Supporting Information

A genetic sensor for strong methylating compounds

Felix Moser, Andrew Horwitz, Jacinto Chen, Wendell A. Lim, and Christopher A. Voigt

Table of Contents

- **I. Cytometry fluorescence distributions**
- **II. Toxicity of alkylating agents on** *E. coli* **and** *S. cerevisiae*
- **III. GC-MS standard curve**
- **IV. Saturation model for MeI activation of the** *S. cerevisiae* **sensor**
- **V. Genetic parts, plasmids, and yeast strains**
- **VI. Supplemental references**

I. Cytometry fluorescence distributions

The responses of the *E. coli* and *S. cerevisiae* sensors to MeI were assessed by flow cytometry. Figure S1 shows the fluorescence distributions of the *E. coli* strains carrying the sensor plasmid pFM45 in response to methyl iodide. The *E. coli* MG1655Δada strain has the *ada* gene knocked out and therefore shows no response to methyl iodide. The wild-type *E. coli* MG1655 populations induce with a bimodal response near the switch point, a behavior that is characteristic of systems containing a genetic positive feedback loop. Interestingly, this bimodal character is largely lost when additional Ada is expressed from a plasmid (pFM141) at all levels of induction. The behavior of the pFM45 sensor in the strain lacking pFM141 is consistent across all concentrations of arabinose. The presence of pFM141 in the MG1655*Δada* strain without arabinose induction is sufficient to rescue activity of the sensor, indicating leakage from the P_{BAD} promoter. Additional expression of Ada from pFM141 via the induction of the arabinose-inducible P_{BAD} promoter lowers the detection threshold of the sensor. High levels of Ada expression raise the basal leakage of the output promoter, which lowers the dynamic range of the sensor.

S. cerevisiae sensors showed a much lower dynamic range and less cooperativity than the *E. coli* response (Figure S2). The response to MeI was dependent on the presence and number of Ada operators in the Cyc1 promoter driving the EGFP reporter. The yeast sensors also showed a much higher basal activity than the *E. coli* sensors.

Figure S1: Cytometry distributions of the *E. coli* **methylation sensor strains in response to MeI.** Shown are the cytometry data for transfer functions of *E. coli* strains MG1655 and MG1655*Δada* carrying the sensor plasmid pFM45 exposed to MeI. Each strain carrying pFM45 is also shown carrying the plasmid pFM141, which expresses the Ada protein from an arabinose-inducible P_{BAD} promoter. Arabinose was added to the cultures represented in the top (0 mM), middle (1 mM), and bottom (10 mM) rows of squares containing cytometry histograms, respectively. The amount of MeI added to each culture, from bottom-most histogram in each square to the top-most, is as follows: 0, 6×10^{-3} , 1.6×10^{-2} , 3.9×10^{-2} , 9.8×10^{-2} , 2.4×10^{-1} , 6.1×10^{-1} , 1.5 , and 9.5 mM. This data corresponds to the data in Figure 1C and 1D of the main text.

Figure S2: Cytometry distributions of the *S. cerevisiae* **methylation sensor in response to MeI.** Shown are the cytometry data for transfer functions of *S. cerevisiae* sensor strains $P_{0x,Cvc1}|P_{Adh1}$ $P_{1x.Cyc1}|P_{Adh1}$, $P_{3x.Cyc1}|P_{Adh1}$, $P_{8x.Cyc1}|P_{Adh1}$, and $P_{8x.Cyc1}|P_{Cyc1}$ in response to MeI. The amount of MeI added to each culture, from bottom-most histogram in each square to the top-most, is as follows: 0, 2.8×10⁻², 6.4×10^{-2} , 1.5×10^{-1} , 3.4×10^{-1} , 7.8×10^{-1} , 1.8 , 4.1 , and 9.5 mM. This data corresponds to the data in Figure 1G and 1H of the main text, which reports the average of the geometric means for three different fluorescence distributions.

II. Toxicity of alkylating agents on *E. coli* **and** *S. cerevisiae*

Both sensors responded to methyl iodide (MeI), methyl methane sulfonate (MMS), dimethyl sulfate (DMS), and 1-methyl-3-nitro-1-nitrosoguanidine (MNNG). The toxic effects of these agents were evident in the cytometry distributions (Figure S1 and S2,). At toxic concentrations, the population distribution widened considerably and lost fluorescence.

The toxicity of MeI, MMS, DMS, and MNNG on *E. coli* and *S. cerevisiae* strains containing methylation sensors was assessed and the LD₅₀ of each alkylating agent was determined (Figure S4). *E. coli* is more robust to growth defects than *S. cerevisiae* at the same concentrations of alkylating agent. No difference in toxicity was observed between wild-type MG1655 and MG1655 with the Δ*ada* mutation. Overproduction of the Ada protein in *E. coli* did not significantly reduce the toxicity of any alkylating agents. Higher levels of N-Ada-Gal4 production the yeast sensor strain also did not reduce the toxicity in that strain.

Figure S3: Cytometry distributions of *E. coli* **and** *S. cerevisiae* **methylation sensors in response to MMS, DMS, and MNNG.** *E. coli* MG1655 carrying plasmid pFM45 and *S. cerevisiae* strain P_{8x.Cyc1}|P_{Adh1} were exposed to MMS, DMS, and MNNG as described in the methods. The amount of MMS and DMS added to each culture, from bottom-most histogram in each square to the top-most, is as follows: 0, 1.2×10^{-2} , 2.7×10^{-2} , 6.4×10^{-1} , 1.5×10^{-1} , 3.4×10^{-1} , 7.8×10^{-1} , 1.8 , 4.1 , and 9.5 mM. The amount of MNNG added was as follows: 0, 1.6×10^{-4} , 2.3×10^{-3} , 5.2×10^{-3} , 1.2×10^{-2} , 2.8×10^{-2} , 6.4×10^{-2} , 1.5×10^{-1} , 3.4×10^{-1} , and 7.8×10 $^{-1}$ mM. This data corresponds to Figure 2 in the main text.

Figure S4: Toxicity of alkylating agents on *E. coli* **and** *S. cerevisiae* **containing methylation** sensors. The strains were exposed to MeI, MMS, DMS, and MNNG. The OD₆₀₀ of *E. coli* was measured 3 hours after exposure. **(A)** *E. coli* strains include: MG1655 (white square, dashed lines), MG1655*Δada* (white circle, dashed lines), MG1655 containing pFM45 and pFM141 (black squares, solid black lines), MG1655*Δada* containing pFM45 and pFM141 (black circles, solid black lines), MG1655 containing pFM45 and pFM141 and induced with 10 mM arabinose (red squares, red lines), and MG1655*Δada* containing pFM45 and pFM141 and induced with 10 mM arabinose (red squares, red lines). **(B)** *S. cerevisiae* strains measured include: SO992 (no Ada sensor, squares), P_{0x.Cyc1} | P_{Adh1} (circles), P_{8x.Cyc1} | P_{Adh1} (diamonds), $P_{8x,Cyc1}$ | P_{Cyc1} (triangles). The OD₆₀₀ of the *S. cerevisiae* cells were measured 12 hours after exposure. For both *E. coli* and yeast cultures, all OD₆₀₀ measurements were normalized to the highest measured value of that day for better comparison between days. Each data point is averaged from three measurements performed on different days. Error bars are one standard deviation from the mean.

III. GC-MS Standard Curve

We generated a standard curve to calculate the MeI produced by yeast cultures expressing methyl halide transferases (MHTs; Figure 3A, main text). To measure this curve, we added a known amount of MeI into a volume of media equivalent to the volume in which sample cultures were grown. Following addition of MeI, the tubes were immediately stoppered. To allow the sample to adequately dissolve and equilibrate between liquid and gas phases in conditions comparable to those of the yeast culture, the standard curve samples were shaken for 30 minutes at 30°C in the same incubator as the MHT yeast cultures. To sample the MeI in each tube, 100 µl of air from the headspace of each tube was injected into the GC-MS. Because some MeI degradation was observed over time, all samples were injected 30 seconds apart in a single long run. Each sample's MeI peaks, clearly differentiable, were integrated by the software. The resulting counts were plotted against the respective known amounts of MeI to generate the standard curve. The standard curve was re-run for each assay on each day and varied widely depending on machine settings. The slope of the standard curve, however, was consistent between days.

Figure S5: Standard curve for GC-MS measurements. Known amounts of MeI were added to sample tubes, equilibrated, and measured with a GC-MS. The measured GC-MS counts of MeI are plotted against the amount of MeI added to each respective tube. A power law fits the data (R^2 = 0.98) and is used to calculate MHT production. The standard curve shown corresponds to the one used to calculate MeI production from the MHT yeast cultures shown in Figure 3A in the main text.

IV. Saturation model for MeI activation of the *S. cerevisiae* **sensor**

A simple model was derived for the activation of the sensors. In this model, the promoter is activated by methylated Ada (Ada*) and responds instantaneously to a change in MeI concentration. The probability that RNA polymerase binds to the reporter promoter is given by,

$$
f(s) = \frac{c_0 + K_d [Ada^*]^n}{1 + c_0 + K_d [Ada^*]^n}
$$
 (S1)

where s is the concentration of the inducing alkylating agent, K_d is the binding constant for activated Ada to its operator, n is the empirically-derived Hill coefficient, and c_0 is the basal level of RNAP binding to the promoter causing leakage. The rate of Ada activation is

$$
\frac{d[Ada^*]}{dt} = k_{met}[Ada][Mel] - k_{deg}[Ada^*] = 0
$$
\n
$$
(S2)
$$

where *kmet* is the methylation rate constant, which is irreversible, and *kdeg* is the degradation rate constant. At steady-state,

$$
[Ada^*] = \frac{k_{met}[Ada][Mel]}{k_{deg}} \tag{S3}
$$

Substituting Equation S3 into S1 produces

$$
f(s) = \frac{c_0 + K_d \left(\frac{k_{met}[Ada]}{k_{deg}}\right)^n [Mel]^n}{1 + c_0 + K_d \left(\frac{k_{met}[Ada]}{k_{deg}}\right)^n [Mel]^n} = \frac{c_0 + K[Mel]^n}{1 + c_0 + K[Mel]^n} \qquad ,
$$
\n(54)

where *c0*, *K* and *n* are treated as fit parameters. This equation was used to fit the measured response functions reported in the main text. The Hill coefficients reported in the Tables in the main text were fit using this equation. The regression line in Figure 3A was also fit using this equation and the data in that chart (resulting in *c0* = 0.15, *K* = 0.056, and *n* = 1.8).

V. Genetic Parts, plasmids, and yeast strains

Genetic parts and plasmids were derived from previous work, the SynBERC Registry¹, or the Registry of Standard Biological Parts². Table S1 lists all the parts used in this work and their sequences, relevant plasmids, and source of each part sequence. Table S2 provides a concise description of all the plasmids used in this work and the GenBank accession #s. Table S3 provides a concise description of the genotype of the yeast strains used.

Table S2: Summary of plasmids used in this work

Table S3: Summary of yeast strains used in this work.

VI. Supplemental References

¹ SynBERC Registry. https://registry.synberc.org/#page=login. SynBERC, 2013.

² Registry of Biological Parts. www.partregistry.org Biobricks Foundation, 2013.

3 Nakamura, T., Tokumoto, Y., Sakumi, K., Koike, G., Nakabeppu, Y., and M. Sekiguchi. Expression of the *ada* Gene of *Escherichia coli* in Response to Alkylating Agents. *J. Mol. Biol*. 202: 483-494 (1988).

 4 Mumberg, D., Muller, R., and M. Funk. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156: 119-122 (1995).

⁵ Takinowaki, H., Matsuda, Y., Yoshida, T., Kobayashi, Y., and T. Ohkubo. The solution structure of the methylated form of the N-terminal 16-kDa domain of Escherichia coli Ada protein. *Protein Science* 15:487-497 (2006).

 6 Cormack, B.P., Bertram, G., Egerton, M., Gow, N.A.R., Falkow, S., and A.J.P. Brown. Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in *Candida albicans. Microbiology* 43: 303-311 (1997).