molecular
systems
biology

Engineering bacterial thiosulfate and tetrathionate sensors for detecting gut inflammation

Kristina Daeffler, Jeffrey Galley, Mr. Ravi Sheth, Ms. Laura Ortiz-Velez, Christopher Bibb, Noah Shroyer, Robert Britton and Jeffrey Tabor

Corresponding author: Jeffrey Tabor, Rice University

Editor: Maria Polychronidou

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision   13 December 2016

Thank you again for submitting your work to Molecular Systems Biology. We have now heard back from the three referees who agreed to evaluate your study. As you will see below, the reviewers acknowledge that the presented sensors are a useful tool for future analyses. They raise however a series of concerns, which we would ask you to address in a revision.

The reviewers' recommendations are rather clear so there is no need to repeat the points listed below. One of the more fundamental issues refers to the need to include additional analyses to provide better support for the function of the sensors in vivo.

REFEREE REPORTS

Reviewer #1:

The manuscript by Däffler et al. describes bacterial thiosulfate and tetrathionate sensors for the detection of gut inflammation.

The paper comprises two parts: first, identification, reconstitution in E. Coli Nissle, and testing of the aforementioned two-component signaling pathways. This part is well-done, detailed and thorough. The results will be of interest to a wide community of microbiologists and synthetic biologists. I do not have criticisms here.

In the second part, the engineered bacteria are used as inflammation sensors in mice. I am afraid that

this second part is lacking. The most striking result is a very weak correlation between histological inflammation markers and the sensor response.

The histological scoring is not a "gold standard", and I am surprised the authors did not use multiple pathologists to score the same samples to increase their confidence. However, this is not the point. The bacteria should detect their respective ligands, thiosulfate and tetrathionate, in the gut. To prove this statement, there must be an independent, direct, chemically-based analytical measurement of these species in the relevant physiological samples. For example, this method might be used https://www.ncbi.nlm.nih.gov/pubmed/19559870
or another similar method.

Granted, the authors performed extensive titrations and measurements in vitro. However, the responses in vivo are something entirely different, as the bacteria might lose their payload (even if the residence time is short), etc. Thus having an independent measurement of the ligand and the correlation between this measurement and the biosensor output is crucial to support the conclusion of this study. The fact that the correlation between inflammation and sensor output is weak, should be worrying.

Minor comments

1) I request that 2D flow cytometry data (Cherry vs GFP) of all the animal-derived bacterial samples be shown in supplementary information.
2) ROC curves should be built to illustrate sensor diagnostic performance; AuROC should be calculated and reported.

Reviewer #2:

The authors used a bioinformatics approach to identify two component systems from Shewannella bacteria capable to sensing thiosulfate and tetrathionate. The authors then implemented these systems first in laboratory E. coli and find that they are functional. The authors go on to implement these systems in E. coli Nissle probiotic bacteria and show that they are again functional. They then show that when these engineered bacteria are administered to colitis model mice by oral gavage, they provide a readout of gut thionate levels and that these readouts correlate with gut inflammation by histology scoring. In general, this manuscript is easy to read, provides compelling arguments, and presents well-supported data.

I believe this work to be significant. To my knowledge, this work is the first example of an engineered bacteria with a thiosulfate sensor that is functional in a mouse model. This builds off previous work. Previous work provided the mouse model for colitis and E. coli Nissle bacterial chassis strain; evidence of correlation of thiosulfate and tetrathionate to gut inflammation; biochemical understandings of thiosulfate, tetrathionate, and two component systems in general; and a previously characterized thiosulfate sensory system to which the current system can be compared. While building off previous work, these authors put the whole system together, with new sensors, for the first time. By providing new and useful thiosulfate and tetrathionate sensors to the community, this work provides methods for implementing and analyzing engineered probiotic systems as biomarker readouts, and lays the groundwork for the system that may be able to detect gut inflammation in humans. The work should be of interest to a large audience of researchers including: synthetic biologists, gastroenterologists, gut microbiologists, and bioinformaticians. I think it is well suited for Molecular Systems Biology, but believe there are some points to be addressed before suitable for publication.

1. In figure 2D, what are the range of concentrations used to examine specificity. I believe I missed it, but clarifying physiological ranges in the text would strengthen the authors' results.

2. It might be beneficial to expose Nissle carrying these sensors to DSS outside of the host organism to determine if they show differences in sfGFP and mCherry response due to this molecule being present in its environment alone. I am wondering if this treatment somehow impacts the sensor.

3. In the comparisons of this work's FACS method, it would be fair to discuss the possible drawbacks of this method such as the maturation time required and the requirement of FACS equipment to provide accurate readouts. The authors could comment on how these issues could be overcome (such as by the use of colorimetric indicators as explained later in the discussion).

4. Judging by the data in panel F of Figure 5 the claim that "in vivo activity of our thiosulfate sensor is proportional to gut inflammation" found on pages 17 and 6 might be better put "in vivo activity of our thiosulfate sensor correlates with gut inflammation." Correlation is clear,, though weak. I find proportionality as too strong of wording and believe this to be one of the weaker conclusions in the paper. Yet, this is an important conclusion.

5. When the authors state that the ThsSR strain operates robustly, I feel this may be a slight overstatement. I feel the data show it works in most instances, but not all.

6. Is there a way for the authors to figure out if the TtrSR strain circuity is broken, or the concentration of tetrathionate is low? Is there any way to test for tetrathionate concentration directly? HPLC methods, etc. Although the article did not claim one or the other conclusion, this experiment would help decouple the question of whether the sensor doesn't work in vivo or if levels are too low to detect -- providing valuable insights to the field. I believe this is important.

-Presentation and style
Label plasmid numbers in Figure 2A and 3C
Indicate k1/2 and n in Figure 2C and 3G, or at least in caption.

Reviewer #3:

Summary
In this study, Daeffler et al. reported the identification and characterization of one thiosulfate and on tetrathionate sensors from marine Shevanella. The two sensoring systems were reconstituted in E.coli and different substrates were tested to show the efficiency and high substrate specificity for these two systems when tested in tube. The authors further optimized the sensoring system and gavaged the engineered strains into mice. The authors show that the activity of the thiosulfate sensor is proportional to gut inflammation claiming that this molecule could be a novel inflammation biomarker. In comparison, tetrathionate sensor shows low in vivo activity at high inflammation levels. Comparing to previous literature, the highlights of the paper lie in that 1. the identification and characterization of a new thiosulfate sensor, 2. the authors have shown that the thiosulfate sensoring system works well, non-invasively, in both health and DSS mouse model as a reporting tool for gut inflammation. In general, the publication in Molecular System Biology is recommended after the comments below have been addressed.

Major points
1. The experiments reported for characterization of both thiosulfate and tetrathionate-sensing TCSs are somewhat similar to each other. It will be better if the authors could combine these two parts into one in the results section.

2. The authors show that the expression of sfGFP levels in thiosulfate sensing system are weakly proportional to histologic score. How does sfGFP level relate to the thiosulfate level in the mice gut? It will be more convincing if the authors could first show that the sfGFP expression level is in good correlation with the concentration of thiosulfate molecules in the (healthy or DSS) mice gut, then shows that it is proportional to inflammation level. This will bridge the gap between in vitro aerobic test of their sensing system and the in vivo outputs (sfGFP expression levels). This will also set up a direct linkage between thiosulfate concentration and inflammation in DSS disease mice (although more experiments are definitely necessary to show the validity of thiosulfate as a biomarker for gut inflammation, these experiments do not have to be included in this study). It is the same for tetrathionate sensing system as well.

3. It might be helpful to normalize the population of the engineered E.coli that has been gavaged with other microbial members in the healthy and DSS disease mice via sequencing. This will show how much E. coli is actually there after 6hrs gavage and euthanization. One reason is because the microbiota sometimes changed a lot in diseased DSS model. An elevated population of E coli in the DSS mice would also lead to the upregulation of sfGFP levels comparing to those healthy subjects, in which the gavaged strain might be cleared off much faster. This might help to explain why the sensing system for tetrathionate barely works in vivo because one of the possibility is the strain gets rapidly cleaned off the gut in DSS mice comparing to the healthy group.

Minor points
1. I would recommend combining Figures 2 and 3 into one figure.

1st Revision - authors' response   14 February 2017

Response to Reviewers
We thank the three Reviewers for reading our manuscript carefully and providing helpful feedback. We respond to each Reviewer criticism below.

Reviewer #1:
 …*In the second part, the engineered bacteria are used as inflammation sensors in mice. I am afraid that this second part is lacking. The most striking result is a very weak correlation between histological inflammation markers and the sensor response.*

We observe a statistically significant increase in ThsSR output when mice have inflammation, with zero false positives (**Fig. 5B**). This response is completely abolished in a negative control experiment where ThsSR is disabled by mutation (**Fig. 5C**). These results demonstrate that we can use our engineered thiosulfate-sensing bacteria to detect a diseased state *in vivo*. We believe this is a significant advance. Due to the complexity of the gut environment, the fact that we measure our bacteria at a single time point, and the subjective nature of pathological scoring, we do not find it too surprising that the correlation between inflammation score and sfGFP levels is weak. Future advances in our sensor or strain design, *in vivo* protocol, or analytical methods, particularly those enabling increased spatial and temporal resolution *in vivo,* all may improve the strength of the correlation between inflammation and ThsSR output. Nonetheless, we have removed the corresponding figure and all claims of proportionality between ThsSR activity and extent of inflammation from the manuscript.

The histological scoring is not a "gold standard", and I am surprised the authors did not use multiple pathologists to score the same samples to increase their confidence.

We thank the Reviewer for this feedback. We have increased confidence in our inflammation scoring by having a second blinded pathologist score the samples (Appendix **Fig. S24**), and described this change in the main text.

However, this is not the point. The bacteria should detect their respective ligands, thiosulfate and tetrathionate, in the gut. To prove this statement, there must be an independent, direct, chemically-based analytical measurement of these species in the relevant physiological samples. For example, this method might be used https://www.ncbi.nlm.nih.gov/pubmed/19559870 or another similar method.

Accurate measurement of many metabolites from animal stool by direct chemical methods is highly challenging. Thiosulfate and tetrathionate are colorless molecules and therefore need to undergo a chemical transformation to be easily distinguished by chromatography, requiring high specificity of reactivity toward the molecule of interest over other molecules in the sample. The method that the Reviewer provides in the above reference is used for urine, which is low in sulfide and other sulfur-donors that would react with their derivatizing agent, thereby enabling their successful quantification. However, the chemical reaction they use is not specific for thiosulfate over other

sulfur-donors likely to be found in fecal samples (sulfide, iron-sulfur clusters, etc.). While it is likely possible with the right columns, running buffers, and analytical equipment to develop a protocol to quantify these molecules, it would require significant development, be very costly, and is outside the scope of this study.

Commercial analysis is another option. However, Metabolon, a leading company in this space, quoted us $17,000 and a four-month turn-around for qualitative analysis of thiosulfate with an unknown lower limit of detection. Furthermore, they cannot measure tetrathionate. Thus, we aim to develop our own quantitative thiosulfate and tetrathionate analytical methods in future work.

Nonetheless, we clearly demonstrate that our engineered bacteria detect and report inflammation, as discussed above. We believe this is an important and non-trivial advance and sufficient for this first report. To address the fact that we do not perform analytical chemical measurements, we have removed all claims that we are directly detecting thiosulfate in the gut. Finally, as we write in the Discussion, our study is the first to *suggest* thiosulfate may be an inflammation biomarker and will motivate future studies into gut thiosulfate levels, gut sulfur metabolism, and inflammation.

Granted, the authors performed extensive titrations and measurements in vitro. However, the responses in vivo are something entirely different, as the bacteria might lose their payload (even if the residence time is short), etc. Thus having an independent measurement of the ligand and the correlation between this measurement and the biosensor output is crucial to support the conclusion of this study. The fact that the correlation between inflammation and sensor output is weak, should be worrying.

We completely agree with the Reviewer that *in vitro* and *in vivo* environments are very different. To address this concern, we have performed additional experiments wherein we harvest colons from healthy mice, inject our sensor bacteria and the relevant ligand (thiosulfate or tetrathionate), incubate for 6 hours, harvest the colon contents, and measure the response of our sensors by flow cytometry (Appendix **Fig. S21** and **S22**). This experiment shows that our bacteria do not lose their payloads (i.e. sensors) after a 6-hour incubation in the colon without antibiotics and that both of our sensors are activated by their respective ligands in the complex colon environment. This *ex vivo* experiment provides strong additional evidence that our sensors function as designed *in vivo*.

Minor comments
1) I request that 2D flow cytometry data (Cherry vs GFP) of all the animal-derived bacterial samples be shown in supplementary information.

We have added all data in Appendix **Fig. S27**.

2) ROC curves should be built to illustrate sensor diagnostic performance; AuROC should be calculated and reported.

We have built ROC curves and calculated AuROC to be 0.8692 for ThsSR *in vivo* due to the strong response to inflammation and very low false positive rate (Appendix **Fig. S25**).

Reviewer #2:
1. In figure 2D, what are the range of concentrations used to examine specificity. I believe I missed it, but clarifying physiological ranges in the text would strengthen the authors' results.

We tested 10 mM of each ligand. We have incorporated this information into the text and the **Fig. 2D** legend for clarity.

2. It might be beneficial to expose Nissle carrying these sensors to DSS outside of the host organism to determine if they show differences in sfGFP and mCherry response due to this molecule being present in its environment alone. I am wondering if this treatment somehow impacts the sensor.

We agree that this is an important experiment. As shown in Appendix **Fig. S23**, DSS does not have any effect on either of our sensors *in vitro*.

3. In the comparisons of this work's FACS method, it would be fair to discuss the possible drawbacks of this method such as the maturation time required and the requirement of FACS equipment to provide accurate readouts. The authors could comment on how these issues could be overcome (such as by the use of colorimetric indicators as explained later in the discussion).

This is an interesting point and we have now added the requested discussion in the text:

"Potential drawbacks of this method include the requirement for flow cytometry equipment to measure fluorescence and the maturation time required for chromophore formation. We show that sfGFP and mCherry maturation is complete after one hour in the presence of oxygen and stable for a minimum of two hours at 37C in the presence of a translation inhibitor (Appendix **Fig. S19**), however other readouts including colorimetric assays could be used if necessary."

4. Judging by the data in panel F of Figure 5 the claim that "in vivo activity of our thiosulfate sensor is proportional to gut inflammation" found on pages 17 and 6 might be better put "in vivo activity of our thiosulfate sensor correlates with gut inflammation." Correlation is clear,, though weak. I find proportionality as too strong of wording and believe this to be one of the weaker conclusions in the paper. Yet, this is an important conclusion.

We agree that may have been too strong of a claim and have removed the sfGFP vs. histopathology score panels from **Fig. 5** and the corresponding text. Instead we give the ROC curve requested by Reviewer #1 as a better indicator of sensor performance as a diagnostic of gut inflammation.

5. When the authors state that the ThsSR strain operates robustly, I feel this may be a slight overstatement. I feel the data show it works in most instances, but not all.

We thank the Reviewer for noticing this issue. We have changed the statement to say: "These results demonstrate that ThsSR can be activated in a living mouse gut and indicate that thiosulfate may be elevated upon DSS treatment.", which we believe is supported by the data.

6. Is there a way for the authors to figure out if the TtrSR strain circuity is broken, or the concentration of tetrathionate is low? Is there any way to test for tetrathionate concentration directly? HPLC methods, etc. Although the article did not claim one or the other conclusion, this experiment would help decouple the question of whether the sensor doesn't work in vivo or if levels are too low to detect -- providing valuable insights to the field. I believe this is important.

Our newly added colon explant experiments show that the sensor is still functional after 6 hours of incubation in the colon environment. We agree that direct quantification of thiosulfate and tetrathionate would further strengthen our conclusions. However, as mentioned above, direct measurement of thiosulfate and tetrathionate in the gut is outside the scope of this study.

-Presentation and style
Label plasmid numbers in Figure 2A and 3C
Plasmid numbers have been labeled in the corresponding figure panel.

Indicate k1/2 and n in Figure 2C and 3G, or at least in caption.
k1/2 and n have been labeled in the corresponding figure panel.

Reviewer #3:
 1. The experiments reported for characterization of both thiosulfate and tetrathionate-sensing TCSs are somewhat similar to each other. It will be better if the authors could combine these two parts into one in the results section.

We thank the Reviewer for this suggestion. Indeed, we originally wrote the manuscript in this way. However, we ultimately decided to separate the two sensor characterization sections in the text. Although the experimental methods are similar, there are considerable differences between the two sensors, which became confusing to discuss together. However, in response to the Reviewer's comment, we have shortened the description of TtrSR characterization to minimize redundant explanations.

2. The authors show that the expression of sfGFP levels in thiosulfate sensing system are weakly proportional to histologic score. How does sfGFP level relate to the thiosulfate level in the mice gut? It will be more convincing if the authors could first show that the sfGFP expression level is in good correlation with the concentration of thiosulfate molecules in the (healthy or DSS) mice gut, then shows that it is proportional to inflammation level. This will bridge the gap between in vitro aerobic test of their sensing system and the in vivo outputs (sfGFP expression levels). This will also set up a direct linkage between thiosulfate concentration and inflammation in DSS disease mice (although more experiments are definitely necessary to show the validity of thiosulfate as a biomarker for gut inflammation, these experiments do not have to be included in this study). It is the same for tetrathionate sensing system as well.

We agree that directly correlating thiosulfate levels with GFP output would strengthen our interpretations of *in vivo* sensor performance. However, as discussed in the response to Reviewer #1, developing a method for this is difficult. We have added the colon explant experiment and ROC analysis to strengthen our interpretations. We have edited the manuscript to avoid conclusions involving *in vivo* thiosulfate concentrations.

3. It might be helpful to normalize the population of the engineered E.coli that has been gavaged with other microbial members in the healthy and DSS disease mice via sequencing. This will show how much E. coli is actually there after 6hrs gavage and euthanization. One reason is because the microbiota sometimes changed a lot in diseased DSS model. An elevated population of E coli in the DSS mice would also lead to the upregulation of sfGFP levels comparing to those healthy subjects, in which the gavaged strain might be cleared off much faster. This might help to explain why the sensing system for tetrathionate barely works in vivo because one of the possibility is the strain gets rapidly cleaned off the gut in DSS mice comparing to the healthy group.

Using flow cytometry we measure whole cell fluorescence from >250 individual bacterial counts and can distinguish our sensor bacteria with strong constitutive mCherry expression from other bacteria in the gut. The sfGFP values reported are the geometric mean of individual cell fluorescence measurements and do not reflect bulk GFP, which would require normalization to cell count. We do see different bacterial counts at the end of the 6 hour experiment between healthy and DSS-treated mice, with higher counts for DSS-treated mice. Enterobacteria blooms have been observed during inflammation, but we don't believe that this will have an effect on sfGFP levels.

Minor points

1. I would recommend combining Figures 2 and 3 into one figure.
We thank the reviewer for the suggestion. In an initial version of the manuscript, we did combine Figures 2 and 3. However, as described above, we believe that this reduces readability because the systems have many different details. Thus, we have chosen to leave them separate.

2nd Editorial Decision   10 March 2017

Thank you for sending us your revised manuscript. We have now heard back from reviewer #2 who agreed to evaluate the study. As you will see below, s/he thinks that the major issues have been satisfactorily addressed. However, s/he raises a couple of remaining concerns, which we would ask you to address in a minor revision.

REFEREE REPORT

Reviewer #2:

I thank the authors for their work in carefully addressing the reviewer comments. I still believe that this manuscript is well suited for Molecular Systems Biology because the engineered thiosulfate-sensing bacteria detect a diseased state in vivo - even though further optimization could improve the system. After considering the revisions, the following points remain:

1. I would still like to see the plot showing GFP readout with histological scoring data included in the supplement. While I do not believe that this data is sufficient to make claims of proportionality, I do believe that it could provide valuable information to the field about the behavior of the sensor and the robustness of the FACS method presented in this work. Including only the ROC curve in S25 seems to obscure this information. Please also clarify if the ROC curve was generated using a "diagnosis" by histological scoring or simply the addition of DSS.

2. I thank the authors for performing ex vivo analysis as shown in S21 and S22. This data brings two points to my attention:

(i) the first is that the D55A mutation seems to only attenuate rather than completely "abolish" the response of the TtrSR both in vivo and in vitro. Am I seeing that correctly? I am not sure as to the significance. If it is attenuated, perhaps the wording should be changed.

(ii) the other is that the authors should use a statistical test to show significance of the increase in fluorescence in the ex vivo WT columns of panel G of S21 and note this result in the figure or in the caption.

In conclusion, I believe that the authors sufficiently addressed the reviewer comments and recommend this article for publication after these minor points are addressed.

2nd Revision - authors' response   14 March 2017

<u>Reviewer #2 comments</u>

1. I would still like to see the plot showing GFP readout with histological scoring data included in the supplement. While I do not believe that this data is sufficient to make claims of proportionality, I do believe that it could provide valuable information to the field about the behavior of the sensor and the robustness of the FACS method presented in this work. Including only the ROC curve in S25 seems to obscure this information.

Thank you. We agree and have now added the GFP fluorescence vs. histology scores to the Appendix (Figure S27).

Please also clarify if the ROC curve was generated using a "diagnosis" by histological scoring or simply the addition of DSS.

We generated the ROC curve based on the addition of DSS. We have clarified this fact in the main text and in the figure caption.

2. (i) The D55A mutation seems to only attenuate rather than completely "abolish" the response of the TtrSR both in vivo and in vitro. If it is attenuated, perhaps the wording should be changed.

The reviewer is correct. We have changed the wording accordingly.

(ii) The other is that the authors should use a statistical test to show significance of the increase in fluorescence in the ex vivo WT columns of panel G of S21 and note this result in the figure or in the caption.

While the flow cytometry histograms clearly show sensor activation in these experiments, we don't believe a statistical test is meaningful for such a small sample size. However, we have now added the p-values for the equal and unequal variance assumptions.

Thank you again for sending us your revised manuscript. We are now satisfied with the modifications made and I am pleased to inform you that your paper has been accepted for publication.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ⬇

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Jeffrey Tabor	
Journal Submitted to: MSB	
Manuscript Number: MSB-16-7416	

Reporting Checklist For Life Sciences Articles (Rev. July 2015)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures
1. Data
The data shown in figures should satisfy the following conditions:

➔ the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
➔ figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
➔ graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
➔ if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
➔ Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

➔ a specification of the experimental system investigated (eg cell line, species name).
➔ the assay(s) and method(s) used to carry out the reported observations and measurements
➔ an explicit mention of the biological and chemical entity(ies) that are being measured.
➔ an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

➔ the exact sample size (n) for each experimental group/condition, given as a number, not a range;
➔ a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
➔ a statement of how many times the experiment shown was independently replicated in the laboratory.
➔ definitions of statistical methods and measures:
 • common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
 • are tests one-sided or two-sided?
 • are there adjustments for multiple comparisons?
 • exact statistical test results, e.g., P values = x but not P values < x;
 • definition of 'center values' as median or average;
 • definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

Please ensure that the answers to the following questions are reported in the manuscript itself. We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

In the pink boxes below, provide the page number(s) of the manuscript draft or figure legend(s) where the information can be located. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable).

B- Statistics and general methods

Please fill out these boxes ⬇ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	Statistical methods were not used. A pilot experiment of two mice per group (-/+ DSS) was initially performed to see if the study had promise. Following this we used another 6 mouse per treatment cohort, which confirmed the findings in the pilot study. One cage of DSS-treated mice had low response and to determine if this was common we performed an additional 6 mice per treatment for the thiosulfate sensor and null control to solidify our findings.
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	Statistical methods were not used. A pilot experiment of two mice per group (-/+ DSS) was initially performed to see if the study had promise. Following this we used another 6 mouse per treatment cohort, which confirmed the findings in the pilot study. One cage of DSS-treated mice had low response and to determine if this was common we performed an additional 6 mice per treatment for the thiosulfate sensor and null control to solidify our findings.
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	Animal samples giving less than 250 counts in 5 minutes by flow cytometry were excluded from the analysis. The criteria were not pre-established, but based on what was required to get a decent distribution in the flow cytometry histogram. Being more stringent would not change the analysis, it would only result in removal of more "healthy" samples, which tended to have lower total counts.
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	Though no specific randomization procedure was undertaken, mice were randomized to treatment according to co-housing within cages. Cages were randomly selected for DSS-Treatment or H2O-Control. Likewise, at the time of gavage with the sensors, cages were randomly selected for the active sensor or the null control.
For animal studies, include a statement about randomization even if no randomization was used.	Though no specific randomization procedure was undertaken, mice were randomized to treatment according to co-housing within cages. Cages were randomly selected for DSS-Treatment or H2O-Control. Likewise, at the time of gavage with the sensors, cages were randomly selected for the active sensor or the null control.
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	No additional blinding was performed
4.b. For animal studies, include a statement about blinding even if no blinding was done	Blinding of histopathology scoring was performed, however no blinding of mouse samples for flow cytometry was performed. Samples from DSS-treated mice were visibly bloody and would have been immediately obvious.
5. For every figure, are statistical tests justified as appropriate?	Statistical tests were used in the determination of statistical significance for in vivo analysis of the thiosulfate and tetrathionate sensors.
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	The Shapiro-Wilk test was administered to ascertain normality of the distributions for the in vivo data. Data that was non-normal did not pass the assumptions for parametric testing, and were compared using Mann-Whitney U tests. Data with a normal distribution were compared with T-tests.
Is there an estimate of variation within each group of data?	Variation was estimated and compared with the Levene's Test.

Is the variance similar between the groups that are being statistically compared?	Variance was similar amongst all groups that had a normal distribution. The non-normally distributed groups had unequal variances.

C- Reagents

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	No antibodies were used in this study
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	Nissle strains were sequenced for the presence of microcin M, a protein found in this strain and not other strains used in the laboratory to ensure we were using the appropriate cell line.

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	C57Bl/6 mice were used. All mice were male, aged 6-9 weeks. Mice were obtained from the Baylor College of Medicine Center for Comparative Medicine Mouse Colony. Animals were housed in AALAC-accredited, pathogen free animal housing at Baylor College of Medicine. They received free access to food and water, and were monitored daily during the course of this experiment.
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	The study complied with all ethical regulations and was approved by the Baylor College of Medicine Institutional Animal Care and Use Committee (IACUC).
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	The ARRIVE guidelines were reviewed. All relevant aspects of our study have been reported.

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	N/A
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	N/A
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	N/A
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	N/A
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	N/A
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	N/A
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	N/A

F- Data Accessibility

18. Provide accession codes for deposited data. See author guidelines, under 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	N/A
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	Bioinformatics search results will be supplied.
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	N/A
21. As far as possible, primary and referenced data should be formally cited in a Data Availability section. Please state whether you have included this section. Examples: **Primary Data** Wetmore KM, Deutschbauer AM, Price MN, Arkin AP (2012). Comparison of gene expression and mutant fitness in Shewanella oneidensis MR-1. Gene Expression Omnibus GSE39462 **Referenced Data** Huang J, Brown AF, Lei M (2012). Crystal structure of the TRBD domain of TERT and the CR4/5 of TR. Protein Data Bank 4O26 AP-MS analysis of human histone deacetylase interactions in CEM-T cells (2013). PRIDE PXD000208	N/A
22. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	N/A

G- Dual use research of concern

23. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	N/A