<table>
<thead>
<tr>
<th>Title</th>
<th>Engineered reversal of function in glycolytic yeast promoters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Rajkumar, Arun S.; Özdemir, Emre; Lis, Alicia V.; Schneider, Konstantin; Qin, Jiufu; Jensen, Michael K.; Keasling, Jay D.</td>
</tr>
<tr>
<td>Publication date</td>
<td>2019-05-03</td>
</tr>
<tr>
<td>Type of publication</td>
<td>Article (peer-reviewed)</td>
</tr>
</tbody>
</table>
| Link to publisher's version | https://pubs.acs.org/doi/10.1021/acssynbio.9b00027  
http://dx.doi.org/10.1021/acssynbio.9b00027  
Access to the full text of the published version may require a subscription. |
| Rights | ©2019 American Chemical Society. This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.  
https://pubs.acs.org/page/policy/authorchoice_termsofuse.html |
| Item downloaded from | http://hdl.handle.net/10468/8545 |

Downloaded on 2019-12-24T14:28:53Z
Engineered Reversal of Function in Glycolytic Yeast Promoters

Arun S. Rajkumar, †,‡ Emre Özdemir, † Alicia V. Lis, † Konstantin Schneider, †,§ Jiufu Qin, † Michael K. Jensen, †,‡,∥ and Jay D. Keasling †,†,‡,∥,§,¶

†Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
‡Joint BioEnergy Institute, Emeryville, California 94608, United States
§Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94704, United States
∥Department of Chemical and Biomolecular Engineering & Department of Bioengineering, University of California, Berkeley, California 94720-1462, United States
¶Center for Synthetic Biochemistry, Institute for Synthetic Biology, Shenzhen Institutes of Advanced Technologies, Shenzhen, China

Supporting Information

ABSTRACT: Promoters are key components of cell factory design, allowing precise expression of genes in a heterologous pathway. Several commonly used promoters in yeast cell factories belong to glycolytic genes, highly expressed in actively growing yeast when glucose is used as a carbon source. However, their expression can be suboptimal when alternate carbon sources are used, or if there is a need to decouple growth from production. Hence, there is a need for alternate promoters for different carbon sources and production schemes. In this work, we demonstrate a reversal of regulatory function in two glycolytic yeast promoters by replacing glycolytic regulatory elements with ones induced by the diauxic shift. We observe a shift in induction from glucose-rich to glucose-poor medium without loss of regulatory activity, and strong ethanol induction. Applications of these promoters were validated for expression of the vanillin biosynthetic pathway, reaching production of vanillin comparable to pathway designs using strong constitutive promoters.

KEYWORDS: transcriptome, synthetic promoter, glycolysis, gluconeogenesis, vanillin

The budding yeast Saccharomyces cerevisiae is a popular cell factory platform, with several genetic and molecular tools available to facilitate the production of compounds of interest for the chemical and pharmaceutical industry. The synthesis of increasingly complex molecules, in turn, require successful construction of large and complex heterologous pathways,1–3 which impose an added metabolic burden by competing for required cellular resources with native metabolism. Competition for a common pool of ATP, cofactors, coenzymes and precursors between native and heterologous pathways can adversely affect growth, lengthen fermentation, and decrease product titers and yields.1–4 One way to avoid this problem is to activate the production pathway after the major growth phase. For a typical process with a fermentable carbon source, a good regulatory trigger would be the diauxic shift following fermentation.5,6 However, most promoters commonly used for heterologous pathway construction are glycolytic (PGK1pr, TPI1pr, or TDH3pr) or constitutive (TEF1pr), and lack the desired induction properties. Native yeast promoters have been identified for gene expression under these conditions and used for production,7–11 yet it is unclear if sufficient native promoters with suitable properties exist to be used for large biosynthetic pathways. Therefore, using synthetic promoters with optimal regulatory output at the diauxic shift would be beneficial for cell factory design. This can be achieved by hybrid promoter fusion or mutagenesis,12,13 but the level of control over regulatory signals or expression levels using these approaches may not be sufficient. In previous work, we developed a promoter engineering workflow to engineer yeast promoters responsive to any environmental condition given transcriptome or transcription factor binding site (TFBS) data, and functional genomics for the condition of interest, if available.14 Other studies have also made use of expression data in a similar manner to engineer synthetic promoters and biosensors.15,16 However, generally most work on synthetic Saccharomyces promoters does not focus on rationally engineering native regulatory elements to generate defined regulatory properties, as has been done in other yeasts. Having used this workflow to make promoters inducible by low extracellular pH,14 we now use the same workflow to design promoters inducible by glucose starvation and alternating carbon sources.

Using S. cerevisiae CEN.PK 113-7D transcriptome data generated during the lag (10% glucose consumption, low

Received: January 22, 2019
Published: May 3, 2019
ethanol production), mid-exponential (75% glucose consumed, increasing ethanol production), and post-exponential phases (>99% glucose consumed, start of ethanol consumption), we identified genes with the desired expression profile: strong induction upon glucose depletion (see Figure 1A). Induction under these conditions involves transcriptional activators such as Cat8, Sip4, Rds2, and Adr1 (Table 1), clustered in carbon source response elements (CSREs) in promoters of interest. Therefore, CSREs are relevant regulatory elements to ensure that promoters are activated at the diauxic shift and remain activated past this shift. Transcriptome analysis revealed 4418 differentially expressed genes (DEGs) from the 3-way comparison, with the majority being induced at, or following, glucose depletion and the start of diauxic shift (26 h, as compared to 14 h), including 17 out of 19 genes (excluding ENO2 and PYC2) associated with gluconeogenesis (term GO:0006094) (see Figures 1B and 1C, as well as Table S1 in the Supporting Information).

Next, we scanned the promoters for clustered TFBSs in CSREs. However, reported sequences for their binding sites contain palindromic or inverted repeats, as well as stretches of ambiguous bases. For further confirmation, we analyzed the discovered TFBSs for overlap with CSREs reported in the literature (see Table 1, as well as Table S2 in the Supporting Information). From this analysis, we selected CSREs from PCK1 and FBP1 as candidate elements for synthetic promoter engineering based on three criteria: (i) promoters had a strong induction and appeared in both gene ontology mappings (see Table 1, as well as Figure S1 in the Supporting Information); (ii) the CSREs have been experimentally validated in the past in multiple studies; and (iii) while other genes had a higher fold-induction, e.g., MLS1, the CSREs from FBP1 and PCK1 promoters also contained the most overlapping TFBSs.
Table 1. Genes Associated with Carbon Metabolism 
Upregulated between the Log Phase and Diauxic Shift

<table>
<thead>
<tr>
<th>gene</th>
<th>fold-induction</th>
<th>binding sites in CSREs</th>
<th>references for binding sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annotated “Glucogenesis” by Gene Ontology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCK1</td>
<td>472.97</td>
<td>Adr1, Cat8, Sip4, Rds2</td>
<td>18, 20, 23, 24</td>
</tr>
<tr>
<td>FBP1</td>
<td>52.67</td>
<td>Adr1, Cat8, Rds2</td>
<td>18, 20, 24</td>
</tr>
<tr>
<td>MDH2</td>
<td>6.55</td>
<td>Adr1, Cat8, Rds2</td>
<td>18, 20, 23</td>
</tr>
<tr>
<td>PYC1</td>
<td>1.87</td>
<td>Cat8</td>
<td>18, 20</td>
</tr>
<tr>
<td>GPM2</td>
<td>1.74</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>TPI2</td>
<td>1.46</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>Annotated “Carbon Metabolism Process” by GO_slim</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLS1</td>
<td>650.68</td>
<td>Adr1, Cat8, Rds2</td>
<td>18, 20, 24</td>
</tr>
<tr>
<td>PCK1</td>
<td>472.97</td>
<td>Adr1, Cat8, Sip4, Rds2</td>
<td>18, 20, 23, 24</td>
</tr>
<tr>
<td>ICL1</td>
<td>145.86</td>
<td>Cat8, Sip4</td>
<td>18, 19, 20, 23</td>
</tr>
<tr>
<td>YGI1</td>
<td>129.24</td>
<td>Adr1, Cat8</td>
<td>23</td>
</tr>
<tr>
<td>FBP1</td>
<td>52.67</td>
<td>Adr1, Cat8, Rds2</td>
<td>18, 20, 24</td>
</tr>
<tr>
<td>MAL12</td>
<td>28.48</td>
<td>Cat8</td>
<td>23</td>
</tr>
<tr>
<td>MAL11</td>
<td>29.81</td>
<td>Cat8</td>
<td>23</td>
</tr>
<tr>
<td>CAT8</td>
<td>22.31</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>MAL32</td>
<td>16.89</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>GAL3</td>
<td>13.67</td>
<td>Cat8</td>
<td>23</td>
</tr>
<tr>
<td>MDH2</td>
<td>6.55</td>
<td>Adr1, Cat8, Rds2</td>
<td>18, 20, 23</td>
</tr>
<tr>
<td>DAK2</td>
<td>6.67</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>PYC1</td>
<td>1.87</td>
<td>Cat8</td>
<td>18, 20</td>
</tr>
<tr>
<td>Fold-Induction of Other Native Promoters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADH2</td>
<td>10.93</td>
<td>Cat8, Sip4</td>
<td>18, 19, 21</td>
</tr>
<tr>
<td>TDIH3</td>
<td>0.163</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>TEF1</td>
<td>0.762</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>PGK1</td>
<td>0.192</td>
<td>Adr1</td>
<td>23</td>
</tr>
<tr>
<td>TPI1</td>
<td>0.0732</td>
<td>none</td>
<td></td>
</tr>
</tbody>
</table>

This potential redundancy of function would allow for more reliable induction and ensure that the promoter activity would remain activated past the diauxic shift.20,21,23 Finally, since native PCK1 and FBP1 promoters have relative high expression levels during the early exponential phase (see Table S1 in the Supporting Information), these native promoters did not themselves have desired glucogenenic responses.

These CSREs were used to engineer a 200 bp upstream activating sequence (UAS) from the glycolytic TDH3 promoter (TDH3pr). We chose this UAS to test the reversal of regulatory function by TFBS exchange, because its main regulatory elements are well-annotated and confirmed by ChIP-chip studies (Figures 1D and 1E).24 In particular, the regulatory elements are well-annotated and confirm the core promoter (TDH3pr) retained 60% of the activity of TDH3pr (see Figure 1D, as well as Figures S2 and S3 in the Supporting Information). Replacing both Gcr1 sites and surrounding neutral sequence with CSREs from either PCK1pr or FBP1pr to yield promoters UAS349-PCK1 and UAS349-FBP1 (Figure 1D) did indeed shift the induction trigger of the UAS from a glucose-rich medium to a glucose-depleted medium, with the induction of UAS349-PCK1 and UAS349-FBP1 triggered by glucose depletion, after ~24 h of culture. Importantly, both promoters had a stronger output than not only UAS349 and but also TDH3pr—one of the strongest yeast promoters in use—which was sustained as long as glucose was not replenished. The replacement of Gcr1 sites with CSREs was sufficient and necessary for the induction. When they were exchanged with Gis1 (a TF implicated in upregulating genes at the stationary phase) binding sites instead (UAS349-Gis1),27 the promoter did not induce upon glucose starvation and remained weak throughout the assay period (Figure 1D). Ultimately, UAS349-PCK1 and UAS349-FBP1 were up to 3-fold stronger than commonly used constitutive promoters over long time periods (Figure 1E). A stronger switch in induction was achieved when the Gcr1 site exchange was made for the triose phosphate isomerase promoter (TP11pr) with CSREs from FBP1pr (TP11-FBP1h; see Figure 1F, as well as Figures S2 and S3).28 Although the orientation, spacing, and sequence context of Gcr1 sites in both TDH3pr and TP11pr were different, we observed near-identical induction behavior. It is likely the change in induction is independent of the local context of Gcr1 sites, provided that no other neighboring TFBSs are modified (as was the case for TDH3pr and TP11pr), and that their substitution could be applied to engineer induction switching in any glycolytic promoter. Finally, the synthetic promoters are stronger when grown on ethanol as a carbon source (Figure S4 in the Supporting Information). The lack of glucose repression, and the higher levels of ethanol (2–3-fold), compared to those built up during the growth on glucose, probably induce them to higher levels.

To demonstrate the usefulness of our synthetic promoters, we applied them for the design of yeast cell factories for vanillin-β-glucoside biosynthesis (Figure 2A).29 The choice of product was motivated by the fact that actual vanillin-β-glucoside synthesis, when using glycolytic promoters, occurs in the ethanol phase.30 In existing cell factories and their fermentations, ~70% of the carbon ends up as toxic intermediates like protocatechucic acid (PCA), continuously produced and largely secreted through both glucose and ethanol phases.31,32 We used UAS349-PCK1 and UAS349-FBP1 instead of constitutive promoters only for the enzymes that convert PCA to vanillin (ACAR and EntD (see Figure 2A, as well as Figure S5 in the Supporting Information), expected improved PCA conversion, faster growth, and ultimately similar or higher extracellular vanillin-β-glucoside production (Figure 2A). Compared to the existing pathway design relying on the use of constitutive (TEF1pr) and glycolytic (PGK1pr) promoters (C-VG) alone,30 the use of synthetic glucogenenic promoters to control expression of the vanillin-β-glucoside biosynthetic pathway (strains ScASR.V001 and ScASR.V003) supported production of vanillin-β-glucoside to similar levels as the C-VG strain, following a 65 h fermentation (Figure 2B). Similarly, the growth and glucose consumption rates were increased, and conversion of PCA toward protocatechucic aldehyde (PAL) was also improved in the engineered strains, compared to C-VG (see Figures 2C and 2D, as well as Figures S6–S8 in the Supporting Information). Ultimately, ScASR.V003 produced a higher amount of vanillin-β-glucoside as a percentage of the total heterologous extracellular metabolites, suggesting that the CSRE from FBP1pr has the best performance for engineering glucogenenic promoters for practical applications (see Figure 2D, as well as Figure S8). We can envisage further improving PCA conversion by placing the enzyme synthesizing it from 3-dehydroshikimate (3-DSD) under the control of another synthetic promoter responsive to
glucose depletion; the current choice (TEF1pr) still shows some decreased activity under glucose depletion or high ethanol (see Figure 1D, as well as Figures S1 and S4). In this way, the entire pathway would only be activated when sufficient biomass has accumulated and PCA is given no time to accumulate and be secreted, as seen in the strain designs tested here. Testing different CSREs in synthetic promoters would also allow us to select more promoters for every gene in the pathway in future configurations. Ultimately, it may be advantageous to have a combination of CSREs (native or synthetic) and glycolytic TFBSs to modulate expression across glucose concentrations. Such a configuration would be beneficial in fermentations where glucose levels are low or periodically replenished, such as a fed-batch cultivation.

The promoters that have been engineered and put to use in this study reaffirm our promoter engineering strategy, which does not require altered transcription factor expression, or orthogonal regulatory systems. However, this strategy may not require induction at the diauxic phase, future engineering efforts will focus on finding an optimal balance and configuration of glycolytic, diauxic, and gluconeogenic transcription factors and their binding sites to achieve constitutively high expression levels ultimately supporting the rational design of next-generation, high-performance yeast cell factories.

**MATERIALS AND METHODS**

**Reporter Strain Construction.** *S. cerevisiae* CEN.PK 113-7D was used as background strain for all experiments. Promoters were amplified from genomic DNA or gBlocks (IDT) using Phusion Master Mix with HF buffer (Thermo-Fisher). YFP and resistance markers were amplified from plasmid pASR0130.12 YFP expression cassettes with different promoters were cloned into the EasyClone site XII-4 using plasmid pASR0130.12 YFP expression cassettes with different promoters were cloned into the EasyClone site XII-4 using in vivo assembly by homologous recombination. In brief, 500 pmol of the relevant promoter, YFP, the VPS13 terminator, and kanMX cassette for G418 resistance marker were transformed to yeast with 500 bp of homology to XII-4 on either end by the lithium acetate method onto YPD agar with 350 ng/μL G418. Parts used in assembly had 50 bp homology with adjacent parts added by PCR. G418-resistant colonies were screened by multiplex colony PCR as previously described, checking for integration and assembly at the correct locus. Colonies with sequence-verified reporter constructs were streaked out on YPD agar, regrown in YPD with 200 ng/μL G418 and preserved as glycerol stocks. All primers used for strain construction are listed in Table S3, and all constructed strains are listed in Table S4.
YFP Reporter Assays. Reporter strains, as well as the background strain, were grown overnight in synthetic drop-out medium minus leucine (SD-leu, Sigma) containing 1.1 g/L monosodium glutamate as a carbon source and 200 ng/µL G418 where appropriate. These were diluted to an OD of ∼0.02 in minimal Delft medium, pH 6 and grown at 30 °C in deep-well microtiter plates with 300 rpm agitation. The culture was sampled at 4, 8, 24, 48, 72, and 96 h for measuring YFP fluorescence by flow cytometry. Cells were suitably diluted in phosphate-buffered saline (PBS) (Life Technologies) and their fluorescence measured on an LSRFortessa flow cytometer (BD Biosciences) equipped with an HTS module for sampling. YFP fluorescence was measured with 488 nm excitation and 530 nm emission. A total of 10,000 events per sample were acquired using FACSDiva software, and the resulting FCS files were analyzed using the FCSelectr utility (http://research.stowers.org/mcm/efg/ScientificSoftware/Utility/FCSelectr/index.htm), R scripts and Origin 9.1 (OriginLab) to extract the mean population fluorescence. Alternatively, the cells sampled at the same time points and diluted in PBS had their OD and YFP fluorescence (488 nm excitation and 527 nm emission) measured on a SynergyMX microtiter plate reader using clear-bottomed black microtiter plates (Thermofisher), and fluorescence values normalized to sample OD following the subtraction of background fluorescence from CEN.PK 113-7D. A qualitative estimate of residual glucose was determined using test strips (VWR) at each sampling point.

Vanillin-β-Glucoside Production Strains. The five-gene vanillin-β-glucoside biosynthetic pathway (Figure 2A) was introduced using CasEMBLR42 into three loci: BGL1/EXG1 and ADH6, to simultaneously knock these genes out and thereby avoid product degradation, and the EasyClone site XII-S. The pathway was integrated into the genome in two steps. In the first step, 4 pmoles of each part (insertion homology arms, promoters, terminators, and genes (ACAR and EntD)) were co-transformed with a gRNA expression plasmid targeting ADH6 (clonNAT resistance) into CEN.PK 113-7D already carrying a Cas9 expression plasmid (G418 resistant). Following genotyping of transformants and sequence verification of the assemblies, colonies with correctly assembled and integrated expression cassettes were cured for loss of the ADH6 plasmid and subsequently transformed with a gRNA plasmid targeting EXG1 and site XII-S, along with parts for expression cassettes of OMT, UGT, and 3-DSD. Colonies were screened as done previously, and those containing all five genes of the pathway successfully integrated and assembled were retained for fermentations. Figure S5 gives the layout of the pathway integration and the genotyping assays. The gRNA plasmids are listed in Table S5 in the Supporting Information.

Bioreactor Cultures and HPLC Analysis of Metabolites. For the transcriptome analysis, an overnight culture of S. cerevisiae grown in Delft medium was inoculated into a 500 mL Delft medium at a starting OD of 0.03. The cultures were carried out in 1 L Biostat Q bioreactors (Sartorius) in triplicate at 30 °C with 800 rpm agitation, with controlled aeration and pH maintained at 6 using 2 M NaOH. Fermentation broth was sampled every 2 h for the monitoring of OD, and glucose and yeast metabolites by high-performance liquid chromatography (HPLC). For the latter, the broth sample was centrifuged at 10,000g for 2 min, and the supernatant was syringe-filtered using a 0.22 µm syringe filter. These were diluted 2-fold or 5-fold for HPLC analysis (UltiMate 3000, Dionex) as previously described, and data acquired and analyzed using Chromeleon (Dionex/Thermo Fisher). For the cultivations of the vanillin-β-glucoside production strains, a single colony of each of the vanillin-β-glucoside production strains C-VG, ScASRV.001 and ScASRV.003 were picked from a YPD plate to inoculate a culture in Delft medium. Overnight cultures were used as inoculum to start duplicate cultures of each strain in Delft medium at pH 6.0 in 250 mL shake flasks with a starting OD600 nm of ∼0.2. Growth rates were calculated based on OD600 nm measurements in the exponential growth phase (Figure S6 in the Supporting Information). For the metabolite profile of vanillin-β-glucoside and precursors, the strains were cultivated in duplicates in 24-deep well plates using 5 mL Delft medium at pH 6.0 (Figure S8). The respective cultures were inoculated at a starting OD600 nm of ∼0.04 with a previously prepared overnight culture, and OD600 nm measured during the course of the cultivation by means of plate reader (SynergyMX). Samples taken at regular time intervals were extracted and quantified by HPLC, as previously described. For glucose and ethanol quantifications, samples taken during cultivations were centrifuged for 5 min at 13,000 rpm and 4 °C. Thereafter, 20 μL of the supernatants were analyzed using an Aminex HPX-87H column (Bio-Rad) in an Ultimate 3000 HPLC system (ThermoFisher Scientific) with a refractive index detector (RI-101 Refractive Index Detector, Shodex) (see Figure S7). The column was set at 45 °C and eluted with 5 mM H2SO4 at a flow rate of 0.6 mL min−1, as previously described.

RNaseq and Transcriptome Data Analysis. Thirty (30) OD units of bioreactor yeast culture grown in 6, 14, and 26 h were harvested for total RNA extraction, the time points corresponding to at lag/early log phase, mid-log phase and late-log phase/diauxic shift. Sampling points were determined from trends in growth curves. Cells were harvested from fermentation broth in chilled 15 mL centrifuge tubes half-filled with ice, and then the total RNA was extracted, as previously described using the RNeasy Mini kit (Qiagen). RNA concentration and quality control were performed using a Qubit 2.0 fluorometer (Life Technologies) and an Agilent 2100 Bioanalyzer with RNA 6000 Nano kit, respectively (Agilent). RNaseq was performed as previously described. TopHat (2.0.14) and the Cufflinks (2.2.1) suite were employed for RNaseq analysis, as described previously, using the reference genome and annotations for S. cerevisiae S288C (NCBI RefSeq GCF_000146045.2). Three biological replicates were used to determine expression levels (Fragments Per Kilobase of exon per Million fragments mapped; FPKM) for each condition. Upper-quartile normalization was preferred and reads mapping to rRNA genes were masked. Cuffdiff is used to obtain fold change differences and to perform statistical testing. A q-value cutoff of <0.05 was used to identify genes that have significant differential expression (Table S1). GO and GO slim mappings were retrieved from the Saccharomyces Genome Database (http://www.yeastgenome.org). When extraction was not possible due to chromosome ends, the largest possible sequence fitting the criteria was used. For transcription factor binding site analysis of CSRE, sequences spanning −700 bp to −125 bp upstream from CDS features (or largest possible sequence fitting the criteria) were extracted and patterns matching those listed in Table S2 were identified using biopython (Version 1.68).

Database for RNA-seq Data. RNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI (www.

RIreland.

Signe Stentoft for help with strain construction, Lars Boje thank Anna Koza for technical assistance with RNAseq, (BioREFINE-2G, Project No. FP7-613771). We wish to 7th Framework program under the European Union

This work was funded by the Novo Nordisk Foundation and

ACKNOWLEDGMENTS

The authors declare the following competing financial interest(s): J.D.K. has a financial interest in Amyris, Lygos, Demetrix, Constructive Biology, Maple Bio, and Napigen.

REFERENCES


