Identification and treatment of heme depletion attributed to overexpression of a lineage of evolved P450 monooxygenases

Joshua K. Michener[a,1], Jens Nielsen[b], and Christina D. Smolke[c,2]

[a]Department of Bioengineering, California Institute of Technology, Pasadena, CA 91125; [b]Department of Chemical and Biological Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden; and [c]Department of Bioengineering, Stanford University, Stanford, CA 94305

Edited by Arnold L. Demain, Drew University, Madison, NJ, and approved October 7, 2012 (received for review July 18, 2012)

Recent advances in metabolic engineering have demonstrated that microbial biosynthesis can provide a viable alternative to chemical synthesis for the production of bulk and fine chemicals. Introduction of a new biosynthetic pathway typically requires the expression of multiple heterologous enzymes in the production host, which can impose stress on the host cell and, thereby, limit performance of the pathway. Unfortunately, analysis and treatment of the host stress response can be difficult, because there are many sources of stress that may interact in complex ways. We use a systems biological approach to analyze the stress imposed by expressing different enzyme variants from a lineage of soluble P450 monooxygenases, previously evolved for heterologous activity in *Saccharomyces cerevisiae*. Our analysis identifies patterns of stress imposed on the host by heterologous enzyme overexpression that are consistent across the evolutionary lineage, ultimately implicating heme depletion as the major stress. We show that the monooxygenase evolution, starting from conditions of either high or low stress, caused the cellular stress to converge to a common level. Overexpression of rate-limiting enzymes in the endogenous heme biosynthetic pathway alleviates the stress imposed by expression of the P450 monooxygenases and increases the enzymatic activity of the final evolved P450 by an additional 2.3-fold. Heme overexpression also increases the total activity of an endogenous cytosolic heme-containing catalase but not a heterologous P450 that is membrane-associated. This work demonstrates the utility of combining systems and synthetic biology to analyze and optimize heterologous enzyme expression.

systems biology | yeast | CYP102A1 | cytochrome P450

The burgeoning field of metabolic engineering offers the promise of efficient, controlled, scalable production of both fine and bulk chemicals (1–4). This approach is particularly important when seeking to make complex molecules with defined stereochemistry, because such molecules are difficult to synthesize chemically. For example, plant secondary metabolites are a rich source of pharmaceuticals but are difficult to produce on an industrial scale. Total chemical synthesis of complex metabolites is prohibitively costly (5), and extraction from plants can be unpredictable (6). Microbial production of plant secondary metabolites offers a powerful alternative to traditional extraction methods (3). However, these biosynthetic pathways can be very complex and, therefore, require simultaneous expression of many heterologous enzymes in the production host. For example, the complete synthesis of morphine from tyrosine requires a total of 14 separate enzymatic reactions (7). These enzymes can potentially interact with each other and with the host cell in undesirable ways (8, 9), and the longer the biosynthetic pathway the more opportunities arise for such interactions. Understanding and alleviating these harmful interactions can significantly improve the pathway productivity and yield (10, 11).

There are many different potential stresses resulting from heterologous enzyme expression. These stresses range from the predictable, including the common stresses attributable to heterologous protein production (12) and byproduct toxicity (13, 14), to the novel, such as the specific depletion of glycyl-tRNA in strains of *Escherichia coli* that overexpress spider silk proteins (15). Individually analyzing each source of stress would be a lengthy process requiring an exhaustive and accurate list of potential stressors. As an alternative to such a bottom-up approach, researchers have used global analyses to measure the host's response to the induced stress, looking, for example, at changes in transcript (16) or protein levels (17). These approaches offer an unbiased analysis of the system and have the potential to discover unexpected sources of stress. However, in many cases, a top-down analysis will identify too many cellular responses. As a result, selecting the cause that is the best target for treatment can still require either a lengthy search of potential targets (10) or intuition regarding likely targets (18). Instead, inspiration can be taken from inverse metabolic engineering, where researchers frequently compare multiple strains with varying levels of productivity to identify the components responsible for the observed variation in productivity (19–21). By analyzing multiple enzyme or pathway variants, rather than only the final optimized strain, a search can quickly be narrowed to identify those select few stress responses that show a consistent pattern across the entire set of variants. Finally, once the sources of stress are identified, pathway productivity can be increased by treating the key stressor, either by eliminating the source of stress or by augmenting the host's ability to respond to the stress.

In this work, we consider the stress imposed by heterologous overexpression of a P450 monooxygenase in *Saccharomyces cerevisiae* (Fig. 1 *A* and *B*). The yeast *S. cerevisiae* is a powerful platform for the expression of a wide range of biosynthetic pathways, particularly those involving membrane-bound proteins such as eukaryal P450s (22). P450s are a major family of enzymes in plant secondary metabolism, participating in the biosynthesis of metabolites ranging from alkaloids and terpenoids to hormones and lipids (23) and, as a result, are important heterologous targets for metabolic engineering. For example, the complete biosynthesis of hydrocortisone, taxol, or vincristine would require the heterologous expression of four (24), eight (25), and at least three P450s (26), respectively. However, although yeast is a good host for heterologous P450 expression, achieving efficient bioconversions using native plant P450s is still challenging (27). This challenge has motivated a search for more convenient alternatives, such as adapting a tractable bacterial P450, CYP102A1, to catalyze the desired reaction (28, 29).

The CYP102A1 variant used in this study had undergone seven rounds of evolution in vivo, during which a synthetic RNA

Author contributions: J.K.M., J.N., and C.D.S. designed research; J.K.M. performed research; J.K.M., J.N., and C.D.S. analyzed data; and J.K.M., J.N., and C.D.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Present address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138.

[2]To whom correspondence should be addressed. E-mail: csmolke@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1212287109/-/DCSupplemental.

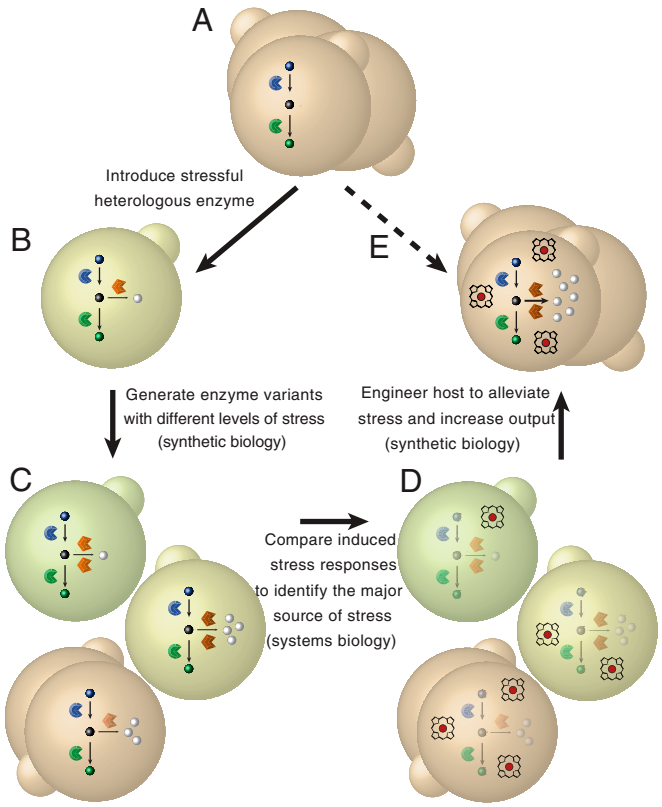

Fig. 1. Systems analysis of multiple pathway variants can identify limiting stresses. (*A*) A wild-type host cell has low stress. Circles, metabolites; polygons, enzymes. (*B*) Addition of a heterologous pathway places a stress on the host that limits growth and pathway productivity. (*C* and *D*) A systems analysis of multiple enzyme or pathway variants (*C*) can identify consistent patterns in gene expression that may indicate major sources of stress, such as heme depletion (*D*). (*E*) Alleviating the stress improves the pathway productivity, yielding a host and pathway that function well in combination with each other.

biosensor was used to screen for improvements in the enzyme's ability to demethylate caffeine to theophylline (29). As a result, we have an entire lineage of enzyme mutants that exhibit a range of enzyme activities and impose different levels of stress on the host (Fig. 1 *C* and *D*). We performed a systems analysis on multiple members of this lineage to track changes in the host's response and, thereby, identify a major cause of cellular stress. Our global analysis indicated that enzyme overexpression starves the cells for heme, leading to a significant change in the host physiology. Targeted modifications increased the host's heme biosynthetic capacity, raising the heme level by up to 60-fold and the product concentration by 2.3-fold (Fig. 1*E*). These results demonstrate that systems and synthetic biology can be successfully combined to introduce and optimize a heterologous enzyme (29), analyze the resulting stress imposed on the host, identify the most significant source of stress, and alleviate that source of stress to ultimately increase the productivity of a biocatalytic process.

Results

Selection for Increased Enzymatic Activity at a Low Expression Level Does Not Result in Increased Activity at High Expression. We previously evolved a caffeine demethylase (CDM) in vivo using a biosensor based on a synthetic RNA switch (29). Through seven rounds of mutagenesis and screening, we improved the enzymatic activity by more than 30-fold. The caffeine demethylase underwent five rounds of evolution while expressed from a high-copy 2μ plasmid to yield the intermediate enzyme yCDM6. We then switched to a low-copy centromeric plasmid and saw no change in product accumulation. Two further rounds of evolution on a low-copy plasmid gave our final enzyme, yCDM8. However, when characterizing the set of enzyme variants, we discovered an unexpected relationship between the selection conditions and the enzyme performance. When we moved the final enzymes, yCDM7 and yCDM8, from low- to high-copy plasmids, we found that they were only marginally more active than yCDM6 expressed from a high-copy plasmid (Fig. 2). In contrast, the enzymes that were evolved on high-copy plasmids, yCDM2 through yCDM6, were equally active when expressed from high- and low-copy plasmids. We hypothesized that the enzymes that were identified on low-copy plasmids, yCDM7 and yCDM8, had evolved to impose a stress on the cell that was tolerable when the enzyme was expressed at low levels but deleterious when the enzyme expression level was increased.

Unfortunately, there are too many potential sources of stress associated with heterologous P450 expression to exhaustively verify each one. For example, the enzyme binds heme as a cofactor, sequestering it from other endogenous enzymes. The monooxygenase consumes NADPH and may produce toxic reaction byproducts. The enzyme may be uncoupled, so many of the electrons taken from the oxidation of NADPH are not transferred to the substrate but instead produce reactive oxygen species (30). Additionally, the relaxed substrate selectivity of the monooxygenase might allow it to oxidize endogenous compounds, consuming important metabolites and producing potentially toxic byproducts. A careful characterization of each possible source of stress would be infeasible.

Microarray Analysis on a Lineage of Evolved P450s Identifies Heme Depletion As the Major Cellular Stress. Rather than individually testing for each potential source of stress, we instead used DNA microarrays to identify the source of the stress by analyzing the global transcriptional response to monooxygenase overexpression. We selected a total of eight strains for analysis: 1–6, yCDM1, yCDM6, and yCDM8, each expressed from high- and low-copy plasmids; 7, the catalytically inactive variant yCDM1-A264H (31) expressed from a high-copy plasmid; and 8, an empty high-copy plasmid. We then used principal component analysis (PCA) to identify common patterns in expression across the different strains. Although previous metabolic engineering analyses have treated

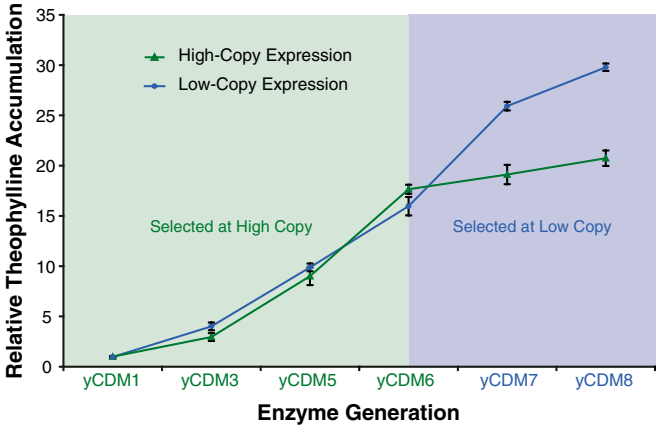

Fig. 2. Comparison of theophylline accumulation for high- and low-copy enzyme expression. The early enzymes, through yCDM6, were identified based on activity when expressed from a high-copy plasmid in yeast. The last two enzymes, yCDM7 and yCDM8, were evolved on a low-copy plasmid backbone. Cells containing each of the enzymes, either at high or low copy, were grown in the presence of 1 mM caffeine. Theophylline accumulation was assayed in the supernatant after 24 h. The data are all normalized to yCDM1 at high copy (relative accumulation, 1.0). The error bars show ±1 SD, calculated from three biological replicates, and the lines are a guide for the eye.

the genes as variables and the samples as observations, applying this analysis method to our data did not produce useful results. Instead, we did the reverse, treating the samples as variables and the genes as observations. This alternative method of PCA has been applied to time-series data to identify patterns in expression over time (32), and we used a similar approach to identify patterns over evolutionary time. By analyzing the genes that are most strongly associated with a given pattern of gene expression, we can identify the specific source of stress that elicits the pattern. Ultimately, this pattern can be interpreted as a measure of the magnitude of the associated stress, because a stronger transcriptional response suggests a greater level of stress.

The loadings for the first principal component identified a pattern of expression that explained ~53% of the variability among samples (Fig. 3). Next, the individual genes were given scores that indicate how well the observed mRNA levels match this pattern of expression (Fig S1*A*). We analyzed the genes with the highest magnitude scores to identify transcription factors whose regulatory targets were overrepresented in the subpopulation (33). These transcription factors were likely involved in coordinating the stress response, and the signals that they respond to are likely causes of stress. The transcription factors identified in this analysis include Rcs1, which responds to iron starvation, as well as Hap1 and Rox1, which regulate genes in a heme-dependent manner. As predicted, a closer look at several Rcs1-, Rox1-, and Hap1-dependent genes shows a similar pattern to the loadings of the first principal component (Fig. S1*B*), suggesting that the loadings are associated with the heme and iron levels, with lower values of the loading correlating to lower intracellular levels of heme.

Increased Heme Biosynthesis Raises both the Intracellular Heme Level and Monooxygenase Activity. Having implicated heme limitation as the major source of cellular stress during heterologous overexpression of the evolved P450 variants, we next sought to restore the intracellular heme level to its native range. In bacteria, the first committed step in heme biosynthesis, producing δ-aminolevulinic acid (ALA), is limiting and feeding additional ALA can increase the expression of functional hemoproteins (34). However, ALA synthesis is not a rate-limiting step in heme biosynthesis in yeast

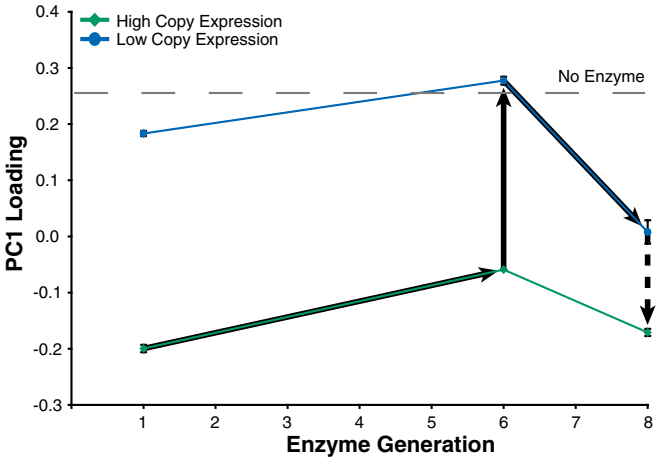

Fig. 3. Microarray analysis identifies heme depletion as the major cellular stress. mRNA levels were measured in triplicate for cells containing different enzyme variants, with either high or low expression. The mRNA expression data were then analyzed using principal component analysis. The loadings for the first principal component are plotted for each of the six enzymes assayed. The dashed black line denotes the loading for the no-enzyme control. The black arrows show the evolutionary trajectory taken to identify the enzymes. The error bars show ±1 SD, calculated from three biological replicates, and the green and blue lines are a guide for the eye.

(35). Therefore, to increase cellular heme levels, we tested a range of conditions designed to increase flux through the heme biosynthetic pathway (Fig. S2*A*): overexpression of a single rate-limiting enzyme in heme biosynthesis, *HEM3*; overexpression of *HEM3* in combination with the next two rate-limiting enzymes, *HEM2* and *HEM12*; or overexpression of all three enzymes in addition to feeding iron (II) and ALA.

We measured total cellular heme and porphyrin levels in yeast with different levels of heme biosynthesis, using combinations of gene overexpression and substrate feeding, and heme use, by varying the P450 expression level. For a given level of heme use, increasing the capacity of the heme biosynthetic pathway led to an increase in the total heme concentration (Fig. S2*B*). However, we only observed large increases in total heme content when heme biosynthesis and use were both elevated (Fig. S2*B*). As expected, increasing the capacity of the initial stages of heme biosynthesis led to accumulation of biosynthetic intermediates (Fig. S2*C*). However, strong overexpression of a heme-containing enzyme led to lower porphyrin levels, suggesting that the cell is actively controlling the conversion of porphyrins to heme, presumably by transcriptional repression of *HEM13* in response to increasing levels of free heme (36). When we measured *HEM13* mRNA levels using quantitative (q)RT-PCR, we found that *HEM13* mRNA expression increased with increasing P450 expression levels and decreased below background levels when both heme and the P450 were highly expressed (Fig. S2*D*). These results are consistent with our hypothesis that P450 overexpression starves the cell of free heme but overexpression of the heme biosynthetic pathway can restore the free heme level.

Heme Depletion Limits the Activity of Cytosolic Hemoproteins. Having shown that increasing the cells' capacity to produce heme leads to an increase in the total heme level, we next considered the effect this extra heme has on production of theophylline by the heterologous P450. The effect of overexpressing heme biosynthesis is not obvious: although our data suggest that heme is limiting in cells that overexpress the P450 variants, total product accumulation depends on a large number of factors. Heme is costly for the cell to produce, and free heme is toxic to the cell. The benefit of alleviating the stress caused by heme limitation may be outweighed by the burden of producing the necessary heme.

We expressed the evolved P450 and the heme-overexpression constructs from high-copy plasmids in *S. cerevisiae* and measured enzymatic production of theophylline. When we increased the host's capacity to produce heme, we observed an increase in the amount of theophylline produced, averaging a 2.3-fold improvement at the optimal heme level (Fig. 4*A*). Unfortunately, the experiments showed significant day-to-day and culture-to-culture variability, likely a result of varying copy numbers of the two plasmids. We also discovered that cells expressing both the P450 and the heme-overexpression constructs grew more slowly than cells with the enzyme and an empty plasmid (Fig. S3*A*). Given the stress and variability associated with plasmid-based heme overexpression, we next sought to optimize the heme-overexpression constructs to minimize this stress. We integrated the heme-overexpression constructs into the yeast genome, eliminating the need to maintain a high-copy plasmid while also presumably lowering the expression of the genes. Unexpectedly, we found that cells with the integrated copies of the overexpression constructs produced more heme than cells with those same constructs on high-copy plasmids. The strain with *HEM2*, *HEM3*, and *HEM12* integrated simultaneously grew very slowly (Fig. S3*B*) and produced more than 200 times as much porphyrin (Fig. S3*C*). However, the strain with only *HEM3* integrated exhibited no growth defect (Fig. S3*B*) and, in combination with yCDM8 expressed from a high-copy plasmid, produced the highest theophylline titers yet observed, reaching 42% conversion (Fig. 4*B*). Notably, *HEM3* overexpression had no effect on theophylline production, either positive or negative, when yCDM8 was

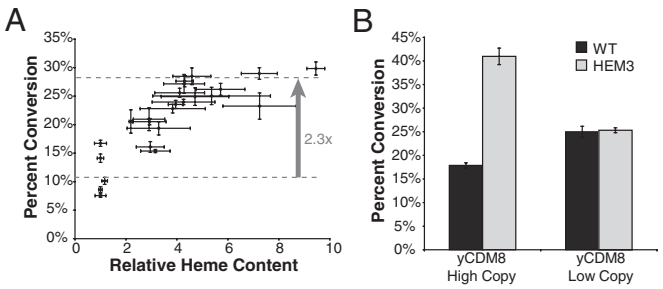

Fig. 4. Increasing the total cellular heme level leads to an increase in total enzymatic activity. (*A*) Cells expressing yCDM8 from a high-copy plasmid were cotransformed with high-copy plasmids containing *HEM3* or *HEM2*/3/12. Cultures were grown in the presence of 1 mM caffeine and varying amounts of iron and ALA. After 48 h, the theophylline concentration was measured in the supernatant, and the cell pellets were assayed for heme content. Each point represents the average of three biological replicates, and error bars show ±1 SD. The data shown are concatenated from four separate experiments to account for day-to-day variability. (*B*) Cells expressing yCDM8 from either a high- or low-copy plasmid were transformed into strains with and without an integrated *HEM3*-expression construct. Strains were grown and assayed as described above. Adding ALA to either the WT or *HEM3* strains did not affect activity. The error bars show ±1 SD calculated from three biological replicates.

expressed from a low-copy plasmid (Fig. 4*B*), demonstrating that integrated *HEM3* does not have an inherent deleterious effect and that the benefits of heme overexpression are dependent on the simultaneous overexpression of the hemoprotein yCDM8. As with the plasmid-based heme-overexpression strategy, integrated expression of *HEM3* also led to a decrease in the *HEM13* mRNA levels (Fig. S4*A*).

Finally, we examined whether the activity of other hemoproteins was limited by heme biosynthesis. We tested two representative enzymes: a cytosolic catalase, CTT1, and a membrane-associated P450, yeast codon–optimized CYP2D6 (2). To test CYP2D6 activity, we expressed both the P450 and heme biosynthetic gene(s) from separate high-copy plasmids in a yeast strain with chromosomally integrated copies of three methyltransferases and a P450 reductase partner (CSY409; Table S1). We fed a substrate, norlaudanosoline, and measured accumulation of the product of the P450, salutaridine (2). In contrast to the soluble CYP102A1 variant, CYP2D6 was not heme-limited and showed a slight decrease in activity with increasing concentrations of heme (Fig. S5*A*). We expect that CYP2D6 expression is limited instead by the ability of the cell to accommodate large quantities of functional membrane proteins (37). However, when we coexpressed the soluble catalase CTT1 with the heme biosynthetic genes, yeast cell lysates exhibited nearly twice the catalase activity when the total heme level was increased (Figs. S5*B* and S6). For comparison, previous work overexpressing *HEM2* alone led to a 40–60% increase in activity from CTT1 (38). The differing responses of these two enzymes demonstrate that our heme-overexpression strategy is neither limited to the specific P450 we tested nor universally beneficial to hemoprotein expression.

Heme Depletion Limits Total Enzyme Expression. We next sought to understand the mechanism underlying the observed increase in enzyme activity as a result of increased heme levels. The increased activity could plausibly be explained either by an increase in total P450 expression or an increase in the specific activity of the enzyme. We previously demonstrated that total P450 expression was roughly constant, irrespective of the enzyme variant or the plasmid copy number (29). We hypothesized that most of the additional proteins produced from the high-copy plasmid were misfolding, leading to increased rates of protein degradation and equivalent steady-state protein levels. When we measured the proteasomal activity of cells expressing high or low levels of the final enzyme yCDM8, we found that cultures expressing yCDM8 from a high-copy plasmid showed increased levels of proteasomal activity compared with cells expressing the same enzyme from a low-copy plasmid (Fig. 5*A*). However, when we increased the heme levels, the proteasomal activity decreased to background levels (Fig. 5*A*) and the total CDM expression increased (Fig. 5*B*). These results suggest that the increase in total enzymatic activity that we observed is attributable to increases in functional enzyme expression rather than specific enzyme activity.

Discussion

In the past, systems biology has generally been applied to metabolic engineering in two distinct fashions. When used in forward metabolic engineering, researchers typically consider a single strain and ask how that strain differs from a control. The difficulty of using a global analysis lies in the sheer volume of data that can be obtained. In one of our engineered strains, for example, the transcription of 953 genes was significantly different from the control ($P < 0.001$), and greater than 20% of the genes measured showed statistically significant variation in at least one of the seven experimental strains. In this work, we adopted approaches from inverse metabolic engineering, where researchers examine multiple improved variants to reverse engineer the causes of the improvement and applied them to a forward engineering problem, to rationally optimize a system. By analyzing multiple closely related enzyme variants, we were able to quickly narrow our focus to only those genes that showed consistent patterns of expression across these variants. This restricted set of genes then allowed us to identify the transcription factors that were producing the coordinated response and the ultimate source of stress. Although, in hindsight, heme depletion is not an unexpected source of stress, this was not the case at the beginning of the project. There were many potential sources of stress, such as formaldehyde toxicity or NADPH depletion, and heme depletion was, at most, only one of several equally likely explanations.

Currently, any indirect analysis of this type is limited by the assumption that cellular stress will be reflected in changes to the levels of endogenous biomolecules, such as proteins or mRNA.

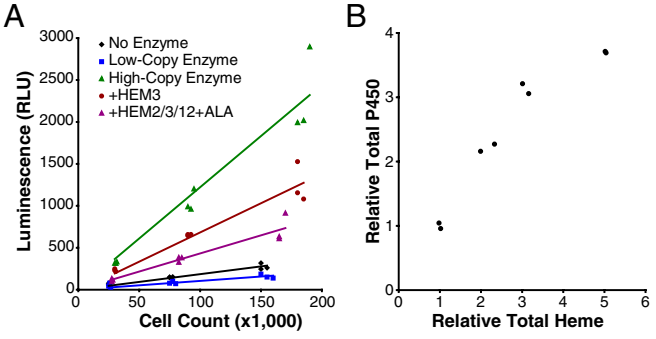

Fig. 5. Heme depletion leads to increased proteasomal activity and low enzyme expression. (*A*) Proteasomal activity of strains expressing different levels of yCDM8. Proteasomal activity was measured using a caged luciferin that is released by the chymotrypsin-like activity of the yeast proteasome. Luminescence is proportional to both the cell number and the specific proteasomal activity. For each condition, three biological replicates were assayed at each of three dilutions. The lines shown are a linear fit to the data. For yCDM8 expressed from a high-copy plasmid, overexpression of heme biosynthetic genes reduces the proteasomal activity. (*B*) Cells expressing yCDM8 from a high copy plasmid were cotransformed with a second heme-overexpression plasmid. Cultures were grown to midlog phase and then assayed for heme content and total P450 expression. Increased enzyme expression correlates with increased heme accumulation. Each data point corresponds to a single biological sample.

SYSTEMS BIOLOGY

When we observe changes in transcript levels, for example, we can generally trust that they result from a disturbance to the host. However, constant transcript levels can mean either that the corresponding stress is absent or that the stress does not produce a change in transcript levels. For example, our global analysis did not identify protein degradation as a major host stress, even though we later found that the proteasomal activity was significantly higher in some strains. This misidentification is likely attributable to posttranslational regulation of the proteasome (39, 40). Conversely, the P450 used in this work produces formaldehyde as a byproduct, and, therefore, we identified formaldehyde toxicity as one potential source of stress. The gene responsible for formaldehyde detoxification, *SFA1*, is not differentially transcribed in the strains that we analyzed. We could only rule out formaldehyde toxicity by verifying in the literature that *SFA1* would be induced were formaldehyde present at high concentrations (41). Global analyses would be more informative if we had a better method of screening through these negative data to separate situations where a lack of response is meaningful, such as *SFA1*, from those where measuring transcript or protein levels does not inform us about the underlying stress, as was the case for proteasomal activity.

Based on our global transcript analysis of a lineage of evolved monooxygenases, we propose that the overexpression of a heme-containing monooxygenase depletes the intracellular pool of heme and the resulting lack of heme places a stress on the cell that limits the total enzymatic activity. High-copy expression of the enzyme sequesters more heme and, therefore, produces a greater stress than low-copy expression. Evolution of the enzyme at high copy reduced the stress, suggesting that the stress imposed a selective pressure to minimize the deleterious effects. However, decreasing the enzyme DNA copy number removed both the stress of heme depletion and the selective pressure to minimize that depletion. Further evolution at low copy led to an increase in enzymatic activity at low copy but also to greater heme use by the monooxygenase. As a result, the enzymes that show improved activity at low expression levels do not show corresponding increases in activity at high expression levels.

Our results demonstrate that heme depletion has several effects on yeast cells. First, lack of heme disrupts the host metabolism. Without heme, the cells show gene expression profiles consistent with anaerobic growth, despite the presence of sufficient oxygen. Second, low heme levels limit production of active P450 enzymes. We propose that when heme levels are low, additional nascent peptides are unable to properly bind heme and, as a result, misfold and are subsequently degraded by the proteasome. Although our results are consistent with this explanation, we cannot rule out an alternate model in which increased proteasomal activity is not directly responsible for limiting P450 expression. In either case, the combination of disrupted aerobiosis and the burden of recycling misfolded protein (42) would reduce the host's growth rate and correspondingly reduce production of the desired compound. In our system, increasing the heme supply, by chromosomally overexpressing an endogenous rate-limiting biosynthetic enzyme, alleviated this source of stress and increased the total enzymatic activity by 2.3-fold. In a system such as ours that is already highly optimized, this level of improvement represents a significant advance (21, 43).

The interplay between heme level, P450 expression, and final product titers is complicated. Focusing only on P450 expression, we see only slight evidence of saturation after a fivefold increase in the heme level (Fig. 5*B*), suggesting that further increases to the heme level could lead to still higher P450 expression. However, product titers reach a point of rapidly diminishing returns after a ~4-fold increase in the heme level (Fig. 4*A*), and the increase in product titer (2.3-fold) is smaller than the increase in P450 expression (~3.8-fold), demonstrating either that enzyme expression is no longer the limiting factor or that an increasing fraction of the enzyme is functionally inactive. We conclude that, in this system, a ~fourfold increase in the heme level represents an optimal balance between the burdens caused by increased heme expression, P450 expression, and enzymatic activity. Therefore, further optimization of the heme level by itself is unlikely to yield significant further improvements.

We have demonstrated that CYP102A1 is not the only enzyme with heme-limited expression in yeast but also that not all hemoproteins are heme-limited. When expression of a soluble hemoprotein is limiting, a similar heme-overexpression strategy may be worthwhile. If optimal hemoprotein expression is required, our heme-overexpression system could be improved by tuning the expression level of the HEM genes. Similarly, although feeding ALA is a straightforward method for increasing precursor availability for the heme biosynthetic pathway, the cost would be prohibitive on an industrial scale, and overexpression of the ALA biosynthetic pathway might be a preferable solution (44). Finally, this coupled strategy of optimizing an enzyme and its host could be applied iteratively, for example, by finding newly beneficial enzyme mutations that take advantage of the increased heme supply or identifying the next limiting stress once the host has sufficient heme.

We anticipate that the approaches described in this work will be generally useful in the optimization of heterologous metabolic pathways. The introduction of novel enzymes into a production host can impose new sources of stress that the host is unable to tolerate. Understanding and alleviating these sources of stress is an important, and frequently unaddressed, challenge in enzyme and pathway engineering. In addition to our specific solution, in which heme overexpression leads to increased expression of a P450 monooxygenase, we propose that a systems analysis of multiple variants of a heterologous pathway, when analyzed in the fashion we have described, will generally provide additional insight into the factors limiting pathway productivity and, therefore, enable further pathway improvement. Such an analysis could study responses at the transcript level, as in this example, or measure other factors, such as protein expression or posttranslational modifications. Finally, after techniques from systems biology are used to identify the key source of stress, methods from synthetic biology can be used to alleviate the stress. Augmenting the host's tolerance to stress, such as by increasing the host's ability to produce heme, requires careful control to alleviate the existing source of stress without replacing it with a new stressor. Too much heme can be worse than too little, for example, and tools from synthetic biology will be particularly useful for this type of delicate optimization. As metabolic engineers build increasingly complex biosynthetic pathways, techniques to optimize the host–pathway interface will become increasingly valuable.

Materials and Methods

Strains and Cultivation. The strains used in this work are derivatives of W303α (MATα leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15). Transcription analysis was conducted in CSY492 (W303 lys2::PTEF-GFP-L2Bulge8-ADH1T). Cultures were grown in shake flasks at 30 °C and 200 rpm in the appropriate synthetic dropout media (Formedium) supplemented with 2% (wt/vol) glucose, an additional 10 mg/L of adenine (Sigma-Aldrich), and 1 mM caffeine (Sigma-Aldrich). Heme-overexpression strains were fed varying amounts of iron (II) citrate and δ-aminolevulinic acid (Sigma-Aldrich). HEM-overexpression constructs were integrated into the lys2 locus of W303. Assays for salutaridine production were fed 4 mM norlaudanosoline (Santa Cruz Biotechnology). Additional details on the plasmid construction process are available in *SI Materials and Methods*.

Metabolite Analysis. Supernatant theophylline and salutaridine were separated by liquid chromatography and assayed by UV-VIS spectroscopy (theophylline) or MS (salutaridine) according to previously published protocols (2). Additional details are available in *SI Materials and Methods*.

DNA Microarray Experiments. Details of the microarray experimental procedures are available in *SI Materials and Methods*.

Proteasomal Activity Measurements. Proteasomal activity was measured using the Proteasome-Glo kit (Promega), according to the manufacturer's directions. Details are available in *SI Materials and Methods*.

Heme Measurements. The intracellular heme concentration was measured using a derivative of the protocol described by Sassa (45). Complete details are included in *SI Materials and Methods*.

Quantitive Western Blots. Quantitative Western blots were performed as described previously (29). Additional details are available in *SI Materials and Methods*.

Catalase Activity Assays. A description of the catalase activity assay is available in *SI Materials and Methods*.

qRT-PCR Measurements. qRT-PCR measurements were performed as described previously (46). Additional details are available in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank R. Kumar for assistance with the microarray experiments and J. L. M. Ruiz and L. Liu, who provided the HEM-overexpression plasmids. This work was supported by the National Science Foundation and Vetenskapsrådet through a Nordic Research Fellowship (to J.K.M.), National Science Foundation Grant CBET-0917638, and an Alfred P. Sloan Foundation fellowship (to C.D.S.).

1. Atsumi S, Hanai T, Liao JC (2008) Non-fermentative pathways for synthesis of branched-chain higher alcohols as biofuels. *Nature* 451(7174):86–89.
2. Hawkins KM, Smolke CD (2008) Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae. *Nat Chem Biol* 4(9):564–573.
3. Ro DK, et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* 440(7086):940–943.
4. Yim H, et al. (2011) Metabolic engineering of Escherichia coli for direct production of 1,4-butanediol. *Nat Chem Biol* 7(7):445–452.
5. Rice KC (1980) Synthetic opium alkaloids and derivatives. A short total synthesis of (.+-.)-dihydrothebainone, (.+-.)-dihydrocodeinone, and (.+-.)-nordihydrocodeinone as an approach to a practical synthesis of morphine, codeine, and congeners. *J Org Chem* 45(15):3135–3137.
6. Hale V, Keasling JD, Renninger N, Diagana TT (2007) Microbially derived artemisinin: A biotechnology solution to the global problem of access to affordable antimalarial drugs. *Am J Trop Med Hyg* 77(6 Suppl):198–202.
7. Liscombe DK, Facchini PJ (2008) Evolutionary and cellular webs in benzylisoquinoline alkaloid biosynthesis. *Curr Opin Biotechnol* 19(2):173–180.
8. Ro DK, et al. (2008) Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. *BMC Biotechnol* 8:83.
9. Kennedy CJ, Boyle PM, Waks Z, Silver PA (2009) Systems-level engineering of nonfermentative metabolism in yeast. *Genetics* 183(1):385–397.
10. Lee KH, Park JH, Kim TY, Kim HU, Lee SY (2007) Systems metabolic engineering of Escherichia coli for L-threonine production. *Mol Syst Biol* 3:149.
11. Park JH, Lee KH, Kim TY, Lee SY (2007) Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc Natl Acad Sci USA* 104(19):7797–7802.
12. Goff SA, Goldberg AL (1985) Production of abnormal proteins in E. coli stimulates transcription of lon and other heat shock genes. *Cell* 41(2):587–595.
13. Zhu MM, Lawman PD, Cameron DC (2002) Improving 1,3-propanediol production from glycerol in a metabolically engineered Escherichia coli by reducing accumulation of sn-glycerol-3-phosphate. *Biotechnol Prog* 18(4):694–699.
14. Zhu MM, Skraly FA, Cameron DC (2001) Accumulation of methylglyoxal in anaerobically grown Escherichia coli and its detoxification by expression of the Pseudomonas putida glyoxalase I gene. *Metab Eng* 3(3):218–225.
15. Xia XX, et al. (2010) Native-sized recombinant spider silk protein produced in metabolically engineered Escherichia coli results in a strong fiber. *Proc Natl Acad Sci USA* 107(32):14059–14063.
16. Kizer L, Pitera DJ, Pfleger BF, Keasling JD (2008) Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl Environ Microbiol* 74(10):3229–3241.
17. Han MJ, Yoon SS, Lee SY (2001) Proteome analysis of metabolically engineered Escherichia coli producing Poly(3-hydroxybutyrate). *J Bacteriol* 183(1):301–308.
18. Choi JH, Lee SJ, Lee SJ, Lee SY (2003) Enhanced production of insulin-like growth factor I fusion protein in Escherichia coli by coexpression of the down-regulated genes identified by transcriptome profiling. *Appl Environ Microbiol* 69(8):4737–4742.
19. Askenazi M, et al. (2003) Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nat Biotechnol* 21(2):150–156.
20. Bro C, Knudsen S, Regenberg B, Olsson L, Nielsen J (2005) Improvement of galactose uptake in Saccharomyces cerevisiae through overexpression of phosphoglucomutase: Example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* 71(11):6465–6472.
21. Hong KK, Vongsangnak W, Vemuri GN, Nielsen J (2011) Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proc Natl Acad Sci USA* 108(29):12179–12184.
22. Siddiqui MS, Thodey K, Trenchard I, Smolke CD (2012) Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS Yeast Res* 12(2):144–170.
23. Mizutani M, Ohta D (2010) Diversification of P450 genes during land plant evolution. *Annu Rev Plant Biol* 61:291–315.
24. Szczebara FM, et al. (2003) Total biosynthesis of hydrocortisone from a simple carbon source in yeast. *Nat Biotechnol* 21(2):143–149.
25. Jennewein S, Wildung MR, Chau M, Walker K, Croteau R (2004) Random sequencing of an induced Taxus cell cDNA library for identification of clones involved in Taxol biosynthesis. *Proc Natl Acad Sci USA* 101(24):9149–9154.
26. Giddings L-A, et al. (2011) A stereoselective hydroxylation step of alkaloid biosynthesis by a unique cytochrome P450 in Catharanthus roseus. *J Biol Chem* 286(19): 16751–16757.
27. Westfall PJ, et al. (2012) Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci USA* 109(3):E111–E118.
28. Dietrich JA, et al. (2009) A novel semi-biosynthetic route for artemisinin production using engineered substrate-promiscuous P450(BM3). *ACS Chem Biol* 4(4):261–267.
29. Michener JK, Smolke CD (2012) High-throughput enzyme evolution in Saccharomyces cerevisiae using a synthetic RNA switch. *Metab Eng* 14(4):306–316.
30. Fasan R, Meharenna YT, Snow CD, Poulos TL, Arnold FH (2008) Evolutionary history of a specialized p450 propane monooxygenase. *J Mol Biol* 383(5):1069–1080.
31. Neeli R, et al. (2005) The dimeric form of flavocytochrome P450 BM3 is catalytically functional as a fatty acid hydroxylase. *FEBS Lett* 579(25):5582–5588.
32. Raychaudhuri S, Stuart JM, Altman RB (2000) Principal components analysis to summarize microarray experiments: Application to sporulation time series. *Pac Symp Biocomput* 455–466.
33. Oliveira AP, Patil KR, Nielsen J (2008) Architecture of transcriptional regulatory circuits is knitted over the topology of bio-molecular interaction networks. *BMC Syst Biol* 2:17.
34. Kraus JP, Kery V (1997) Method of increasing the yield and heme saturation of cystathione β-synthase. Patent US 5,635,375.
35. Hoffman M, Góra M, Rytka J (2003) Identification of rate-limiting steps in yeast heme biosynthesis. *Biochem Biophys Res Commun* 310(4):1247–1253.
36. Keng T (1992) HAP1 and ROX1 form a regulatory pathway in the repression of HEM13 transcription in Saccharomyces cerevisiae. *Mol Cell Biol* 12(6):2616–2623.
37. Schunck WH, et al. (1991) Comparison of two cytochromes P-450 from Candida maltosa: Primary structures, substrate specificities and effects of their expression in Saccharomyces cerevisiae on the proliferation of the endoplasmic reticulum. *Eur J Cell Biol* 55(2):336–345.
38. Mattoon JR, Bajszar G (1998) Method for enhancing the production of hemoproteins. US Patent 5,824,511.
39. Mason GG, Hendil KB, Rivett AJ (1996) Phosphorylation of proteasomes in mammalian cells. Identification of two phosphorylated subunits and the effect of phosphorylation on activity. *Eur J Biochem* 238(2):453–462.
40. Zhang F, et al. (2003) O-GlcNAc modification is an endogenous inhibitor of the proteasome. *Cell* 115(6):715–725.
41. Wehner EP, Rao E, Brendel M (1993) Molecular structure and genetic regulation of SFA, a gene responsible for resistance to formaldehyde in Saccharomyces cerevisiae, and characterization of its protein product. *Mol Gen Genet* 237(3):351–358.
42. Geiler-Samerotte KA, et al. (2011) Misfolded proteins impose a dosage-dependent fitness cost and trigger a cytosolic unfolded protein response in yeast. *Proc Natl Acad Sci USA* 108(2):680–685.
43. Santos CN, Xiao W, Stephanopoulos G (2012) Rational, combinatorial, and genomic approaches for engineering L-tyrosine production in Escherichia coli. *Proc Natl Acad Sci USA* 109(34):13538–13543.
44. Kang Z, Wang Y, Gu PF, Wang Q, Qi QS (2011) Engineering Escherichia coli for efficient production of 5-aminolevulinic acid from glucose. *Metab Eng* 13(5):492–498.
45. Sassa S (1976) Sequential induction of heme pathway enzymes during erythroid differentiation of mouse Friend leukemia virus-infected cells. *J Exp Med* 143(2):305–315.
46. Babiskin AH, Smolke CD (2011) A synthetic library of RNA control modules for predictable tuning of gene expression in yeast. *Mol Syst Biol* 7:471.