Estimating the protein burden limit of yeast cells by measuring the expression limits of glycolytic proteins

Yuichi Eguchi et al
Figure 1. Overexpression of most glycolytic proteins using a strong promoter and a multicopy plasmid causes growth defects. (A) The plasmid used in this study. Each glycolytic gene was cloned into a 2μm-based multicopy plasmid (pTOW40836) and expressed from the TDH3 promoter (TDH3pro).

Figure 1 continued on next page.
(with the exception of PYK1 which was expressed from the PYK1 promoter [PYK1pro], represented in the figure as (p)). In +leucine conditions, the copy number of the plasmid is relatively low (~30). In –leucine conditions, the copy number goes up to 150 copies per cell due to the bias of leu2-89 (Moriya et al., 2012). Here, we designate these conditions low- and high-copy conditions, respectively. (B and C) Maximum growth rate of yeast cells harboring the plasmid overexpressing each glycolytic protein in the indicated growth conditions. The unit is min$^{-1} \times 10^{-4}$. (D and E) Copy number of the plasmid overexpressing each glycolytic protein in the indicated growth conditions. The unit is copy number per haploid genome. The error bars shows the standard deviation of at least three independent biological measurements.

DOI: https://doi.org/10.7554/eLife.34595.003
Figure 2. Expression limits of glycolytic proteins. (A) Measurement of the expression level of an overexpressed glycolytic protein. Whole cellular proteins were stained by a fluorescent dye and separated by SDS-PAGE. The overexpression of the target protein (TBS), estimated from the intensity of the target band of TBS. (B) Protein expression levels in low-copy conditions. (C) Protein expression levels in high-copy conditions.
its band on the gel and its molecular weight, was compared with the corresponding expression of the vector control (TBV), after normalization using control bands (CBS and CBV) to calculate protein expression (AU). Measurement of Gpm1 expression level is shown as an example. (B and C) The expression level of glycolytic proteins overexpressed using the experimental system shown in Figure 1 in the indicated conditions. The TDH3 promoter was used for the expression of all genes except where (p) indicates the use of the PYK1 promoter. (D) Relationship between copy number and protein level in low-copy conditions. The copy number data are the same as in Figure 1D, and the protein level data are the same as in Figure 2B. The error bars show the standard deviation of at least three independent biological measurements.

DOI: https://doi.org/10.7554/eLife.34595.005
Figure 2—figure supplement 1. Relationship between percentage of total protein in (Kintaka et al., 2016) and protein level (AU) in this study. The equation derived by the linear regression of the data and the R-squared value are shown. Each circle indicates an independent measurement from Gpm1 (six samples), Pgk1 (three samples), and Adh4 (four samples) experiments.

DOI: https://doi.org/10.7554/eLife.34595.006
Figure 2—figure supplement 2. SDS-PAGE-separated total cellular proteins of the cells overexpressing glycolytic proteins. Proteins were analyzed in low-copy conditions. Red dots indicate the expected size of the target proteins. The TDH3 promoter was used for the expression of all genes except where (p) indicates the use of the PYK1 promoter.

DOI: https://doi.org/10.7554/eLife.34595.007
Estimated fold overexpression of glycolytic proteins in this study. Native expression level of each protein (A) and the percentage per total *S. cerevisiae* protein were obtained and calculated using published data (Kulak et al., 2014). The percentage of each overexpressed protein (B) was calculated using the data in Figure 2B and the equation in Figure 2—figure supplement 1. Fold overexpression of each target protein over the native level was then calculated as B/A.

DOI: https://doi.org/10.7554/eLife.34595.008
Figure 3. Effects of mutations on the expression limits of glycolytic proteins. (A and B) Expression levels of wild-type and CC mutant glycolytic proteins in the indicated conditions. Each CC mutant has a mutation in the position shown in Supplementary file 1. (C) SDS-PAGE gel images of whole cellular proteins overexpressing Adh3 and ΔMTS-Adh3 in low-copy conditions. Red dots indicate the expected sizes of the target proteins. (D) Protein expression levels of Adh3 and ΔMTS-Adh3 in low-copy conditions. The error bars indicate the standard deviation of the mean. *p<0.05; **p<0.01 in Welch’s t-test.

DOI: https://doi.org/10.7554/eLife.34595.010
<table>
<thead>
<tr>
<th>Protein</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adh1</td>
<td>MSIPETQKVGYESHGKLEYKDVIPVKPKANE</td>
</tr>
<tr>
<td>Adh5</td>
<td>MPSQVIPEKQKAIVFETDGKLEYKDVIPVEPKPE</td>
</tr>
<tr>
<td>Adh3</td>
<td>MLRTSTLFTRVQPSLSRNLQLSTAAIPKTQKVGYENKGLHYKDVIPVEPKPE</td>
</tr>
<tr>
<td>ΔMTS-Adh3</td>
<td>MSTAAIPKTQKVGYENKGLHYKDVIPVEPKPE</td>
</tr>
</tbody>
</table>

**Figure 3—figure supplement 1.** Comparison of the N-terminal amino-acid sequences of Adh1, Adh5, Adh3, and Adh3 mutants without the mitochondrial targeting sequence (ΔMTS-Adh3). Red letters indicate the predicted mitochondrial targeting sequence.  
DOI: https://doi.org/10.7554/eLife.34595.011
Figure 4. Metabolic perturbations triggered by overexpression of glycolytic proteins in silico. (A–D) Metabolic change triggered by overexpression of the indicated glycolytic protein in a kinetic model of glycolytic metabolism (Smallbone et al., 2013). Log₁₀ fold-change in each metabolite level in a simulation with 128-fold overexpression of each glycolytic protein compared to that in the wild-type is shown. In (D), Pfk1 and Pfk2 are simultaneously overexpressed.

DOI: https://doi.org/10.7554/eLife.34595.013
Figure 4—figure supplement 1. Schematic representation of the kinetic model of glycolytic metabolism used in this study (Smallbone et al., 2013). The diagram is drawn using the Systems Biology Graphic Notation (Le Novère et al., 2009).

DOI: https://doi.org/10.7554/eLife.34595.014
Figure 5. Metabolic changes triggered by overexpression of glycolytic proteins in vivo. (A–C) The bar graph shows the log2-fold change in each metabolite in the cells overexpressing wild-type and mutant Pfk2, Pfk1, and Tdh3 over the vector controls. The red circle shows the log2 fold-difference.

Figure 5 continued on next page.
in each metabolite between the wild-type and the CC mutant measurements. The metabolites were measured in exponentially growing cells cultured in low-copy conditions. The error bars indicates the standard deviations of the mean for three (Pfk2) and two (Pfk1, Tdh3) biological replicates. *p<0.05 in Welch’s t-test.
DOI: https://doi.org/10.7554/eLife.34595.016
**Figure 5—figure supplement 1.** Comparison of metabolic changes triggered by overexpression of glycolytic proteins in silico and in vivo. The line graph shows the simulation result, and the data are the same as those used in Figure 4. Circles and diamonds show the metabolic measurements in vivo, and the data are the same as those used in Figure 5.

DOI: https://doi.org/10.7554/eLife.34595.017
Figure 6. Codon usage affects the expression level, but not the synthesis timing, of Glk1. (A) Expression levels of Glk1, Pyk2, Pdc6, and codon-optimized GLK1 (CoGlk1) in the indicated conditions. (B) Relationship between the tAlg and the expression level of each glycolytic protein in low-copy conditions. Protein level data are the same as those shown in Figure 2B. (C) Growth curves and GFP fluorescence of cells expressing codon-optimized GFPs. oG-GFP (tAlg = 0.40): a GFP gene whose codons were optimized for the GLK1 codon usage. oT-GFP (tAlg = 0.64): a GFP gene whose codons were optimized for the TDH3 codon usage. (D) Lag time between the timings with the maximum GFP fluorescence and the maximum growth rate.

Figure 6 continued on next page
Timings of the maximum GFP fluorescence and the maximum growth rate are the time points with maximum second derivatives of GFP fluorescence and growth curves. The error bars indicate the standard deviations of the means. **p<0.01; ***p<0.001 in Welch’s t-test.

DOI: https://doi.org/10.7554/eLife.34595.019
Figure 6—figure supplement 1. Scatter plot showing the tA1g and native expression level of S. cerevisiae protein. The tA1g was calculated on the basis of the method described by Tuller et al. (2010). For the native expression level of S. cerevisiae protein, a published dataset was used (Kulak et al., 2014). Red points indicate glycolytic genes/proteins that are analyzed in this study. The tA1gs of all yeast proteins are listed in Figure 6—source data 1.

DOI: https://doi.org/10.7554/eLife.34595.020
Figure 7. Overexpressed Eno2 and Pgk1 form protein aggregates. (A) SDS-PAGE-separated total cellular proteins from cells overexpressing the indicated proteins. (B) SDS-PAGE-separated total cellular proteins from cells overexpressing Eno2 after treatment with (+) or without (–) the reducing agent DTT. (C and D) Enriched proteins in the high molecular bands from cells overexpressing the indicated proteins. Proteins enriched in the high molecular bands > 1.5-fold over the vector control are shown. Proteins were analyzed in low-copy conditions. The red point indicates the expected molecular weight of overexpressed proteins. The asterisk indicates the high-molecular-weight band specifically observed upon overexpression of each glycolytic protein. Gel images were contrasted so that high-molecular-weight bands were visible. CC mut.: catalytic center mutant. DOI: https://doi.org/10.7554/eLife.34595.022
Figure 7—figure supplement 1. Aggregation of Eno2 and Pgk1 requires cysteine residues in Eno2 and Pgk1. (A and C) SDS-PAGE-separated total proteins of cells overexpressing the indicated proteins. Three biological replicates for each experiment are shown. C248S: substitution of cysteine

Eguchi et al. eLife 2018;7:e34595. DOI: https://doi.org/10.7554/eLife.34595

Figure 7—figure supplement 1 continued on next page
248 to serine; C98S: substitution of cysteine 98 to serine. The arrow from A to B indicates the direction of the pixel intensity analysis shown in (B) and (D). (B and D) The pixel intensity analysis of the gel corresponding to the location containing the protein aggregate shown as the arrow between A and B. Image Quant TL software was used for the image analysis. The asterisk indicates the high-molecular-weight band specifically observed upon overexpression of each glycolytic protein.

DOI: https://doi.org/10.7554/eLife.34595.023
Figure 7—figure supplement 2. Effect of cysteine substitutions on the expression levels of Eno2 and Pgk1. C248S: substitution of cysteine 248 with serine; C98S: substitution of cysteine 98 to serine. Proteins were analyzed in low-copy conditions.

DOI: https://doi.org/10.7554/eLife.34595.024
Figure 8. Overexpressed Tpi1 forms protein aggregates. (A) SDS-PAGE-separated total proteins from cells overexpressing indicated proteins, and their Western blot imaging using anti-Tpi1 antibodies. The red points indicate the expected molecular weight of overexpressed proteins. (B) Effect of

Eguchi et al. eLife 2018;7:e34595. DOI: https://doi.org/10.7554/eLife.34595
cysteine substitutions on the expression level of Tpi1. *p<0.05; **p<0.01; ***p<0.001 in Welch’s t-test. C41S: substitution of cysteine 41 to serine; C126S: substitution of cysteine 126 to serine. Proteins were analyzed in low-copy conditions.

DOI: https://doi.org/10.7554/eLife.34595.026