ACS SyntheticBiology

pubs.acs.org/synthbio

Letter

A Reporter System for Cytosolic Protein Aggregates in Yeast

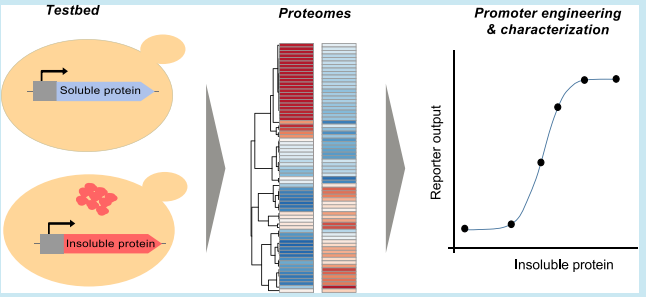
David Romero-Suarez, Tune Wulff, Yixin Rong, Tadas Jakočiūnas, Satoshi Yuzawa, Jay D. Keasling, and Michael K. Jensen*

Cite This: *ACS Synth. Biol.* 2021, 10, 466−477

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Protein misfolding and aggregation are linked to neurodegenerative diseases of mammals and suboptimal protein expression within biotechnology. Tools for monitoring protein aggregates are therefore useful for studying disease-related aggregation and for improving soluble protein expression in heterologous hosts for biotechnology purposes. In this work, we developed a promoter−reporter system for aggregated protein on the basis of the yeast native response to misfolded protein. To this end, we first studied the proteome of yeast in response to the expression of folded soluble and aggregation-prone protein baits and identified genes encoding proteins related to protein folding and the response to heat stress as well as the ubiquitin-proteasome system that are over-represented in cells expressing an aggregation-prone protein. From these data, we created and validated promoter−reporter constructs and further engineered the best performing promoters by increasing the copy number of upstream activating sequences and optimization of culture conditions. Our best promoter−reporter has an output dynamic range of approximately 12-fold upon expression of the aggregation-prone protein and responded to increasing levels of aggregated protein. Finally, we demonstrate that the system can discriminate between yeast cells expressing different prion precursor proteins and select the cells expressing folded soluble protein from mixed populations. Our reporter system is thus a simple tool for diagnosing protein aggregates in living cells and should be applicable for the health and biotechnology industries.



KEYWORDS: *protein aggregates, misfolding, yeast, biosensor, promoter engineering, proteome*

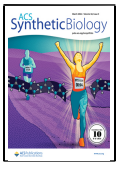
INTRODUCTION

Proteins need to fold into their native conformations to achieve their functions, as misfolding results in impaired protein function. Misfolding also has potential toxic effects for cells. Partially folded and misfolded proteins often expose hydrophobic residues that are otherwise buried in folded proteins, leading to the formation of potentially toxic aggregates.[1,2] It has been proposed that the toxicity of protein aggregates arises from aberrant interactions with the cell proteome[3] or the impairment of membrane structural integrity,[4] and the accumulation of protein aggregates has been associated with the pathogenesis of Alzheimer's, Parkinson's, prion-related diseases, and aging.[2,5]

Besides misfolding and aggregation observed in the context of disease, proteins can misfold and aggregate as a consequence of changes in their expression levels and expression host as well as expression conditions (*e.g.*, pH, temperature, exposure to chemicals).[6] This is particularly relevant in the context of protein production,[8] the directed evolution of new protein variants,[7,9] and the construction of cell factories for biotechnological purposes,[10] as these endeavors often involve expression of proteins under non-native and potentially suboptimal conditions (i.e heterologous expression, mutant variants of proteins, high expression). In addition, protein overexpression can result in metabolic burden[11−13] by draining resources from other cellular processes as well as causing cellular stress responses.[14,15] It is therefore of interest to develop methods to assess misfolding, aggregation, and cell stress responses to these processes for applied and basic research purposes.

Several methods have been developed in *E. coli*, mainly to screen proteins with enhanced stability, solubility, and expression in this host.[16−20] In eukaryotes, the focus has been on studying proteins in the context of disease. Tools using fusions of proteins of interest to engineered transcription factors have been used to study aggregation-prone proteins.[21,22] Similarly, tripartite fusions, proteins of interest inserted in antibiotic resistance markers, have been shown to be useful tools to test protein stability and aggregation propensities.[23] The stress response produced by the heterologous production of metabolites has been leveraged to develop biosensors and promoter−reporters,[24−26] an approach used to develop a

Received: September 16, 2020
Published: February 12, 2021



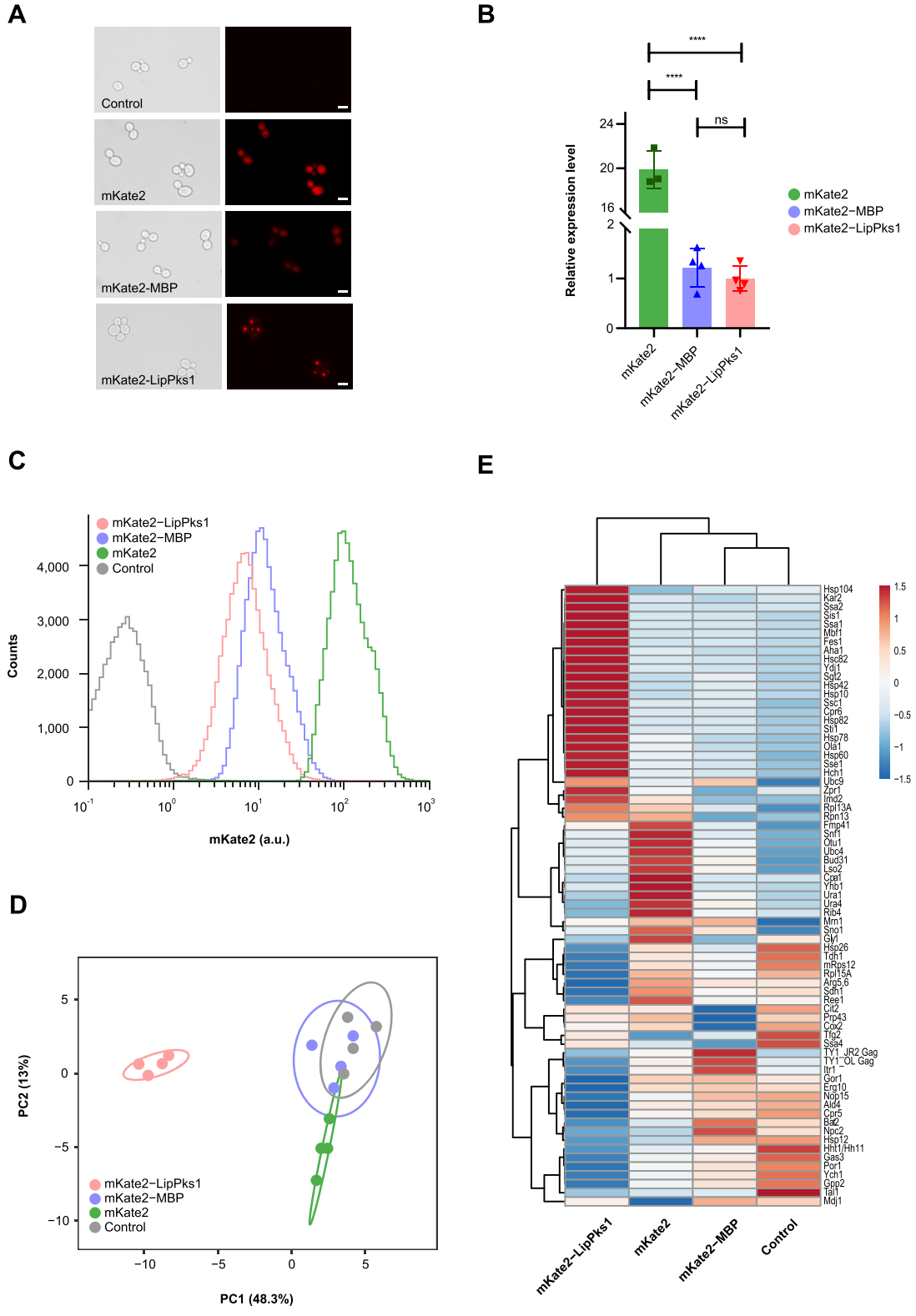


Figure 1. Expression of protein baits in yeast cytosol. (A) Microscopy images of strains expressing protein baits mKate2, mKate2-MBP, and mKate2-LipPks1 and the control strain (no heterologous protein). Cultures were grown in SC medium. Images were taken with different exposure times; thus, fluorescence is not directly comparable. (B) Relative expression levels of protein baits as compared to synthetic peptides specific for each of the three protein baits. $n = 3–4$. Bars represent the mean and error bars, the standard deviation of all data points. Asterisks represent the statistical significance between conditions indicated by horizontal bar ends: **** = $p < 0.0001$, *** = $p < 0.001$, ** = $p < 0.001$, * = $p < 0.05$, and ns = not significant ($p > 0.05$). (C) Flow cytometry measurements of strains expressing the bait proteins mKate2, mKate2-MBP, and mKate2-LipPks1 and the control strain. $n = 6 \times 10\,000$ events. (D) PCA analysis of differentially expressed proteins in strains expressing protein baits as well as the yeast strain CEN.PK2-1C *pep4*Δ *prb1*Δ (control). (E) Hierarchical clustering analysis of differentially expressed proteins in response to the expression of protein baits compared to wild-type yeast CEN.PK2-1C (control) ($n = 3–4$, $p < 0.01$).

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

promoter−reporter for insoluble protein in *E. coli*.[20] In this work, we developed an optimized promoter−reporter system that responds to cytosolic expression of aggregation-prone protein in yeast.

We developed the promoter−reporter for aggregated protein on the basis of the cytosolic unfolded protein response (cytosolic UPR) in yeast as inferred from proteome analysis of yeast expressing soluble and aggregation-prone proteins. From native promoter characterization, we further present the optimization of the dynamic output range as well as the characterization of the operational range of the promoter−reporter system and apply the promoter−reporter to sort cells expressing soluble proteins from mixed populations of cells expressing soluble and aggregation-prone proteins.

RESULTS

Cytosolic Unfolded Protein Response Is Triggered upon Expression of Aggregation-Prone Protein. As a first step toward developing promoter−reporters (hereafter reporters) for the expression of an aggregation-prone protein in the cytosol, we first sought to characterize the cellular responses at the proteome level of yeast expressing folded soluble or aggregation-prone proteins in the cytosol. To do so, we expressed the monomeric fluorescent protein mKate2 as well as tagged fusion proteins of maltose-binding protein (labeled here as mKate2-MBP) and the engineered lipomycin polyketide synthase L2(KR null)[27] (labeled here as mKate2-LipPks1). MBP is a widely used solubility tag that has been observed to improve the expression of fusion proteins in *E. coli*[28] and yeast.[29] Although functional polyketide synthases have been expressed successfully in standard heterologous hosts like *E. coli* and *S. cerevisiae*,[30,31] in some cases, PKSs are expressed as misfolded insoluble protein.[32,33] For our system-level analysis, all three protein baits were expressed from the strong glycolytic promoter *TDH3*p in a background strain with deletions of genes *PEP4* and *PRB1*, encoding vacuolar proteases known to improve heterologous protein production.[34] Fusions to fluorescent proteins have been widely used to assess expression patterns of folded and misfolded proteins,[35,36] and for this reason, we first used this approach to qualitatively assess the expression of mKate2, mKate2-MBP, and mKate2-LipPks1. For mKate2 and mKate2-MBP, fluorescence was observed throughout the cytosol, while in the case of mKate2-LipPks1, fluorescence was localized to distinct foci (Figure 1A), indicating protein aggregation as previously observed when expressing other misfolded proteins in yeast.[35,36] Targeted proteomics, using peptides specific for mKate2, MBP, and LipPks1 proteins, revealed similar expression levels of mKate2-MBP and mKate2-LipPks1 and higher expression levels of mKate2 (Figures 1B and S1). The expression of mKate2 was 19.9× higher than mKate2-LipPks1, while expression levels of mKate2-MBP and mKate2-LipPks1 showed no significant difference between them (Figure 1B). Furthermore, the fluorescence of mKate2 in the strains expressing mKate2-MBP and mKate2 was 1.8× and 16.8× higher, respectively, relative to mKate2-LipPks1 (Figure 1C), corroborating the expression levels obtained from targeted proteomics.

Next, we used untargeted proteomics to study the yeast proteome response to the expression of each of the three bait proteins, compared to a control strain without heterologous protein expression. When analyzing the proteomes from the four different strains, we detected a total of 1445 proteins (Supplementary Data File 1). Of those, 71 proteins were differentially expressed upon the expression of protein baits compared to the control strain ($p < 0.01$); 31 of them were upregulated in response to the mKate2-LipPks1 expression (Supplementary Data File 1). Hierarchical clustering and PCA analysis of the differentially expressed proteins showed that the strain expressing mKate2-LipPks1 had a distinct expression profile compared to the profiles observed for the control strain as well as for the strains expressing mKate2-MBP and mKate2 (Figure 1D,E). As for the strains expressing soluble proteins, the mKate2-MBP strain had a similar expression profile as that of the control strain, while the proteome of the mKate2 strain showed more variance between replicate samples, coinciding with less similarity to the control and mKate2-MBP strains (Figure 1D,E).

Furthermore, when gene ontology enrichment analysis of the genes encoding proteins upregulated in response to mKate2-LipPks1 compared to the control strain was performed, biological processes related to protein folding and responses to heat stress were over-represented (False Discovery Rate (FDR) $< 1 \times 10^{-3}$) (Figure S2, Supplementary Data File 1). Of the 31 upregulated proteins in response to mKate2-LipPks1 expression, 18 are chaperones and chaperone cofactors including chaperones of the families HSP90 (*e.g.*, Hsc82, Hsp82), HSP70 (*e.g.*, Ssa1, Ssa2), HSP40 (*e.g.*, Sis1, Ydj1), and HSP60 (Hsp60), small heat shock proteins (Hsp42), and cochaperones (*e.g.*, Aha1, Sti1).[37] We also observed that three of the proteins upregulated in response to mKate2-LipPks1 expression are associated with the ubiquitin-proteasome system (Rpn13, Ubc4, Ubc9)[38] (Supplementary Data File 1). Among the 31 proteins upregulated in response to mKate2-LipPks1, we observed 17/25 proteins that were observed previously to upregulate upon expression of misfolded insoluble protein in the cytosol.[15]

In summary, on the basis of the expression pattern in the cytoplasm, expression levels (Figure 1A−C), and the stress response observed by proteomics analysis upon heterologous protein expression (Figure 1D,E), our data highlight that mKate2-MBP and mKate2-LipPks1 are suitable baits to compare in order to further identify and engineer promoters sensitive to aggregated protein in the cytosol.

Characterizing Promoters as Reporters for Aggregated Protein. From the list of differentially expressed proteins, we selected 12 promoters from genes encoding the upregulated proteins that had the highest differential expression in response to mKate2-LipPks1 compared to mKate2-MBP (Figure S3). For all promoters, we cloned the region up to 1000 bp delimited by the start codon of the gene of interest and the closest upstream open reading frame, equaling promoter sizes ranging from 254 to 1000 bp (Table S5). Although Ubc4 and Ubc9 were among the proteins with the highest differential expression between mKate2-LipPks1 and control strain, they were not considered for further characterization as they did not show differential expression when comparing strains expressing mKate2-LipPks1 with strains expressing mKate2-MBP or mKate2 (Figure S3).

Among the native yeast promoters driving the expression of proteins upregulated upon mKate2-LipPks1 expression, we next looked for over-represented transcription factor binding sites (TFBSs). Here, we found that 28 out of the 31 upregulated promoters are regulated by the heat shock factor 1 (Hsf1) transcription factor (Table S7) and that Hsf1 binding sites (heat shock element, HSE) are over-represented in those promoters (Table S8). Hsf1 responds to protein folding stress and triggers

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

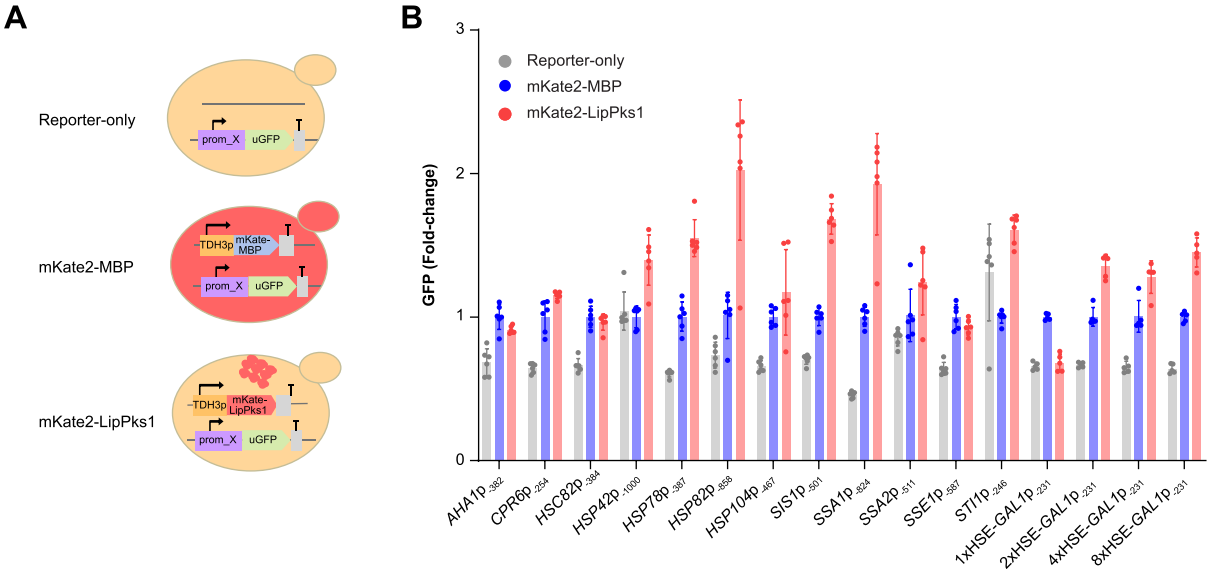


Figure 2. Promoter screening of native and synthetic promoters as agregated protein reporters. (A) Strains used for screening promoters as insoluble protein reporters. (B) Flow cytometry measurements of yeast strains harboring reporter constructs were grown for 6 h and assessed by flow cytometry (n = (5−6) × 10 000 events). For each promoter, the fold change was calculated by normalizing GFP geometric means for each replicate (10 000 events) to the GFP geometric mean of all mKate2-MBP replicates (50 000 or 60 000 events). Bars represent the mean and error bars, the standard deviation of all data points.

the expression of chaperones to restore protein homeostasis, proteostasis.[39] Under nonstress conditions, the Ssa1 and Ssa2 chaperones inhibit Hsf1-mediated gene expression by binding to Hsf1.[40] Upon the presence of misfolded protein, Ssa1 and Ssa2 chaperones are titrated away from Hsf1, which results in the activation of Hsf1 targets.[40] Furthermore, it has been observed that depleting the Hsf1 transcription factor from the nucleus leads to a decreased expression of chaperones and concomitant protein aggregation at physiological temperatures.[39] Hence, we hypothesized that, in addition to reporter systems derived from native promoters, synthetic promoters containing HSEs could be developed as reporters of protein aggregates in the cytosol. To this end, we constructed synthetic promoters by attaching 1, 2, 4, or 8 copies of HSEs to a minimal *GAL1*p promoter without Gal4 binding sites (*GAL1*p$_{-231}$). HSEs were placed 90 bp upstream of the *GAL1*p TATA box (Figure S4). We have previously used this location in the *GAL1*p to insert 6× *trpO* binding sites and successfully develop a tryptophan biosensor on the basis of the engineered transcription factor Gal4$_{AD}$-trpR.[41]

Next, all 12 native promoters and 4 synthetic HSE-*GAL1*p promoters were screened by fluorescence reporter assays, using the unstable green fluorescent protein variant yEGFP3-CLN2-(PEST) (labeled uGFP herein) as a reporter;[42] a GFP variant has been used to study promoter expression kinetics.[42] Reporter constructs were integrated in strains expressing mKate2-MBP and mKate2-LipPks1 and a control strain (Figure 2A). Initially, strains were assessed by flow cytometry during exponential growth (6 h of cultivation) (Figure S5). We observed that 13/16 promoters responded to mKate2-MBP (Figure 2B). We also observed that, for 10/16 screened promoters (*CPR6*p$_{-254}$, *HSP42*p$_{-1000}$, *HSP78*p$_{-387}$, HSP82p$_{-858}$, *SIS1*p$_{-501}$, *SSA1*p$_{-824}$, *STI1*p$_{-246}$, 2xHSE-*GAL1*p$_{-231}$, 4xHSE-*GAL1*p$_{-231}$, and 8xHSE-*GAL1*p$_{-231}$), the fluorescence output of their respective reporters when expressing mKate2-LipPks1 was higher than when expressing mKate2-MBP (p < 0.05), indicating such reporters were responsive to the expression of aggregation-prone protein. In order to select for the best candidate promoter responding to the expression of the aggregation-prone protein, we calculated the fold-change ratio of reporter fluorescence outputs in strains expressing mKate2-LipPks1 compared to strains expressing mKate2-MBP. Fold changes for the 11 reporters ranged from 1.2 to 2.0 (Figure 2B). The reporters with the highest fold changes based on the native promoters were *HSP82*p$_{-858}$:uGFP (2.0×), *SIS1*p$_{-501}$:uGFP (1.7×), and *SSA1*p$_{-824}$:uGFP (1.9×). As for the synthetic promoters, although 2xHSE-*GAL1*p$_{-231}$, 4xHSE-*GAL1*p$_{-231}$, and 8xHSE-*GAL1*p$_{-231}$ reporters were responsive to the expression of aggregation-prone protein, an increase in the HSE copy number overall did not result in a significant (p < 0.05) improvement in their fold change (Figure 2B). Hence, as *HSP82*p$_{-858}$:uGFP, *SIS1*p$_{-501}$:uGFP, and *SSA1*p$_{-824}$:uGFP reporters had the highest fold changes among all tested promoters, we focused our efforts on further optimization using these reporters.

Optimization of Aggregated Protein Reporters. Given the modest fold changes observed for the three native promoters upon the expression of aggregation-prone *vs* folded soluble protein expression, we sought to increase the copy number of the upstream activating sequence (UAS), as previously successfully applied for improving output and dynamic range of synthetic promoters and reporters.[26,43] Before increasing the copy number of the UASs, we truncated the three native promoters to a minimal size while conserving all the UASs with all the HSEs as inferred from the literature and databases of transcription factor binding sites.[44,45] This was done to avoid reduced UAS-mediated activation as the effect of the UAS in promoter activation decreases when the distance to the TATA box increases.[46] With this approach, promoter sizes were reduced from the initial sizes of 501−858 bp to 324−365 bp (Table S5). Compared to the designs using full-length promoters, truncated uGFP-based reporters based on *HSP82*p and *SSA1*p (1xUAS-*HSP82*p$_{-324}$:uGFP and 1xUAS-*SSA1*p$_{-365}$:uGFP) showed a minimal variation in fold change, while the truncated *SIS1*p reporter (1xUAS-*SIS1*p$_{-345}$:uGFP) showed a decreased (∼50%) fold change compared to its WT

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

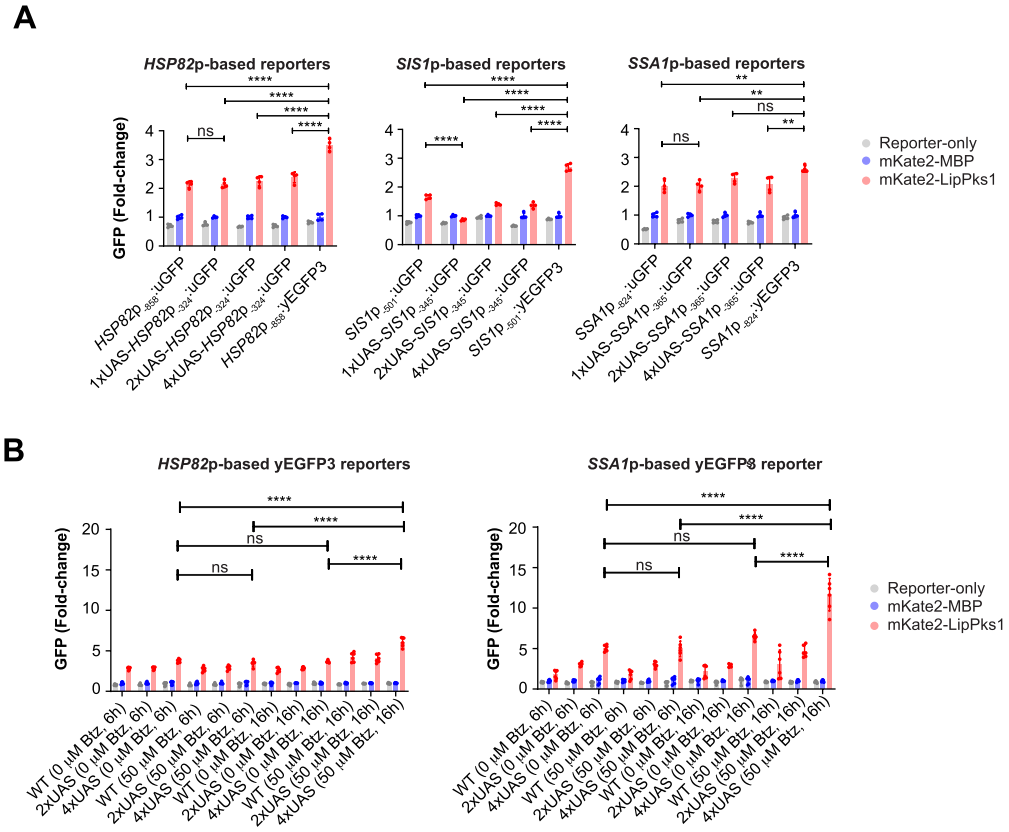


Figure 3. Promoter−reporters for agregated protein optimization. (A) Initial optimization: promoter truncation, UAS copy number, and GFP variants. (B) Optimization of yEGFP3 reporters: UAS copy number, sampling time, and proteasome inhibition. Yeast strains harboring reporter constructs were grown for 6 or 16 h and assessed by flow cytometry (n = (4−6) × 10 000 events). For each promoter, the fold change was calculated by normalizing GFP geometric means for each replicate (10 000 events) to the GFP geometric mean of all mKate2-MBP replicates (40 000 or 60 000 events). Bars represent the mean and error bars, the standard deviation of the data points. Asterisks represent statistical significance between conditions indicated by horizontal bar ends: **** = $p < 0.0001$, *** = $p < 0.001$, ** = $p < 0.001$, * = $p < 0.05$, and ns = not significant ($p > 0.05$).

version (Figure 3A). An increase in the the copy number of the UAS in the synthetic promoters did not result in significant improvements in fold changes when compared to reporters based on WT promoters (Figure 3A).

As we did not observe significant improvements in fold changes upon increasing UAS copy number, we next sought to boost the reporters fold change by changing the reporter protein uGFP to yeast-enhanced GFP mutant 3 (yEGFP3),[47] a variant with a half-life of ∼7 h, ∼13 times the half-life of uGFP.[42] It has been observed that promoter−reporters having yEGFP3 as the reporter can result in higher fluorescence and better signal-to-noise ratio than promoter−reporters with uGFP.[48] Indeed, when uGFP was replaced with yEGFP3 for reporters based on *HSP82*p$_{-858}$, *SIS1*p$_{-501}$, and *SSA1*p$_{-824}$, increased fold changes of 31−64% were observed (Figure 3A).

Next, since the previous efforts related to increasing UAS copy numbers in combination with uGFP were unsuccessful, we further tested if the combination of yEGFP3 together with 2x and 4x copies of UASs could improve the output of the reporter systems based on *HSP82*p and *SSA1*p. Combinations of yEGFP3 and an increased UAS copy number for *SIS1*p-based reporters were not tested as the truncated variants already had lowered fold changes compared to WT *SIS1*p (Figure 3A). Additionally, for these assays, we compared reporter fold changes between control cultures and cultures supplemented with bortezomib (Btz), a proteasome inhibitor,[49] known to result in the accumulation of misfolded protein aggregates,[50] following both 6 and 16 h of cultivation. Here, on the basis of the microscopy analysis of fluorescence distribution in yeast cells under all four conditions (6 and 16 h, ±50 μM Btz), we observed that only mKate2-LipPks1 formed aggregates (Figure S7A,B). More specifically, these aggregates were observed to colocalize with Hsp104-GFP (Figure S7B), a marker for Q-bodies, and the insoluble protein deposit (IPOD), protein deposition sites for soluble and insoluble protein aggregates, respectively.[35,36,51] Furthermore, cells of 16 h cultures treated with Btz had larger single inclusions than those observed for 6 h cultures, consistent with mKate2-LipPks1 residing at the IPOD under these conditions (Figure S7B).

When comparing the different reporter designs, we observed that, for most conditions, the 4xUAS-*HSP82*p$_{-324}$:yEGFP3 and 4xUAS-*SSA1*p$_{-365}$:yEGFP3 constructs had 1.3−4.4× higher fold changes ($p < 0.05$) than full-length and 2x UAS constructs (Figure 3B). Also, reporters with a 4x UAS for both promoters sampled at 16 h following Btz treatment had higher fold changes than such reporters in any other condition ($p < 0.0001$) (Figure 3B). Specifically, for the 4xUAS-*HSP82*p$_{-324}$:yEGFP3 reporter treated with Btz, this resulted in a fold change up to 6.1×, while the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter showed the highest fold change (11.5×).

Characterization of the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 Reporter. In order to further characterize the operational range of the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter, we assessed the change in fluorescence output when protein baits mKate2-

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

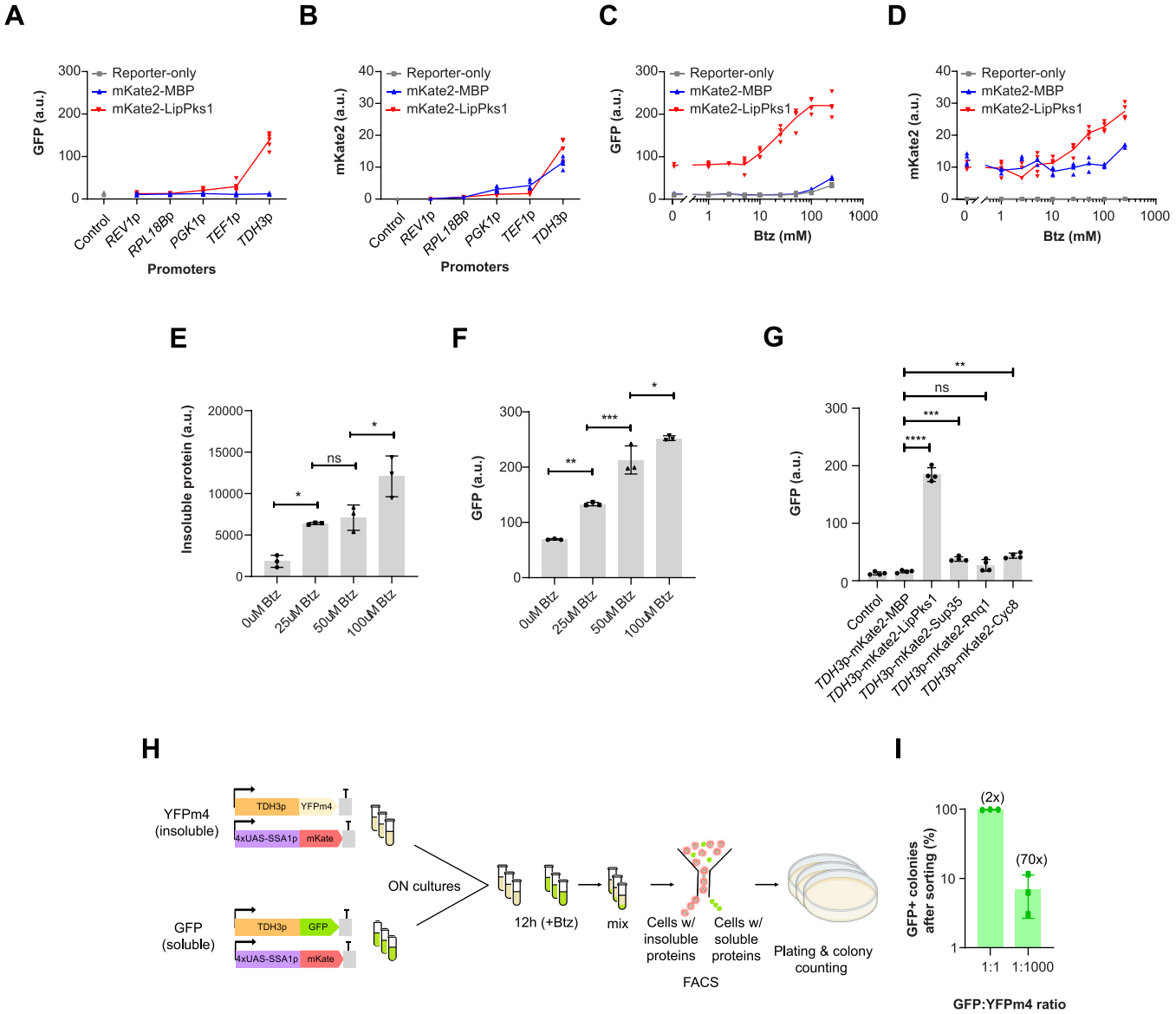


Figure 4. 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter characterization. (A) GFP reporter output and (B) mKate2 fluorescence of strains expressing soluble (mKate2-MBP) and insoluble (mKate2-LipPks1) protein baits from different promoters. Cultures were supplemented with 50 μM Btz and sampled after 16 h of growth. n = 6 × 10 000 events. (C) GFP reporter output and (D) mKate2 fluorescence of strains expressing protein baits. Cultures supplemented with 0−250 μM Btz and sampled after 12 h of growth. n = (3−4) × 10 000 events. (E, F) Effect of Btz treatment on the insoluble fraction of mKate2-LipPks1 (E) and fluorescence output (F) from the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter. Cultures were supplemented with 0−100 μM Btz and sampled after 16 h of growth. n = 3. (G) GFP reporter output in strains expressing prion precursor proteins. Cultures supplemented with 50 μM Btz and sampled after 16 h of growth. n = 6 × 10 000 events. (H) Scheme for the enrichment of strains expressing soluble protein from mixed populations. (I) Proportion of strains expressing soluble protein following FACS. Asterisks represent statistical significance between conditions indicated by horizontal bar ends: **** = $p < 0.0001$, *** = $p < 0.001$, ** = $p < 0.001$, * = $p < 0.05$, and ns = not significant ($p > 0.05$). For all panels, bars represent the mean and error bars, the standard deviation of all data points.

MBP and mKate2-LipPks1 were expressed from five promoters of different strengths. Here, we did not observe changes in the reporter fluorescence output when increasing the expression level of mKate2-MBP, whereas reporter fluorescence output increased when increasing the expression levels of mKate2-LipPks1 (Figure 4A,B, Figure S10). More specifically, reporter fluorescence outputs increased only when expressing mKate2-LipPks1 from *TEF1*p and *TDH3*p, compared to the mKate2-LipPks1 expression from the weak promoter *REV1*p ($p < 0.05$). Notably, reporter fluorescence outputs increased almost 5-fold when mKate2-LipPks1 was expressed from *TDH3*p compared to when it was expressed from *TEF1*p, whereas a 2.3-fold increase was observed when comparing *TEF1*p to *REV1*p, thus yielding an 11-fold increase in reporter output upon varying the expression level of mKate2-LipPks1 (Figure 4A). When the reporter outputs in strains expressing mKate2-LipPks1 or mKate2-MBP under the control of any of the five promoters used in this assay were compared, fold changes ranged from 1.1 to 11.5 (Supplementary Data File 1).

On the basis of our observations that Btz promotes accumulation of mKate2-LipPks1 in the IPOD while mKate2-MBP does not form aggregates (Figure S7) and that the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter responds to increasing expression levels of insoluble protein (Figure 4A,B), we sought to

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

investigate if the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter could respond to different concentrations of Btz. For this experiment, we used the strains expressing the protein baits from the *TDH3*p promoter and Btz concentrations ranging from 0 to 250 μM. We used concentrations of up to 250 μM, as the vehicle (DMSO) largely inhibits cell growth when using concentrations higher than 2.5% (equivalent to 250 μM Btz) (Figure S11). For the strain expressing mKate2-LipPks1, Btz concentrations of 25−250 μM resulted in increased fluorescence of mKate2-LipPks1 compared to the strains grown without Btz ($p < 0.05$), just as increases in the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter output increased when treating cultures with 25 to 250 μM Btz compared to untreated cultures ($p < 0.0001$) (Figure 4C,D, Supplementary Data File 1). Importantly, Btz treatment did not result in a significant increase in the fluorescence of mKate2-MBP compared to cultures without proteasome inhibition, whereas we observed that Btz concentrations of 100 and 250 μM resulted in increased 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter fluorescence output for strains expressing mKate2-MBP and the reporter-only strain ($p < 0.001$) (Figure 4C, Supplementary Data File 1), suggesting these concentrations disrupt cell proteostasis, even in the absence of insoluble protein expression. Furthermore, we observed that Btz treatment increased the amount of insoluble protein fraction for mKate2-LipPks1 but not for mKate2-MBP (Figure S12) and that increasing concentrations of Btz increased the amount of the mKate2-LipPks1 insoluble fraction together with reporter output (Figures 4E,F and S14). Therefore, using Btz increased the performance of the reporter system by increasing the amount of the insoluble protein, while it did not affect the behavior of the soluble protein bait.

We also characterized the performance of the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter in the background strain CEN.PK2-1C, which expresses the proteases Pep4 and Prb1, which were initially deleted for the development of our reporter system. We observed that the reporter 4xUAS-*SSA1*p$_{-365}$:yEGFP3 was able to discriminate between the expression of mKate2-MBP and mKate2-LipPks1 in CEN.PK2-1C following both 6 and 16 h cultivations, albeit with lower induction of the reporter upon mKate2-LipPks1 expression compared to those observed in CEN.PK2-1C *pep4*Δ *prb1*Δ (Figure S15A,B). This data suggests that Pep4 and Prb1 proteases are involved in processing mKate2-LipPks1 aggregates.

We further tested the reporter system by assessing the output in response to the expression of proteins known to produce protein aggregates. To this end, we constructed strains expressing mKate2 fusions to prion precursor proteins Sup35, Rnq1, and Cyc8[52] controlled by the strong promoter *TDH3*p. Here, we observed that the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter responded to Sup35 and Cyc8 overexpression by 2.4- and 2.7-fold increases compared to the response to mKate2-MBP, respectively, whereas the expression of Rnq1 only lead to a modest, yet not significant, 1.6-fold change (Figure 4G).

Finally, to investigate if the 4xUAS-*SSA1*p$_{-365}$ engineered promoter would be able to be used as a tool for discriminating cells producing soluble proteins from those expressing aggregation-prone proteins, we performed an experiment with the aim of sorting strains heterologously expressing soluble protein from strains expressing aggregation-prone protein. For this experiment, we used strains expressing either soluble yEGFP3 or an insoluble variant of yellow fluorescent protein (YFPm4)[15] in background strains harboring the 4xUAS-*SSA1*p$_{-365}$:mKate2 reporter (Figure S16). Both strains were initially grown separately in medium supplemented with Btz. Following 12 h of growth, cells were mixed in 1:1 and 1:1000 ratios of yEGFP3:YFPm4 strains. Next, mixed populations were sorted using fluorescence-activated cell sorting (FACS) by selecting the top 1% of the total events with the lowest mKate2 fluorescence. Next, sorted cells were plated before counting yEGFP3 fluorescent colonies (Figure 4H). Importantly, for the culture with the 1:1 starting ratio, ∼98% of the colonies were fluorescent, while the culture with the 1:1000 starting ratio had ∼7% fluorescent colonies following FACS enrichment, equivalent to approximately 2× and 70× enrichments of cells expressing soluble yEGFP3, respectively, from a single round of cocultivation and sorting (Figure 4I). This result demonstrates the potential of the reporter as a tool for facile screening and selection of cells expressing soluble proteins in living cells.

DISCUSSION

The assessment of the expression of misfolded proteins and the related cellular stress is of interest for both medicine and biotechnology as protein misfolding and eventual aggregation have been linked to disease[2,5] and low enzymatic activity.[29,53] In this context, several methods have been developed both for the assessment of protein solubility and aggregation toward studying disease-related proteins[21−23] and for the directed evolution of enzymes and antibodies with improved solubility for biotechnological purposes.[16−20]

In this study, we have constructed and optimized a reporter, 4xUAS-*SSA1*p$_{-365}$:yEGFP3, for assessing the expression of aggregation-prone protein in yeast cytosol on the basis of the characterization of the proteomic response to the misfolded protein. We have successfully optimized the reporter and growth conditions to optimize the dynamic range of the reporter obtaining up to an ∼11.5-fold change comparing strains expressing folded soluble with those expressing aggregation-prone proteins. An advantage of our system is that it does not rely on protein fusions and is complementary and compatible with other methods developed recently such as yTRAP[21] and tripartite fusions.[23] Specifically, we believe our system would be advantageous to use when target proteins are of large sizes, such as PKSs, as assays based on yTRAP and tripartite fusions could be limited by the size of the fusion proteins. For instance, tripartite fusions in yeast are currently limited to proteins up to ∼100 kDa,[23] while yTRAP synthetic transcription factor fusions may be limited by nuclear pore size, generally considered to allow proteins of up to 60 kDa, yet proteins up to 230 kDa have been observed to diffuse to the nucleus.[54,55]

Beyond the bait used for testing the reporter system (*i.e.*, mKate2-LipPks1), we observed induction from the reporter system upon strong expression of two out of three prion precursor proteins, indicating that the reporter is suitable for the detection of a number of aggregation-prone proteins. For the prion precursor Rnq1, which did not yield significant induction in our assays, we speculate that Rnq1 might interact with yeast proteome and produce stress through a mechanism that is not detected by the 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter, as it has been recently observed that yeast proteome responds differently when human pathogenic aggregation-prone proteins are overexpressed.[56] Furthermore, while the developed reporter responds to increasing amounts of insoluble protein, the lower detection limit of the reporter may make the system unsuitable for all applications as reporter outputs were only detected when expressing aggregation-prone protein baits at high expression

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

levels. Hence, we see the need to choose the most suitable tool according to the experiment to be performed in an effort toward understanding and engineering protein folding and proteostasis. Likewise, since misfolded proteins are localized to different compartments in yeast cells[35,36] and protein aggregates elicit different proteome responses,[56] we consider that several mechanisms and pathways could be further harnessed to develop more proteostasis reporters. Still, we envision that the reporter system developed in this study can be applied to studies focused on the optimization of protein solubility and expression and the assessment of cytosolic proteostasis as well as to studies of aggregation-prone pathogenic proteins.

MATERIALS AND METHODS

Strains, Media, and Chemicals. Yeast strains were derivatives of the strain CEN.PK2-1C (MATa; his3Δ1; leu2-3,112; ura3-52; trp1-289; MAL2-8c; SUC2) obtained from Peter Kötter (Johann Wolfgang Goethe-University Frankfurt, Germany). *E. coli* strain DH5α was used for cloning. All *E. coli* strains with plasmids were grown in LB or 2xYT media supplemented with 100 μg/mL ampicillin.

Yeast strains were grown in Yeast Synthetic Drop-out (SC) medium (pH 5.6) or Mineral medium (MM; pH 6.0) supplemented with amino acids. SC medium (g/L) is composed of 6.7 g of yeast nitrogen base without amino acids (Sigma-Aldrich), an adequate amount of Yeast Synthetic Drop-out medium Supplements (Sigma-Aldrich), 20 g of glucose and, when necessary, 20 g of agar. MM was prepared as described previously.[57]

Cloning and Strain Construction. Plasmids, *in vivo* assembled cassettes, and strains and plasmids of this work are listed in Tables S1, S2, and S3. Plasmids were constructed by USER Cloning (New England Biolabs). *E. coli* transformations were performed using chemically competent cells.[58] PCR assays were performed using Phusion U Hot Start and Phusion Hot Start II Polymerases (Thermo Fisher Scientific). Synthesized DNA was obtained from Integrated DNA Technologies.

In vivo assembly of expression cassettes was performed using the CasEMBLR method[59] using overlapping regions of at least 30 bp. Integrations of the cassettes were performed in yeast chromosomal sites reported previously.[60] Yeast transformations were performed according to previously described methods.[61]

Microscopy. Yeast precultures were grown overnight (~20 h) in SC medium from single colonies. Next day cultures were diluted ~1:20 in SC or MM and grown for 6 or 16 h before imaging in medium with or without 50 μM bortezomib (Btz). Yeast cells were observed in a Leica DM4000 B microscope equipped with a DFC 300 FX R2 camera and a EL600 light source (Leica Microsystems). Excitation/emission filters used for GFP and mKate2 fluorescence were EX:470/40, EM:525/50, and EX:545/26, EM:605/70, respectively.

Untargeted Proteomics. Single colonies were inoculated in 2 mL of SC-Trp medium and grown for 20 h. The next day, 250 mL flasks with 25 mL of Mineral medium + Ura,His,Leu,Trp were inoculated to an OD of ~0.1. Cultures were grown to exponential phase for 6 h at 30 °C and 250 rpm. ~2×10^8 cells from each flask were harvested and frozen at −80 °C. At least three replicates per strain, each one corresponding to a different flask, were used for data-dependent proteomics analysis. Yeast cells were resuspended in 75 μL of 6 M guanidinium hydrochloride (GdnHCl), 5 mM tris(2-carboxyethyl)phosphine (TCEP), 10 mM chloroacetamide (CAA), and 100 mM Tris−HCl (pH 8.5) at 95 °C and lysed using a Mixer Mill MM 400 (Retsch) set to 25 Hz for 5 min at room temperature, followed by incubation in a ThermoMixer (Eppendorff) at 95 °C and 2000 rpm. Samples were centrifuged at 15 000*g* for 10 min, and 50 μL of each sample supernatant was taken and mixed with 50 mM ammonium bicarbonate. The protein concentration was measured with the Micro BCA Protein Assay kit (Thermo Fisher Scientific) using BSA as the standard. A volume containing 100 μg of protein was used for digestion with 1 μg of trypsin. Samples were digested for 8 h at 37 °C under constant shaking at 400 rpm. Afterward, 10 μL of 10% TFA was added, and the samples were cleaned and concentrated with stage tips using a C18 resin (Empore, 3M, USA). Samples were used for both targeted and untargeted proteomics analyses. One μg of digested peptides for each sample was injected to a Cap-LC system (Thermo Fisher Scientific) coupled to a 75 μm × 15 cm, 2 μm C18 easy spray column (Thermo Fisher Scientific). Flow rate was set to 1.2 μL/min. Prior to implementing a stepped gradient starting at 4% and ending in 40% acetonitrile in water over 60 min, the peptides were trapped on a precolumn (μ-precolumn C18 PepMap 100, 5 μm, 100 Å). Samples were sprayed into an Orbitrap Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific). The parameters used for mass spectrometry (MS)-level scans were: Orbitrap resolution, 60 000; AGC target, 1.0×10^6; maximum injection time, 50 ms; intensity threshold, 5.0×10^3; dynamic exclusion, 25 s. The top 12 mode with the HCD collision energy set to 28% (AGC target, 1.0×10^4; maximum injection time, 22 ms) was used for data-dependent MS2 selection. Data was analyzed in Proteome Discoverer 2.4 (Thermo Fisher Scientific). The following settings were used: fixed modifications, carbamidomethyl (C); variable modifications, oxidation of methionine residues; first search mass tolerance, 20 ppm; MS/MS tolerance, 20 ppm. One missed cleavage was set for trypsin. The false discovery rate was set at 0.1%. The *S. cerevisiae* database retrieved from Uniprot with proteome ID AUP000002311 was used for searching proteins in the samples. Sequences of heterologous proteins were added to this database. Label-free quantification was used, and the abundance of each protein is reported as the relative intensity with respect to total intensity of all identified peptides.

Statistical analysis was performed in Proteome Discoverer 2.4. ANOVA (individual proteins) was done to compare the protein abundances among strains. Adjusted *p*-values referred to in the text were calculated using the the Benjamini-Hochberg method. Hierarchical clustering and principal component analysis were done in ClustVis.[62] Complete-linkage clustering and correlation distance were used for hierarchical clustering of rows and columns. The mass spectrometry proteomics raw data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository[63] with the data set identifier PXD021470.

Targeted Proteomics. After cleaning the samples with stage tips, the samples were dried and resuspended in 0.1% formic acid containing synthetic heavy labeled peptides to a final concentration of 50 fmol for each synthetic peptide. The synthetic peptides representing a unique peptide for each of the proteins were MPGVYYVDR for mKate2, LYPFTWDAVR for MBP, and ELLDSAPVFAR for LipPks1. One 1 μg of total protein for each sample was used for injections for mass spectrometry according to the protocol described for untargeted proteomics. The unscheduled PRM method used the following parameters for the mass spectrometry: The full scan was collected at a resolution of 120 000, and MS2 scans were collected at resolution of 30 000 with a maximum injection time

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

of 100 ms. Data was analyzed using Skyline-daily,[64] and ratios were calculated for the synthetic peptide on the basis of the MS2 scans.

Flow Cytometry. Single colonies of strains tested were grown in 150 μL of SC-Trp medium at 30 °C and 300 rpm overnight (16−20 h) in 96-well microtiter plates in at least 3 replicates. Plates were covered with VIEWseal sealers (Greiner Bio-One). They were diluted ∼1:20 in MM+Ura+His+Leu +Trp in 96-well microtiter plates. Bortezomib (Btz; 10 mM) in DMSO (ApexBIO) and DMSO to the desired concentration were added at this point in experiments that required it. Cultures were grown at 30 °C and 300 rpm and covered with VIEWseal sealers for 6 or 16 h unless otherwise indicated. Samples were analyzed immediately or kept at 4 °C until analyzed. If necessary, samples were diluted in PBS prior to flow cytometry analysis. Flow cytometry measurements were done in a MACSQuant Analyzer VYB Flow Cytometer (Miltenyi Biotec). Cells populations were gated for singlets (FSC-H *vs* FSC-A), and this data was used for further analyses. At least 10 000 events were measured per sample. A 488 nm laser and 525/50 nm filter was used for the GFP measurements, and a 561 nm laser and 661/20 nm filter was used for the mKate2 measurements. Flow cytometry data was analyzed using Cytoflow (https://bpteague.github.io/cytoflow/), and GraphPad Prism version 8.4 flow cytometry measurement data (geometric mean of each replicate) are in Supplementary Data File 1. The histograms for all recorded events are in Figures S6, S8−S10, and S13. Fold-change values are based on the geometric means calculated for all events across all replicates. Statistical analysis was done based on the geometric means of each replicate using the Holm-Sidak Method and Tukey and Dunnet tests.

FACS. Single colonies of strains were grown overnight in 96-well microtiter plates with 150 μL of SC-TRP medium as indicated previously. Cultures were diluted ∼1:20 to a total volume of 150 μL in MM+URA+HIS+LEU+TRP+ Btz (100 μM) and grown for 12 h. After, cultures were mixed in 1:1 and 1:1000 (GFP:YFPm4 strains) ratios according to the OD. Mixed cultures were then sorted in a SH800 FACS Cell Sorter (Sony) by selecting the ∼1% of the cultures with the lower mKate2 reporter fluorescence values. Cells populations were gated for singlets (FSC-H *vs* FSC-A), and 200 cells per sample were collected. Collected cells were plated in SC medium agar and incubated at 30 °C. Total and GFP colonies were counted after 2 days. Enrichment fold changes were calculated by dividing the initial ratio of cells expressing GFP (soluble protein) by the final ratio of GFP colonies to total colonies plated after sorting ($n = 3$).

Insoluble Protein Extraction and Quantification. Single colonies of the tested strains were grown in 2 mL of SC-TRP medium in 14 mL tubes at 30 °C and 250 rpm overnight (16−20 h). Cultures were diluted ∼1:20 to a total volume of 4 mL in 14 mL tubes with MM+Ura+His+Leu+Trp medium with the indicated amount of Btz and grown for 16 h at 30 °C and 250 rpm. Cultures were sampled for flow cytometry and for insoluble protein extraction. Insoluble protein extraction was done according to a procedure described previously.[65] Briefly, ∼2 × 10^8 cells were collected by centrifugation at 2500*g* for 1 min at RT. Cell pellets were washed in 1 mL of ice-cold buffer S0 (120 mM KCl, 2 mM EDTA, 20 mM HEPES-KOH, pH 7.4) and centrifuged for 30 s at 5000*g* and 4 °C. Cell pellets were resuspended in 150 μL of buffer S (120 mM KCl, 2 mM EDTA, 20 mM HEPES-KOH, 0.5 mM DTT, 1:100 protease inhibitor cocktail IV [Calbiochem], 1 mM PMSF, pH 7.4). 75 μL of samples was flash frozen with liquid nitrogen (LN) in a 2 mL Eppendorf "Safe-Lock" tube containing a 7 mm stainless steel ball. Cells were lysed in a Mixer Mill MM 400 (Retsch) with 4 × 90 s × 30 Hz pulses by placing the 2 mL tubes in an adapter (Retsch 22.008.0005) chilled in LN before and in between milling cycles. Lysates were resuspended in 400 μL of ice-cold buffer S, and thawed lysates were centrifuged at 3000*g* for 30 s. 200 μL of the clarified supernatant was centrifuged at 100 000*g* for 20 min at RT. Pellets (insoluble fractions) were washed in 200 μL of buffer S and centrifuged again at 100 000*g* for 20 min at RT. Pellets were resuspended in 50 μL of buffer S and sonicated for 5 s.

Samples were quantified by Western blotting. Briefly, aliquots of the insoluble protein pellets were mixed with 4× NuPAGE LDS Sample Buffer (Invitrogen) and DTT to a final concentration of 12.5 μM and heated for 5 min at 95 °C before loading on NuPAGE Novex 4−12% Bis-Tris Gels (Invitrogen). Samples were transferred to PVDF membranes using PVDF iBlot 2 Transfer Stacks (Invitrogen) and detected by a Penta-His HRP Conjugate kit (Qiagen). Western blot bands were quantified using ImageJ Software.[66]

Bioinformatics Tools. Transcription factors that bind to promoters of genes of interest were retrieved using YEASTRACT[67] by only considering the TFBSs with simultaneous DNA binding and expression evidence and where the TF acted as an activator. The retrieved data is in Table S7. Over-represented TFBSs in upregulated genes were identified by YeTFaSCo using the rank sum (Mann−Whitney−Wilcoxon) test against the YeTFaSCo curated database of the TFBSs; TFBSs with $p < 0.01$ were retrieved.[45] The retrieved data is in Table S8. Gene ontology analysis (biological process) was done using PANTHER version 14.[68] Significance is reported as False Discovery Rates (FDR).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00476.

Quantification of bait proteins mKate2, mKate2-MBP, and mKate2-LipPks1 expression by targeted proteomics; gene ontology enrichment analysis upon mKate2-LipPks1 expression; comparison of proteome responses to protein baits; synthetic HSE-*GAL*p1-231 reporter designs; growth curves of parental strains of strains used to screen promoter−reporters; flow cytometry histograms of screen of native and synthetic promoters as aggregated protein reporters; localization of protein baits in cellular compartments; flow cytometry histograms of reporters optimization by truncations, UAS copy numbers, and GFP variants; flow cytometry histograms of yEGFP3 reporters optimization by UAS copy numbers, and sampling time and proteasome inhibition; flow cytometry histograms of 4xUAS-SSA1p:yEGFP3 reporter output and mKate2 fluorescence protein baits from different promoters; DMSO toxicity in yeast strains expressing protein baits and 4xUAS-SSA1p:yEGFP3 reporter; effect of bortezomib (Btz) in mKate2-MBP and mKate-LipPks1 insoluble fractions; 4xUAS-SSA1p:yEGFP3 reporter output and mKate2 fluorescence of strains expressing protein baits and in cultures supplemented with 1−250 μM Btz; effect of bortezomib at different concentrations on the mKate2-LipPks1 insoluble fraction; effect of *pep4*Δ and *prb1*Δ

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

deletions on 4xUAS-*SSA1*p$_{-365}$:yEGFP3 reporter performance; yEGFP3 expressed as a soluble protein in yeast cytosol; plasmids, *in vivo* assembled cassettes, yeast strains, gene and promoter sequences, and sequences of representative plasmids used in this study; TFs that regulate genes upregulated upon mKate2-LipPks1 expression; over-represented TFBS in genes upregulated upon mKate2-LipPks1 expression

Yeast proteomes upon expression of protein baits, differentially expressed proteins, GO analysis of differentially expressed genes, and flow cytometry measurements (XLSX)

AUTHOR INFORMATION

Corresponding Author

Michael K. Jensen − *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby 2800, Denmark;* orcid.org/0000-0001-7574-4707; Phone: +45 6128 4850; Email: mije@biosustain.dtu.dk; Fax: +45 4525 8001

Authors

David Romero-Suarez − *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby 2800, Denmark*

Tune Wulff − *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby 2800, Denmark*

Yixin Rong − *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby 2800, Denmark*

Tadas Jakočiūnas − *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby 2800, Denmark*

Satoshi Yuzawa − *Joint BioEnergy Institute, Emeryville, California 94608, United States; Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* orcid.org/0000-0001-8358-2787

Jay D. Keasling − *Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby 2800, Denmark; Joint BioEnergy Institute, Emeryville, California 94608, United States; Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Chemical and Biomolecular Engineering, Department of Bioengineering, University of California, Berkeley, California 94720, United States; Center for Synthetic Biochemistry, Institute for Synthetic Biology, Shenzhen Institutes of Advanced Technologies, Shenzhen, Guangdong 518055, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00476

Author Contributions

D.R.-S., J.D.K., and M.K.J. conceived the study. D.R.-S. and M.K.J. designed the experiments. D.R.-S., Y.R., T.J., and S.Y. conducted all experimental work related to strain designs, plasmid and strain constructions, and cultivations. D.R.-S. and T.W. performed and analyzed proteomics data. D.R.-S. and M.K.J. wrote the manuscript.

Notes

The authors declare the following competing financial interest(s): J.D.K. has a financial interest in Amyris, Lygos, Demetrix, Maple Bio, Napigen, Apertor Laboratories, Ansa Biotechnology, and Berkeley Brewing Sciences. The authors declare that they have no other competing interests.

ACKNOWLEDGMENTS

This work was supported by the Novo Nordisk Foundation, Novo Nordisk Copenhagen Bioscience Ph.D. grant No. NNF16CC0020908. The authors thank Jie Zhang for input on the strain design and experimental work, Alex Toftgaard Nielsen for advice regarding the system design, and Jonathan Dahlin for advice on the protein work.

REFERENCES

(1) Balchin, D., Hayer-Hartl, M., and Hartl, F. U. (2016) In Vivo Aspects of Protein Folding and Quality Control. *Science 353* (6294), aac4354.

(2) Chiti, F., and Dobson, C. M. (2006) Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem. 75*, 333−366.

(3) Olzscha, H., Schermann, S. M., Woerner, A. C., Pinkert, S., Hecht, M. H., Tartaglia, G. G., Vendruscolo, M., Hayer-Hartl, M., Hartl, F. U., and Vabulas, R. M. (2011) Amyloid-like Aggregates Sequester Numerous Metastable Proteins with Essential Cellular Functions. *Cell 144* (1), 67−78.

(4) Butterfield, S. M., and Lashuel, H. A. (2010) Amyloidogenic Protein-Membrane Interactions: Mechanistic Insight from Model Systems. *Angew. Chem., Int. Ed. 49* (33), 5628−5654.

(5) Hipp, M. S., Kasturi, P., and Hartl, F. U. (2019) The Proteostasis Network and Its Decline in Ageing. *Nat. Rev. Mol. Cell Biol. 20* (7), 421−435.

(6) Schramm, F. D., Schroeder, K., and Jonas, K. (2020) Protein Aggregation in Bacteria. *FEMS Microbiol. Rev. 44* (1), 54−72.

(7) Bloom, J. D., and Arnold, F. H. (2009) In the Light of Directed Evolution: Pathways of Adaptive Protein Evolution. *Proc. Natl. Acad. Sci. U. S. A. 106*, 9995−10000.

(8) Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., Arrowsmith, C., Hui, R., Ming, J., dhe-Paganon, S., Park, H.-W., Savchenko, A., Yee, A., Edwards, A., Vincentelli, R., Cambillau, C., Kim, R., Kim, S.-H., Rao, Z., Shi, Y., Terwilliger, T. C., Kim, C.-Y., Hung, L.-W., Waldo, G. S., Peleg, Y., Albeck, S., Unger, T., Dym, O., Prilusky, J., Sussman, J. L., Stevens, R. C., Lesley, S. A., Wilson, I. A., Joachimiak, A., Collart, F., Dementieva, I., Donnelly, M. I., Eschenfeldt, W. H., Kim, Y., Stols, L., Wu, R., Zhou, M., Burley, S. K., Emtage, J. S., Sauder, J. M., Thompson, D., Bain, K., Luz, J., Gheyi, T., Zhang, F., Atwell, S., Almo, S. C., Bonanno, J. B., Fiser, A., Swaminathan, S., Studier, F. W., Chance, M. R., Sali, A., Acton, T. B., Xiao, R., Zhao, L., Ma, L. C., Hunt, J. F., Tong, L., Cunningham, K., Inouye, M., Anderson, S., Janjua, H., Shastry, R., Ho, C. K., Wang, D., Wang, H., Jiang, M., Montelione, G. T., Stuart, D. I., Owens, R. J., Daenke, S., Schütz, A., Heinemann, U., Yokoyama, S., Büssow, K., Gunsalus, K. C., and Structural Genomics Consortium; Architecture et Fonction des Macromolécules Biologiques; Berkeley Structural Genomics Center; China Structural Genomics Consortium; Integrated Center for Structure and Function Innovation; Israel Structural Proteomics Center; Joint Center for Structural Genomics; Midwest Center for Structural Genomics; New York Structural GenomiX Research Center for Structural Genomics; Northeast Structural Genomics Consortium; Oxford Protein Production Facility; Protein Sample Production Facility, Max Delbrück Center for Molecular Medicine; RIKEN Structural Genomics/Proteomics Initiative; SPINE2-Complexes (2008) Protein Production and Purification. *Nat. Methods 5* (2), 135−146.

(9) Tokuriki, N., and Tawfik, D. S. (2009) Stability Effects of Mutations and Protein Evolvability. *Curr. Opin. Struct. Biol. 19* (5), 596−604.
(10) Nielsen, J., and Keasling, J. D. (2016) Engineering Cellular Metabolism. *Cell 164* (6), 1185−1197.
(11) Glick, B. R. (1995) Metabolic Load and Heterologous Gene Expression. *Biotechnol. Adv. 13* (2), 247−261.
(12) Georgiou, G. (1988) Optimizing the Production of Recombinant Proteins in Microorganisms. *AIChE J. 34* (8), 1233−1248.
(13) Ceroni, F., Algar, R., Stan, G.-B., and Ellis, T. (2015) Quantifying Cellular Capacity Identifies Gene Expression Designs with Reduced Burden. *Nat. Methods 12* (5), 415−418.
(14) Gill, R. T., Valdes, J. J., and Bentley, W. E. (2000) A Comparative Study of Global Stress Gene Regulation in Response to Overexpression of Recombinant Proteins in Escherichia Coli. *Metab. Eng. 2* (3), 178−189.
(15) Geiler-Samerotte, K. A., Dion, M. F., Budnik, B. A., Wang, S. M., Hartl, D. L., and Drummond, D. A. (2011) Misfolded Proteins Impose a Dosage-Dependent Fitness Cost and Trigger a Cytosolic Unfolded Protein Response in Yeast. *Proc. Natl. Acad. Sci. U. S. A. 108* (2), 680−685.
(16) Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999) Rapid Protein-Folding Assay Using Green Fluorescent Protein. *Nat. Biotechnol. 17* (7), 691−695.
(17) Foit, L., Morgan, G. J., Kern, M. J., Steimer, L. R., von Hacht, A. A., Titchmarsh, J., Warriner, S. L., Radford, S. E., and Bardwell, J. C. A. (2009) Optimizing Protein Stability in Vivo. *Mol. Cell 36* (5), 861−871.
(18) Wang, T., Badran, A. H., Huang, T. P., and Liu, D. R. (2018) Continuous Directed Evolution of Proteins with Improved Soluble Expression. *Nat. Chem. Biol. 14* (10), 972−980.
(19) Asial, I., Cheng, Y. X., Engman, H., Dollhopf, M., Wu, B., Nordlund, P., and Cornvik, T. (2013) Engineering Protein Thermostability Using a Generic Activity-Independent Biophysical Screen inside the Cell. *Nat. Commun. 4*, 2901.
(20) Lesley, S. A., Graziano, J., Cho, C. Y., Knuth, M. W., and Klock, H. E. (2002) Gene Expression Response to Misfolded Protein as a Screen for Soluble Recombinant Protein. *Protein Eng., Des. Sel. 15* (2), 153−160.
(21) Newby, G. A., Kiriakov, S., Hallacli, E., Kayatekin, C., Tsvetkov, P., Mancuso, C. P., Bonner, J. M., Hesse, W. R., Chakrabortee, S., Manogaran, A. L., Liebman, S. W., Lindquist, S., and Khalil, A. S. (2017) A Genetic Tool to Track Protein Aggregates and Control Prion Inheritance. *Cell 171* (4), 966−979.E18.
(22) Zeng, Y., Jones, A. M., Thomas, E. E., Nassif, B., Silberg, J. J., and Segatori, L. (2018) A Split Transcriptional Repressor That Links Protein Solubility to an Orthogonal Genetic Circuit. *ACS Synth. Biol. 7* (9), 2126−2138.
(23) Sachsenhauser, V., Deng, X., Kim, H.-H., Jankovic, M., and Bardwell, J. C. A. (2020) Yeast Tripartite Biosensors Sensitive to Protein Stability and Aggregation Propensity. *ACS Chem. Biol. 15* (4), 1078−1088.
(24) Dahl, R. H., Zhang, F., Alonso-Gutierrez, J., and Baidoo, E. (2013) Engineering Dynamic Pathway Regulation Using Stress-Response Promoters. *Nat. Biotechnol. 31*, 1039.
(25) Rajkumar, A. S., Liu, G., Bergenholm, D., Arsovska, D., Kristensen, M., Nielsen, J., Jensen, M. K., and Keasling, J. D. (2016) Engineering of Synthetic, Stress-Responsive Yeast Promoters. *Nucleic Acids Res. 44* (17), e136.
(26) Zhang, J., Sonnenschein, N., Pihl, T. P. B., Pedersen, K. R., Jensen, M. K., and Keasling, J. D. (2016) Engineering an NADPH/NADP Redox Biosensor in Yeast. *ACS Synth. Biol. 5*, 1546−1556.
(27) Yuzawa, S., Deng, K., Wang, G., Baidoo, E. E. K., Northen, T. R., Adams, P. D., Katz, L., and Keasling, J. D. (2017) Comprehensive in Vitro Analysis of Acyltransferase Domain Exchanges in Modular Polyketide Synthases and Its Application for Short-Chain Ketone Production. *ACS Synth. Biol. 6* (1), 139−147.
(28) Kataeva, I., Chang, J., Xu, H., Luan, C.-H., Zhou, J., Uversky, V. N., Lin, D., Horanyi, P., Liu, Z. J., Ljungdahl, L. G., Rose, J., Luo, M., and Wang, B.-C. (2005) Improving Solubility of Shewanella Oneidensis MR-1 and Clostridium Thermocellum JW-20 Proteins Expressed into Esherichia Coli. *J. Proteome Res. 4* (6), 1942−1951.
(29) Reider Apel, A., d'Espaux, L., Wehrs, M., Sachs, D., Li, R. A., Tong, G. J., Garber, M., Nnadi, O., Zhuang, W., Hillson, N. J., Keasling, J. D., and Mukhopadhyay, A. (2017) A Cas9-Based Toolkit to Program Gene Expression in Saccharomyces Cerevisiae. *Nucleic Acids Res. 45* (1), 496−508.
(30) Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C., and Barr, P. J. (1998) Production of a Polyketide Natural Product in Nonpolyketide-Producing Prokaryotic and Eukaryotic Hosts. *Proc. Natl. Acad. Sci. U. S. A. 95* (2), 505−509.
(31) Pfeifer, B. A., Admiraal, S. J., Gramajo, H., Cane, D. E., and Khosla, C. (2001) Biosynthesis of Complex Polyketides in a Metabolically Engineered Strain of E. Coli. *Science 291* (5509), 1790−1792.
(32) Gao, X., Wang, P., and Tang, Y. (2010) Engineered Polyketide Biosynthesis and Biocatalysis in Escherichia Coli. *Appl. Microbiol. Biotechnol. 88* (6), 1233−1242.
(33) Lowell, A. N., DeMars, M. D., Slocum, S. T., Yu, F., Anand, K., Chemler, J. A., Korakavi, N., Priessnitz, J. K., Park, S. R., Koch, A. A., Schultz, P. J., and Sherman, D. H. (2017) Chemoenzymatic Total Synthesis and Structural Diversification of Tylactone-Based Macrolide Antibiotics through Late-Stage Polyketide Assembly, Tailoring, and C—H Functionalization. *J. Am. Chem. Soc. 139*, 7913−7920.
(34) Jones, E. W. (1991) [31] Tackling the Protease Problem in Saccharomyces Cerevisiae. *Methods Enzymol. 194*, 428−453.
(35) Kaganovich, D., Kopito, R., and Frydman, J. (2008) Misfolded Proteins Partition between Two Distinct Quality Control Compartments. *Nature 454* (7208), 1088−1095.
(36) Escusa-Toret, S., Vonk, W. I. M., and Frydman, J. (2013) Spatial Sequestration of Misfolded Proteins by a Dynamic Chaperone Pathway Enhances Cellular Fitness during Stress. *Nat. Cell Biol. 15* (10), 1231−1243.
(37) Rizzolo, K., Huen, J., Kumar, A., Phanse, S., Vlasblom, J., Kakihara, Y., Zeineddine, H. A., Minic, Z., Snider, J., Wang, W., Pons, C., Seraphim, T. V., Boczek, E. E., Alberti, S., Costanzo, M., Myers, C. L., Stagljar, I., Boone, C., Babu, M., and Houry, W. A. (2017) Features of the Chaperone Cellular Network Revealed through Systematic Interaction Mapping. *Cell Rep. 20* (11), 2735−2748.
(38) Finley, D., Ulrich, H. D., Sommer, T., and Kaiser, P. (2012) The Ubiquitin−Proteasome System of Saccharomyces Cerevisiae. *Genetics 192* (2), 319−360.
(39) Solís, E. J., Pandey, J. P., Zheng, X., Jin, D. X., Gupta, P. B., Airoldi, E. M., Pincus, D., and Denic, V. (2016) Defining the Essential Function of Yeast Hsf1 Reveals a Compact Transcriptional Program for Maintaining Eukaryotic Proteostasis. *Mol. Cell 63* (1), 60−71.
(40) Krakowiak, J., Zheng, X., Patel, N., Feder, Z. A., Anandhakumar, J., Valerius, K., Gross, D. S., Khalil, A. S., and Pincus, D. (2018) Hsf1 and Hsp70 Constitute a Two-Component Feedback Loop That Regulates the Yeast Heat Shock Response. *eLife 7*, e31668.
(41) Zhang, J., Petersen, S. D., Radivojevic, T., Ramirez, A., Pérez-Manríquez, A., Abeliuk, E., Sánchez, B. J., Costello, Z., Chen, Y., Fero, M. J., Martin, H. G., Nielsen, J., Keasling, J. D., and Jensen, M. K. (2020) Combining Mechanistic and Machine Learning Models for Predictive Engineering and Optimization of Tryptophan Metabolism. *Nat. Commun. 11* (1), 4880.
(42) Mateus, C., and Avery, S. V. (2000) Destabilized Green Fluorescent Protein for Monitoring Dynamic Changes in Yeast Gene Expression with Flow Cytometry. *Yeast 16* (14), 1313−1323.
(43) Blazeck, J., Garg, R., Reed, B., and Alper, H. S. (2012) Controlling Promoter Strength and Regulation in Saccharomyces Cerevisiae Using Synthetic Hybrid Promoters. *Biotechnol. Bioeng. 109* (11), 2884−2895.
(44) Erkine, A. M., Magrogan, S. F., Sekinger, E. A., and Gross, D. S. (1999) Cooperative Binding of Heat Shock Factor to the Yeast HSP82 Promoter in Vivo and in Vitro. *Mol. Cell. Biol. 19* (3), 1627−1639.
(45) de Boer, C. G., and Hughes, T. R. (2012) YeTFaSCo: A Database of Evaluated Yeast Transcription Factor Sequence Specificities. *Nucleic Acids Res. 40*, D169−D179.

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477

(46) Dobi, K. C., and Winston, F. (2007) Analysis of Transcriptional Activation at a Distance in Saccharomyces Cerevisiae. *Mol. Cell. Biol. 27* (15), 5575−5586.
(47) Cormack, B. P., Bertram, G., Egerton, M., Gow, N. A. R., Falkow, S., and Brown, A. J. P. (1997) Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in Candida albicans. *Microbiology 143*, 303−311.
(48) Peng, B., Williams, T. C., Henry, M., Nielsen, L. K., and Vickers, C. E. (2015) Controlling Heterologous Gene Expression in Yeast Cell Factories on Different Carbon Substrates and across the Diauxic Shift: A Comparison of Yeast Promoter Activities. *Microb. Cell Fact. 14*, 91.
(49) Adams, J., Behnke, M., Chen, S., Cruickshank, A. A., Dick, L. R., Grenier, L., Klunder, J. M., Ma, Y. T., Plamondon, L., and Stein, R. L. (1998) Potent and Selective Inhibitors of the Proteasome: Dipeptidyl Boronic Acids. *Bioorg. Med. Chem. Lett. 8* (4), 333−338.
(50) Samant, R. S., Livingston, C. M., Sontag, E. M., and Frydman, J. (2018) Distinct Proteostasis Circuits Cooperate in Nuclear and Cytoplasmic Protein Quality Control. *Nature 563* (7731), 407−411.
(51) Spokoini, R., Moldavski, O., Nahmias, Y., England, J. L., Schuldiner, M., and Kaganovich, D. (2012) Confinement to Organelle-Associated Inclusion Structures Mediates Asymmetric Inheritance of Aggregated Protein in Budding Yeast. *Cell Rep. 2* (4), 738−747.
(52) Liebman, S. W., and Chernoff, Y. O. (2012) Prions in Yeast. *Genetics 191* (4), 1041−1072.
(53) Siewers, V., Chen, X., Huang, L., Zhang, J., and Nielsen, J. (2009) Heterologous Production of Non-Ribosomal Peptide LLD-ACV in Saccharomyces Cerevisiae. *Metab. Eng. 11* (6), 391−397.
(54) Wang, R., and Brattain, M. G. (2007) The Maximal Size of Protein to Diffuse through the Nuclear Pore Is Larger than 60 kDa. *FEBS Lett. 581* (17), 3164−3170.
(55) Popken, P., Ghavami, A., Onck, P. R., Poolman, B., and Veenhoff, L. M. (2015) Size-Dependent Leak of Soluble and Membrane Proteins through the Yeast Nuclear Pore Complex. *Mol. Biol. Cell 26* (7), 1386−1394.
(56) Melnik, A., Cappelletti, V., Vaggi, F., Piazza, I., Tognetti, M., Schwarz, C., Cereghetti, G., Ahmed, M. A., Soste, M., Matlack, K., de Souza, N., Csikasz-Nagy, A., and Picotti, P. (2020) Comparative Analysis of the Intracellular Responses to Disease-Related Aggregation-Prone Proteins. *J. Proteomics 225*, 103862.
(57) Jensen, N. B., Strucko, T., Kildegaard, K. R., David, F., Maury, J., Mortensen, U. H., Forster, J., Nielsen, J., and Borodina, I. (2014) EasyClone: Method for Iterative Chromosomal Integration of Multiple Genes Saccharomyces Cerevisiae. *FEMS Yeast Res. 14* (2), 238−248.
(58) Inoue, H., Nojima, H., and Okayama, H. (1990) High Efficiency Transformation of Escherichia Coli with Plasmids. *Gene 96* (1), 23−28.
(59) Jakociunas, T., Rajkumar, A. S., Zhang, J., Arsovska, D., Rodriguez, A., Jendresen, C. B., Skjødt, M. L., Nielsen, A. T., Borodina, I., Jensen, M. K., et al. (2015) CasEMBLR: Cas9-Facilitated Multiloci Genomic Integration of in Vivo Assembled DNA Parts in Saccharomyces Cerevisiae. *ACS Synth. Biol. 4* (11), 1226−1234.
(60) Mikkelsen, M. D., Buron, L. D., Salomonsen, B., Olsen, C. E., Hansen, B. G., Mortensen, U. H., and Halkier, B. A. (2012) Microbial Production of Indolylglucosinolate through Engineering of a Multi-Gene Pathway in a Versatile Yeast Expression Platform. *Metab. Eng. 14* (2), 104−111.
(61) Gietz, R. D., and Schiestl, R. H. (2007) Quick and Easy Yeast Transformation Using the LiAc/SS Carrier DNA/PEG Method. *Nat. Protoc. 2* (1), 35−37.
(62) Metsalu, T., and Vilo, J. (2015) ClustVis: A Web Tool for Visualizing Clustering of Multivariate Data Using Principal Component Analysis and Heatmap. *Nucleic Acids Res. 43* (W1), W566−W570.
(63) Vizcaíno, J. A., Côté, R. G., Csordas, A., Dianes, J. A., Fabregat, A., Foster, J. M., Griss, J., Alpi, E., Birim, M., Contell, J., O'Kelly, G., Schoenegger, A., Ovelleiro, D., Pérez-Riverol, Y., Reisinger, F., Ríos, D., Wang, R., and Hermjakob, H. (2012) The PRoteomics IDEntifications (PRIDE) Database and Associated Tools: Status in 2013. *Nucleic Acids Res. 41*, D1063−D1069.
(64) MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: An Open Source Document Editor for Creating and Analyzing Targeted Proteomics Experiments. *Bioinformatics 26* (7), 966−968.
(65) Wallace, E. W. J., Kear-Scott, J. L., Pilipenko, E. V., Schwartz, M. H., Laskowski, P. R., Rojek, A. E., Katanski, C. D., Riback, J. A., Dion, M. F., Franks, A. M., Airoldi, E. M., Pan, T., Budnik, B. A., and Drummond, D. A. (2015) Reversible, Specific, Active Aggregates of Endogenous Proteins Assemble upon Heat Stress. *Cell 162* (6), 1286−1298.
(66) Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods 9* (7), 671−675.
(67) Monteiro, P. T., Oliveira, J., Pais, P., Antunes, M., Palma, M., Cavalheiro, M., Galocha, M., Godinho, C. P., Martins, L. C., Bourbon, N., Mota, M. N., Ribeiro, R. A., Viana, R., Sá-Correia, I., and Teixeira, M. C. (2020) YEASTRACT+: A Portal for Cross-Species Comparative Genomics of Transcription Regulation in Yeasts. *Nucleic Acids Res. 48* (D1), D642−D649.
(68) Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P. D (2019) PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res. 47*, D419−D426.

https://dx.doi.org/10.1021/acssynbio.0c00476
ACS Synth. Biol. 2021, 10, 466−477