Genetic Sensor for Strong Methylyating Compounds

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

†Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
‡Howard Hughes Medical Institute and Department of Cellular and Molecular Pharmacology, University of California—San Francisco, San Francisco, California 94158, United States

Supporting Information

ABSTRACT: Methylating chemicals are common in industry and agriculture and are often toxic, partly due to their propensity to methylate DNA. The Escherichia coli Ada protein detects methylating compounds by sensing aberrant methyl adducts on the phosphoester backbone of DNA. We characterize this system as a genetic sensor and engineer it to lower the detection threshold. By overexpressing Ada from a plasmid, we improve the sensor’s dynamic range to 350-fold induction and lower its detection threshold to 40 μM for methyl iodide. In eukaryotes, there is no known sensor of methyl adducts on the phosphoester backbone of DNA. By fusing the N-terminal domain of Ada to the Gal4 transcriptional activation domain, we built a functional sensor for methyl phosphothriester adducts in Saccharomyces cerevisiae. This sensor can be tuned to variable specifications by altering the expression level of the chimeric sensor and changing the number of Ada operators upstream of the Gal4-sensitive reporter promoter. These changes result in a detection threshold of 28 μM and 5.2-fold induction in response to methyl iodide. When the yeast sensor is exposed to different S₉.1 and S₉.2 alkylating compounds, its response profile is similar to that observed for the native Ada protein in E. coli, indicating that its native function is retained in yeast. Finally, we demonstrate that the specifications achieved for the yeast sensor are suitable for detecting methylating compounds at relevant concentrations in environmental samples. This work demonstrates the movement of a sensor from a prokaryotic to eukaryotic system and its rational tuning to achieve desired specifications.

KEYWORDS: synthetic biology, genetic circuit, genetic device, fumigant, Gal4

A transcriptional genetic sensor is a unit of DNA that contains all of the necessary parts to convert an input stimulus to the up- or down-regulation of a promoter.†‡§ Following this paradigm, the output promoter of a sensor can be used as the input promoter of a genetic circuit, which can implement signal-processing functions. Genetic sensors have been constructed that respond to many environmental signals, including light,3,4 temperature,5,6 gases,7,8 toxins (e.g., arsenic),9,10 and chemicals (e.g., industrial products, pollutants or explosives).11–14 Many of these sensors are based on the transfer of parts from one organism to another; for example, moving a TNT sensor from cyanobacteria and plants to E. coli15 and a redox sensor from Streptomyces to mammalian cells.16 Such transfers often require sensor re-engineering and the substitution of parts to make the sensor functional in the new host.

Different applications require different performance specifications of a genetic sensor, which can be achieved by tuning the response function of the sensor. The response function is defined by how the sensor output (promoter transcription) changes as a function of the input stimulus. The shape of this function captures the responsiveness of the sensor to the input and provides information that aids its connection to a downstream circuit.18,19 There are several descriptors of the response function that are particularly useful: the basal activity, cooperativity, dynamic range, detection threshold (lowest input concentration sensed above background), and sensitivity (the slope during the transition).20 Additionally, it is useful to determine the specificity of a sensor to understand how a sensor will respond in a mixture of ligands or complex environment. Various approaches, including directed evolution, have been applied to alter the properties of genetic sensors.21,22 Synthetic biology has also developed “tuning knobs” to control transcription and translation that could be applied to altering sensor response23–27.

Here, we design and characterize a sensor of methylating compounds, transfer it from E. coli to yeast, and tune its response characteristics. Methylating agents are relevant to human health because they can induce the aberrant methylation of DNA, which can lead to mutations, misregulation, and ultimately disease. Many methylating agents leave methyl phosphothriester (PTE) adducts on DNA. These adducts are very stable, long-lasting moieties in eukaryotic cells due to their innocuous nature and resistance to DNA repair.28 Because of their stability, methyl PTE adducts have been proposed as a biomarker for cumulative genotoxic exposure.29,30 Methylating agents that generate these adducts are common in industrial and agricultural processes and often produce other more
Figure 1. Mechanism and activity of *E. coli* and *S. cerevisiae* methylation sensors. (A) Diagram of the *E. coli* methylation sensor. The Ada protein is expressed from either the genome or a low-copy plasmid (inset in lower right). (B) Overlaid histograms show the induction of *E. coli* MG1655 carrying the plasmid pFM45. The line colors correspond to different concentrations of MeI: 0 (light blue), 98 μM (dark blue), 244 μM (violet), 610 μM (pink), and 3.8 mM (red). (C) Response functions for the methylation sensor are shown in *E. coli* MG1655. Solid lines are strains expressing inducible Ada from plasmid pFM141 with variable amounts of arabinose: 0 mM (circles), 1 mM (triangles), and 10 mM (diamonds). *E. coli* containing only the reporter plasmid pFM45 is shown for comparison (black squares, dashed line). (D) Response functions for the methylation sensor are shown for strain *E. coli* MG1655Δada. The same symbols and arabinose concentrations are used as in C. (E) Diagram of the *S. cerevisiae* methylation sensor. The sensor protein (Gal4-N-Ada) is expressed from the genome (inset in lower right). (F) The histogram shows the induction of *S. cerevisiae* carrying the P8x.Cyc1|PAdh1 methylation sensor. The line colors correspond to different concentrations of MeI: 0 (light blue), 148 μM (dark blue), 783 μM (violet), 9.5 mM (red). (G) Response functions of the methylation sensor in *S. cerevisiae* with variable number of operators upstream of the reporter promoter. Yeast sensor strains with promoter P0x.Cyc1 (circles), P1x.Cyc1 (triangles), P3x.Cyc1 (diamonds), and P8x.Cyc1 (squares) contain corresponding numbers of Ada operators upstream of reporter promoter PCyc1. The inset shows which symbols correspond to the number of operators in the promoter. The Gal4-N-Ada sensor protein in these strains is driven by the PAdh1 promoter. (H) Response functions of the methylation sensor in *S. cerevisiae* with Gal4-N-Ada expression driven by either the strong promoter PAdh1 (dark blue squares) or the weaker promoter PCyc1 (light green squares). Both strains contain the EGFP reporter driven by promoter PAdh1. Diagonal gray dashes indicate MeI concentrations observed to be toxic (SI Figure S4). The P4x.Cyc1 (squares) data is the same in both G and H for ease of comparison. Lines were generated by the sensor activation model described in the Supporting Information (SI) and are matched to their respective data by color. For all data, error bars represent one standard deviation from three independent experiments performed on different days.
Research Article

ACS Synthetic Biology

Table 1. Performance of E. coli Methylation Sensors in Response to Mel

<table>
<thead>
<tr>
<th>organism</th>
<th>sensor</th>
<th>arabinose (mM)</th>
<th>detection threshold (μM)</th>
<th>sensitivity ± (au/ μM)</th>
<th>basal activity ±</th>
<th>dynamic range (fold-change)</th>
<th>cooperativity ±</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli</td>
<td>a,b</td>
<td>0</td>
<td>100</td>
<td>1.3 ± 0.4</td>
<td>0.97 ± 0.01</td>
<td>260 ± 86</td>
<td>2.3 ± 0.6</td>
</tr>
<tr>
<td>E. coli</td>
<td>a,c</td>
<td>0</td>
<td>40</td>
<td>1.9 ± 0.8</td>
<td>1.0 ± 0.03</td>
<td>350 ± 39</td>
<td>3.2 ± 0.7</td>
</tr>
<tr>
<td>E. coli</td>
<td>a,c</td>
<td>1</td>
<td>16</td>
<td>1.9 ± 0.8</td>
<td>1.6 ± 0.2</td>
<td>230 ± 34</td>
<td>2.4 ± 0.2</td>
</tr>
<tr>
<td>E. coli</td>
<td>a,c</td>
<td>10</td>
<td>6</td>
<td>2.3 ± 0.7</td>
<td>16 ± 6.7</td>
<td>49 ± 18</td>
<td>1.5 ± 0.6</td>
</tr>
<tr>
<td>E. coli</td>
<td>a,c</td>
<td>0</td>
<td>40</td>
<td>0.2 ± 0.04</td>
<td>1.2 ± 0.2</td>
<td>51 ± 14</td>
<td>1.5 ± 0.4</td>
</tr>
<tr>
<td>E. coli</td>
<td>a,c</td>
<td>1</td>
<td>16</td>
<td>0.8 ± 0.2</td>
<td>1.5 ± 0.2</td>
<td>155 ± 81</td>
<td>1.3 ± 0.7</td>
</tr>
<tr>
<td>E. coli</td>
<td>a,c</td>
<td>10</td>
<td>6</td>
<td>2.0 ± 0.2</td>
<td>13 ± 2.5</td>
<td>40 ± 8</td>
<td>1.3 ± 0.6</td>
</tr>
</tbody>
</table>

“E. coli MG1655 containing pFM45. "E. coli MG1655 expressing sensor protein from a native locus on the genome. "E. coli MG1655 expressing sensor protein from Pbad promoter on pFM141. "E. coli MG1655 containing the Δada mutation. Sensitivity is the difference between the minimum and maximum outputs divided by difference in inducer concentration at the detection threshold and the maximum output. “Basal activity is reported in multiples of the background fluorescence of cells containing no reporter GFP. "Cooperativity is reported as the Hill coefficient in the fit equations detailed in the SI.

Damage to DNA is a common adaptive response to methylating compounds. For example, phenyl glycidyl ether (PGE) and N-nitrosodiethanolamine (NDEA) are used in the manufacture of paints, resins, and rubber. Damaging DNA lesions. For example, phenyl glycidyl ether (PGE) and N-dimethylformamide (DMF) is a common alkylating agent used in kiloton quantities in a variety of industries. Methylthioformamide (MTHF) and methyl methanesulfonate (MMS) are used in laboratories to study DNA damage and repair. Methyl halides such as methyl chloride and methyl iodide (MeI) are methylating agents that are controversially used as soil fumigants and as intermediates to various chemical processes, including silicon rubber production. All of these agents methylate the bases of DNA. Due to the ubiquity and potency of genotoxic methylating agents, a sensor for DNA methylation damage could be a tool for environmental biosensing or a diagnostic system for long-term genotoxic exposure.

Our lab previously engineered E. coli and S. cerevisiae to produce methyl halides by introducing a methyl halide transferase (MHT) gene. Screening for MHT activity is tedious and slow throughput because it is based on a GC-MS assay. Cell-based sensors have been used as a tool to screen libraries of mutant enzymes and pathways for increased activity or titer. To facilitate easier screening of MHT activities, we aimed to develop a sensor for methyl halide production in Saccharomyces cerevisiae, the organism in which the greatest Mel yield was achieved.

Escherichia coli has a strong adaptive response to methylating agents such as methyl halides. These agents are sensed via the Ada protein, which is either directly methylated by S$_2$ methylating agents or indirectly by S$_1$ methylating agents via methyl PTE DNA adducts. Ada moves along DNA, detects, and then transfers a single DNA Sp methyl PTE adduct onto its Cys38 residue (Figure 1A). The methylation of Ada’s Cys38 residue activates Ada as a transcription factor. Ada then upregulates transcription of various DNA repair proteins, including its own ada gene. This positive feedback loop turns the very low basal expression of Ada into a strong, sustained response to the exposure of genotoxic methylating agents. Ada has been used as a sensor for DNA methylation toxicity of genotoxic compounds to complement the Ames test, the gold standard for assaying mutagenicity of a compound.

No comparable, specific sensor of DNA methyl PTEs is known in eukaryotes. To develop such a sensor in eukaryotes, we fused the N-terminal domain of Ada (N-Ada) to the Gal4 trans-activation domain. This Gal4 domain, taken from yeast, is functional in a broad range of hosts, including yeast, flies, plants, and human cells. We demonstrate that the resulting Gal4-N-Ada fusion protein acts as a specific and strong sensor of methylating compounds in Saccharomyces cerevisiae. We show that the sensor retains Ada’s characteristic specificity for methylating compounds, indicating that Gal4-N-Ada can detect and remove DNA methyl PTE adducts in S. cerevisiae. To demonstrate tuning the S. cerevisiae sensor to different specifications, we change the detection threshold of the sensor by changing expression of the sensor protein and change its sensitivity by altering the number of operators in the promoter driving the reporter. Finally, we demonstrate the utility of the tuned S. cerevisiae sensor to detect MeI in an MHT-expressing culture and in a complex soil sample.

RESULTS

Construction of a Methylation Sensor in E. coli. The native E. coli ada promoter was used to measure the sensor response to methylating compounds. The ada promoter region, which includes a single Ada operator upstream of the −45 site, was transcriptionally fused to a green fluorescent protein (GFP) reporter on a p15A plasmid backbone (Figure 1A). In the first design, Ada is expressed from its native locus in the E. coli MG1655 genome. When uninduced, it has been estimated that there are 2−4 Ada proteins per cell. Upon induction with Mel, this sensor shows a strong 250-fold activation and detection threshold of 100 μM Mel (Table 1; Figure 1C). Near the switch point of the response function, the population of cells exhibits a bimodal distribution of fluorescence (Figure 1B, Supporting Information (SI) Figure S1). This is characteristic of positive feedback loops, as in the case of the native autoregulatory control of Ada expression.

A challenge in the design of genetic sensors is the tuning of their detection threshold to respond to different target levels of stimulus. To this end, we sought to lower the detection threshold of the Ada sensor to respond to lower concentrations of Mel. This was achieved by increasing the expression level of the Ada protein. A plasmid was constructed in which the ada gene was placed under control of the arabinose-inducible Pbad promoter on a low-copy incW origin plasmid. Even in the absence of inducer, the basal expression of Ada from Pbad lowers the detection threshold of the sensor and increased its dynamic range (Table 1; Figure 1C). When the Ada concentration was further increased via arabinose induction, the detection threshold decreased from 100 μM to 6 μM. At intermediate levels of Ada, the OFF state of the sensor stayed at a constant level. However, when Ada was maximally expressed from Pbad (10 mM arabinose) the basal activity of the OFF state increased significantly, which attenuated the dynamic range of the sensor.
The impact of knocking out the native ada gene on the sensor was investigated. This knockout eliminated the positive feedback loop. As expected, the sensor was nonresponsive to MeI when ada is knocked out (Figure 1D). This response was rescued when Ada was expressed from P_BGD. The detection threshold of the sensor was similar to when Ada was genomically expressed. However, the response function was impacted in several ways that are consistent with the disruption of a positive feedback loop. First, the cooperativity of the sensor was similar to when Ada was expressed from P_BGD. Second, the highest ON state of the sensor depended more on the level of Ada expression, increasing by 10.7-fold from basal expression to full induction of Ada. The bimodality of the response was also disrupted, which diminished the variability in the population near the threshold. Third, the highest ON state of the sensor depended more on the level of Ada expression, increasing by 10.7-fold from basal expression to full induction of Ada. The bimodality of the response was also disrupted, which diminished the variability in the population near the threshold. When the positive feedback loop is disrupted, the detection threshold did not affect the variability in the population near the threshold. The detection threshold of the sensor did not change with the number of operators in the P_Cyc1 promoter, but no significant difference was observed between 1 and 3 operators (Table 2). The detection threshold of the sensor did not change with the number of operators in the P_Cyc1 promoter, but no significant difference was observed between 1 and 3 operators (Table 2).

### Table 2. Performance of S. cerevisiae Methylation Sensors in Response to MeI

<table>
<thead>
<tr>
<th>organism</th>
<th>sensor</th>
<th>no. operators</th>
<th>detection threshold (μM)</th>
<th>sensitivity (au/μM)</th>
<th>basal activity (fold-change)</th>
<th>cooperativity</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. cerevisiae (SO992)</td>
<td>a</td>
<td>1</td>
<td>28</td>
<td>0.22 ± 0.01</td>
<td>2.9 ± 0.05</td>
<td>1.4 ± 0.4</td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>3</td>
<td>28</td>
<td>0.33 ± 0.09</td>
<td>2.4 ± 0.04</td>
<td>1.5 ± 0.8</td>
</tr>
<tr>
<td></td>
<td>c</td>
<td>8</td>
<td>28</td>
<td>0.42 ± 0.03</td>
<td>2.8 ± 0.10</td>
<td>1.6 ± 0.3</td>
</tr>
<tr>
<td></td>
<td>d</td>
<td>8</td>
<td>340</td>
<td>0.35 ± 0.02</td>
<td>2.7 ± 0.13</td>
<td>1.8 ± 0.9</td>
</tr>
</tbody>
</table>

*aContains P_Adh1 driving the expression of Gal4-N-Ada. *bContains P_Cyc1 driving the expression of Gal4-N-Ada. *cOperators in the P_Cyc1 promoter driving EGFP. *dSensitivity is the difference between the minimum and maximum outputs divided by difference in inducer concentration at the detection threshold and the maximum output. *eBasal activity is reported in multiples of the background fluorescence of cells containing no reporter GFP. *fCooperativity is reported as the Hill coefficient in the fit equations detailed in the SI.

To move the Ada sensor into yeast, we built a chimeric protein that contains the N-terminal domain of Ada (N-Ada, residues 1 to 180) fused to the Gal4 trans-activator (residues 767 to 881; Figure 1E). N-Ada is the site of DNA binding and methyltransferase activity and is necessary and sufficient to induce the adaptive response in *E. coli*. The Gal4 trans-activation domain is a native yeast protein that upregulates transcription when localized to the PCyc1 promoter. We modified the P_Cyc1 promoter to include eight Ada operators (P8xCyc1) and placed it upstream of an enhanced GFP (EGFP) reporter (Figure 1E). The strong, constitutive P_Adh1 promoter was placed upstream of the Gal4-N-Ada chimera. A S. cerevisiae strain was built on the basis of the completed sensor (P8xCyc1, P_Adh1) by integrating the P_Adh1-driven Gal4-N-Ada expression cassette and the P8xCyc1-driven EGFP reporter cassette into the genome (Materials and Methods).

We hypothesized that exposure of this strain to methylating agents would lead to methylation and subsequent activation of the N-Ada domain. This would localize the Gal4-N-Ada fusion protein to P8xCyc1 and upregulate expression of the EGFP reporter. Upon exposure to MeI, the completed yeast sensor strain P8xCyc1(P_Adh1) showed a maximal 5.2-fold induction of the EGFP reporter (Figure 1F). The population’s fluorescence changed gradually with the concentration of MeI and no bimodality in the population’s fluorescence distribution was observed (SI Figure S2). Additionally, the response function was more linear than the native system in *E. coli* (Table 2). Both of these observations are consistent with the response observed when the positive feedback loop in *E. coli* is disrupted (Figure 1D and SI Figure S1). The detection threshold of this yeast sensor to MeI is 28 μM, which is lower than the uninduced *E. coli* sensors (Table 2).

In building the sensor, variations of the P_Cyc1 promoter containing different numbers of Ada operators were tested (Figure 1G and SI Figure S2). The level of expression from a Gal4-driven promoter is a function of how many Gal4-containing proteins are recruited to the promoter. Therefore, increasing the number of operators upstream of the target promoter can tune the response of the sensor. Variations of the P_Cyc1 promoter containing 0, 1, 3, and 8 copies of the Ada operator were built. The presence of one copy of the operator upstream of P_Cyc1 was sufficient to upregulate transcription from the promoter, even in the uninduced state (Figure 1G). Both the dynamic range and the sensitivity increased when 8 operators were included in the promoter, but no significant difference was observed between 1 and 3 operators (Table 2). The detection threshold of the sensor did not change with the number of operators in the P_Cyc1 promoter, and no significant increase in cooperativity was observed when the number of operators was increased (Table 2). These results show that changing the number of operators driving the P_Cyc1 promoter enables tuning of the dynamic range and sensitivity of the response.

The impact of varying the expression level of Gal4-N-Ada was also tested. The sensor uses a constitutive promoter to drive the expression of Gal4-N-Ada. When the P_Adh1 promoter is used, the detection threshold is 28 μM MeI and is independent of the number of operators. We hypothesized that, similar to the Ada sensor in *E. coli*, the detection threshold was dependent on the level of Gal4-N-Ada expression. To test this, we replaced P_Adh1 with the 20 to 100-fold weaker P_Cyc1 promoter. This replacement resulted in a 10-fold higher detection threshold of 340 μM (Figure 1H). Notably, this change in the threshold did not affect the magnitude of the ON or OFF states.

**Comparison of Sensor Responses to Different Methylation Compounds.** To test whether the Gal4-N-Ada sensor retained its native activity following species transfer, both the *E. coli* and yeast methylation sensors were exposed to a panel of different S1 and S2 alkylating agents. S1 and S2 agents react via different mechanisms and have different affinities for methylating DNA. The specificity of the yeast methylation sensor’s response to these agents was expected to be comparable to the *E. coli* sensor. S1 agents, such as MNNG, are known to promiscuously methylate the phosphoester backbone of DNA and have not been observed to methylate Ada directly. In nature, nitrosoamines similar to MNNG are produced via endogenous chemistry and are thought to be the source of naturally occurring DNA methyl PTE adducts. As such, MNNG is
highly toxic to both organisms (SI Figure S4). The *E. coli* and yeast sensors both responded strongly to MNNG, showing the lowest observed detection thresholds (Figure 2A, Table 3). Because MNNG is only known to activate Ada indirectly through methylation of DNA, this supports the hypothesis that Gal4-N-Ada is detecting and removing methyl PTE adducts from the DNA backbone. Interestingly, the *E. coli* sensor is less cooperative in its response to MNNG as compared to MeI and other SN2 compounds (Table 3).

SN2 agents such as MeI, MMS, and DMS readily activate Ada in *E. coli* (Figures 1 and 2). Though these agents have not been observed to attack the phosphoester backbone of DNA, MeI has been observed to methylate Ada directly *in vitro*. It is not known to what extent SN2 agents activate Ada directly or indirectly via scant DNA phosphoester methylation. The yeast sensor responded to all of the SN2 methylation agents. Compared to the *E. coli* sensor, it responded to MMS and DMS with a lower detection threshold and a more graded, less cooperative response (Table 3). DMS can donate two methyl groups and is more toxic than MMS (SI Figure S4). Both sensors detected DMS at lower concentrations than MMS (Table 3).

Ada is also sensitive to the size of the alkyl group of PTE adducts on the DNA backbone. Larger alkyl groups sterically hinder the mechanism of detection and activate Ada poorly. EMS, an analogue of MMS that donates a larger ethyl group, was added to the *E. coli* and yeast sensors to test for the retention of this specificity. As expected, neither sensor responded to EMS (Figure 2D). The fact that the yeast sensor responded to the same range of alkylating agents as the native *E. coli* sensor suggests that the Gal4-N-Ada sensor retains much of its native activity in yeast, including its ability to detect and remove methyl PTE adducts on DNA.

**Biosensing Applications.** We previously reported the construction and screening of a library of 70 homologous methyl halide transferases (MHTs) and subsequent engineering of an MHT-expressing *S. cerevisiae* strain for production of high titers of MeI. To enable faster screening of MHT libraries and further engineering of productive yeast strains, we sought to use the yeast methylation sensor as a reporter of MHT activity. Different MHT enzymes have been shown to produce 0.3–1.3 mM MeI/h in *S. cerevisiae*. Because this range is consistent with the thresholds obtained for the genetic methylation sensors, we predicted that it would be possible to use them as a cell-based screen.

We carried out an experiment to determine if the methylation sensor could respond to MeI produced in yeast and whether the linear range is sufficient to distinguish enzymes of different activity. For this experiment, the *S. cerevisiae* P_{desCyc1} P_{Adh1} sensor strain was transformed with plasmids encoding a set of 7 MHT homologues (Figure 3A). Each strain was then grown to high density and MeI production was induced by adding NaI (Methods). The cells were grown for 1 h, after which each culture was analyzed by cytometry and the MeI titer was measured using tandem gas chromatography-mass spectrometry (GC-MS). The sensor output correlated with the activities produced by the different MHT homologues and saturated at high titers (Figure 3A).

Another potential application for the genetic sensor is as a biosensor for environmental samples. In particular, methyl bromide and methyl iodide are used in agriculture as soil fumigants. MeI is typically used at initial concentrations of 0.4–0.6 mM during fumigation but dissipates quickly due to
evaporation and subsequent light-induced decay. However, decay rates vary with soil composition, and MeI can be found in some soils for up to several days after exposure. On-site measurement of MeI levels with advanced instrumentation is impractical. The development of biosensors for fast, cheap on-site detection of compounds is a valuable alternative.

We sought to assess the sensor’s utility as a biosensor for the presence of MeI in soil. To test this, we added an aqueous solution of MeI to soil and then monitored MeI levels in the soil over time using the yeast P8x.Cyc1|PAdh1 sensor (Methods).

At different time points, the soil samples were fractionated by centrifugation and the runoff was collected. The runoff was then added to the culture and grown for 3 h, after which the cells were assayed by cytometry. Due to reaction and evaporation, MeI is lost exponentially from the soil (t½ = 1.5–2.0 h), which is consistent with the degradation of MeI in soils with high organic content. To control for MeI loss due to evaporation from the soil sample, half the sample tubes were closed during the assay, but this was found to have minimal impact on sensor activity. The sensor showed no activation when MeI was omitted, indicating that it responded specifically to the MeI present in the soil runoff. Because soil runoff is a complex mixture of compounds, this also demonstrated the sensor’s specificity and robustness.

### DISCUSSION

Our results demonstrate that the Ada methylation sensor is functional after its transfer from *E. coli* into yeast. Additional engineering was required to convert the Ada response into a transcriptional signal. MNNG, an Sn1 methylating agent, is only known to activate Ada indirectly by methylating the phosphoester backbone of DNA. The Sn2 reagents MeI, MMS, and DMS are hypothesized to methylate Ada’s Cys38 residue either directly or indirectly via undetectable amounts of methylation of the PTE backbone of DNA. The fact that the yeast sensor responded to the entire array of methylating agents supports the hypothesis that the N-terminal domain of Ada retained its native functions in the transfer from *E. coli* to yeast and that it detects and removes methyl PTE adducts from eukaryotic chromosomal DNA. We know of no other system for the detection or removal of DNA methyl PTE adducts in eukaryotic cells. It should be noted that while the sensor is operational in yeast, there was no impact on resistance to methylating toxins (SI Figure S4).

---

Table 3. Response of the Methylation Sensors to Different Agents

<table>
<thead>
<tr>
<th>organism</th>
<th>detection threshold (μM)</th>
<th>dynamic range (fold-change)</th>
<th>cooperativity*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MMS</td>
<td>DMS</td>
<td>MNNG</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td>783</td>
<td>340</td>
<td>1</td>
</tr>
<tr>
<td><em>S. cerevisiae</em></td>
<td>28</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

*E. coli* MG1655 containing pFM45. *S. cerevisiae* containing P8x.Cyc1 and PAdh1 driving the expression of Gal4-N-Ada. *Cooperativity is reported as the Hill coefficient in the fit equations detailed in the SI.*
Several tuning strategies were effective at changing the performance of the sensors in both organisms. Tuning the expression of the sensor protein consistently changed the detection threshold. However, in E. coli this came with a trade-off where at very low detection thresholds the basal activity of the sensor increased. This high basal activity also occurred for the engineered sensor in yeast and was constant irrespective of the strength of Gal4-N-Ada expression. We also found that the dynamic range and sensitivity of the yeast sensor could be improved by increasing the number of Ada operators in the output promoter, but this effect attenuates after three operators, consistent with previous observations. Although the yeast sensor was unable to achieve the same response as observed in proteins per cell to one of thousands of Ada proteins per cell, this came with a trade-off where at very low detection thresholds the basal activity of the sensor increased. This high basal activity also occurred for the engineered sensor in yeast and was constant irrespective of the strength of Gal4-N-Ada expression. We also found that the sensor increased. This high basal activity also occurred for the engineered sensor in yeast and was constant irrespective of the strength of Gal4-N-Ada expression. We also found that the sensor increase.

Additional network architectures can impact sensor performance. The native E. coli Ada system contains a strong positive feedback loop that amplifies a basal state of only 2–4 Ada proteins per cell to one of thousands of Ada proteins per cell. This has been shown to result in a more ultrasensitive, digital response. An ultrasensitive response could be replicated in yeast by either building such a loop or through the inclusion of interactions that sequester the regulator. On the other hand, to engineer a more linear, faster, or pulsed response in the sensor, one could implement negative genetic feedback. More complex architectures, such as feed-forward loops, can also be used to engineer complex dynamics and robustness to noise. Also useful for sensing systems would be the engineering of “scale-free sensing”, a characteristic of some complex networks that enables those networks to detect changes in the environment regardless of the level of the background signal. Clearly, many modes of action remain open to further engineer the sensors presented here to altered dynamics and specifications.

A persistent challenge in synthetic biology has been the development of well-characterized sensors that can respond to environmental signals. The design of a functional sensor and its tuning to a particular performance specification is often more difficult than building genetic circuits, in part because a ligand-binding event has to be converted into a transcriptional output. The movement of sensors between organisms has provided a successful avenue for the production of novel sensors. However, such movement often requires extensive re-engineering of a sensor and can be prone to unpredictable context effects. Sensors for strong methylating compounds present a novel sensory input that can be harnessed in genetic engineering. Beyond the applications outlined in this paper, the MHT enzymes and Ada sensor also offer new parts that can have other uses. For example, these modules may act as sensor and receiver devices for engineering communication between cells where the volatile methyl halide signal acts in the gas phase. Also, the slow decay and orthogonality of the methyl PTE adducts of DNA in eukaryotes may enable a route to engineering epigenetic memory, though this application will be limited by the dilution of the methyl PTE adducts by half with each successive generation. The specific parts from this study as well as the generalizable design and tuning strategies can be broadly applied to problems in design, species-transfer, and tuning of novel genetic sensors.

## MATERIALS AND METHODS

### Strains and Media.

Cloning was performed in E. coli DH10B and plasmids were transformed into E. coli MG1655 or E. coli MG1655 Δada for measurement. E. coli transfer function assays were performed in 1 mL of supplemented M9 media, containing 0.2% casamino acids (BD #228820), 1 mM thiamine HCl (Sigma-Aldrich T4625), and antibiotics. The E. coli MG1655 Δada strain was made by deleting the ada CDS (2307363–2308427) from the E. coli MG1655 chromosome using the technique of Datsenko and Wanner. To maintain plasmids in E. coli, we used antibiotic concentrations of 100 μg/mL for kanamycin and 100 μg/mL for spectinomycin. For all yeast experiments, we used S. cerevisiae strain SO992 (W303-derived, MATα, trp1, leu2, ura3, his3, ade2, can1 (s2)), modified as follows to create the sensor strains. Yeast sensor strains were made by integrating the Gal4-N-Ada expression cassette (contained in pJAC90/pJAC91) and the EGFP reporter cassette (contained in pJAC92, pJAC93, pJAC98, and pJAC100) into the his3 and trp1 loci, respectively. SI Table S3 summarizes the genotypes of the yeast strains used in this work. Yeast sensor strains were grown on standard dextrose (SD) complete media (Difco) for transfer function and soil detection assays. Yeast strains were grown in SD-Ura media for MHT experiments to maintain the MHT plasmids.

### Plasmid Construction.

All plasmids were constructed using the Chew-Back, Anneal, and Repair ( Gibson) method. The E. coli Ada sensor reporter plasmid pFM45 was derived from pSB3K (80,81) and contained the native ada promoter (−80 to +1) driving GFPmut3b fluorescent protein, a p15A origin of replication, and a kanamycin resistance marker. This plasmid is comparable to the standard promoter reference plasmid created by Kelly, et al. (2009). Plasmid pFM141 contains ada downstream of the PBAD promoter and medium strength RBS (B0032) as well as the araC gene, a spectinomycin resistance marker, and the incW origin of replication. Plasmids pJAC90 and pJAC91 were derived from the shuttle vector pNH603 (derived from pRS303, Addgene) and contained promoter P5′dha and P5′cycl, respectively, driving Gal4-N-Ada fusion expression as well as his3 homology regions flanking the Gal4-N-Ada expression cassette, the E. coli colE1 origin of replication, and an ampicillin resistance marker. The Gal4-N-Ada sequence in pJAC90/pJAC91 is a fusion of the Gal4 activation domain (amino acids 768–881), an intervening GSXSXS linker, and the N-terminal domain of Ada (amino acids 1–180). Yeast sensor reporter cassette plasmids pJAC100, pJAC92, pJAC93, and pJAC98 were derived from the pNH604 vector and contained 0, 1, 3, and 8 Ada operator sequences (AAATTTAAAGCCGCAA: consensus underlined), respectively, upstream of a P5′cycl promoter driving yeast-optimized enhanced green fluorescent protein (EGFP). Ada operator repeats were generated by iteratively cutting and ligation two annealed, 5′-phosphorylated oligos (5′-GGGGCCGAAAAATTAAGCG-CAAGATGC-3′ and 5′-GGGGCCGATCTTGCGCTT-TAATTTCG-3′) into pJAC92 with enzyme PspOMI. The PspOMI site is fully reconstituted on the 5′ end of the double stranded oligo but broken on the 3′ end such that iterative insertion of the oligo then digestion with PspOMI allows expansion of the number of operators. Plasmids pJAC90/pJAC91 and pJAC92/pJAC93/pJAC98/pJAC100 were transformed into S. cerevisiae SO992 using a standard lithium acetate technique, and their flanked expression cassettes were integrated into the his3 and trp1 loci, respectively.
plasmids, their components, and GenBank accession numbers are listed and described in detail in the Supporting Information.

Preparation of Alkylating Agents. Alkylating agents used for induction included methyl iodide (MeI; Sigma-Aldrich #289566), methyl methanesulfonate (MMS; Aldrich #129925), ethyl methanesulfonate (EMS; Sigma M0880), dimethyl sulfate (DMS; Sigma-Aldrich #D186309), and 1-methyl-3-nitro-1-nitrosoguanidine (MNNG; Aldrich #129941). MNNG was dissolved in DMSO. When sensor cultures were exposed to DMSO alone, no induction of fluorescence was observed (data not shown). To make accurate dilutions of MeI, it was important to first make a 1:100 water dilution of pure MeI in 1.5 mL Eppendorf tubes and then vortex the solution several times over 5 min to thoroughly dissolve the MeI before adding it to the 96-well plate. Higher concentrations of MeI required direct addition of MeI to the cultures, which must be done quickly and carefully given the compound’s volatility.

E. coli Response Function Assays. These assays were performed in 96-well plates (PlateOne #1896-2000). Triplicate cultures of E. coli MG1655 carrying plasmids were grown overnight (~18 h) in 3 mL of supplemented M9 media plus antibiotics and were diluted back 1:100 into 1 mL of media into the wells of the 96-well plate. The plate was covered with a breathable membrane (USA Scientific #9123-6100). Cultures were grown for 3 h at 37 °C while shaken at 900 rpm in a plate incubator in a fume hood until early exponential phase (OD600 = 0.2) and were then induced. For induction, 50× solutions of alkylating agents were first prepared in wells of a 96-well plate, so that a 12-channel pipet could be used to pipet 20 μL of the 50X solution in parallel into the rows of the 96-well culture plate. After alkylating agents were added, the 96-well plate was covered with an airtight sealing mat (Genesee Scientific #22-517) to prevent excessive evaporation of the alkylating agents. Once induced, cells were grown for 3 h, as described above. Cells were collected by pipetting 2 μL of each culture into 200 μL of cold phosphate buffered saline (PBS; pH 7) and 2 mg/mL kanamycin (to stop translation) in a 96-well cytometry plate (Costar #3363). These samples were then analyzed by cytometry as described.

S. cerevisiae Response Function Assays. S. cerevisiae transfer function assays were carried out similarly to E. coli assays in 1 mL cultures in 96-well plates. Triplicate cultures of S. cerevisiae were grown overnight in Standard Dextrose (SD) media on a rotator (New Brunswick TC7) at 80 rpm at 30 °C. The next day, cultures were diluted back 1:100 in SD, and grown to OD600 of 0.04 on a shaker (Eppendorf MixMate) at 800 rpm. Methylating agents were added to the cultures, as described above, and growth was continued for an additional 3 h. Cells were collected by adding 10 μL of culture to 200 μL of cold phosphate buffered saline (PBS; pH 7) and 5 μg/mL cyclohexamide (Sigma C1988) to arrest translation in 96-well plates. These samples were then analyzed by cytometry, as described.

Cytometry and Data Analysis. Cells were analyzed by flow cytometry on a BD LSRII using a 488 nm laser and 510/20 nm band-pass filter to collect GFP and EGFP fluorescence. Samples of up to 40 μL of cells in cold PBS were analyzed at a flow rate of 0.5 μL/s until 50 000 gated counts were collected. FSC-H and SSC-H thresholds were set to exclude background events. For accurate, reproducible fluorescence measurements, it was critical that cells were diluted at least 100-fold (OD600 < 0.04) so the event rate was low enough for individual cells to be measured. Data was analyzed using FlowJo software (Treestar). The cell populations were gated by time and forward/side scatter to exclude read-through from previous wells and residual background events. The final analyzed populations included >90% of collected events. The geometric mean of the fluorescence histogram of each gated population was calculated and is reported here as the fluorescence value of a sample in arbitrary units (au). Modeling and fitting of response functions was done in Matlab using the nlinfit function applied to the model presented in the Supporting Information. Fit data sets excluded data points where cells experienced toxicity.

Detection of Mel Production by MHTs. The S. cerevisiae PβEcCyc1::PAdh1 sensor strain was transformed with the methyl halide transferase (MHT) expression plasmids previously described.37 Plasmids expressing MHTs from the following organisms were tested: Batis maritima, Burkholderia pseudomallei, Burkholderia xenovorans, Vitis vinifera, Burkholderia thailandensis, Brassica rapa, and Oryza sativa. Transformants were grown overnight in 2 mL SD-Ura selective media to retain the MHT plasmids. The following day, cultures were added to 100 mL of fresh SD-Ura and grown for 24 h. Cultures were centrifuged to pellet the cells, then resuspended in 8 mL YPD (final OD600 = 50) and 1 mL of 1 M NaI as a source of iodide. Cells were grown for 1 h in 14 mL Falcon tubes (Becton Dickinson #35209) sealed with Septa Seal rubber stoppers (Sigma #124605). Gas chromatography–mass spectrometry (GC-MS) was conducted using a model 6850 Series II Network GC system and model 5973 Network mass-selective system (Agilent). GC-MS measurements were done as previously described,37 except for the following changes: the oven temperature was set at 55 °C and increased to 70 °C over a period of 9 min so as to process all samples, including the standard curve, in one run. Samples were injected 30 s apart so that their MeI GC peaks were clearly separated and identifiable with respect to the air peak. A sample of the remaining cells was diluted 1:1000 in SD media and grown for an additional 3 h before being assayed for fluorescence by cytometry as described. To better illustrate how fluorescence changed with MeI production, the fluorescence background of the PβEcCyc1::PAdh1 sensor strain was subtracted from the fluorescence measurements of each of the experimental strains measured. This was only done for the data in Figure 3A; all other data does not have the background subtracted.

Detection of Mel Contamination in Soil Samples. Garden soil was added to the 1 mL fill line (~230 mg) in a 1.5 mL Eppendorf tube. Then, 800 μL of distilled water was added to the soil and the sample was briefly vortexed. To half the samples, Mel was added to a final concentration of 0.6 mM with respect to the water, simulating the amount added to the soil in agriculture.59 No Mel was added to control samples. After addition of water and Mel, all samples were vortexed for 20 s and then placed in a dark fume hood at room temperature (20 °C). Half the sample tubes were left closed and the other half open. Samples containing Mel and control (water only) soil samples were set up 0, 1, 2, 4, 8, and 24 h prior to processing. Processing was done as follows: all tubes were closed, the tubes were tapped to bring the soil sample to the top, and a hot 26 gauge needle was used to pierce the bottom of the sample tubes. Pierced tubes were then placed in 1.5 mL Eppendorf collection tubes and centrifuged at 10 000 rpm for 15 s. This served to separate the majority of the liquid fraction from the soil. Samples were then centrifuged further in closed caps for 5 min at 13 500 rpm, after which 400 μL of liquid was removed, taking care to avoid picking up solid material with the pipet. This supernatant was then diluted 1:10 in water and 40
µL of this dilution was added to 1 mL cultures of *S. cerevisiae* PrLAdh1 reporter cells in a 96-well plate in triplicate. The cells were shaken at 800 rpm at 30 °C for 3 h and then measured by flow cytometry as described.

## ASSOCIATED CONTENT

### Supporting Information

Data describing the impact of alkylating agents on cell growth and cytometry fluorescence distributions, the standard curve for the GC-MS data, a detailed description of the sensor activation model used to describe the response functions, and tables listing the yeast stain genotypes and genetic parts and plasmids used in this work. This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*E-mail: cavvoigt@gmail.com.*

### Author Contributions

$F$.M. and A.H. contributed equally to this work.

### Author Contributions

F.M. designed and performed the experiments involving *E. coli* and *S. cerevisiae*. A.H. designed and performed the experiments involving *S. cerevisiae*. J.C. assisted with *S. cerevisiae* assays and conceived of the soil contamination experiments. F.M. and C.A.V. wrote the manuscript. W.A.L. and C.A.V. managed the project.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

The authors thank P. R. O. de Montellano and J. Johnston for technical assistance with the GC-MS. F.M. was supported by a National Science Foundation (NSF) Graduate Research Fellowship. C.A.V. is supported by Defense Advanced Research Projects Agency Agency Chronical of Lineage Indicative of Origins (DARPA CLIO) N66001-12-C-4018, Office of Naval Research (ONR) N00014-10-1-0245, ONR N00014-13-1-0074, NSF 557686-2117, and the NSF Synthetic Biology Engineering Research Center (SynBERC) SA5284-11210. W.A.L. is supported by the NSF Synthetic Biology Engineering Research Center (SynBERC) EEC 0540879, National Institutes of Health (NIH) P50—5PS0GM081879-03 (Exploring Design Principles of Cellular Control Circuits), NIH PN2—SPN2EY016546-09 (Cellular Control: Synthetic Signaling/ Motility (RMI)), and the Howard Hughes Medical Institute. A.H. is a Fellow of The Jane Coffin Childs Memorial Fund for Medical Research. This investigation has been aided by a grant from The Jane Coffin Childs Memorial Fund for Medical Research.

### REFERENCES


Development of a mevalonate biosensor.

J. Am. Chem. Soc. 131(10), 1

expression.


optimization of gene dosage in


Carcinogenesis 31, 59–70.


