*Cell*, Volume *137*

# Supplemental Data

# A Systematic Survey Identifies Prions and Illuminates Sequence Features of Prionogenic Proteins

**Simon Alberti, Randal Halfmann, Oliver King, Atul Kapila, and Susan Lindquist**

## SUPPLEMENTAL FIGURES AND LEGENDS

### Figure S1. Comparison of cPrD-EYFP expression levels

Yeast cells carrying expression plasmids for cPrD-EYFP fusion proteins were grown under inducing conditions for 24 h. Subsequently, cell lysates were prepared and analyzed by Western blotting using an anti-GFP antibody. Asterisks denote cPrDs with low expression levels that could only be detected with longer exposure times (data not shown). Endogenous Rnq1p was detected with a Rnq1p-specific antibody and served as a loading control.

### Figure S2. Fluorescence microscopy of cPrD-EYFP proteins

cPrD-EYFP fusion proteins were expressed from a galactose-regulatable expression plasmid in yeast cells containing the [*RNQ*+] prion. Cells were subjected to fluorescence microscopy after 24 hours. The fluorescence of representative cells was recorded. Fluorescence images are shown together with DIC pictures (insets). Arrows point to aggregates in the yeast cytosol. Expression

levels of EYFP fusion proteins after 24 h were assessed by Western blotting (Figure S1).

**Figure S3. Most candidates tested positively in multiple assays**

A Venn diagram is shown to illustrate the degree of overlap between the different assays employed in this study. The blue circle represents the number of candidates that tested positively in the *in vitro* assembly assay, the green circle represents the Sup35C prion assay data, the grey circle the microscopy data and the red circle the SDD-AGE *in vivo* aggregation data after 48 h of expression.

**Figure S4. Using Sup35p for prion detection**

(A) Wild-type yeast cells were compared to *SUP35*-deleted yeast cells carrying a plasmid for expression of Sup35p or NM-Sup35C from the *SUP35* promoter. All three strains maintained the red colony color of the [*psi*-] state and the white colony color of the [*PSI*+] state in a stable manner. This indicates that translation termination is fully functional in the [*psi*-] state and that the [*PSI*+] prion can be stably propagated, despite plasmid-based expression and the presence of a linker between the NM and C domains.

(B) NM-Sup35C or (C) N-Sup35C fusion proteins were expressed from plasmids carrying different promoters (*SUP35*, *ADH1*, *TEF2* and *GPD*) to determine optimal conditions for the Sup35p-based prion assay. Expression of the fusion proteins was determined by Western blotting and detection with an anti-Sup35p antibody. The blot was stripped and reprobed with an anti-Rnq1p antibody to

confirm equal loading. Colony colors of the corresponding strains are shown for comparison.

**Figure S5. Candidate *PrD-SUP35C* strains display a variety of colony colors**

Colonies of 90 cPrD-Sup35C-expressing strains growing on YPD plates (see Supplemental Experimental Procedures for details of strain construction). [*psi*-] and [*PSI*+] cells are shown for comparison (upper left corner). Some cPrDs, such as Nrp1p, New1p, Lsm4p and Nsp1p showed colony color switching under these non-inducing conditions. The low basal activity of the cPrD-Sup35C fusions of Def1p, Nup116p, Psp2p, Nup100p, Ddr48p and Rbs1p (indicated by the white colony color) prevented their subsequent use in prion-induction and selection assays.

**Figure S6. Comparison of cPrD-Sup35C expression levels**

Candidate *PrD-SUP35C* strains (Figure S5) were grown to mid-log phase and analyzed for the expression of cPrD-Sup35C fusion proteins by Western Blotting and probing with a Sup35p-specific antibody. Endogenous Rnq1p was detected with an anti-Rnq1p-antibody and served as a loading control. Colony colors of the corresponding strains growing on YPD are shown for comparison.

**Figure S7. SDD-AGE of *cPrD-SUP35C* strains**

Candidate *PrD-SUP35C* strains (Figure S5) were grown to mid-log phase and analyzed by SDD-AGE and Western blotting with a Sup35p-specific antibody.

Colony colors of the corresponding strains growing on YPD are shown for comparison. Only the cPrDs of Rnq1p, Lsm4p, Nup100p and Nsp1p showed detectable SDS-resistant aggregation under these conditions. The aggregation of Rnq1p was by far the strongest, consistent with it being the only pre-exising prion in this strain background. Lysates from wild-type [*psi-*] and [*PSI+*] cells are shown for comparison.

**Figure S8. Induction of the prion state in *cPrD-SUP35C* strains**

Yeast strains expressing cPrD-Sup35C fusion proteins were transformed with galactose-regulatable plasmids coding for corresponding cPrD-EYFP protein chimeras. The transformants were grown in galactose-containing medium for 24 hours and were then plated on adenine-deficient medium (↑). The same strain grown under non-inducing conditions (raffinose-containing medium) served as a control. The candidates highlighted in red showed a detectable increase in the number of Ade+ colonies under inducing conditions.

**Figure S9. Comparison of [*prd-c-*] and [*PrD-C+*] strains on YPD**

$Ade^{+}$ colonies from prion induction experiments (Figure S8) were re-streaked on YPD (bottom half of the plates). [*prd-c-*] cells (top half of the plates) and wild-type [*psi-*] and [*PSI+*] strains (plate in the top left corner) are shown for comparison.

**Figure S10. SDD-AGE and GdnHCl curing of additional [*PrD*-C+] strains**

(A) Cells lysates were prepared from [*prd-c-*] and [*PrD-C+*] strains and analyzed by SDD-AGE and Western Blotting. The cPrD-Sup35C fusion proteins were detected using an anti-Sup35p antibody. Colony colors of the corresponding [*prd-c-*] and [*PrD-C+*] strains growing on YPD are displayed above the SDD-AGE Western blots.

(B) [*PrD-C+*] strains were passaged three times on plates containing 5 mM GdnHCl and then spotted onto YPD plates ('cured'). The corresponding [*prd-c-*] and [*PrD-C+*] strains are shown for comparison.

**Figure S11. Plots of amino acid biases in core (A) and full-length (B) cPrDs**

For each candidate the results from all four assays (see Table S2) were combined into a cumulative score. Only candidates that could be tested in all four assays were used for analysis. For the fluorescence microscopy and Sup35 prion assay results, positive candidates received 2 points and 0 points if they were negative. Results from SDD-AGE (48 hours of induction) and *in vitro* assembly were treated as follows: All candidates received points according to the scheme used in Table S2 (- = 0 points, + = 1 point, ++ = 2 points, +++ = 3 points). Therefore, the maximum combined score is: 2 + 2 + 3 + 3 = 10. The combined score (x-axis) was plotted against several other parameters on the y-axis. Parameters investigated were the relative frequencies of single amino acids or combinations of amino acids (percent X or percent XY), length of the cPrD, the number of occurrences of amphiphilicity patterns (count PHPH,HPHP, PHPHP and HPHPH, where H stands for a hydrophobic amino acid and P for a polar

amino acid according to the Kyte-Doolittle hydropathy index (Kyte and Doolittle, 1982)). The graphs of the form log2(HP/eHP) show the log of the actual counts of HP patterns divided by the expected counts given the amino acid composition and length (0.01 was added to the numerator and denominator to avoid infinities). The graphs "non P" and "mask P" examine the influence of proline abundance and spacing: "non P" shows the number of residues other than proline, and "mask P" shows the number or residues remaining when prolines and any intervening sequences of length <5 are removed. Spearman rank-correlations were computed using R, and p-values were computed by randomly permuting the gene labels 10000 times. For 95% of these random permutations, none of the 35 associated p-values was <0.002, so p-values <0.002 remain significant at the 0.05 level after accounting for the testing of multiplenon-independent hypotheses.

**SUPPLEMENTAL TABLES AND LEGENDS**

Table S1 is provided as a separate PDF file. Tables S3 and S4 are provided as separate Excel sheet files.

**Table S1. Prediction of cPrDs in the yeast proteome**

Output format of the cPrD prediction algorithm for the top 179 prion candidates. The amino acid sequence of the candidate proteins are shown at the bottom in single letter code. The core region of the cPrDs is highlighted in orange and the extended region in pink. The top panel shows the probability of each residue

belonging to the HMM state “cPrD” (red) and “background” (black); the tracks “MAP” and “Vit” illustrate the Maximum a Posteriori and the Viterbi parses of the protein into these two states. The lower panel shows sliding averages over a window of width 60 of net charge (pink), hydropathy (blue), and predicted disorder (gray) as in FoldIndex, along with a sliding average based on cPrD amino acid propensities (red).

**Table S2. Comparison of different aggregation and prion assays**

Prion candidates are ranked according to their highest scoring core cPrDs (left column). Data from fluorescence microscopy, SDD-AGE analysis (after 48 hours of expression), Sup35C prion assay and *in vitro* assembly assay are shown to allow comparison of the results. N/A (not applicable) indicates that these experiments were not conducted for the following reasons: the cPrD could not be cloned (candidates shown in grey), were recalcitrant to purification (*in vitro* aggregation), expression levels were too low (microscopy and SDD-AGE) or the basal activity of the cPrD-Sup35C fusion protein was too low for selection. We used a simple scheme (-, +, ++, +++) to categorize cPrDs depending on the extent and kinetics of aggregation. Candidates in red are considered to be most promising for further characterization (they either showed switching behavior or strong amyloid formation). Known prions are underlined.

**Table S3. Oligonucleotides used in this study**

**Table S4. Yeast strains used in this study**

**Table S5. Comparison of median amino acid frequencies**

Candidates PrDs were split into a positive (cPrDs that satisfied the criteria of the assay) and a negative (cPrDs that did not satisfy the criteria) set. The sequences of the complete PrDs and the core PrDs of the two groups were used to calculate the median amino acid frequencies. Amino acids are shown in the left column in single letter code. QN refers to the combined median frequency of Qs and Ns.

## SUPPLEMENTAL RESULTS AND DISCUSSION

### Pilot experiments using the Sup35p-based phenotype switching assay

Before we tested our cPrD library for switching behavior we conducted a set of control experiments to determine optimal conditions for the *SUP35*-based prion assay. First, we compared plasmid-expressed wild-type Sup35p to a Gateway® *NM-SUP35C* expression construct that carries a linker between NM and the C region (Figure S4A). This linker contains a stretch of glycines and prolines and a translated portion of the recombinogenic sites (see Supplemental Experimental Procedures for details). The linker-containing Sup35p was indistinguishable from wild-type Sup35p in all respects (Figure S4A). In a second set of control experiments we tested our recombination-based *SUP35C* fusion system with different promoters (Figure S4B). It turned out that too high expression levels

interfere with translation termination, probably by titrating out other factors that are important for translation termination.

The *SUP35* assay is based on a variety of previous studies that successfully used the Sup35p protein to test the prion character of cPrDs, including the PrDs of Rnq1p and New1p (Sondheimer and Lindquist, 2000; Osherovich and Weissman, 2001; Osherovich et al., 2004). In these studies, either the N or the NM domains of Sup35p were replaced by the cPrD to be tested and the resulting cells were then assayed for [*PSI*+]-like phenotypes. We generated Gateway® tagging plasmids for the expression of fusions to the MC or C domain of Sup35p. To test these plasmids, we inserted the N domain of Sup35p, resulting in fusion proteins with the linker located between the N and MC or the N and C domains of Sup35p. Only the fusion protein lacking the M domain showed sufficient translation termination activity (data not shown and Figure S4C), which lead us to focus our efforts on experiments with *SUP35C* tagging plasmids. The expression level of N-C Sup35p was lower than that of wild-type Sup35p (Figure S4C), but the overall translation termination activity seemed to be increased (as judged by the colony color). Notably, the colony color of the strain expressing N-C Sup35p from an *ADH1* promoter construct was comparable to that of wild-type Sup35p. Based on these results, we used an *ADH1*-promoter expression plasmid that enables fusions to the C domain for subsequent prion switching assays with the cPrD library.

We acknowledge that using the Sup35C domain rather than both the M and C domains may not detect some prions that, like Sup35, require a highly charged

domain to maintain a soluble non-prion state. However, prion proteins like Ure2p and Rnq1p do not contain a highly charged domain and thus arguably would be more closely mimicked by PrD fusions to the C domain only, as in our construct. Additionally, we did include the M domain in the *in vitro* assay, explained below, in part to increase our ability to detect prion properties using these diverse assays.

**[*NEW1-C*+] forms Hsp104p-independent prion particles**

Contrary to other [*PrD-C*+] strains, the [*NEW1-C*+] strain was not cured after passaging on GdnHCl-containing plates (Figure S10B). In agreement with this finding, we were able to induce the [*NEW1-C*+] prion in a *HSP104*-deleted strain (data not shown). We performed an SDD-AGE analysis on the Hsp104p-independent [*NEW1-C*+] strain and found that the pattern of aggregation was similar to that of the wild-type [*NEW1-C*+] strain (data not shown), indicating that the New1p PrD is able to undergo the necessary structural changes without assistance from the prion remodeling factor Hsp104p. These results do not agree with the findings of an earlier study (Osherovich and Weissman, 2001; Osherovich et al., 2004) that also investigated the prion properties of New1p using the Sup35p-based system. The fusion protein tested by Osherovich et al. gave rise to a prion state, [*NU*+], that could be cured by GdnHCl treatment. We noted a high spontaneous induction and loss rate for [*NEW1-C*+], suggesting that [*NEW1-C*+] has a high rate of nucleation and polymerization but an insufficient rate of aggregate shearing. In contrast, Osherovich et al. reported a much higher

stability for the [*NU*+] prion. We suggest that these differences arose because of the different constructs that were used. We fused the New1p PrD to the C domain of Sup35p, whereas Osherovich et al. used the MC domain. In addition, the sizes of the New1p PrDs used in each study were different. While the initial study focused on a region spanning amino acids 1-153, our prediction set the New1p PrD between amino acids 39-115. We think that additional sequences in the [*NU*+] prion could provide binding sites for Hsp104p that enable shearing of the prion particles and therefore more stable inheritance. Moreover, the M domain of Sup35p is important for stable inheritance of [*PSI*+] and could provide additional binding sites for Hsp104p (Liu et al., 2002).

## SUPPLEMENTAL EXPERIMENTAL PROCEDURES

### Computational prion prediction

A hidden Markov Model (HMM) was used for identifying candidate prion domains. The HMM had two hidden states, PrD and background, and the output symbols were the 20 amino acids (AAs). The output probabilities for the PrD state were constructed based on the AA frequencies in the PrDs of Sup35p, Rnq1p, Ure2p, and New1p (Masison and Wickner, 1995; Edskes et al., 1999; Li and Lindquist, 2000; Santoso et al., 2000; Sondheimer and Lindquist, 2000; Osherovich and Weissman, 2001; King and Diaz-Avalos, 2004; Osherovich et al., 2004), each normalized to have 100 AA counts total (so that the longer PrDs would not have a disproportionate influence), along with 1 pseudocount for each

AA. The output probabilities for the background state were the AA frequencies for the entire *S. cerevisiae* proteome. The transition probability from PrD to background was set to 0.02 and the transition probability from background to PrD state was set to 0.001. Initial probabilities for PrD and background were set to 0.05 and 0.95, respectively.

Each *S. cerevisiae* protein was parsed using the Viterbi algorithm, and regions of at least 60 consecutive AAs classified in the cPrD state were considered as candidate prion domains. For each candidate prion domains we identified a core of 60 consecutive AAs for which the ratio of the probability of the sequence under the PrD AA frequencies relative to the "background" AA frequencies was maximized; we ordered the candidate prion domains according to this ratio of probabilities.

**Cloning procedures and vector construction**

We recently generated a set of Gateway® System cloning vectors that are derived from the commonly used pRS yeast shuttle vectors (see Alberti et al., 2007 and http://www.addgene.org/). These pAG vectors have a Gateway® chloramphenicol/ccdB resistance cassette inserted into the single SmaI restriction site of each of the 24 pRS plasmids (Mumberg et al., 1995). The basic pAG vector set was modified to allow for the use of different promoters and C-terminal *SUP35* fusion tags. The coding sequences for the tags and the *SUP35* promoter were amplified from genomic S288C DNA using the primers listed in Table S3. The *TEF2* and *ADH1* promoter were derived from the pRS4xx series of

plasmids (Mumberg et al., 1995). To generate pAG4xxADH-ccdB, pAG4xxTEF-ccdB and pAG4xxSUP-ccdB (please see Alberti et al., 2007, for nomenclature), the *GPD* promoter sequence of pAG4xxGPD-ccdB was deleted through digestion with SpeI and SacI, followed by the insertion of the *TEF2*, *ADH1* and *SUP35* promoter sequences. A PCR product (see Table S3 for oligonucleotide sequences) coding for the Sup35C domain (amino acids 234-685) was cloned between the HindIII and XhoI sites producing pAG4xxADH-ccdB-SUP35C, pAG4xxGPD-ccdB-SUP35C, pAG4xx-TEF-ccdB-SUP35C and pAG4xxSUP-ccdB-SUP35C. The 5' oligonucleotides used for generating the *SUP35C*-tagging vectors contained a linker that translates into the sequence GGPGGG. This sequence was included to increase the flexibility and reduce any adverse effects of the PrDs on the translation termination function of Sup35C.

Multiple Gateway® compatible destination vectors were generated for expression of His-tag fusions in *E. coli*. pRH1 was generated from pAED4-SCNM-his7 (Osherovich et al., 2004) as follows. The NM ORF was cut out with HpaI and NdeI and replaced with the Gateway® cassette RfB (Invitrogen, CA). Quick-Change mutagenesis was then applied to remove the Shine-Dalgarno site and add a codon for tryptophan (Trp) prior to the 7xHis sequence, to allow for spectrophotometric quantitation of purified proteins. To create pRH2, an NdeI site was added to pRH1 just prior to the Trp codon, into which the M domain (amino acids 125 - 320 of Sup35p) was inserted. Finally, the NdeI sites were deleted. For N-terminal His-tagged proteins, pRH3 was generated by inserting the M

domain into the NdeI site of pDest17 (Invitrogen, CA), followed by deletion of the resulting 3' NdeI site. All vector modifications were verified by sequencing.

Candidate PrDs and other amyloidogenic gene fragments, such as the N and NM domains of Sup35p and the Q25, Q72 and Q103 length variants of huntingtin were amplified using a two step PCR with overlapping primer sets (see Table S3 for a list of the oligonucleotides used). A two step PCR was chosen as a cost-effective alternative that avoids the synthesis of very long oligonucleotides. In the first PCR, gene-specific primers were used to amplify the region of interest from genomic DNA of the yeast strain S288C or plasmid DNA. Universal primers (Table S3) were then used in the second round to attach the recombinogenic attB1 and attB2 sites at the 5' and 3' ends of the primary PCR product, respectively. Proof-reading Platinum Pfx DNA polymerase (Invitrogen, CA) was used in both reactions to guarantee accurate DNA synthesis. The correct amplification and integration of the DNA sequences into pDONR221 (Invitrogen, CA) was confirmed by sequencing. Several cPrDs contained a single cysteine at the N or C terminus to allow the incorporation of dyes in follow-up experiments. In addition, we added an N terminal serine residue to internal PrDs to increase PrD stability. Below is an illustration of the sequence features that were incorporated into the oligonucleotides to allow Gateway® recombination and dual expression in bacteria and yeast (the sequences overlapping in the gene-specific and the universal primers is underlined):

attB1 site Shine-Dalgarno Kozak

ACA AGT TTG TAC AAA AAA GCA GGC TTC GAA GGA GAT AAC AAA ATG --

attB2 site

AC CAC TTT GTA CAA GAA AGC TGG GTT --

**Yeast strains and media**

The yeast strains used in this study were derived from YJW509 (MATα, leu2-3,112; his3-11,-15; trp1-1; ura3-1; ade1-14; can1-100; [*psi*-]; [*pin-]*) and YJW584 (MATa, leu2-3,112; his3-11,-15; trp1-1; ura3-1; ade1-14; can1-100; [*psi*-]; [*PIN+]*) (see Osherovich et al. (2004) for details on strain generation). A detailed list of the strains generated in this study is shown in Table S4. The media used were complete standard synthetic media or media lacking particular amino acids containing either 2 % D-glucose (SD), 2 % D-galactose (SGal), 2 % D-raffinose (SRaf) or a mix of raffinose and galactose (SRafGal, 1 % each). Plates used for prion curing contained 5 mM guanidine hydrochloride (GdnHCl). For determining drug-resistance prion phenotypes, the indicated concentrations of drugs were added to YPD media after autoclaving, prior to solidification.

A PCR-generated (Baudin et al., 1993; Wach et al., 1994; Goldstein and McCusker, 1999) deletion strategy was used to systematically replace a yeast open reading frame from its start to stop codon with a kanMX4 or hphMX4 module. 64 bp forward and 67 bp reverse primers were used to amplify the kanMX4 gene from pFA6-kanMX4 or the hphMX4 gene from pAG32 DNA. These

primers consisted of 45 bp gene-specific sequences and sequences that recognize regions in the kanMX and hphMX4 modules - a 19 bp sequence (CAGCTGAAGCTTCGTACGC) for the forward primer and a 22 bp sequence (GCATAGGCCACTAGTGGATCTG) for the reverse primer. A standard lithium-acetate transformation protocol (Gietz et al., 1992; Gietz et al., 1995) was used to introduce the gene disruption cassettes into yeast cells followed by selection of colonies on G418 (200 µg/ml) or Hygromycin (300 µg/ml) containing agar plates. The knockouts were confirmed using the KanB primer (Wach et al., 1994) which binds in the kanMX4 cassette and individual 20-22 bp primers several 100 bp upstream of the start codons of the candidate genes.

A PCR-based tagging strategy (for details see Wach et al., 1997 and http://depts.washington.edu/yeastrc/pages/plasmids.html) was used to integrate the coding sequences for Cerulean at the 3' end of genes. 60 bp primers were used to amplify a Cerulean-hphMX4 cassette from pBS10 DNA. The oligonucleotides consisted of a 40 bp homology to the gene of interest and the sequence GGTCGACGGATCCCCGGG for the forward primer and ATCGATGAATTCGAGCTCG for the reverse primer.

The *dan1::URA3* strain for studies of [*MOT3*+] was constructed as follows. A *URA3-HIS5* cassette was amplified from a modified pUG27 (Gueldener et al., 2002) plasmid (bearing *URA3* (-160 to +78 from pRS306) inserted between the HindIII and SalI sites), a kind gift from Sherwin Chan and Gerald Fink, Whitehead Institute for Biomedical Research, Cambridge, MA. Forward and reverse primers for amplification of the cassette contained 44 and 49 bp, respectively of

homology to the immediate upstream and downstream sequence at the *DAN1* ORF. The PCR product was gel-purified and transformed using the methodology described above, followed by selection on SD-his media and confirmation of correct integration using primers dan1upseq and URA3-1-3, which amplify across the 5' integration junction.

**Fluorescence microscopy**

The library of 94 cPrDs was cloned into the vector pAG424GAL-ccdB-EYFP. Each of the resulting constructs was then introduced into the YJW584 strain. The cells were grown for 24 hours in galactose-containing medium and subjected to fluorescence microscopy using an Axiovert 200M microscope (Carl Zeiss, Jena, Germany). Fluorescence and DIC images were acquired and processed using Axiovision (Carl Zeiss, Jena, Germany) image analysis software. To evaluate the expression levels of the cPrD-EYFP fusions, cell lysates were prepared after 24 hours of expression and analyzed by Western blotting and detection with a GFP-specific antibody (Figure S1).

***In vitro* aggregation assays**

In pilot experiments, we found that many cPrDs expressed with only a His-tag (without any additional solubilizing features) could not be completely solubilized prior to the assembly experiments. This was likely due to the absence of solubilizing elements normally found in the native full-length proteins. Prion proteins are able to switch between two states because of a balance between

insolubility – driven by the prion domain – and solubility – driven by non-prion domains. In the case of Sup35, this latter activity is largely provided by a highly charged, natively unfolded middle region, M, which is located adjacent to the prion domain. Thus, to more accurately mimic the solubility of the native proteins while still avoiding the necessity of a renaturing step after purification, we incorporated the M region between the cPrD and poly-His coding sequences..Most were purified with M and a poly-Histidine tag at the C-terminus, with a few exceptions, as explained below, due to expression difficulties. We successfully purified 91 of the 94 cPrDs under fully denaturing conditions, as have been employed for previously characterized prions.

Candidate PrDs were recombined into pRH2 and transformed into BL21AI (Invitrogen, CA), an *E. coli* strain optimized for toxic protein expression. Some cPrDs (Ksp1p, Wwm1p, Nup49p, Nup100p, and Med2p) were coexpressed with pRARE2 (EMD Biosciences, NJ) to increase expression levels. The Swi1p and Nup100p cPrDs were expressed in pRH1. Cbk1p, Def1p, and Psp2p PrDs were expressed in pRH3. 300 ml 2 x YT cultures were induced with 1 mM IPTG at an OD of ~ 0.4. After 3 hours, cells were sedimented and resuspended in lysis buffer (7 M GndHCl; 100 mM $K_2HPO_4$, pH 8.0; 5 mM imidazole; 300 mM NaCl; 5 mM 2-mercaptoethanol) buffer for 1 hr at RT. Lysates were then cleared for 20 min at 20,000 rcf and loaded onto a BioRobot8000 (Qiagen, CA) for purification using TALON® Superflow resin (Clontech, CA), according to the manufacturers' instructions. Proteins were eluted (8 M urea; 100 mM NaOAc/HOAc, pH 4; 5 mM β-mecaptoethanol) and precipitated with 5 volumes of methanol.

For assembly reactions, methanol-precipitated proteins were resuspended in 10-50 µl of resuspension buffer (7 M GndHCl; 100 mM $K_2HPO_4$, pH 5.0; 300 mM NaCl, 5 mM EDTA, 5 mM TCEP). Protein concentrations were determined by measuring absorption at 280 nm using calculated extinction coefficients. Protein stocks were heated for 5 min at 95°C before being diluted to 20 µM in assembly buffer (5 mM $K_2HPO_4$, pH 6.6; 150 mM NaCl; 5 mM EDTA; 2 mM TCEP) plus 0.5 mM ThT. Assembly reactions were performed in black nonbinding microplates (Corning, NY) with 100 µl per well, with 1400 rpm agitation at 30°C in an iEMS incubator/shaker (Thermo Scientific). Fluorescence measurements (450 nm excitation, 482 nm emission) were made with a Sapphire

II plate reader (Tecan, NC). In order to minimize variation arising from sporadic sampling of very large fluorescent aggregates, values for each time point represent the averages of 10 readings taken over a five minute period.

After the final fluorescence time point, reactions were applied to nitrocellulose and cellulose acetate membranes as described by Boye-Harnasch and Cullin (Boye-Harnasch and Cullin, 2006). Ponceau S was used to detect immobilized proteins.

**Semi-denaturing detergent-agarose gel electrophoresis (SDD-AGE)**

A library of pAG424GAL-cPrD-EYFP constructs in YJW584 was used to investigate the amyloid propensities of cPrDs *in vivo*. The cPrD-EYFP proteins were expressed for 24 or 48 hours and then processed for SDD-AGE analysis. Cells were harvested by centrifugation and resuspended in buffer A (50 mM Hepes, pH 7.5; 150 mM NaCl; 2.5 mM EDTA; 1 % (v/v) Triton X-100) containing 30 mM NEM, 1 mM PMSF and 1 x Complete Protease Inhibitor (Roche). Cells were lysed using glass beads and were briefly spun at 3,000 rcf to sediment cell debris. The protein concentrations of the cell lysate were adjusted and mixed with 4 x sample buffer (2 x TAE; 20 % (v/v) glycerol; 4 % (w/v) SDS; bromophenol blue). The samples were incubated at room temperature for 15 minutes and loaded onto a 1.8 % agarose gel containing 1 x TAE and 0.1 % SDS. The gel was run in running buffer (1 x TAE, 0.1 % SDS) at 50 V, followed by blotting onto Hybond-C membrane (Amersham Biosciences), as described in Halfmann and Lindquist (2008).

### *SUP35*-based prion assay

We used *SUP35C* or *SUP35MC* tagging plasmids (see section 'cloning procedures and vector construction' for details) to generate an array of cPrD-Sup35C or cPrD-Sup35MC-expressing strains (see table S4 for a list of the strains). First, the strain YSR100 was transformed with a *cPrD-SUP35C* or a *cPrD-SUP35MC* expression plasmid. Then, a plasmid shuffle was performed by plating the transformants on 5-FOA. This produced a strain that expressed a particular cPrD-Sup35C/MC fusion protein as the only source of functional Sup35p. All strains were examined by Western blotting and SDD-AGE to evaluate expression and aggregation levels of the fusion proteins (see Figure S4, S6 and S7).

To induce the prion state, the *cPrD-SUP35C* strains were transformed with a corresponding pAG424GAL-cPrD-EYFP expression plasmid. The transformants were grown in SRafGal-Trp medium for 24 hours and then plated on YPD and SD-Ade plates with a cell number of 200 and 50,000 per plate, respectively. The same strains grown in SRaf-Trp served as a control. We compared the number of $Ade^+$ colonies on plates with cells grown in raffinose/galactose-containing medium to plates with cells grown in raffinose. A greater number of colonies under inducing conditions suggested that expression of cPrD-EYFP induced a prion switch. In these cases $Ade^+$ colonies were re-streaked on YPD plates and analyzed by SDD-AGE and curing on GdnHCl plates.

## Fiber transformation

Stock Mot3PrD protein (from expression of MOT3 cPrD in pRH2) in resuspension buffer was diluted to 10 μM in 1 ml assembly buffer and rotated end-to-end at 20 rpm for 3 days. The protein was confirmed to have converted to SDS-resistant aggregates by SDS-PAGE analysis of samples incubated for ten minutes in sample buffer at either 23°C or 95°C, and further confirmed to have a fibrillar appearance by transmission electron microscopy. Transformation of *dan1::URA3* yeast strains was performed according to (Tanaka and Weissman, 2006), except that DNA was omitted and URA+ ([*MOT3*+]) spheroplasts were selected directly on SD-ura plates containing 1 M sorbitol.

## SUPPLEMENTAL REFERENCES


Alberti, S., Gitler, A.D., and Lindquist, S. (2007). A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. Yeast *24*, 913-919.

Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., and Cullin, C. (1993). A simple and efficient method for direct gene deletion in Saccharomyces cerevisiae. Nucleic acids research *21*, 3329-3330.

Boye-Harnasch, M., and Cullin, C. (2006). A novel in vitro filter trap assay identifies tannic acid as an amyloid aggregation inducer for HET-s. Journal of biotechnology *125*, 222-230.

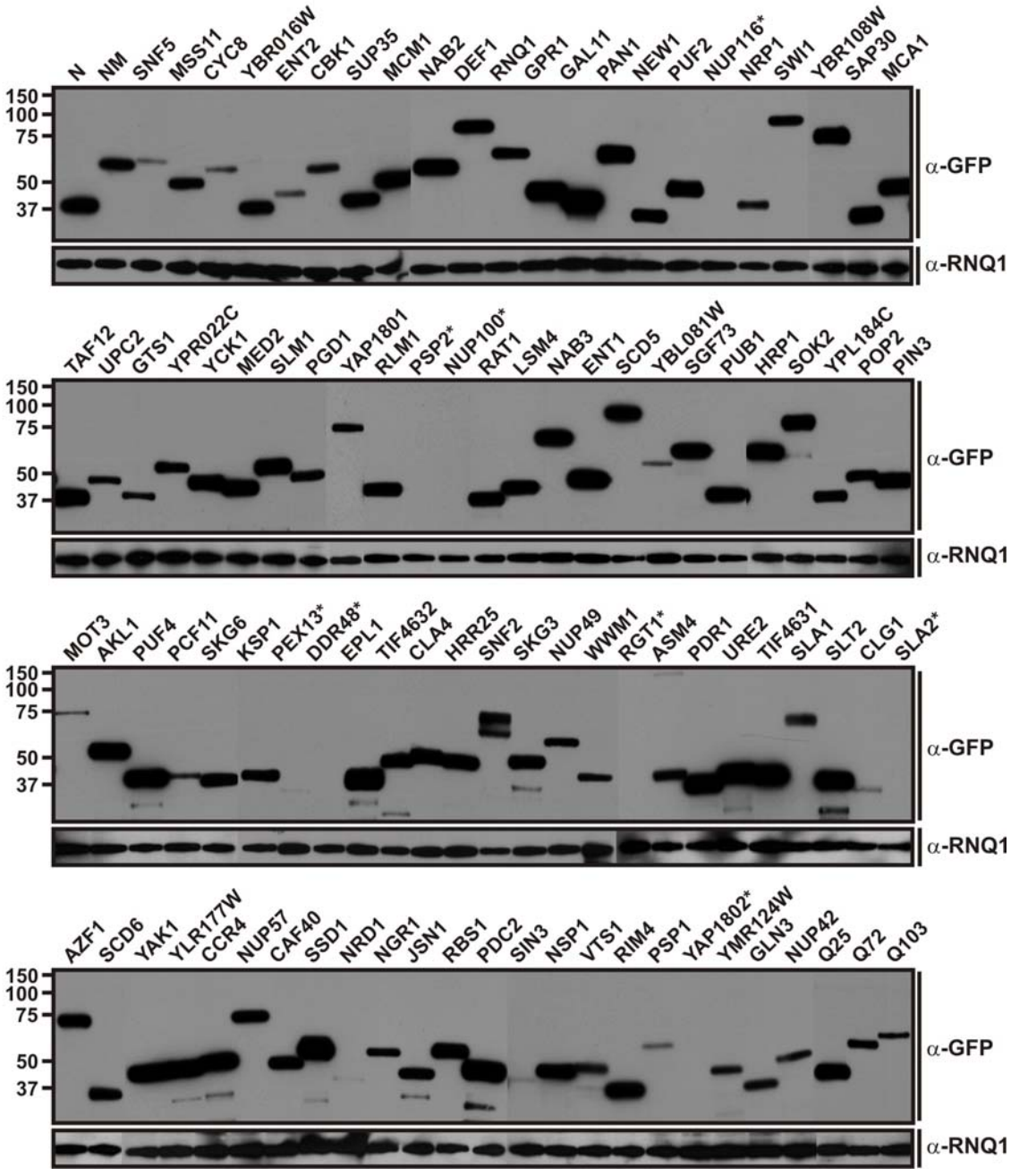
Edskes, H.K., Gray, V.T., and Wickner, R.B. (1999). The [URE3] prion is an aggregated form of Ure2p that can be cured by overexpression of Ure2p fragments. Proceedings of the National Academy of Sciences of the United States of America *96*, 1498-1503.

Gietz, D., St Jean, A., Woods, R.A., and Schiestl, R.H. (1992). Improved method for high efficiency transformation of intact yeast cells. Nucleic acids research *20*, 1425.

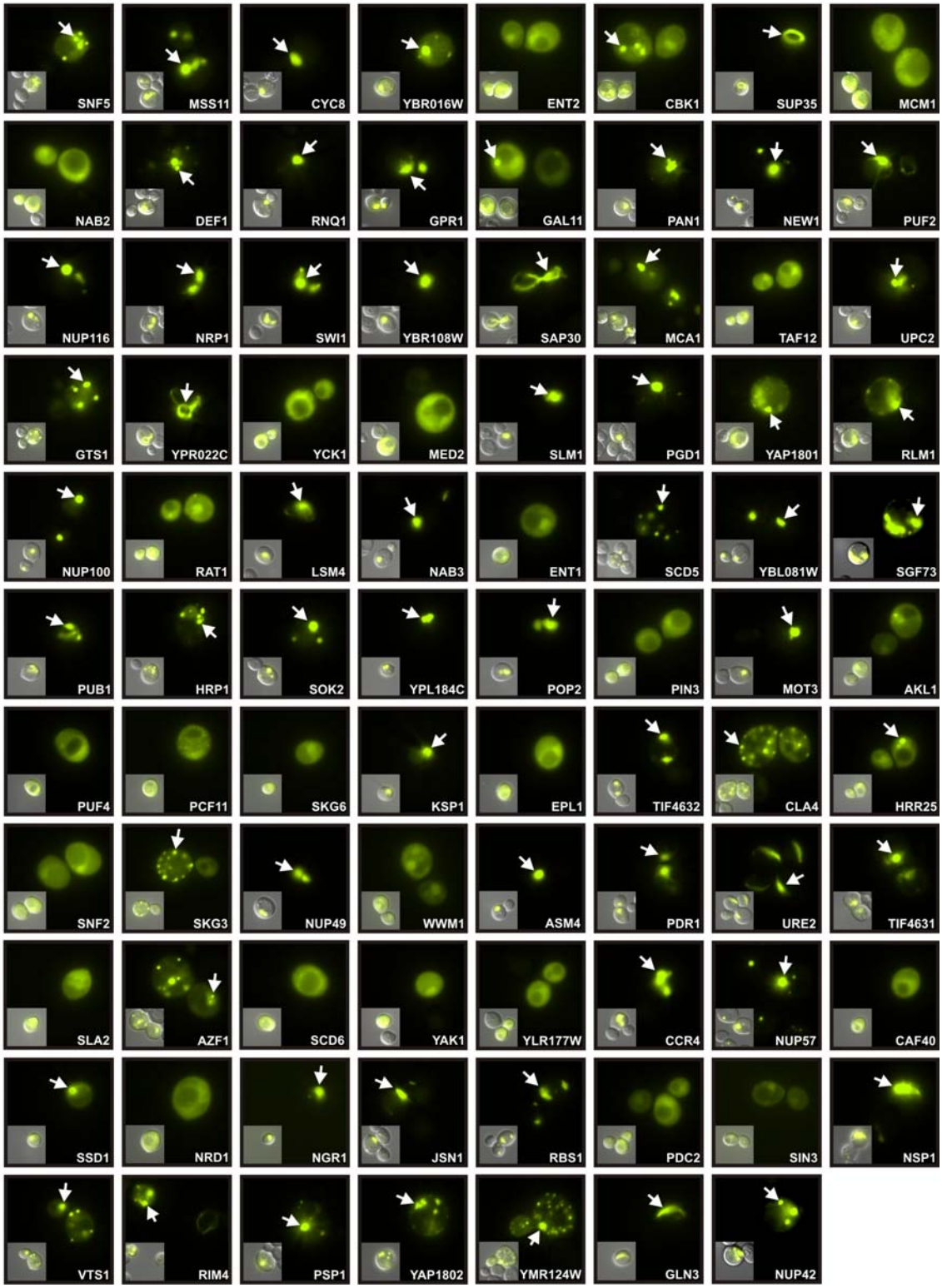
Gietz, R.D., Schiestl, R.H., Willems, A.R., and Woods, R.A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast *11*, 355-360.

Goldstein, A.L., and McCusker, J.H. (1999). Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. Yeast *15*, 1541-1553.
Gueldener, U., Heinisch, J., Koehler, G.J., Voss, D., and Hegemann, J.H. (2002). A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. Nucleic acids research *30*, e23.
Halfmann, R., and Lindquist, S. (2008). Screening for amyloid aggregation by semi-denaturing detergent-agarose gel electrophoresis. J Vis Exp.
King, C.Y., and Diaz-Avalos, R. (2004). Protein-only transmission of three yeast prion strains. Nature *428*, 319-323.
Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of molecular biology *157*, 105-132.
Li, L., and Lindquist, S. (2000). Creating a protein-based element of inheritance. Science (New York, NY *287*, 661-664.
Liu, J.J., Sondheimer, N., and Lindquist, S.L. (2002). Changes in the middle region of Sup35 profoundly alter the nature of epigenetic inheritance for the yeast prion [PSI+]. Proceedings of the National Academy of Sciences of the United States of America *99 Suppl 4*, 16446-16453.
Masison, D.C., and Wickner, R.B. (1995). Prion-inducing domain of yeast Ure2p and protease resistance of Ure2p in prion-containing cells. Science (New York, NY *270*, 93-95.
Mumberg, D., Muller, R., and Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. Gene *156*, 119-122.
Osherovich, L.Z., Cox, B.S., Tuite, M.F., and Weissman, J.S. (2004). Dissection and design of yeast prions. PLoS Biol *2*, E86.
Osherovich, L.Z., and Weissman, J.S. (2001). Multiple Gln/Asn-rich prion domains confer susceptibility to induction of the yeast [PSI(+)] prion. Cell *106*, 183-194.
Santoso, A., Chien, P., Osherovich, L.Z., and Weissman, J.S. (2000). Molecular basis of a yeast prion species barrier. Cell *100*, 277-288.
Sondheimer, N., and Lindquist, S. (2000). Rnq1: an epigenetic modifier of protein function in yeast. Mol Cell *5*, 163-172.
Tanaka, M., and Weissman, J.S. (2006). An efficient protein transformation protocol for introducing prions into yeast. Methods in enzymology *412*, 185-200.
Wach, A., Brachat, A., Alberti-Segui, C., Rebischung, C., and Philippsen, P. (1997). Heterologous HIS3 marker and GFP reporter modules for PCR-targeting in Saccharomyces cerevisiae. Yeast *13*, 1065-1075.
Wach, A., Brachat, A., Pohlmann, R., and Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast *10*, 1793-1808.

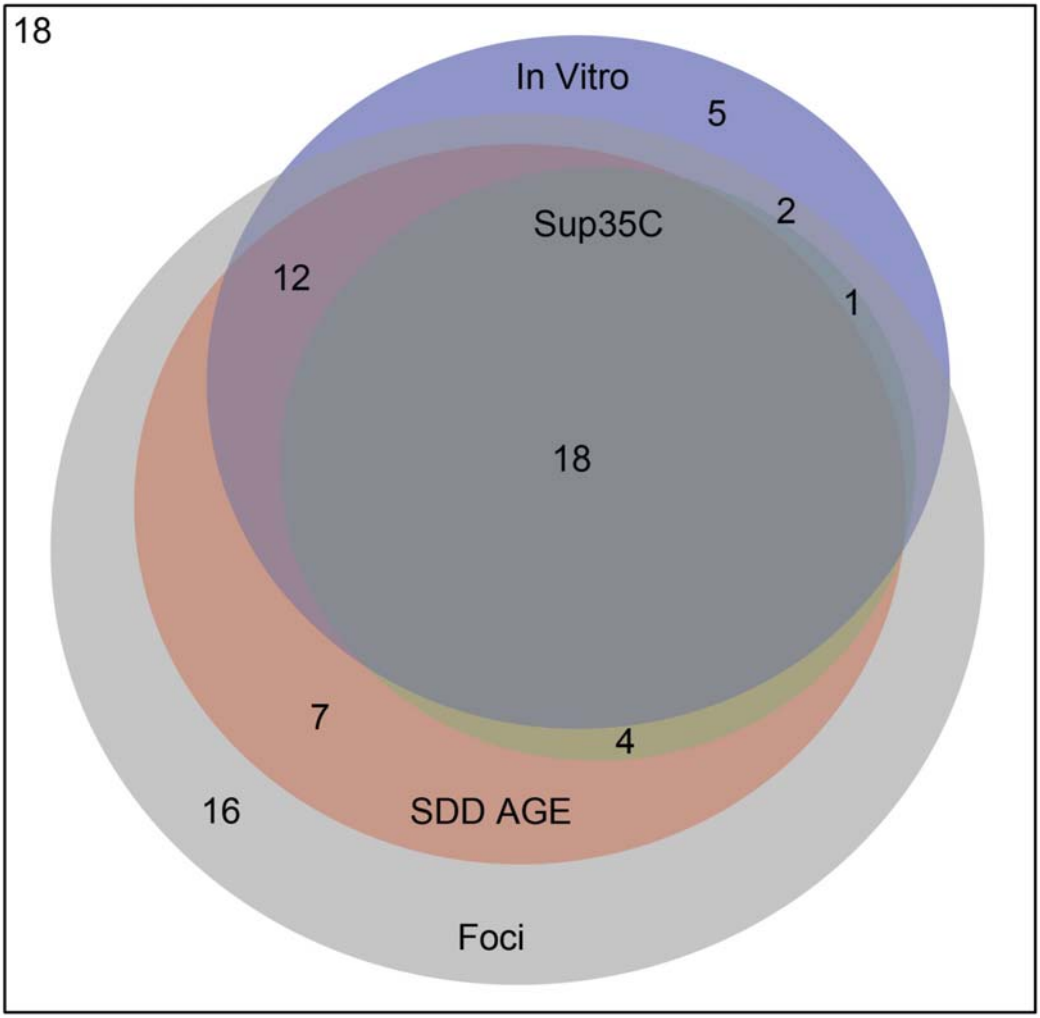
**Supplemental Figure S1**



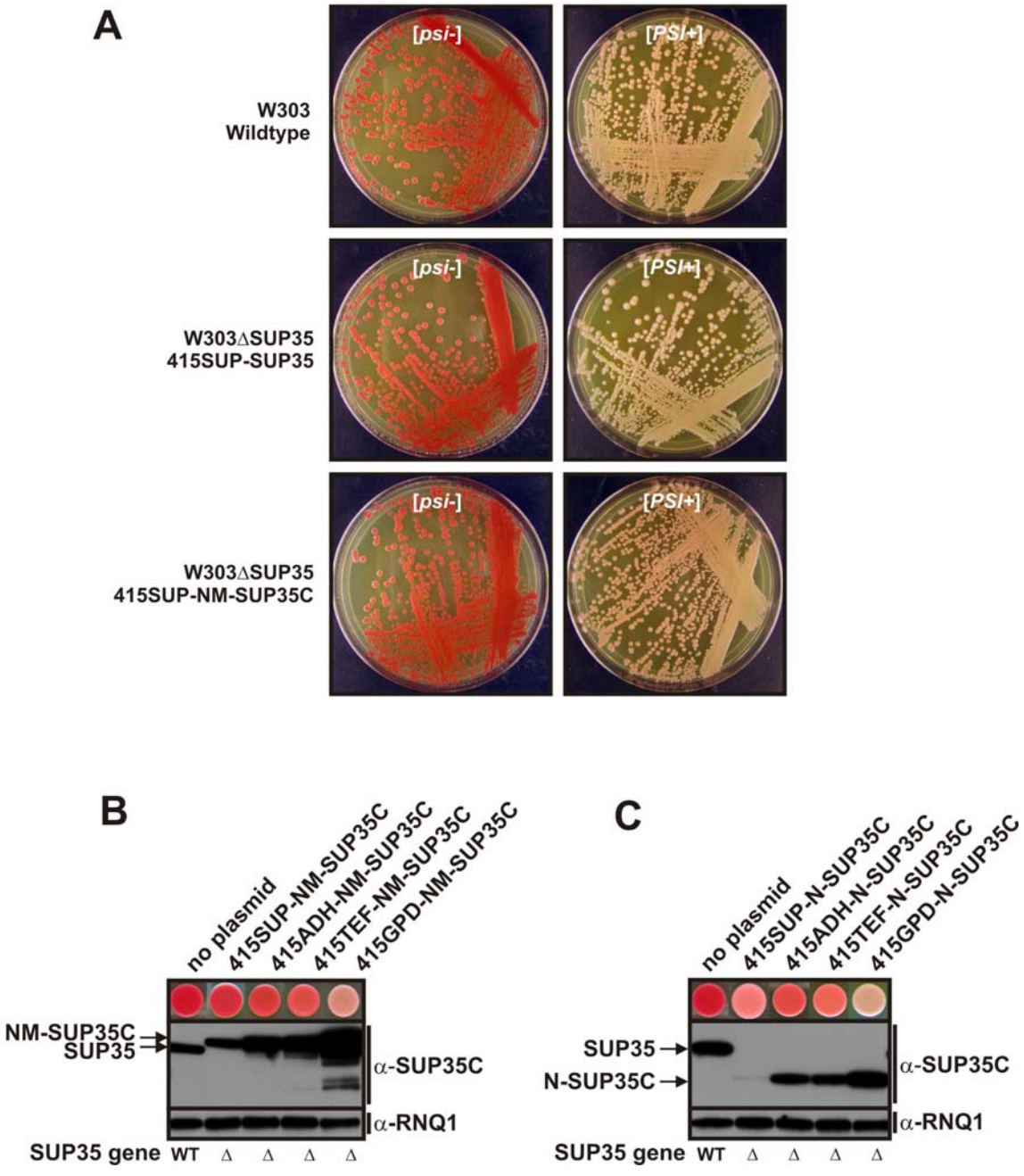
**Supplemental Figure S2**



**Supplemental Figure S3**

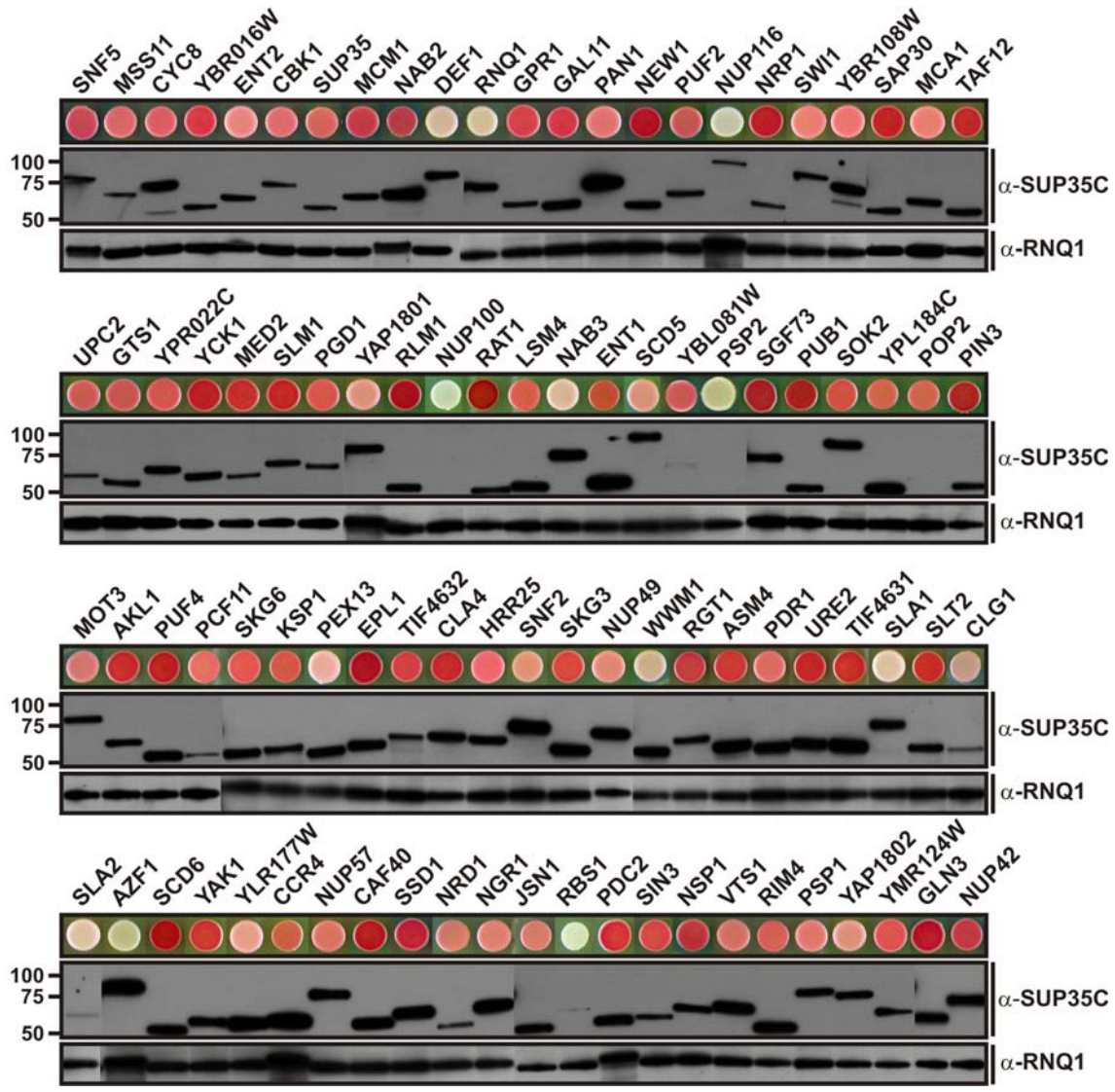


**Supplemental Figure S4**

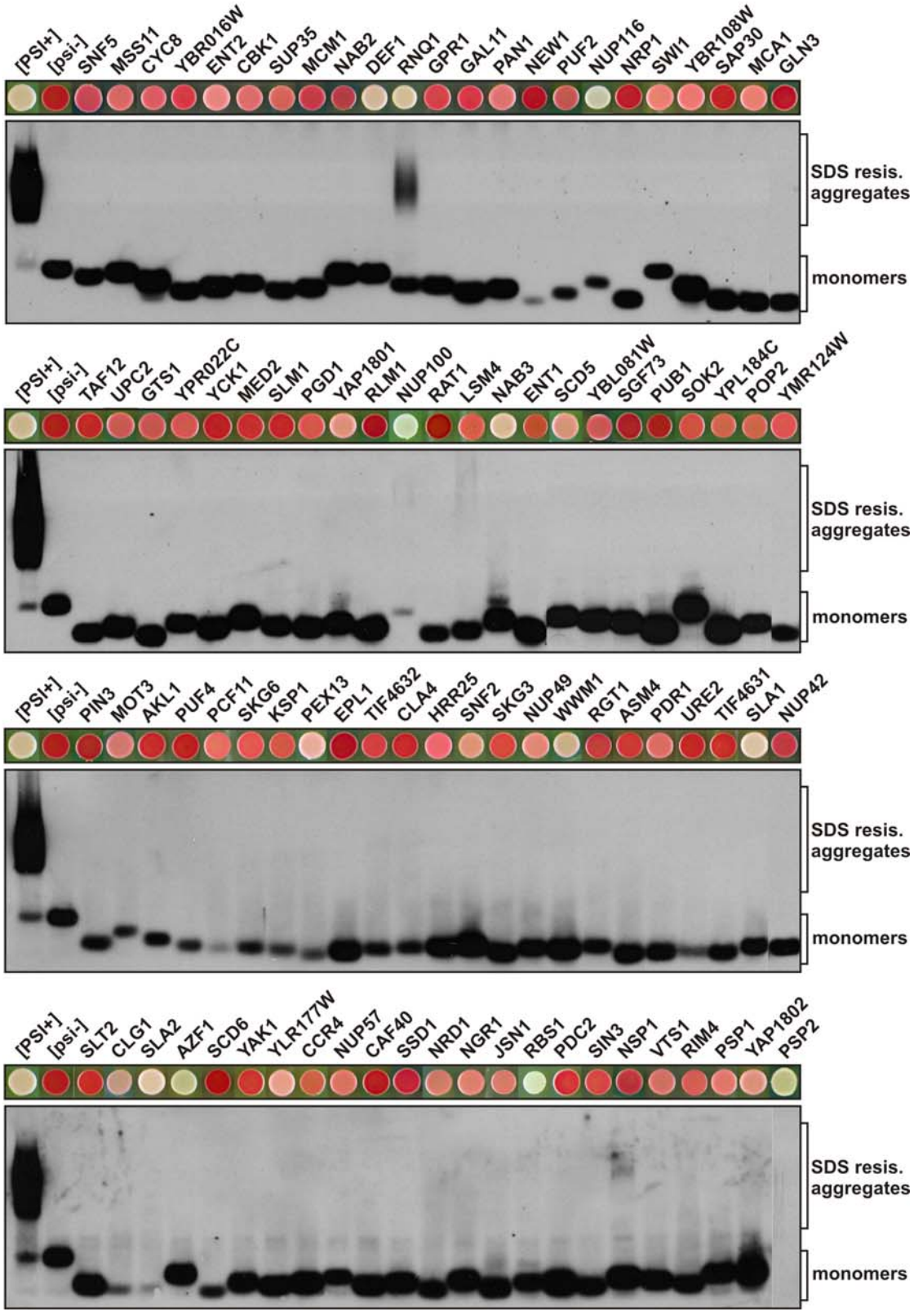


**Supplemental Figure S5**

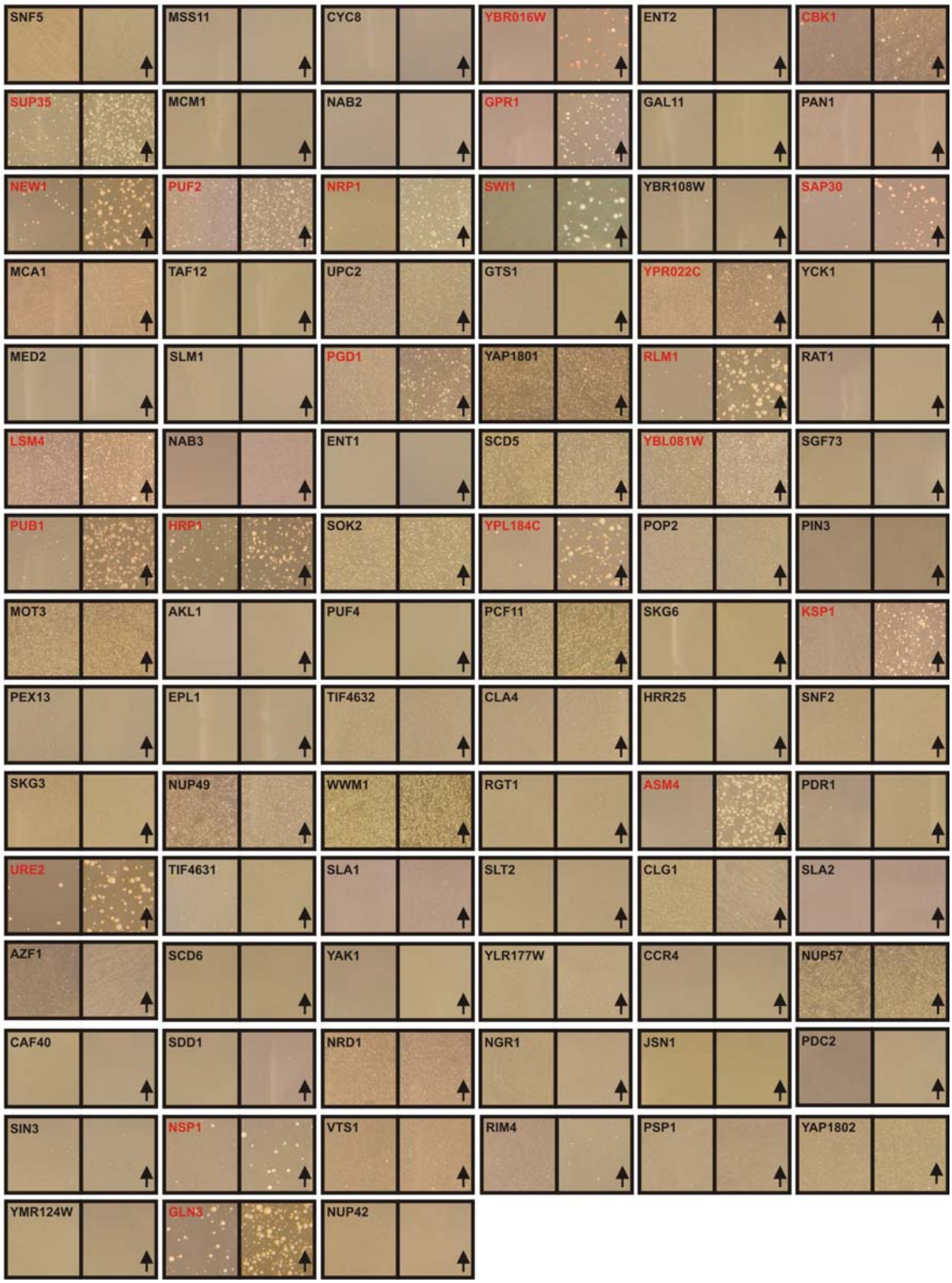
**Supplemental Figure S6**



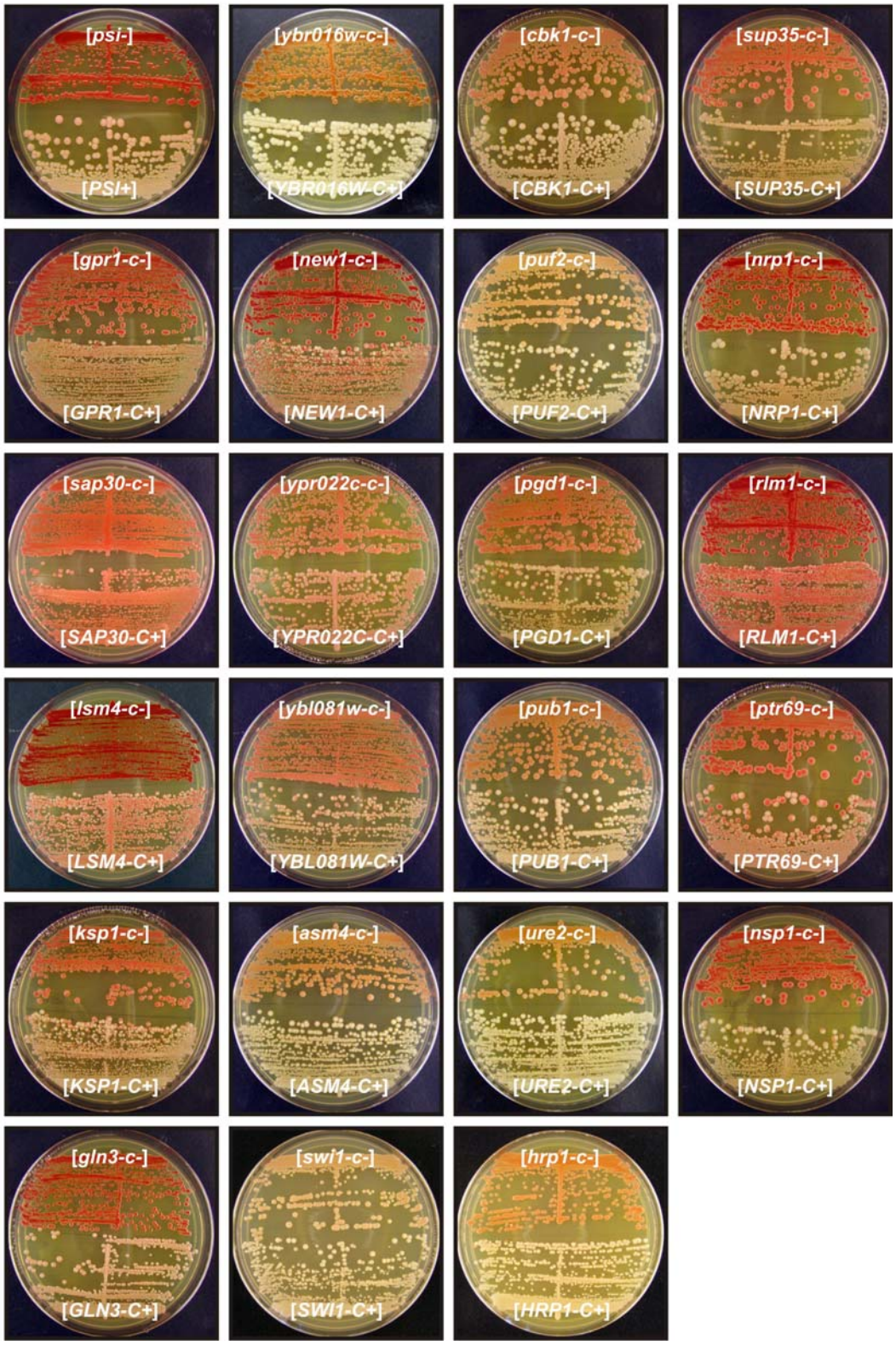
**Supplemental Figure S7**



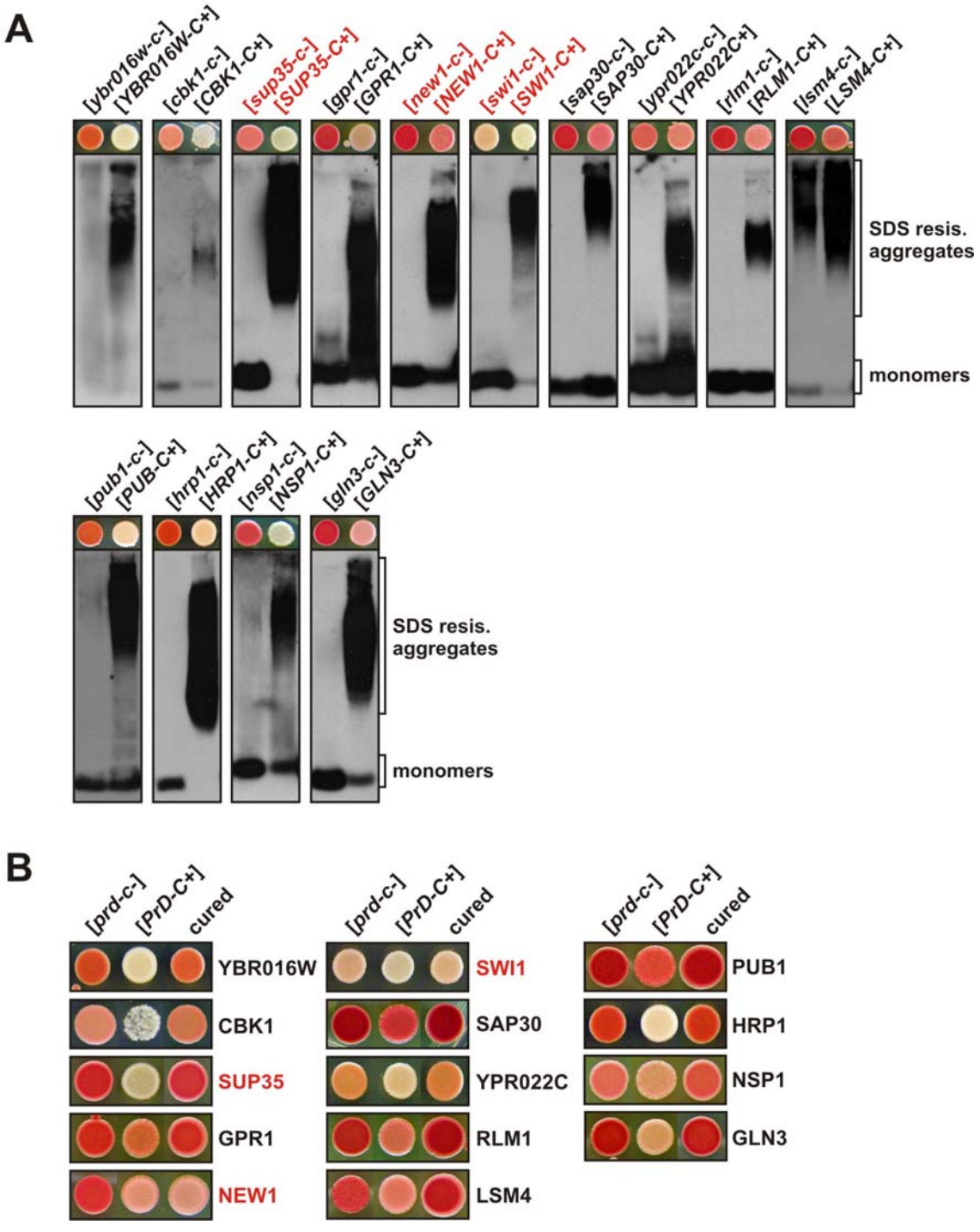
**Supplemental Figure S8**



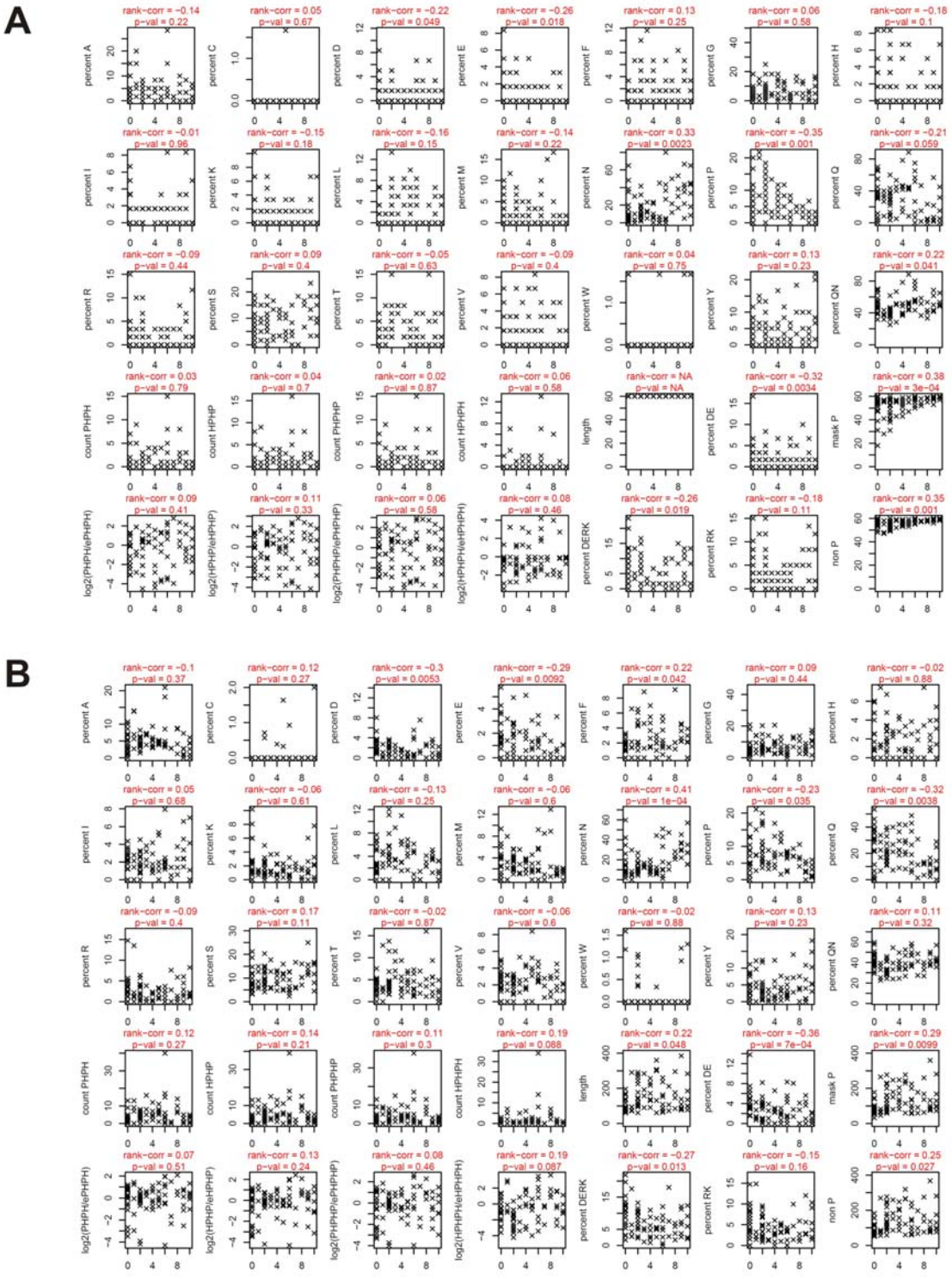
**Supplemental Figure S9**



**Supplemental Figure S10**



**Supplemental Figure S11**



| *Ranking* | *Gene name* | *Fluorescent foci* | *SDD-AGE (48 h ind.)* | *SUP35C* assay | *In vitro* |
|---|---|---|---|---|---|
| 1 | SNF5 | ✓ | ++ | - | + |
| 2 | MSS11 | ✓ | ++ | - | - |
| 3 | CYC8 | ✓ | +++ | - | + |
| 4 | YBR016W | ✓ | + | ✓ | ++ |
| 5 | ENT2 | - | - | - | - |
| 6 | CBK1 | ✓ | + | ✓ | + |
| 7 | SUP35 | ✓ | +++ | ✓ | +++ |
| 8 | MCM1 | - | - | - | - |
| 9 | NAB2 | - | - | - | - |
| 10 | DEF1 | ✓ | + | N/A | + |
| 11 | *YIL130W* | N/A | N/A | N/A | N/A |
| 12 | RNQ1 | ✓ | +++ | ✓ | +++ |
| 13 | *IXR1* | N/A | N/A | N/A | N/A |
| 14 | GPR1 | ✓ | ++ | ✓ | - |
| 15 | GAL11 | ✓ | + | - | - |
| 16 | PAN1 | ✓ | + | - | + |
| 17 | NEW1 | ✓ | +++ | ✓ | +++ |
| 18 | PUF2 | ✓ | +++ | ✓ | ++ |
| 19 | NUP116 | ✓ | ++ | N/A | + |
| 20 | NRP1 | ✓ | +++ | ✓ | +++ |
| 21 | SWI1 | ✓ | ++ | ✓ | +++ |
| 22 | *YEL007W* | N/A | N/A | N/A | N/A |
| 23 | YBR108W | ✓ | ++ | - | + |
| 24 | SAP30 | ✓ | + | ✓ | - |
| 25 | MCA1 | ✓ | - | - | ++ |
| 26 | TAF12 | - | - | - | - |
| 27 | UPC2 | ✓ | - | - | - |
| 28 | *YDL012C* | N/A | N/A | N/A | N/A |
| 29 | GTS1 | ✓ | +++ | - | + |
| 30 | YPR022C | ✓ | + | ✓ | + |
| 31 | YCK1 | - | - | - | - |
| 32 | MED2 | - | - | - | - |
| 33 | SLM1 | ✓ | + | - | - |
| 34 | PGD1 | ✓ | ++ | ✓ | - |
| 35 | YAP1801 | ✓ | - | - | + |
| 36 | RLM1 | ✓ | + | ✓ | ++ |
| 37 | PSP2 | N/A | N/A | N/A | N/A |
| 38 | NUP100 | ✓ | ++ | N/A | + |
| 39 | RAT1 | - | - | - | + |
| 40 | LSM4 | ✓ | +++ | ✓ | +++ |
| 41 | NAB3 | ✓ | - | - | - |
| 42 | ENT1 | ✓ | ++ | - | - |
| 43 | SCD5 | ✓ | ++ | - | + |
| 44 | YBL081W | ✓ | +++ | ✓ | +++ |
| 45 | SGF73 | ✓ | + | - | + |
| 46 | PUB1 | ✓ | ++ | ✓ | ++ |
| 47 | HRP1 | ✓ | - | ✓ | +++ |
| 48 | *YNL208W* | N/A | N/A | N/A | N/A |
| 49 | SOK2 | ✓ | + + | - | + |
| 50 | YPL184C | ✓ | + + | ✓ | - |

Supplemental Table S2, part 1

| *Ranking* | *Gene name* | *Fluorescent foci* | *SDD-AGE (48 h ind.)* | *SUP35C assay* | *In vitro* |
|---|---|---|---|---|---|
| 51 | POP2 | ✓ | +++ | - | - |
| 52 | PIN3 | - | - | - | + |
| 53 | MOT3 | ✓ | +++ | - | +++ |
| 54 | AKL1 | - | - | - | - |
| 55 | PUF4 | - | - | - | - |
| 56 | PCF11 | - | - | - | - |
| 57 | SKG6 | - | - | - | - |
| 58 | KSP1 | ✓ | ++ | ✓ | +++ |
| 59 | PEX13 | N/A | - | - | - |
| 60 | DDR48 | N/A | N/A | N/A | N/A |
| 61 | EPL1 | - | - | - | - |
| 62 | TIF4632 | ✓ | - | - | - |
| 63 | CLA4 | ✓ | - | - | - |
| 64 | HRR25 | ✓ | - | - | - |
| 65 | SNF2 | - | - | - | - |
| 66 | SKG3 | ✓ | - | - | - |
| 67 | NUP49 | ✓ | ++ | - | - |
| 68 | WWM1 | - | - | - | + |
| 69 | RGT1 | N/A | N/A | - | - |
| 70 | ASM4 | ✓ | +++ | ✓ | ++ |
| 71 | PDR1 | ✓ | + | - | +++ |
| 72 | URE2 | ✓ | ++ | ✓ | +++ |
| 73 | TIF4631 | ✓ | - | - | - |
| 74 | SLA1 | - | - | - | + |
| 75 | SLT2 | ✓ | - | - | - |
| 76 | CLG1 | N/A | N/A | - | + |
| 77 | SLA2 | N/A | N/A | - | - |
| 78 | AZF1 | ✓ | - | - | - |
| 79 | SCD6 | - | - | - | - |
| 80 | YAK1 | - | - | - | - |
| 81 | YLR177W | - | - | - | - |
| 82 | CCR4 | ✓ | - | - | - |
| 83 | NUP57 | ✓ | - | - | - |
| 84 | CAF40 | - | - | - | - |
| 85 | SSD1 | ✓ | - | - | - |
| 86 | NRD1 | - | - | - | - |
| 87 | NGR1 | ✓ | +++ | - | +++ |
| 88 | JSN1 | ✓ | + | - | + |
| 89 | RBS1 | ✓ | + | N/A | +++ |
| 90 | PDC2 | - | - | - | - |
| 91 | SIN3 | - | - | - | + |
| 92 | NSP1 | ✓ | +++ | ✓ | + |
| 93 | VTS1 | ✓ | - | - | - |
| 94 | RIM4 | ✓ | + | - | N/A |
| 95 | PSP1 | ✓ | - | - | - |
| 96 | YAP1802 | ✓ | - | - | - |
| 97 | YMR124W | ✓ | - | - | - |
| 98 | NPL3 | N/A | N/A | N/A | N/A |
| 99 | GLN3 | ✓ | ++ | ✓ | +++ |
| 100 | NUP42 | ✓ | + | - | - |

Supplemental Table S2, part 2

**Supplemental Table S5**

**Median amino acid frequencies of cPrDs**

| AA | Fl. foci + | Fl. foci - | SDD-AGE (24 h) + | SDD-AGE (24 h) - | SDD-AGE (48 h) + | SDD-AGE (48 h) - | SUP35C assay + | SUP35C assay - | *In vitro* + | *In vitro* - |
|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.042 | 0.042 | 0.042 | 0.04 | 0.042 | 0.041 | 0.028 | 0.043 | 0.041 | 0.042 |
| C | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D | 0.009 | 0.017 | 0.007 | 0.016 | 0.007 | 0.017 | 0.009 | 0.015 | 0.01 | 0.017 |
| E | 0.007 | 0.014 | 0.003 | 0.015 | 0.006 | 0.014 | 0 | 0.012 | 0.009 | 0.011 |
| F | 0.022 | 0.014 | 0.025 | 0.017 | 0.023 | 0.017 | 0.022 | 0.021 | 0.022 | 0.019 |
| G | 0.069 | 0.053 | 0.069 | 0.064 | 0.069 | 0.06 | 0.072 | 0.061 | 0.069 | 0.06 |
| H | 0.016 | 0.011 | 0.01 | 0.016 | 0.01 | 0.016 | 0.009 | 0.014 | 0.012 | 0.015 |
| I | 0.015 | 0.021 | 0.015 | 0.017 | 0.015 | 0.017 | 0.015 | 0.016 | 0.017 | 0.017 |
| K | 0.014 | 0.013 | 0.014 | 0.013 | 0.014 | 0.012 | 0.016 | 0.014 | 0.012 | 0.016 |
| L | 0.036 | 0.024 | 0.032 | 0.034 | 0.033 | 0.033 | 0.023 | 0.039 | 0.03 | 0.039 |
| M | 0.02 | 0.021 | 0.018 | 0.024 | 0.018 | 0.026 | 0.015 | 0.021 | 0.018 | 0.024 |
| N | 0.153 | 0.09 | 0.198 | 0.103 | 0.178 | 0.102 | 0.282 | 0.11 | 0.177 | 0.108 |
| P | 0.068 | 0.071 | 0.055 | 0.073 | 0.062 | 0.074 | 0.044 | 0.073 | 0.065 | 0.063 |
| Q | 0.165 | 0.287 | 0.127 | 0.224 | 0.142 | 0.217 | 0.112 | 0.214 | 0.154 | 0.22 |
| R | 0.018 | 0.022 | 0.013 | 0.024 | 0.014 | 0.023 | 0.02 | 0.019 | 0.014 | 0.025 |
| S | 0.103 | 0.08 | 0.114 | 0.085 | 0.112 | 0.082 | 0.122 | 0.094 | 0.111 | 0.09 |
| T | 0.039 | 0.032 | 0.039 | 0.035 | 0.039 | 0.034 | 0.027 | 0.035 | 0.033 | 0.038 |
| V | 0.023 | 0.022 | 0.022 | 0.023 | 0.023 | 0.022 | 0.019 | 0.023 | 0.024 | 0.022 |
| W | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Y | 0.033 | 0.037 | 0.035 | 0.035 | 0.033 | 0.035 | 0.039 | 0.034 | 0.047 | 0.028 |
| QN | 0.383 | 0.397 | 0.4 | 0.362 | 0.389 | 0.352 | 0.411 | 0.352 | 0.385 | 0.386 |

**Median amino acid frequencies of cPrD core**

| AA | Fl. foci + | Fl. foci - | SDD-AGE (24 h) + | SDD-AGE (24 h) - | SDD-AGE (48 h) + | SDD-AGE (48 h) - | SUP35C assay + | SUP35C assay - | *In vitro* + | *In vitro* - |
|---|---|---|---|---|---|---|---|---|---|---|
| A | 0.033 | 0.033 | 0.033 | 0.033 | 0.017 | 0.033 | 0.017 | 0.033 | 0.033 | 0.033 |
| C | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D | 0 | 0.017 | 0 | 0.017 | 0 | 0.017 | 0 | 0 | 0 | 0.017 |
| E | 0 | 0.017 | 0 | 0.017 | 0 | 0.017 | 0 | 0 | 0 | 0 |
| F | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 |
| G | 0.05 | 0.033 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| H | 0 | 0.017 | 0 | 0.017 | 0 | 0.017 | 0 | 0.017 | 0 | 0.017 |
| I | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 |
| K | 0 | 0.017 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| L | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0 | 0.017 | 0 | 0.033 |
| M | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0 | 0.017 | 0.017 | 0.017 |
| N | 0.15 | 0.083 | 0.233 | 0.1 | 0.217 | 0.1 | 0.35 | 0.1 | 0.175 | 0.1 |
| P | 0.05 | 0.05 | 0.033 | 0.05 | 0.033 | 0.067 | 0.033 | 0.05 | 0.05 | 0.05 |
| Q | 0.267 | 0.333 | 0.183 | 0.317 | 0.233 | 0.317 | 0.067 | 0.317 | 0.225 | 0.317 |
| R | 0.017 | 0.017 | 0 | 0.017 | 0 | 0.017 | 0.017 | 0.017 | 0.008 | 0.017 |
| S | 0.083 | 0.067 | 0.1 | 0.067 | 0.083 | 0.067 | 0.1 | 0.067 | 0.083 | 0.075 |
| T | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 | 0.017 |
| V | 0.017 | 0.017 | 0 | 0.017 | 0.017 | 0.017 | 0 | 0.017 | 0.017 | 0.017 |
| W | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Y | 0.033 | 0.033 | 0.033 | 0.033 | 0.017 | 0.033 | 0.05 | 0.033 | 0.05 | 0.017 |
| QN | 0.483 | 0.45 | 0.5 | 0.433 | 0.517 | 0.417 | 0.5 | 0.433 | 0.483 | 0.45 |