Supporting Information to article entitled “Ultra-Sensitive Reporter Protein Detection in Genetically-Engineered Bacteria”

Authors: Mona Wells,* Michael Gösch, Rudolf Rigler, Hauke Harms, Theo Lasser, and Jan Roelof van der Meer

(†Swiss Federal Institute of Technology (EPFL), Laboratories of Soil Science and Biomedical Optics, CH-1015, Lausanne, Switzerland; §Centre for Environmental Research, Department of Microbiology, D-04318, Leipzig, Germany; ‡‡University of Lausanne, Department of Fundamental Microbiology, CH-1015, Lausanne, Switzerland)

Contents: This Supporting Information contains additional information on calculations (rate of egfp production/nmol analyte in solution, maturation-limited minimum induction time, pPR-arsR-ABS intra-cellular baseline egfp concentration, quantum yield/fluorophore extinction coefficient corrections for comparison of gfp literature results to egfp data), additional FOMs (Table S-1) and a figure (Figure S-1) showing the relationship of arsenite MDLs vs time as determined by LIF-CS.
Supporting Information – Experimental, Calculations

Rate of egfp production. For the time-dependent response data, linear (zeroth order kinetic) and exponential (first order kinetic) fits of response vs. time were not statistically differentiable except for very high egfp production. For high egfp production (i.e. high arsenite concentration), an exponential does fit the data better (F-test), however, this is not definitive since \( \text{OD}_{660} \)-based measurements of growth may be influenced by the growth stage (e.g. a smaller number of large bacteria in linear growth may have the same OD as a larger number of small bacteria in the stationary phase), and since these higher arsenic concentrations influence growth more than the lower ones (e.g. for \( t > 4 \) h ODs for 10.0 \( \mu M \) arsenite may be up to 25% lower than for lower concentrations; such effects then constituting an influence by growth stage, as described). For linear fits of time dependent data, \( R^2 \) values ranged from 0.95-0.98. Thus though linear behavior may not obtain over all induction times and arsenite concentrations, for the experimental conditions here assumption of linear behavior is sufficient to estimate the rate of nM egfp formation in lysates, as reported in the Results and Discussion. Values for nmol/h egfp per nmol arsenite in solution are forthcoming given the total volume of cells/solution harvested (2 ml) and the volume of resuspension for lysis (200 \( \mu l \)). Rates of egfp production can also be estimated from concentration-dependent induction data, by taking the total amount of egfp formed in \( t \) hours, subtracting background, and performing the necessary unit transformations. Estimated rates from data presented here, as well as from other similar or identical experiments performed in the course of this work, are variable, but typically agree to within ± 50% of the average rates for all experiments.

Maturation-limited minimum induction time. The maturation-limited minimum induction time for detection is estimated using the first order rate expression \( \ln[\text{egfp}(t)/\text{egfp}_0] = kt \), with \( \text{egfp}(t)/\text{egfp}_0 = 1.2 \), and the first order maturation constant \( k = 1.5 \text{ h}^{-1} \). When the initial egfp concentration is above the \( \text{MDL}_{\text{egfp}}^{\text{LIF-CS}} \), we find that a 20% increase in signal over baseline is statistically significant, hence the value of 1.2 for \( \text{egfp}(t)/\text{egfp}_0 \). Clearly, if the initial lysate gfp concentration were below the \( \text{MDL}_{\text{egfp}}^{\text{LIF-CS}} \), more time would be required for egfp concentrations to build before a response could be measured.
**pPR-arsR-ABS intra-cellular baseline egfp concentration.** The pPR-arsR-ABS intra-cellular baseline egfp concentration estimated from *lysate data* was obtained from the average LIF-CS measured lysate concentration, the total lysate volume of 200 µl, the average number of cells harvested \((1.5 \times 10^9)\), and the approximate volume per cell (1 fl), according to

\[
C_{\text{egfp, cellular}} = \frac{C_{\text{lysat}}}{} \text{lysate \ V_{\text{lysat}}} \text{cells} \text{cells} \text{N}
\]

(with appropriate unit conversions), where \(C\) is concentration, \(V\) is volume, and \(N\) is a bacterial number count. Because average spike recoveries are 600% for \(t = 0\) samples, \(C_{\text{egfp, cellular}}\) is further divided by a factor of 6. The intra-cellular baseline egfp expression estimated from whole cells is obtained by first estimating the egfp molecules in the spectroscopic focal volume using the experimentally observed detector counts and the average counts/egfp molecule determined from analysis of auto-correlation curves measured during instrument calibration with egfp. The egfp cellular concentration is obtained from the molecules egfp/focal volume, after calibrating to egfp standards and integrating the variable possible bacterial geometries with respect to the fixed geometry of the focal volume element (average overlap \(\sim 33\%\)). We anticipate that the number so obtained is low for reasons mentioned in the Results and Discussion.

**Quantum yield/fluorophore extinction coefficient corrections for comparison of gfp literature results to present data.** As noted, enhanced gfp was produced by the bacteria employed here. Assuming identical source and detector geometries, and identical intensity of source radiation, concentrations of egfp can be related to gfp via

\[
C_{\text{egfp}} = \frac{\varepsilon_{\text{gfp}}}{\varepsilon_{\text{egfp}}} \phi_{\text{gfp}} \phi_{\text{egfp}} C_{\text{gfp}}^{1},
\]

where \(C\) is concentration, \(\varepsilon\) is the fluorophore extinction coefficient, and \(\phi\) is the quantum yield. The quantum yield and extinction of gfp are 0.8 and 9,500 cm\(^{-1}\)M\(^{-1}\), respectively, as compared to 0.6 and 55,000 cm\(^{-1}\)M\(^{-1}\), respectively for egfp, thus the theoretical

\[
\text{MDL}_{\text{egfp}} = \frac{\varepsilon_{\text{gfp}} \phi_{\text{gfp}}}{\varepsilon_{\text{egfp}} \phi_{\text{egfp}}} \text{MDL}_{\text{gfp}} = 0.23 \text{MDL}_{\text{gfp}},
\]

\[2\]
Table S1. Additional figures of merit associated with the linear response of pPR-arsR-ABS to arsenite exposure as a function of concentration at different induction times.

<table>
<thead>
<tr>
<th></th>
<th>0.5 h</th>
<th>1 h</th>
<th>2 h</th>
<th>3 h</th>
<th>4 h</th>
<th>0.5 h</th>
<th>1 h</th>
<th>2 h</th>
<th>3 h</th>
<th>4 h</th>
<th>0.5 h</th>
<th>1 h</th>
<th>2 h</th>
<th>3 h</th>
<th>4 h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R^2 ) (^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LIF-CS (^f)</td>
<td>0.98</td>
<td>0.95</td>
<td>1.00</td>
<td>1.00</td>
<td>0.99</td>
<td>6.8</td>
<td>14</td>
<td>1.4</td>
<td>3.9</td>
<td>6.9</td>
<td>146</td>
<td>790</td>
<td>1100</td>
<td>690</td>
<td>650</td>
</tr>
<tr>
<td>EFM (^g)</td>
<td>0.97</td>
<td>0.95</td>
<td>0.97</td>
<td>0.99</td>
<td>0.96</td>
<td>9.3</td>
<td>14</td>
<td>11</td>
<td>6.5</td>
<td>11</td>
<td>550</td>
<td>920</td>
<td>2700</td>
<td>2900</td>
<td>930</td>
</tr>
<tr>
<td>SSF</td>
<td>0.99</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>28</td>
<td>1.8</td>
<td>4.9</td>
<td>3.4</td>
<td>5.6</td>
<td>55</td>
<td>610</td>
<td>600</td>
<td>730</td>
<td>710</td>
</tr>
<tr>
<td>( s_c ) (^b) (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDR (^c) (nM arsenite)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) correlation coefficient for linear portion of response curve

\(^b\) \( s_c \) is the standard error (relative) associated with readings from the linear part of the calibration curve assuming the number of replicate measurements (3 for LIF-CS, 3 for EFM, 8 for SSF) and a measurement midway in the LDR

\(^c\) LDR = LOL-LOQ, LOQ = 3*MDL, LDR is linear dynamic range. LOL is limit of linearity. Increasing sensitivity and comparatively low relative standard deviations (RSDs) for LIF-CS are advantageous in increasing the LDR, but an optimum is reached at \( t \approx 2 \) hours as a result of the LOL decreasing at induction times above 2 hours. This optimum is apparent for EFM also, but is less marked since the LOL for this technique is more robustly sustained; the optimum is not observed for SSF. The whole cell LDR \(( t = 1.5 \) h\) is ca. 1100 nM, in agreement with the LIF-CS lysate LDRs at \( t = 1-2 \) h.
Figure S1. Exponential fit to LIF-CS MDL data.

Supporting Information, References