling factor Swi1 in Saccharomyces cerevisiae." Nat Genet **40**(4): 460-5. Epub 2008 Mar 23.

Fernandez-Escamilla, A. M., F. Rousseau, et al. (2004). "Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins." Nat. Biotechnol. **22**(10): 1302-6.

Ferreira, P. C., F. Ness, et al. (2001). "The elimination of the yeast [PSI+] prion by guanidine hydrochloride is the result of Hsp104 inactivation." Mol. Microbiol. **40**(6): 1357-69.

Jung, G. and D. C. Masison (2001). "Guanidine hydrochloride inhibits Hsp104 activity in vivo: a possible explanation for its effect in curing yeast prions." Curr. Microbiol. **43**(1): 7-10.

Kelly, J. W. (1998). "The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways." Curr. Opin. Struct. Biol. **8**(1): 101-6.

Koehl, P. and M. Levitt (1999). "Structure-based conformational preferences of amino acids." Proc Natl Acad Sci U S A **96**(22): 12524-9.

Kushnirov, V. V. and M. D. Ter-Avanesyan (1998). "Structure and replication of yeast prions." Cell **94**(1): 13-6.

Linding, R., J. Schymkowitz, et al. (2004). "A Comparative Study of the Relationship Between Protein Structure and [beta]-Aggregation in Globular and Intrinsically Disordered Proteins." J. Mol. Biol. **342**(1): 345-353.

Masel, J. and C. K. Griswold (2009). "The strength of selection against the yeast prion [PSI+]." Genetics **181**(3): 1057-63.

Masison, D. C., M. L. Maddelein, et al. (1997). "The prion model for [URE3] of yeast: spontaneous generation and requirements for propagation." Proc. Natl. Acad. Sci. USA **94**(23): 12503-8.

Masison, D. C. and R. B. Wickner (1995). "Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells." Science **270**(5233): 93-5.

Michelitsch, M. D. and J. S. Weissman (2000). "A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions." Proc. Natl. Acad. Sci. USA **97**(22): 11910-5.

Nakayashiki, T., C. P. Kurtzman, et al. (2005). "Yeast prions [URE3] and [PSI+] are diseases." Proc Natl Acad Sci U S A **102**(30): 10575-80. Epub 2005 Jul 15.

Nemecek, J., T. Nakayashiki, et al. (2009). "A prion of yeast metacaspase homolog (Mca1p) detected by a genetic screen." Proc Natl Acad Sci U S A **106**(6): 1892-6. Epub 2009 Jan 27.

Osherovich, L. Z., B. S. Cox, et al. (2004). "Dissection and design of yeast prions." PLoS Biol. **2**(4): E86.

Patel, B. K., J. Gavin-Smyth, et al. (2009). "The yeast global transcriptional co-repressor protein Cyc8 can propagate as a prion." Nat Cell Biol **11**(3): 344-9. Epub 2009 Feb 15.

Pawar, A. P., K. F. Dubay, et al. (2005). "Prediction of "Aggregation-prone" and "Aggregation-susceptible" Regions in Proteins Associated with Neurodegenerative Diseases." J. Mol. Biol. **350**(2): 379-92.

Perutz, M. F., B. J. Pope, et al. (2002). "Aggregation of proteins with expanded glutamine and alanine repeats of the glutamine-rich and asparagine-rich domains of Sup35 and of the amyloid beta -peptide of amyloid plaques." Proc. Natl. Acad. Sci. USA **99**(8): 5596-5600.

Peters, T. W. and M. Huang (2007). "Protein aggregation and polyasparagine-mediated cellular toxicity in Saccharomyces cerevisiae." Prion **1**(2): 144-53.

Pierce, M. M., U. Baxa, et al. (2005). "Is the prion domain of soluble Ure2p unstructured?" Biochemistry **44**(1): 321-8.

Prilusky, J., C. E. Felder, et al. (2005). "FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded." Bioinformatics **21**(16): 3435-8.

Roseman, M. A. (1988). "Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds." J. Mol. Biol. **200**(3): 513-522.

Ross, E. D., U. Baxa, et al. (2004). "Scrambled Prion Domains Form Prions and Amyloid." Mol. Cell. Biol. **24**(16): 7206-7213.

Ross, E. D., H. K. Edskes, et al. (2005). "Primary sequence independence for prion formation." Proc. Natl. Acad. Sci. USA **102**(36): 12825-12830.

Santoso, A., P. Chien, et al. (2000). "Molecular basis of a yeast prion species barrier." Cell **100**(2): 277-88.

Serio, T. R., A. G. Cashikar, et al. (2000). "Nucleated conformational conversion and the replication of conformational information by a prion determinant." Science **289**(5483): 1317-21.

Sherman, F. (1991). "Getting started with yeast." Methods Enzymol. **194**: 3-21.

Shewmaker, F., E. D. Ross, et al. (2008). "Amyloids of shuffled prion domains that form prions have a parallel in-register beta-sheet structure." Biochemistry **47**(13): 4000-7. Epub 2008 Mar 7.

Sondheimer, N. and S. Lindquist (2000). "Rnq1: an epigenetic modifier of protein function in yeast." Mol. Cell **5**(1): 163-72.

Song, Y., Y. X. Wu, et al. (2005). "Role for Hsp70 chaperone in Saccharomyces cerevisiae prion seed replication." Eukaryot Cell **4**(2): 289-97.

Street, A. G. and S. L. Mayo (1999). "Intrinsic beta-sheet propensities result from van der Waals interactions between side chains and the local backbone." Proc Natl Acad Sci U S A **96**(16): 9074-6.

Tartaglia, G. G., A. P. Pawar, et al. (2008). "Prediction of aggregation-prone regions in structured proteins." J Mol Biol **380**(2): 425-36.

Ter-Avanesyan, M. D., A. R. Dagkesamanskaya, et al. (1994). "The SUP35 omnipotent suppressor gene is involved in the maintenance of the non-Mendelian determinant [psi+] in the yeast Saccharomyces cerevisiae." Genetics **137**(3): 671-6.

Toyama, B. H., M. J. Kelly, et al. (2007). "The structural basis of yeast prion strain variants." Nature **449**(7159): 233-7. Epub 2007 Sep 2.

Trovato, A., F. Seno, et al. (2007). "The PASTA server for protein aggregation prediction." Protein Eng Des Sel **20**(10): 521-3.

True, H. L. and S. L. Lindquist (2000). "A yeast prion provides a mechanism for genetic variation and phenotypic diversity." Nature **407**(6803): 477-83.

Tuite, M. F., C. R. Mundy, et al. (1981). "Agents that cause a high frequency of genetic change from [psi+] to [psi-] in Saccharomyces cerevisiae." Genetics **98**(4): 691-711.

Weathers, E. A., M. E. Paulaitis, et al. (2004). "Reduced amino acid alphabet is sufficient to accurately recognize intrinsically disordered protein." FEBS Letters **576**(3): 348-352.

Wickner, R. B. (1994). "[URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae." Science **264**(5158): 566-9.

Zibaee, S., O. S. Makin, et al. (2007). "A simple algorithm locates beta-strands in the amyloid fibril core of alpha-synuclein, Abeta, and tau using the amino acid sequence alone." Protein Sci **16**(5): 906-18.

Zoghbi, H. Y. and H. T. Orr (2000). "Glutamine repeats and neurodegeneration." Annu Rev Neurosci **23**: 217-47.

Chapter 4: Conclusion

Sup35p and its prion form, [*PSI*$^+$], has been the most useful and most widely studied yeast prion model system since its discovery 15 years ago. The strength and versatility of this model system has elucidated numerous properties of prions regarding their structure, kinetics and molecular processes, and has even broadened our understanding of their biological functions. One of the most critical discoveries however, came several year ago when it was determined that prion formation was dependent on composition, and not primary sequence (Ross, Baxa et al. 2004; Ross, Edskes et al. 2005).

This revelation of primary sequence independence for prion formation starkly contrasted the understanding of the Sup35p PFD wherein the oligopeptide repeats are essential for efficient prion propagation (Liu and Lindquist 1999; Parham, Resende et al. 2001; Crist, Nakayashiki et al. 2003; Osherovich, Cox et al. 2004; Shkundina, Kushnirov et al. 2006). While truncation studies clearly showed that the region of the Sup35p ORD was important to prion propagation, the repeats per se were likely unimportant since prion formation and propagation was not inhibited by scrambling the primary sequence of the PFD. Despite this however, several ex post facto experiments have made conclusions about the function of the PrP oligopeptide repeats based on the assumption that the oligopeptide repeats of Sup35p are critical to [*PSI*$^+$] propagation.

Additionally, the discovery that composition is the primary determinant for prion formation has provided researchers with a powerful method for the identification of new prions: searching the genome for regions of compositional homology to known PFDs. Since yeast prions have a distinctly unique compositional bias within their PFDs, search algorithms have used these biases for their training sets. Unfortunately, there have not yet been any comprehensive experiments aimed at understanding how composition affects prion formation, more specifically, what physiochemical properties of the amino acid side chains either promote or inhibit prion formation. Gaining this knowledge will simultaneously increase the ability of bioinformatic programs to identify new prions and broaden the general knowledge regarding the function of yeast prions.

Therefore, the overriding goal of my research has been to expand on the finding that composition, not primary sequence, drives prion formation by determining what the implications for this discovery are. With regard to the ORD of Sup35p, we have concluded from our experiments that the primary sequence of the repeats within this domain is not necessary for [*PSI*$^+$] propagation. Furthermore, the sequence requirements for the Sup35p ORD are likely only to be compositional; prion domains are characterized by their intrinsic disorder and prion-forming propensity as defined in Chapter 2 of this dissertation.

I. Primary sequence of ORD not involved in [*PSI*$^+$] maintenance

The molecular architecture of the Sup35p PFD has been defined as modular; the nucleation and elongation abilities reside in within the Q/N rich nucleation

domain (DePace, Santoso et al. 1998; Osherovich, Cox et al. 2004; Tanaka, Chien et al. 2004) while the Hsp104p-dependent [*PSI*$^+$] maintenance ability resides in the ORD (Liu and Lindquist 1999; Parham, Resende et al. 2001; Osherovich, Cox et al. 2004; Shkundina, Kushnirov et al. 2006). Although there exists a significant amount of evidence for these conclusions, it is unlikely that primary sequence plays any role in [*PSI*$^+$] formation or maintenance since scrambling the order of the Sup35 PFD does not inhibit these processes (Ross, Edskes et al. 2005). Despite this revelation however, there still exists a wide spread belief that the ORD is necessary for Hsp104p-dependent [*PSI*$^+$] maintenance.

In Chapter 1, we designed a series of experiments to elucidate the role of the Sup35p ORD. These experiments point to the conclusion that in contrast to what is currently believed in the field, the ORD is not required for Hsp104p-dependent [*PSI*$^+$] maintenance. Curing of [*PSI*$^+$] fibers composed of mutant Sup35p in which the ORD was disrupted, occurred indistinguishably from wild type [*PSI*$^+$]; both inhibition and overexpression of Hsp104p lead to the disappearance of the [*PSI*$^+$] phenotype. Additionally, the specific composition of the ORD is not necessary for Hsp104p-dependent [*PSI*$^+$] maintenance since a scrambled nucleation domain, which has a distinct composition, is able to replace the ORD.

We hypothesize that the ORD is not a distinct domain with specific functions within the Sup35p PFD, but is just a part of the entire domain. Therefore, replacing this domain with any random sequence (Crist, Nakayashiki et al. 2003) will inhibit [*PSI*$^+$] if the sequence is not prone to amyloid formation. In contrast, replacing this domain with sequences capable of amyloid formation (our Fusion Protein, ScrORD,

ScrPrP and ScrNuc mutatnts) allows for [*PSI*$^+$] formation and maintenance. This new model of the Sup35p PFD wherein the ORD is not required for [*PSI*$^+$] maintenance makes since. Hsp104p is a universal chaperone protein that is required to recognize many types of aggregates (prion and non-prion) and break them up, so if Hsp104p requires an inherent recognition sequence in the aggregated protein then it could only function on aggregates made from certain proteins. However, if there are no sequence requirements for Hsp104p recognition and binding, then Hsp104p could universally function to break up a broad spectrum of aggregates, including prion fibers. This model best fits with the flexibility of Hsp104p in facilitating fiber cleavage for many different yeast prion aggregates.

II. Compositional determinants of yeast prion formation

Amino acid composition is the primary driving force behind yeast prion formation: scrambling the primary sequence of the Ure2p (Ross, Baxa et al. 2004) and Sup35p (Ross, Edskes et al. 2005) PFDs does not inhibit prion formation, while additionally, the prion proteins Rnq1p (Sondheimer and Lindquist 2000), New1p (Michelitsch and Weissman 2000) and Mot3p (Alberti, Halfmann et al. 2009) were all identified based on their compositional similarities to other known prions. However, there is no current understanding of how composition drives or inhibits prion formation.

This lack of knowledge is most apparent in the bioinformatic programs designed to identify prion domains through genomic searches, which have less than a 20% success rate. Currently, these programs simply penalize residues that are not

found in yeast PFDs and reward residues that are. Specifically, hydrophobic and charged residues are under-represented in known prion domains so they are both penalized despite the fact that their reason for being under-represented are not known. Hydrophobic residues are predicted to be strong amyloid promoting residues while charged residues are predicted to inhibit amyloid formation (Pawar, Dubay et al. 2005); therefore, the bias against these two types of residues are likely for entirely different reasons.

In Chapter 2, we designed a random mutagenesis experiment to determine how amino acid composition affects prion formation in yeast. As expected, we found a bias against prolines and charged residues. Surprisingly, we found a bias for hydrophobic residues and no bias for or against Q/N residues. Although these biases were determined from mutants in which only an eight amino acid segment within the entire 114 amino acid Sup35-27p PFD was randomly mutated, they nevertheless can be used to predict the prion propensity of a large domain, as shown by Figures 3.5 and 3.6. When accounting for the prion propensities determined through our mutagenesis experiments along with the fold index, we can predict the ability of a given domain to form a prion with greater than 90% accuracy.

Yeast prion domains are unique from most other amyloid forming domains in the fact that they are intrinsically disordered in the native conformation of the protein. Therefore, we hypothesize that hydrophobic residues are excluded from these domains to prevent them from being overly aggregation prone and rarely existing in the non-prion state. Furthermore, the over-representation of Q/N

residues in the known prion domains are likely due to their propensity towards intrinsic disorder, and of the intrinsically disordered prone residues, they are also the most amyloid prone.

The relatively small region to which our random mutagenesis was targeted is the most probable reason that our screen found no bias for Q/N residues. Since the rest of the PFD has a large concentration of Q/N residues, more Q/N residues are unnecessary to provide intrinsic disorder to this domain. Therefore, there was very little selection for Q/N residues since they are not strong amyloid promoters like hydrophobic residues.

III. Summary/Future Directions

Overall, the experiments described in this dissertation represent a unique perspective on the molecular determinants of yeast prion formation. While it was previously hypothesized that prion formation is driven by amino acid composition, there was still a widely held belief in the field that primary sequence motifs within the Sup35p PFD were necessary for Hsp104p dependent [*PSI*$^+$] maintenance. Additionally, there was no knowledge of how amino acid composition affects prion formation, severely limiting the accuracy of bioinformatic programs designed to identify new prion domains based on composition.

Although the ORD of Sup35p appears to be a separate subdomain within the PFD, it does not contribute specific functions to the [*PSI*$^+$] phenotype. In our assays, neither de novo prion formation nor Hsp104p-dependent prion formation is affected by disruption of the ORD. The repeats are likely present because they

represent a simple method for genetically expanding the PFD of Sup35p, providing it with either enough length, or enough total aggregation propensity to allow for prion formation. Overall, amino acid composition is the most important determinant for prion formation of a given domain. Therefore, the biases found in our mutagenesis experiments represent the most useful criterion for accurately predicting prion formation.

Using our prion prediction algorithm based on the prion propensities we found in our experiments in conjunction with the Fold Index, we were able to accurately predict prion formation in yeast with great accuracy. The definitive test of our prediction algorithm will be its ability to predict prion-forming domains in genomic scans. Aside from yeast, this algorithm could be used for scanning other organisms including *C. elegans, Mus musculus* and the human genome. If strong prion domains are identified with these searches, cloning and testing these proteins for legitimate prion activity will be necessary for the definitive conformation of a new prion.

Additionally, by creating artificial sequences that contain intrinsic disorder and mild prion-forming propensities, we may be able flush out the relevance of Q/N residues to these domains; Q/N residues may or may not be absolutely necessary to acquire the characteristics that allow for prion formation. Furthermore, we may even be able to accurately construct prion domains from scratch. Artificial prions with specific desired characteristics could perhaps be designed by tweaking the intrinsic disorder and prion propensity of an artificial peptide. Also, subtler parameters could be tweaked, including the overall length of a strong amyloid

forming core or the patterns of strong and weak amyloid forming regions. These experiments represent the first step in a whole new era of prion research.

IV. References

Alberti, S., R. Halfmann, et al. (2009). "A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins." Cell **137**(1): 146-158.

Crist, C. G., T. Nakayashiki, et al. (2003). "[PHI+], a novel Sup35-prion variant propagated with non-Gln/Asn oligopeptide repeats in the absence of the chaperone protein Hsp104." Genes Cells **8**(7): 603-18.

DePace, A. H., A. Santoso, et al. (1998). "A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion." Cell **93**(7): 1241-52.

Liu, J. J. and S. Lindquist (1999). "Oligopeptide-repeat expansions modulate 'protein-only' inheritance in yeast." Nature **400**(6744): 573-6.

Michelitsch, M. D. and J. S. Weissman (2000). "A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions." Proc. Natl. Acad. Sci. USA **97**(22): 11910-5.

Osherovich, L. Z., B. S. Cox, et al. (2004). "Dissection and design of yeast prions." PLoS Biol. **2**(4): E86.

Parham, S. N., C. G. Resende, et al. (2001). "Oligopeptide repeats in the yeast protein Sup35p stabilize intermolecular prion interactions." EMBO J. **20**(9): 2111-9.

Pawar, A. P., K. F. Dubay, et al. (2005). "Prediction of "Aggregation-prone" and "Aggregation-susceptible" Regions in Proteins Associated with Neurodegenerative Diseases." J. Mol. Biol. **350**(2): 379-92.

Ross, E. D., U. Baxa, et al. (2004). "Scrambled Prion Domains Form Prions and Amyloid." Mol. Cell. Biol. **24**(16): 7206-7213.

Ross, E. D., H. K. Edskes, et al. (2005). "Primary sequence independence for prion formation." Proc. Natl. Acad. Sci. USA **102**(36): 12825-12830.

Shkundina, I. S., V. V. Kushnirov, et al. (2006). "The role of the N-terminal oligopeptide repeats of the yeast sup35 prion protein in propagation and transmission of prion variants." Genetics **172**(2): 827-35.

Sondheimer, N. and S. Lindquist (2000). "Rnq1: an epigenetic modifier of protein function in yeast." Mol. Cell **5**(1): 163-72.

Tanaka, M., P. Chien, et al. (2004). "Conformational variations in an infectious protein determine prion strain differences." Nature **428**(6980): 323-8.

Appendix A: Reporter assay systems for [URE3] detection and analysis

Andreas Brachmann,[a,*] James A. Toombs,[b] and Eric D. Ross[b]

I contributed to the testing of these reporter assay systems described in Figure A.3

A modified version of this work was published in Methods. 2006 May;39(1):35-42

Originally, the random mutagenesis project was designed for the [*URE3*] model system. This work laid the basis for a novel selection assay for this system.

Abstract

The *Saccharomyces cerevisiae* prion [URE3] is the infectious amyloid form of the Ure2p protein. [URE3] provides a useful model system for studying amyloid formation and stability *in vivo*. When grown in the presence of a good nitrogen source, [URE3] cells are able to take up ureidosuccinate, an intermediate in uracil biosynthesis, while cells lacking the [URE3] prion can not. This ability to take up ureidosuccinate has been commonly used to assay for the presence of [URE3]. However, this assay has a number of practical limitations, affecting the range of experiments that can be performed with [URE3]. Here, we describe recently developed alternative selection methods for the presence or absence of [URE3]. They make use of the Ure2p-regulated *DAL5* promoter in conjunction with *ADE2*, *URA3*, *kanMX*, and *CAN1* reporter genes, and allow for higher stringency in selection, selection against [URE3], nonselective assay of prion variants, and direct transformation of prion filaments. We discuss advantages and limitations of each of these assays.

1. Introduction

[URE3] and [*PSI*$^+$] are prion (infectious protein) forms of the *Saccharomyces cerevisiae* proteins Ure2p and Sup35p, respectively [(Wickner 1994)]. Formation of both prions involves conversion of the native proteins into an infectious, amyloid form [reviewed in (Ross, Minton et al. 2005)]. Formation of amyloid fibrils is associated with a wide range of human diseases, including Alzheimer's disease, Type II diabetes and the transmissible spongiform encephalopathies (TSEs). Because of the ease of genetic manipulation of yeast, [URE3] and [*PSI*$^+$] provide powerful model systems for studying *in vivo* amyloid fibril formation and the factors the cause certain amyloids to be infectious.

For both Ure2p and Sup35p, conversion to the prion form results in loss of the normal cellular function of the protein (Wickner 1994). Sup35p is a subunit of the translation termination factor; when Sup35p is in the prion form, stop codons are recognized with lower efficiency (Cox 1965). This nonsense suppression can be used to monitor prion formation by inserting an artificial stop codon into a selectable gene. The most commonly used marker for [*PSI*$^+$] formation is *ade2-1*, in combination with the weak nonsense suppressor tRNA *SUQ5* (*SUP16*) (Cox 1965). In the absence of [*PSI*$^+$], these cells are unable to grow without adenine and form red colonies when grown with limiting adenine due to the accumulation of a pigment derived from the substrate of Ade2p. However, the presence of [*PSI*$^+$] allows the cells to grow without adenine and to form white or pink colonies in the presence of limiting adenine.

Ure2p is involved in nitrogen catabolite repression [reviewed in (Cooper 2002)]. Ure2p blocks the uptake of poor nitrogen sources in the presence of a good nitrogen source by preventing the transcription factor Gln3p from entering the nucleus [reviewed in (Cooper 2002)]. One of the genes activated by Gln3p is the allantoate

permease gene, *DAL5*. Due to the structural similarity between allantoate and ureidosuccinate (USA), an essential intermediate of uracil biosynthesis, Dal5p can also take up USA (Turoscy and Cooper 1987). Loss of Ure2p activity, either due to deletion of the *URE2* gene or because of the presence of the [URE3] prion, allows *DAL5* transcription and uptake of USA in the presence of a good nitrogen source. Therefore, the ability to uptake USA can be used as an assay for [URE3] formation.

[URE3] has a number of practical advantages over [*PSI*$^+$] for the examination of prion formation. Efficient [*PSI*$^+$] formation requires the presence of another yeast prion, [*PIN*$^+$] (Derkatch, Bradley et al. 2001), which complicates interpretation of certain experiments. For example, when assessing how specific mutations in Sup35p influence prion formation, it is difficult to know whether the mutations affect the intrinsic ability of Sup35p to form amyloid aggregates, or whether they simply alter the interaction of Sup35p with [*PIN*$^+$]. Additionally, unlike Ure2p, Sup35p is an essential protein, making some experiments more challenging. Unfortunately, assaying [URE3] by USA selection has a number of distinct disadvantages. It is not particularly stringent, with cross-feeding occurring at high cell densities (Moriyama, Edskes et al. 2000). Additionally, unlike the color phenotype for [*PSI*$^+$], this method for monitoring [URE3] formation is always selective and does not allow detection of different [URE3] variants. It also requires the strains to be *ura2* and *URA3,* which limits the availability of the useful *URA3* locus for further genetic manipulations. Finally, neither the [*PSI*$^+$] or [URE3] systems offer a simple way to select against the presence of the prion.

To overcome these problems, alternative selection methods for [URE3] have been developed recently (Schlumpberger, Prusiner et al. 2001; Brachmann, Baxa et al. 2005). These selections take advantage of Ure2p regulation of the *DAL5* promoter. Because [URE3] formation results in derepression of the *DAL5* promoter, any gene put under control of the *DAL5* promoter will be similarly regulated. This

paper discusses some such reporter systems, the advantages and limitations of each, and the application of these reporters to the transformation of yeast with Ure2p amyloid fibrils.

2. Description of methods

2.1 Generation of reporter strains for [URE3] detection

A number of different alternative reporter systems for [URE3] detection have been established. Schlumpberger *et al.* (Schlumpberger, Prusiner et al. 2001) were the pioneers in this field, successfully using *ADE2* under control of a 561 bp fragment of *DAL5* promoter (P_{DAL5}:*ADE2*) to distinguish between [ure-o] (lacking the [URE3] prion) and [URE3] states, and even between two different [URE3] variants. Introduction of the reporter construct on plasmids showed extensive leakiness of the promoter, even when *CEN*-based, single copy vectors were used. This highlights the need for chromosomal integration of promoter constructs for efficient repression in the [ure-o] state. Their construct was introduced into the *URA3* locus of an *ade2* strain. [ure-o] strains were Ade^- and formed red colonies on ½ YPD plates, whereas [URE3] and *ure2* strains were Ade^+ and white or pink colonies on ½ YPD plates [(Schlumpberger, Prusiner et al. 2001) and Fig. 1]. In addition, the P_{DAL5}:*ADE2* reporter system allows the detection of different [URE3] variants (Fig. 4D). Although the basis of these variants still remains unknown, they are thought to represent different structures of the [URE3] amyloid core, resulting in differences in filament growth rates and stability (Schlumpberger, Prusiner et al. 2001; Brachmann, Baxa et al. 2005). This leads to variation in the amount of functional native Ure2p, observed by colony color (white or different shades of pink), as well

as by different rates of spontaneous [URE3] loss, indicated by the reappearance of red colonies (Fig. 4D).

In addition to the P_{DAL5}:*ADE2* reporter, we have established several other reporters, each of which has different advantages and limitations. P_{DAL5}:*URA3* allows for the selection for, as well as against, [URE3] or [ure-o]. P_{DAL5}:*kanMX* offers an alternative selection for the presence of [URE3] that permits a certain adjustment of the sensitivity of the system: whereas [URE3] cells with the before mentioned reporters are just Usa^+, Ade^+, or Ura^+, the degree of G418 resistance varies depending on the [URE3] variant present. Finally, P_{DAL5}:*CAN1* enables efficient selection against [URE3] or for functional Ure2p constructs or mutants (see below). In addition to these reporter constructs, a P_{DAL5}:*HIS3* construct has also been employed successfully in screens (T.B. Roberts and R.B. Wickner, pers. communication)

We followed a strategy similar to that of Schlumpberger *et al.* (Schlumpberger, Prusiner et al. 2001) for integration of these reporter constructs into the genome. In the case of P_{DAL5}:*ADE2* the most 3' 500 bp of ADE2 promoter were replaced by 586 bp *DAL5* promoter. The most convenient way to generate this reporter strain is to amplify the complete P_{DAL5}:*ADE2* construct [e.g. from strain YMS23 (Schlumpberger, Prusiner et al. 2001)], transform it into an *ade2 ure2* strain, and select for Ade^+ colonies. In the case of P_{DAL5}:*URA3,* promoter replacement was performed in a *URA3 URE2* [ure-o] strain by transformation of a PCR product containing the *DAL5* promoter flanked at the 5' end by 50 bp of *URA3* promoter sequence (position -256 to -207 relative to the translation start point of *URA3*) and at the 3' end by the first 50 bp of the *URA3* gene sequence. Successful transformants were selected by growth on 5-fluororotic acid (5-FOA) containing media, which selects for loss of *URA3* activity. Similarly, P_{DAL5}:*CAN1* was introduced by transformation of a PCR product with flanking sequences from the *CAN1* promoter

Figure A.1

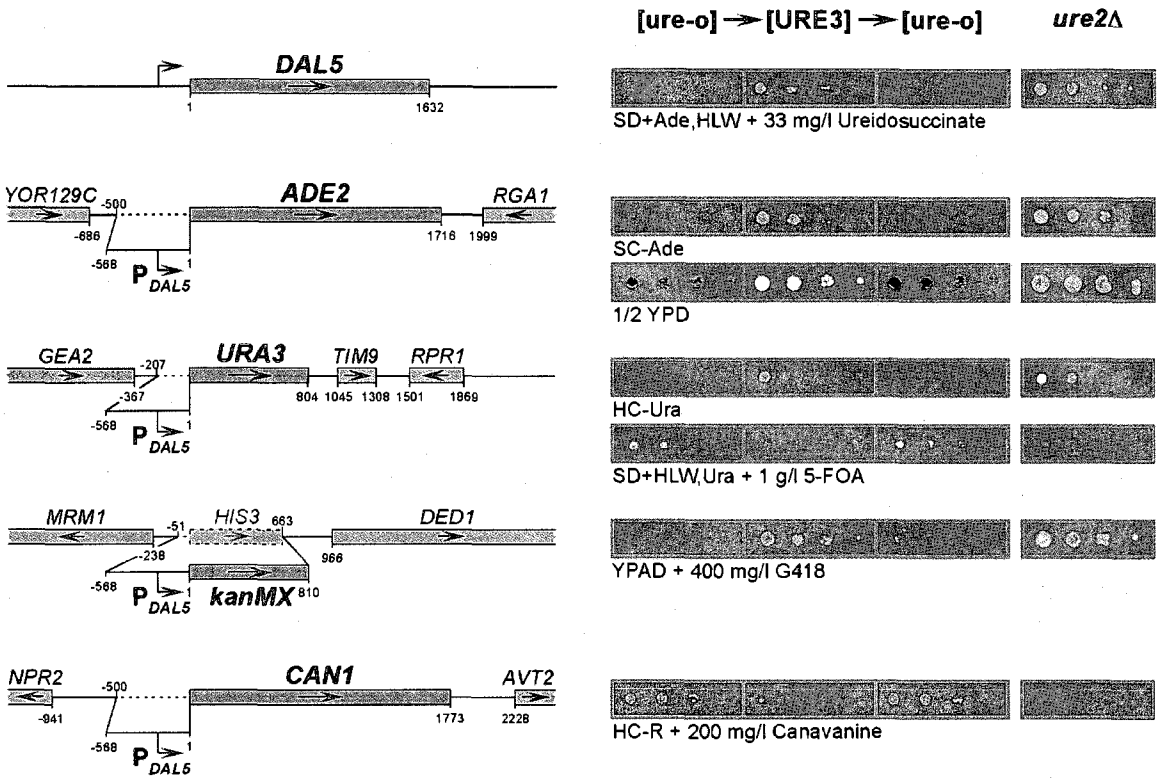


Figure A.1: Reporter assay systems for the detection of [URE3]. On the left, schematic representation of the genomic *DAL5* locus and of the modified P_{DAL5}:*ADE2*, P_{DAL5}:*URA3*, P_{DAL5}:*kanMX* (situated in the *HIS3* locus), and P_{DAL5}:*CAN1* loci. Dashed lines and rectangles indicate deleted genomic regions. On the right, serial 10-fold dilutions of representative reporter strains after 3 days at 30°C on diagnostic media. Cells lacking [URE3] ("[ure-o]") were converted to [URE3] by cytoduction of the original [URE3-1] variant isolate from Lacroute [18], and subsequently cured by growth on guanidine-containing ½ YPD plates for 3 days. On the far right, *ure2*D strains are shown for comparison. The following strains were used: *DAL5*, BY334 and 4132 [(Schlumpberger, Prusiner et al. 2001; Brachmann, Baxa et al. 2005)].; P_{DAL5}:*ADE2*, BY334 and BY256; P_{DAL5}:*URA3*, BY221 and BY194; P_{DAL5}:*kanMX*, BY180 and BY188; P_{DAL5}:*CAN1*, BY334 and BY256.

and gene into a *CAN1 URA2* [ure-o] strain and selected for growth on medium containing canavanine. To obtain P_{DAL5}:*kanMX,* a PCR product containing 50 bp of *HIS3* promoter sequence (position -100 to -51 relative to the translation start point of *HIS3*), the *DAL5* promoter, the *kanMX* gene, and 50 bp *HIS3* terminator sequence were transformed into a *HIS3 ura2* strain and selected for growth on G418 containing media.

Figure 1 gives a schematic representation of the modified reporter loci and shows the appearance of the reporter strains on different selective media (see below). These reporter constructs can easily be transferred into different genetic backgrounds by mating, sporulation, and selection for the desired genotype. Presence of the respective constructs can be verified by analytical whole-cell PCR. Conversely, [URE3] can be transmitted between different reporter strains by cytoduction. Representative reporter strains are given in Table 1.

Table A.1: representative reporter strains

Reporter	Strain	Genotype
P_{DAL5}:*ADE2*	BY327	*MAT***a** *trp1 ura2* P_{DAL5}:*ADE2* P_{DAL5}:*CAN1 kar1*
	BY334	*MAT*α *leu2 ura2* P_{DAL5}:*ADE2* P_{DAL5}:*CAN1 kar1*
	BY256	*MAT***a** *his3 leu2 trp1 ure2*Δ::*kanMX* P_{DAL5}:*ADE2* P_{DAL5}:*CAN1 kar1*
P_{DAL5}:*URA3*	BY221	*MAT***a** *his3 leu2 trp1* P_{DAL5}:*URA3* P_{DAL5}:*CAN1 kar1*
	BY194	*MAT*α *his3 leu2 trp1 ure2*Δ::*kanMX* P_{DAL5}:*URA3 kar1*
P_{DAL5}:*kanMX*	BY180	*MAT***a** *ade2 his3 leu2-3,112 ura2*Δ P_{DAL5}:*kanMX* P_{DAL5}:*CAN1 kar1*
	BY188	*MAT***a** *ade2 his3 leu2-3,112 ura2*Δ *ure2*Δ P_{DAL5}:*kanMX* P_{DAL5}:*CAN1 kar1*

2.2 Assay conditions and media factors influencing [URE3] detection

Media composition has a significant impact on the level of *DAL5* expression, and therefore affects the observed frequency of [URE3] formation. For example, presence of leucine in the growth medium represses *DAL5* activity [(Forsberg, Gilstring et al. 2001)]. Therefore, [URE3] cells carrying *CAN1* under control of the *DAL5* promoter grow more efficiently in the presence of canavanine when leucine is present in the media (Fig. 2A), presumably due to lower *DAL5* transcript levels. Perhaps as a result of this repression of *DAL5*, leucine also inhibits *de novo* [URE3] formation (data not shown).

Other media components also clearly affect *DAL5* activity and [URE3] formation, as can be observed in strains carrying a P_{DAL5}:*ADE2* reporter. Cells grown in Hartwell's complete (HC) medium have lower *DAL5* activity (Fig. 2B) and reduced frequencies of [URE3] formation (Fig. 3A) relative to cells grown in synthetic complete (SC) medium. This effect may be in part due to the presence of glutamate in HC medium, as addition of glutamate to ammonia-containing medium has been shown to repress [URE3] formation (Sekito, Liu et al. 2002).

The sensitivity of the assays to media conditions has to be kept in mind in the design of experiments. Yeast strains vary significantly in their dependence on media components; therefore testing of different media might be necessary to find optimal selection conditions. The media given in Figure 1 and Table 2 worked well for the majority of strains tested and should present a good starting point.

2.3 [URE3] induction experiments

When [ure-o] cells are plated on USA containing medium to select for spontaneous prion formation, between 1 and 100 cells per 1,000,000 plated

Figure A.2

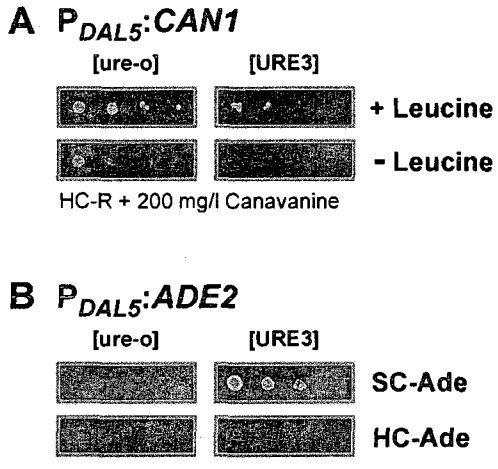


Figure A.2: Examples of media component influence on reporter activity. Negative regulation of *DAL5* promoter by (A) leucine in the growth medium and (B) different complete media. The difference between HC and SC most likely is due to the presence of both nitrogen sources, ammonia and glutamate, in SC, whereas HC contains only ammonia. Serial 10-fold dilution series after 3 days at 30°C. Reporter strains were BY327 and BY334.

Table A.2: Media for [URE3] detection and analysis

Reporter	Selection medium*
DAL5	SD medium + 33 mg/l ureidosuccinate (USA)
P_{DAL5}:*ADE2*	HC or SC medium lacking adenine or SD medium ½ YPD plates for nonselective determination of [URE3]
P_{DAL5}:*URA3*	SD medium or HC medium lacking uracil for counterselection: SD medium + uracil + 1 g/l fluororotic acid (5-FOA)
P_{DAL5}:*kanMX*	YPAD medium + 400 mg/l G418
P_{DAL5}:*CAN1*	HC or SC medium lacking arginine + 200 mg/l canavanine

* Hartwell's complete (HC), synthetic complete (SC), synthetic dextrose minimal (SD), YPAD, and ½ YPD media are according to [19].

spontaneously acquire the prion, with the exact frequency dependent on the strain of yeast used for selection (Wickner 1994). Overexpression of the Ure2p prion domain or full-length protein increases the frequency of prion formation, presumably because the misfolding event that initiates prion formation is a random occurrence, and therefore increasing the pool of prion domains within a cell increases the chances of this initial misfolding (Wickner 1994; Masison and Wickner 1995).

When similar selection for prion formation was done using the P_{DAL5}:*kanMX*, P_{DAL5}:*URA3*, and P_{DAL5}:*ADE2* reporters, both the reporter and the plating conditions used to select for prion formation significantly influenced the observed frequency of prion formation (Fig. 3A). This can be seen in strain BY334, which allows both adenine and USA selection. The observed frequency of prion formation when using USA selection was 10- to 50-fold lower than when selecting on synthetic dextrose medium lacking adenine, but was similar to the frequency on HC medium lacking adenine. Additional experiments were done to confirm that the increased frequency of Ade^+ colony formation observed on SD medium lacking adenine was a result of prion formation. Specifically, the majority of these colonies were demonstrated to also be USA^+ (data not shown). Additionally, growth of [URE3] cells on 5 mM guanidine results in loss of the prion [(Tuite, Mundy et al. 1981)]; most of the cells selected on SD medium lacking adenine lost both the USA^+ and Ade^+ phenotypes upon growth on medium containing guanidine (data not shown).

For the P_{DAL5}:*URA3* reporter (BY221), no colonies were observed when cells were selected on HC medium lacking uracil (data not shown), while a significant number were observed when selecting on SD medium lacking uracil (Fig. 3A). Surprisingly, when we used the P_{DAL5}:*kanMX* reporter (BY180) to select for [URE3] formation, the frequency of prion formation was slightly reduced upon overexpression of the prion domain (Fig. 3A). The reasons for this are unclear,

Figure A.3

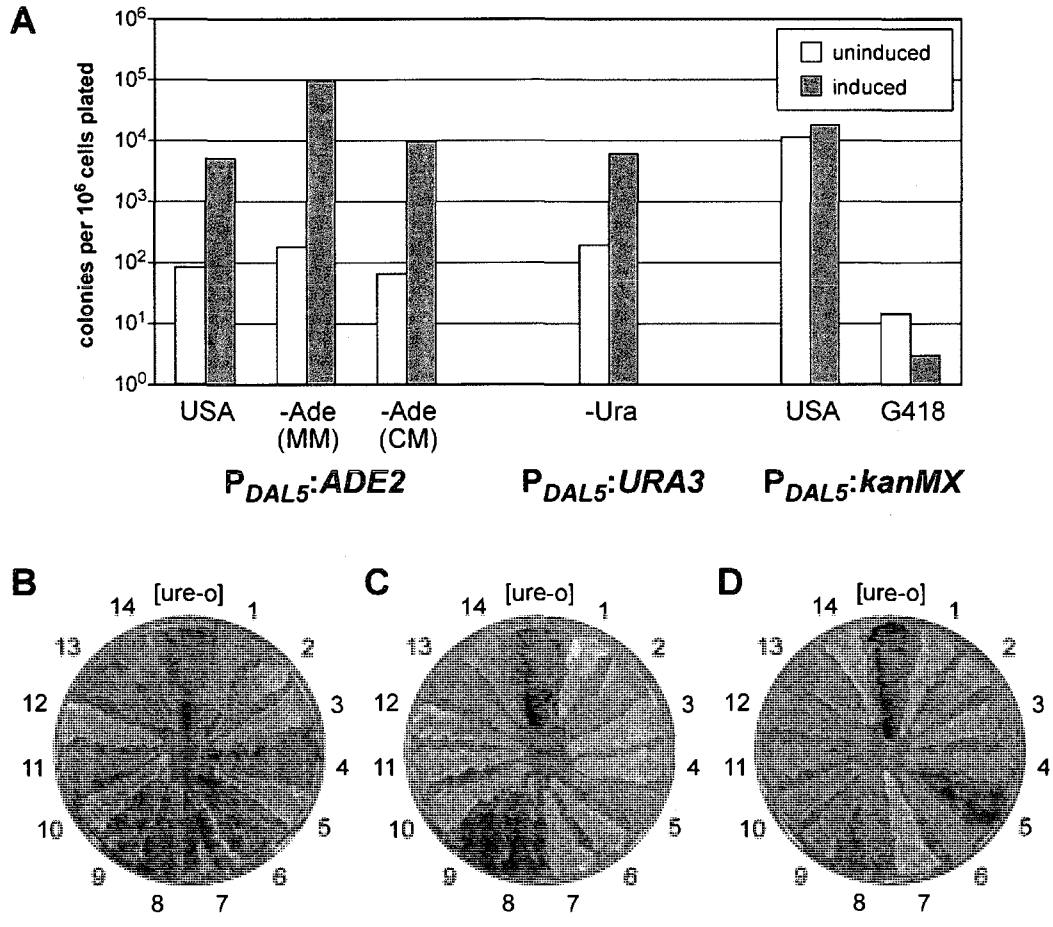


Figure 3: Induction of [URE3] by overexpression. (A) The indicated strains were transformed with a *LEU2* plasmid containing the *GAL1* promoter (uninduced; pH317) or a modified version of pH317 expressing the prion domain of Ure2p from the *GAL1* promoter (induced). Cells were grown for 3 days in galactose-raffinose dropout medium and plated to select for [URE3] formation. Shown are the frequencies of [URE3] formation for each strain. BY334 was plated tested on SD+Ade + 33 mg/l USA ("USA"), SD+Ura ("-Ade (MM)") and HC-Ade,L ("-Ade (CM)"). BY221 was tested on SD+HLW ("-Ura"). BY180 was tested on SD+Ade,HL + 33 mg/l USA ("USA") and YPAD + 400 mg/l G418 ("G418"). (B-D) Testing the array of [URE3] variants in BY334 cells selected on (B) SD+Ade + 33 mg/l USA, (C) SD+Ura, or (D) HC-Ade,L. After selection for [URE3], cells were grown on YPAD and YPAD supplemented with 5 mM guanidine. Cells that maintained the selected phenotype after growth on YPAD, but lost the phenotype after growth in the presence of guanidine were streaked onto 1/2 YPD. On each plate was streaked BY334 ([ure-o]), BY334 selected for [URE3] without induction (1-7) and BY334 selected for [URE3] with induction (8-14).

although it has previously been observed that overexpression of the prion domain both increases the frequency of prion formation and decreases the stability of established prions (Edskes, Gray et al. 1999). It should also be noted that when selected for prion formation on USA containing medium, this strain showed a high frequency of prion formation, but only weak induction upon overexpression of the prion domain, indicating that the lack of induction on G418 medium may be due to the particular strain background.

It is perhaps not surprising that different selection methods and plating conditions would yield different frequencies of prion formation, as each method is likely to have different sensitivity thresholds. One possibility is that more stringent selections such as USA selection allow only the strongest [URE3] variants to grow, while less stringent selections such as adenine selection allows growth of a broader range of variants. However, this does not seem to be the case. The P_{DAL5}:*ADE2* reporter allows for differentiation between weak and strong [URE3] variants; when cells are grown on medium containing limiting adenine, strong strains of [URE3] form white colonies, while weaker strains form pink colonies. Cells selected for [URE3] using either USA or adenine selection showed a similar array of prion variants (Fig. 3B-D).

These results make evident that the rate of [URE3] formation is highly dependent on the specific assay conditions, i.e. the reporter system as well as the media conditions used. Therefore, especially in [URE3] induction experiments, only relative values can be obtained. Since there does not seem to be a bias in the [URE3] variants that can be selected with a specific reporter system, the reason for the apparent differences in colony numbers between experiments remains unclear.

2.4 Transformation with [URE3] filaments

Direct transformation of prion filaments into yeast cells is a very useful technique, not only to confirm that amyloid is the molecular basis for yeast prions, but also for structural investigation of the infective units. Such experiments have been performed successfully with [*PSI*$^+$] using the *ade2-1* marker [15, 16]. In the case of [URE3], transformation of amyloid Ure2p filaments into *ura2*D strains and selecting for Usa$^+$ cells did not result in [URE3] transformants [9]. Only by using P_{DAL5}:*ADE2* reporter strains could a successful transformation protocol be established (Table 3, [9]): sonicated *in vitro* formed Ure2p filaments (Fig. 4A) as well as cell extracts from [URE3] strains are transformed together with a *LEU2* plasmid into protoplasts, and subsequently plated onto complete medium containing sorbitol and lacking adenine and leucine (Fig. 4B). Parallel plating onto medium lacking only leucine (Fig. 4C) allows monitoring of transformation efficiency. Depending on the amount and type of filaments transformed, up to 90% of Leu$^+$ transformants can be infected with the prion and become [URE3] [9]. [URE3] formation was routinely verified by Ade$^+$ phenotype of the transformants and reversion to Ade$^-$ after growth on plates containing guanidine. The P_{DAL5}:*ADE2* reporter has the additional advantage of direct determination of distinct [URE3] variants among the transformants when streaked on ½ YPD plates (Fig. 4D).

One of the most important parameters for successful transformation is the size of the amyloid filaments to be transformed (Fig. 4A). Fractionation experiments by size filtration revealed that the highest specific infectivity lies in fractions with particle sizes between 20 and 200 nm, with the upper limit most likely due to the maximum size of particles that can be taken up by yeast protoplasts [9]. Therefore, vigorous sonication of filament solutions prior to transformation is necessary (Table 3).

Figure A.4

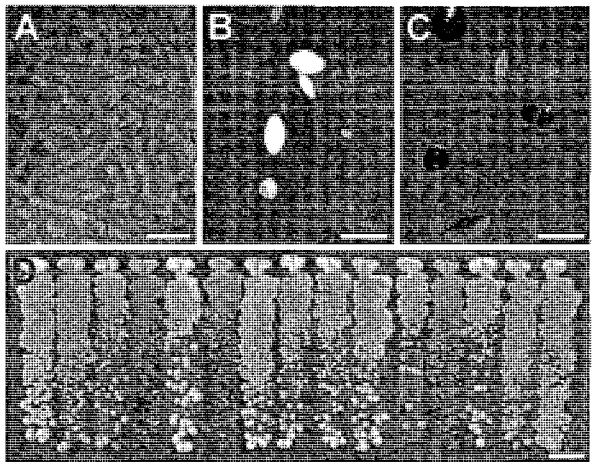


Figure A.4: Transformation of [URE3] filaments into P_{DAL5}:*ADE2* reporter strains. (A) Electron micrograph of negatively stained Ure2p filaments after sonication for 15 s at 60 W. Bar, 50 nm. (B) Transformation plates with 0.1 mg/l adenine after incubation at 30°C for 6 days. Note the appearance of small red Ade^- [ure-o] clones among the white Ade^+ [URE3] clones. Bar, 1 mm. (C) Same as (B) but with 5 mg/l adenine in the transformation plate. Buried colonies do not turn red but stay white due to anoxic conditions. Bar, 1 mm. (D) Spectrum of [URE3] variants after transformation of Ure2p amyloid filaments into P_{DAL5}:*ADE2* reporter strain. Randomly chosen transformants were streaked onto ½ YPD plates and incubated at 30°C for 3 days. Bar, 1 cm.

Table A.3:

Protocol for transformation of yeast cells with Ure2p filaments

Spheroplast preparation

1. Grow yeast strains in 50 ml YPAD* at 30°C to OD_{600} = 0.5.
2. Pellet cells at 1500 *g* for 5 min at room temperature and wash twice with 25 ml ST buffer.
3. Resuspend cells in 5 ml ST buffer with 100 U lyticase (Sigma, L-5263) and spheroplast for 40 min at 30°C.
4. Pellet spheroplasts at 250 *g* for 5 min at room temperature and wash twice with 10 ml ST buffer.
5. Resuspend cells in 1 ml STC buffer.

Transformation of Ure2p filaments or cell extracts

1. Sonicate filament solution or cell extract two times for 15 s at 60 W in ice water.
2. Transfer 100 ml spheroplast suspension into a new reaction tube, add 5 ml freshly denatured salmon sperm DNA (2 mg/ml, Sigma, D-1626), 1 ml *LEU2* plasmid (0.5 mg/ml), and 5 ml solution containing prions (*in vitro* formed filaments or whole cell extracts). Incubate for 10 min at room temperature.
3. Induce spheroplast fusion by addition of 900 ml PTC buffer and incubation for 20 min at room temperature.
4. Pellet spheroplasts at 500 *g* for 5 min at room temperature.
5. Resuspend spheroplasts gently in 200 ml SOS+HLUW and incubate for 30 min at 30°C.

Plating of transformants

1. Prepare two bottom plates with 20 ml HCS+A.1-L medium and one bottom plate with 20 ml HCS+A5-L medium
2. Fill three 15 ml round-bottom tubes, two with 10 ml HCS+A.1-L medium and one with 10 ml HCS+A5-L medium, and place the tubes in a 65°C water bath.
3. Add 1/10 and 1/2 of the transformation mixture to the tubes with HCS+A.1-L medium, mix by flipping over two times and pour onto the two HCS+A.1-L bottom plates.
4. Add 1/10 of the transformation mixture to the tube with HCS+A5-L medium, mix by flipping over two times and pour onto the HCS+A5-L bottom plate.
5. Incubate for 6 days at 30°C

Determination and verification of conversion efficiency

1. Determine rate of conversion to [URE3] by comparing colony numbers on HCS+A.1-L and HCS+A5-L plates.
2. Choose 48 potential transformants from the HCS+A.1-L plates, make little spots on SC-Ade* plates, and incubate for 2-3 days at 30°C. It is important to transfer as few cells as possible to obtain stringent selection for Ade^+ cells.
3. Replica plate onto ½ YPD* and ½ YPD* plates with 3 mM guanidine, incubate for 2 days at 30°C.

4. Check for reversion to red colony color on the guanidine containing plate and correct the conversion efficiency for potential spontaneous *ure2* mutants (colony color stays white).

Anticipated results

1. Transformation efficiency of the *LEU2* plasmid should be on the order of 5 x 10^4 Leu^+ colonies per mg plasmid.
2. Conversion efficiency should be between 10-30% [URE3] transformants when transforming 1 mM Ure2p filaments (calculated as theoretical Ure2p monomer concentration in the transformation mixture). This value is very yeast strain dependent.

Stock solutions

ST buffer (1 M sorbitol, 10 mM Tris-Cl, pH 7.5)

STC buffer (1 M sorbitol, 10 mM Tris-Cl, 10 mM $CaCl_2$, pH 7.5)

PTC buffer (20% (w/v) PEG 8000, 10 mM Tris-Cl, 10 mM $CaCl_2$, pH 7.5)

SOS+HLUW (1 M sorbitol, 7 mM $CaCl_2$, 1/3 YPAD*, 20 mg/l histidine, 100 mg/l leucine, 20 mg/l uracil, 20 mg/l tryptophan)

HCS+A.1-L (solid HC* medium with 1 M sorbitol and 0.1 mg/l adenine)

HCS+A5-L (solid HC* medium with 1 M sorbitol and 5 mg/l adenine)

---

* Hartwell's complete (HC), YPAD, and ½ YPD media are according to [19].

Interestingly, successful transformation of Ure2p filaments into P_{DAL5}:*ADE2* reporter strains is critically dependent on minute amounts of adenine in the transformation plates. There seems to be a strain dependent optimum adenine concentration for transformation efficiency: very low as well as higher adenine concentrations lead to a drastic decrease in the number of [URE3] transformants [9]. It seems that transformants initially need a certain supply of adenine to survive the time period until sufficient amounts of native Ure2p within the cell is transformed into amyloid, thereby allowing P_{DAL5}:*ADE2* transcription. On the other hand, higher adenine concentrations might result in rapid loss of the prion in this initial phase, due to insufficient selection.

2.5 Monitoring Ure2p activity in protein fusion constructs

Since the prion domain of Ure2p is not necessary for Ure2p activity, it is possible to replace the Ure2p prion domain with other potential amyloid forming domains and test for their ability to form prions. Similar substitutions of the Sup35p prion domain have been used to identify potential prion forming proteins. However, because [URE3] cells have the same phenotype as *ure2* cells, it is essential for the fusion proteins to be active. The various reporters are useful as a quick assay of the activity of Ure2p fusions.

For example, when the NM domain of Sup35p is fused to the Ure2p nitrogen regulation domain, the resulting fusion has partial, but not complete activity as assayed by the P_{DAL5}:*ADE2* reporter (*NM-URE2* in Fig. 5). If this NM-Ure2p fusion were assayed for [URE3] formation using the traditional USA assay, the [ure-o] cells would show weak growth on USA, thereby causing high background. In that case, the P_{DAL5}:*HIS3* reporter might be useful, as addition of 3-amino-1,2,4-triazole (3-AT) can be used to modulate the sensitivity of the assay.

Figure A.5

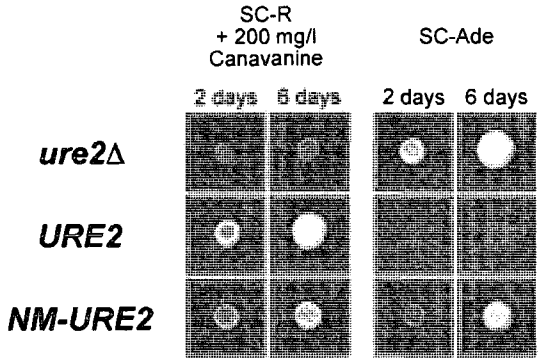


Figure A.5: Repressor activity of a Sup35p$^{1-251}$-Ure2p$^{66-354}$ fusion protein. Isogenic strains with both, the P$_{DAL5}$:*ADE2* and P$_{DAL5}$:*CAN1* constructs and containing either a *ure2* deletion, or expressing Ure2p or a fusion protein between the NM domains of Sup35p and Ure2p$^{66-354}$ were incubated on the indicated media for 2 and 6 days.

Additionally, the P_{DAL5}:*CAN1* reporter can be used to eliminate inactive fusions in library screens. One obvious way to identify new potential prions would be to insert a library of sequences in place of the sequence coding for the Ure2p prion domain and screen for [URE3] formation. However, any such library will invariably result in some fraction of inactive fusions. Because [URE3] cells and cells expressing inactive fusions would have the same phenotype, and because the frequency of prion formation is so low, isolating prion forming fusions would be difficult. The P_{DAL5}:*CAN1* reporter offers the potential to select against inactive fusions, thereby making such library screens more feasible.

3. Concluding remarks

We have described a variety of new reporters of *URE2* activity. In each case, a different gene is inserted under control of the *DAL5* promoter. These reporters each offer distinct advantages over the traditional USA assay for [URE3] formation. P_{DAL5}:*ADE2* offers a non-selective color phenotype of Ure2p activity, thereby making it possible to distinguish among [URE3] variants. Similarly, the degree of G418 resistance in cells expression the P_{DAL5}:*kanMX* reporter is dependent on the prion variant. The P_{DAL5}:*CAN1* reporter allows for selection for Ure2p activity and for loss of [URE3]. Finally, the P_{DAL5}:*URA3* reporter allows for selection both for and against the presence of [URE3]. However, it is important to note that for each of these reporters, the exact media conditions used for the experiments has a profound impact on the efficiency of selection and on the observed frequency of prion formation.

Acknowledgements

We would like to thank Tibor B. Roberts for the original P_{DAL5} containing plasmid, Ulrich Baxa for help with the transformation experiments and Figure 4A, and Reed B. Wickner for his continuous advice and helpful discussions. This work was supported by the NIH and Colorado State University.

References

Brachmann, A., U. Baxa, et al. (2005). "Prion generation in vitro: amyloid of Ure2p is infectious." EMBO J. **24**(17): 3082-3092.

Cooper, T. G. (2002). "Transmitting the signal of excess nitrogen in Saccharomyces cerevisiae from the Tor proteins to the GATA factors: connecting the dots." FEMS Microbiol. Rev. **26**(3): 223-38.

Cox, B. S. (1965). "PSI, a cytoplasmic suppressor of super-suppressor in yeast." Heredity **26**: 211-232.

Derkatch, I. L., M. E. Bradley, et al. (2001). "Prions affect the appearance of other prions: the story of [PIN(+)]." Cell **106**(2): 171-82.

Edskes, H. K., V. T. Gray, et al. (1999). "The [URE3] prion is an aggregated form of Ure2p that can be cured by overexpression of Ure2p fragments." Proc. Natl. Acad. Sci. USA **96**(4): 1498-503.

Forsberg, H., C. F. Gilstring, et al. (2001). "The role of the yeast plasma membrane SPS nutrient sensor in the metabolic response to extracellular amino acids." Mol Microbiol **42**(1): 215-28.

Masison, D. C. and R. B. Wickner (1995). "Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells." Science **270**(5233): 93-5.

Moriyama, H., H. K. Edskes, et al. (2000). "[URE3] prion propagation in Saccharomyces cerevisiae: requirement for chaperone Hsp104 and curing by overexpressed chaperone Ydj1p." Mol. Cell. Biol. **20**(23): 8916-22.

Ross, E. D., A. Minton, et al. (2005). "Prion domains: sequences, structures and interactions." Nature Cell Biol. **7**(11): 1039-1044.

Schlumpberger, M., S. B. Prusiner, et al. (2001). "Induction of distinct [URE3] yeast prion strains." Mol. Cell. Biol. **21**(20): 7035-46.

Sekito, T., Z. Liu, et al. (2002). "RTG-dependent mitochondria-to-nucleus signaling is regulated by MKS1 and is linked to formation of yeast prion [URE3]." Mol. Biol. Cell. **13**(3): 795-804.

Tuite, M. F., C. R. Mundy, et al. (1981). "Agents that cause a high frequency of genetic change from [psi+] to [psi-] in Saccharomyces cerevisiae." Genetics **98**(4): 691-711.

Turoscy, V. and T. G. Cooper (1987). "Ureidosuccinate is transported by the allantoate transport system in Saccharomyces cerevisiae." J. Bacteriol. **169**(6): 2598-600.

Wickner, R. B. (1994). "[URE3] as an altered URE2 protein: evidence for a prion analog in Saccharomyces cerevisiae." Science **264**(5158): 566-9.