Multi-scale model of the proteomic and metabolic consequences of reactive oxygen species

Laurence Yang[1,†], Nathan Mih[1,2,†], James T. Yurkovich[1,2], Joon Ho Park[3], Sangwoo Seo[6], Donghyuk Kim[7], Jonathan M. Monk[3], Colton J. Lloyd[1], Justin Tan[1], Ye Gao[4], Jared T. Broddrick[1,4], Ke Chen[1], David Heckmann[1], Adam M. Feist[1,5], Bernhard O. Palsson[1,2,5,*] [1]Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA. [2]Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA, USA. [3]Department of NanoEngineering, University of California, San Diego, La Jolla, CA, USA. [4]Division of Biological Sciences, University of California, San Diego, La Jolla, CA, USA. [5]Novo Nordisk Foundation Center for Biosustainability, The Technical University of Denmark, Kgs. Lyngby, Denmark. [6]School of Chemical and Biological Engineering and Institute of Chemical Process, Seoul National University, Seoul, Republic of Korea. [7]Department of Genetic Engineering, Kyung Hee University, Yongin, Republic of Korea. [†]These authors contributed equally to this work. [*]Corresponding email: palsson@ucsd.edu

All aerobically growing microbes must deal with oxidative stress from intrinsically-generated reactive oxygen species (ROS), or from external ROS in the context of infection. To study the systems biology of microbial ROS response, we developed a genome-scale model of proteome damage and maintenance in response to ROS, by extending a genome-scale metabolism and macromolecular expression (ME) model of *E. coli*. This OxidizeME model recapitulated measured microbial oxidative stress response including metalloenzyme inactivation by ROS and amino acid auxotrophies. OxidizeME also correctly predicted differential expression under ROS stress. We used OxidizeME to investigate how environmental context affects the flexibility of ROS stress response. The context-dependency of microbial stress response has important implications for infectious disease. OxidizeME provides a computational resource for model-driven experiment design in this direction.

The human immune system includes phagocytes that use reactive oxygen species (ROS) to combat pathogens. Most microbes are considerably weakened by this oxidative stress, whereas certain pathogens can grow inside the phagosome [1]. How microbes, including pathogens, evolve to tolerate such intense oxidative stress remains a long-standing question. To answer it requires a systems-level understanding of the many interlinked biochemical and physiological adaptations for responding to ROS damage. Here, we approach this question through *in silico* experiments subjecting *E. coli* to varying degrees of oxidative stress in a multitude of environments. To this end, we extended a genome-scale model of *E. coli* metabolism and protein expression by incorporating ROS damage and response mechanisms.

By their chemical nature, ROS can react with a broad array of macromolecules, often with deleterious consequences. In particular, a number of protein targets of ROS have been experimentally verified in detail; however, most of the potential targets of ROS have not been characterized individually. Thus, to better understand ROS stress, we must answer two questions: which biomolecules are damaged to what extent, and what is the cellular response to this damage? Here, we address these questions, focusing on ROS damage to metalloproteins. To predict cellular response to damage, we developed a novel model called OxidizeME. OxidizeME accounts for ROS damage to iron and iron-sulfur cluster cofactors of over 40 enzymes. We also modeled enzyme-catalyzed repair of oxidized iron-sulfur clusters. OxidizeME is able to predict cellular response at the metabolic and gene expression levels.

RESULTS

A multiscale model of reactive oxygen species damage and response.

We developed the OxidizeME model by extending the iLE1678-ME model of *E. coli* Metabolism and macromolecular Expression (ME model) [2] to account for damage to proteins by ROS, and cellular response mechanisms (Fig. 1). The major processes we added include the following: demetallation of mononuclear Fe^{2+} proteins, mismetallation of these proteins by divalent metal ions (Zn^{2+}, Mn^{2+}, Co^{2+}), and damage and repair of Fe-S clusters. We considered protein damage by H_2O_2(hydrogen peroxide) or O_2^-(superoxide). Details of the model are described fully in Supporting Information Methods.

OxidizeME recapitulates measured oxidative stress responses.

For each protein, OxidizeME computes the the fraction of inactivated enzyme due to cofactor oxidation as a function of intracellular ROS concentrations. The computed fraction of active enzyme recapitulates measured values [3] (Fig. 2A).

OxidizeME computes the network-level effects of individual enzyme inactivation. We thus asked if higher-level physiology could be predicted from molecular-level

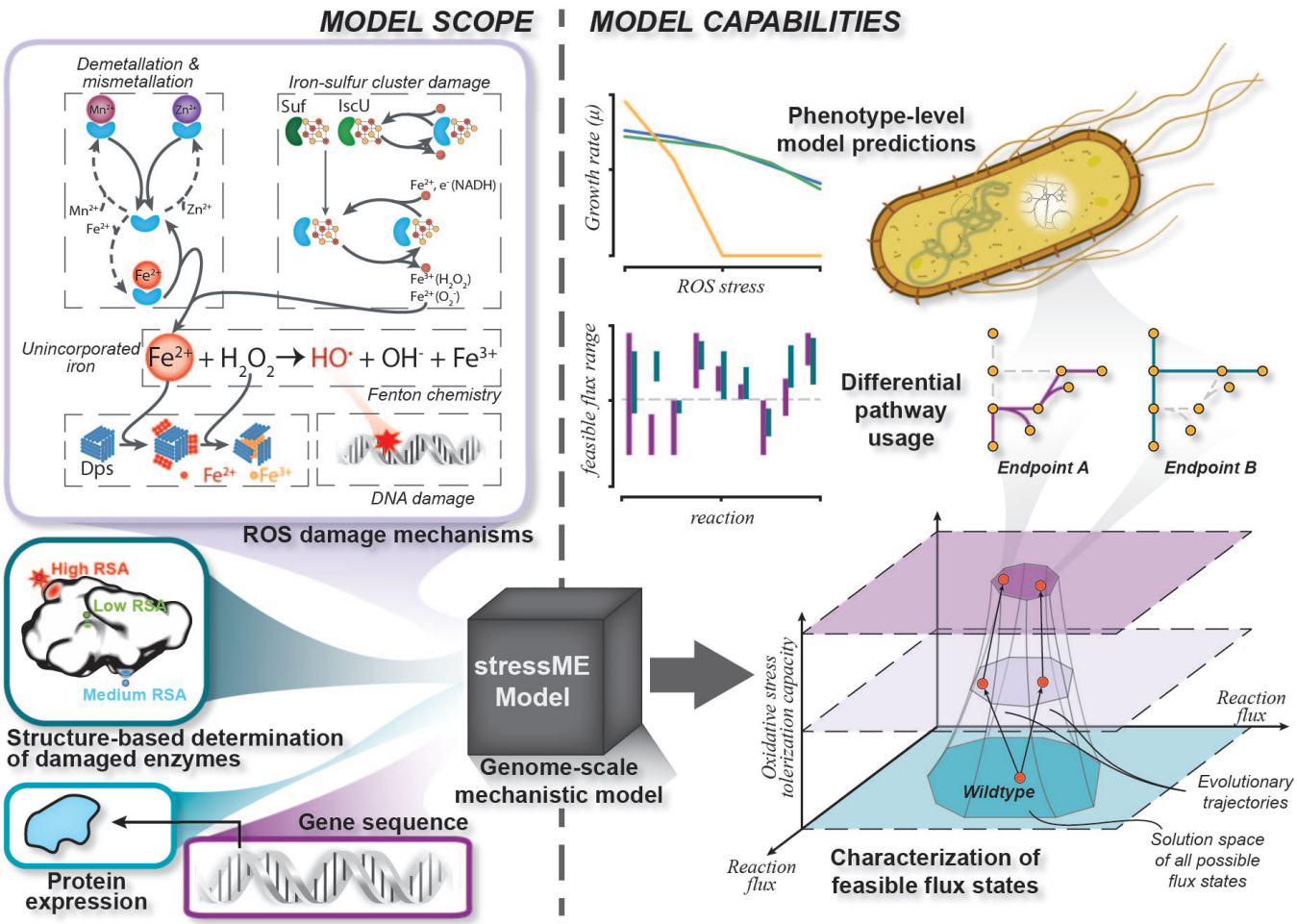

Figure 1: Principles of oxidative stress tolerance are revealed by integrating sequence, transcriptome, fluxome, and physiology. A OxidizeME model is reconstructed from sequence and its reactome expanded through known ROS biochemistry. Specifically, reactive oxygen species (ROS) damage metalloproteins through demetallation, mismetallation and iron-sulfur cluster oxidation. The 3D structure of metalloproteins affects how susceptible they are to damage. We compute these properties and use them to refine our model scope. OxidizeME predicts ROS tolerance capacity as a function of growth rate and proteome composition. By calibrating the model using omics from evolved endpoints, OxidizeME helps uncover insights into cellular responses to improve ROS tolerance.

mechanisms. A striking example is the auxotrophy of superoxide dismutase (SOD) mutants for certain amino acids (aromatic, branched-chain, and sulfur-containing) [4]. When grown on glucose minimal medium, *E. coli* endogenously produces superoxide at approximately 5 μM/s [4]. Without SOD, this amount of superoxide is sufficient to deactivate enzymes within amino acid biosynthesis pathways. To investigate this phenomenon, we simulated growth under increased cytosolic superoxide concentration and varied the supplementation of the different classes of amino acids. OxidizeME correctly recapitulated the amino acid auxotrophies (Fig. 2B). Furthermore, OxidizeME correctly predicted how severely fitness was reduced when specific classes of amino acids were not supplemented relative to the other classes.

We next investigated if OxidizeME could provide insight into the transcriptomic response to ROS stress. We thus measured the transcriptome of *Escherichia coli* under oxidative stress, by adding 0.25 mM paraquat to induce intracellular superoxide generation. We compared this stressed transcriptome against that of *E. coli* grown without stress. We found 648 differentially expressed genes (DEGs): 121 up- and 527 down-regulated ($|\log_2(\text{Fold change})| \geq 0.6$, FDR-adjusted $P \leq 0.02$). An immediate consequence of ROS stress is lowered exponential growth rate: *E. coli*'s growth rate was halved at 300 nM of paraquat. Thus, we first asked which of the 648 differentially expressed genes (DEGs) were due simply to lowered growth. This question is generally not easy to answer. However, using OxidizeME, we could discriminate stress-associated from growth-associated DEGs (see Methods). Using this approach,

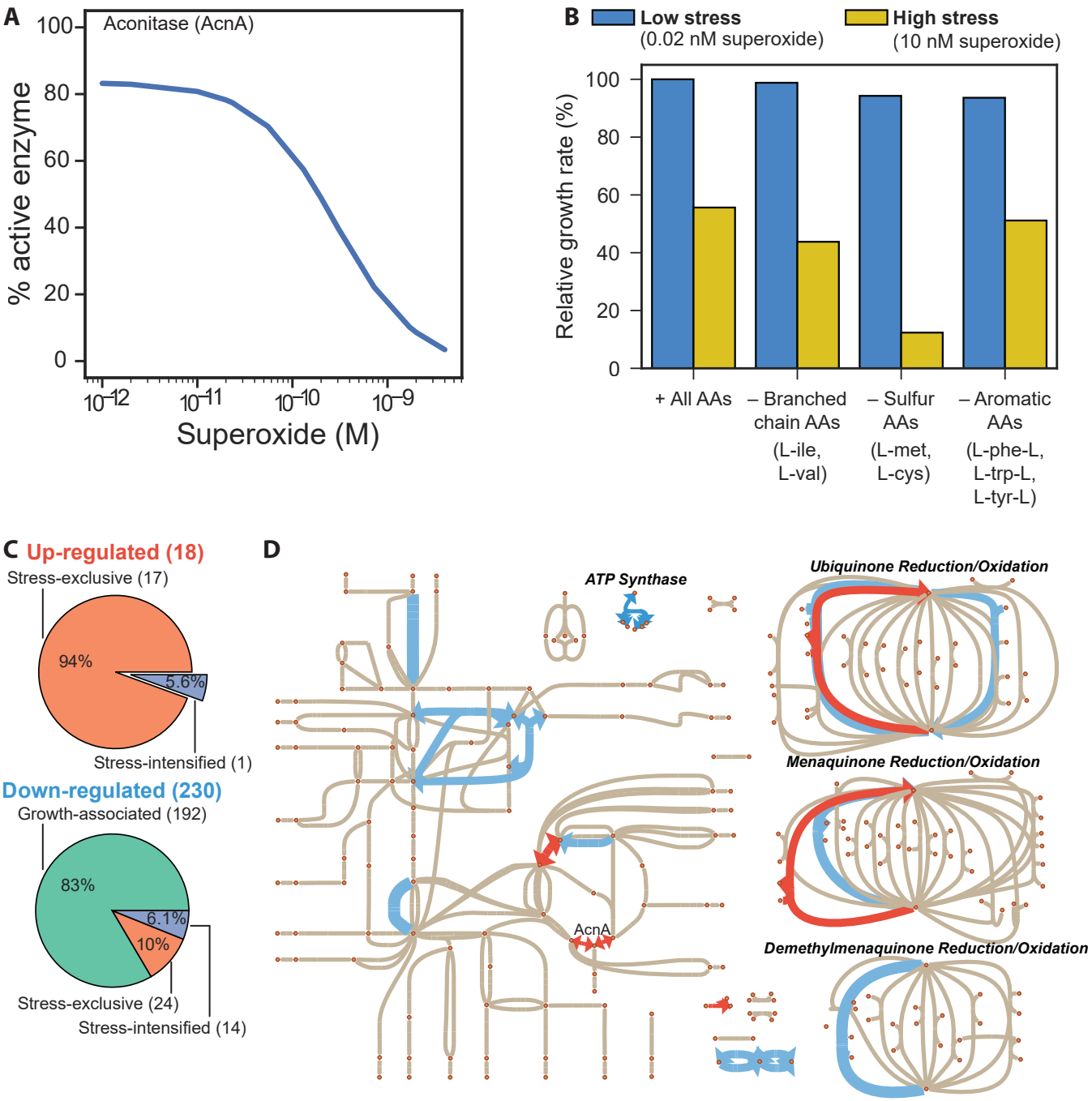

Figure 2: Validation of OxidizeME. (A) Predicted percent of active enzyme as a function of superoxide concentration for aconitase (AcnA). (B) Simulated relative growth rate as a function of superoxide stress under different amino acid supplementation media. (C) Significantly up- and down-regulated genes upon PQ addition that were correctly predicted by OxidizeME. (D) The up- and down-regulated genes displayed on a pathway map.

we identified 32 stress-associated genes, of which 14 and 18 genes were up- and down-regulated, respectively (Fig. 2C-D). These DEGs included processes in amino acid biosynthesis, protoheme biosynthesis, TCA cycle and aerobic respiration (nuoMN, atpABCDEFGH), metal ion transport, ROS detoxification, global translational efficiency (def, prfA), and stress-associated gene regulation (RpoS). Overall, OxidizeME correctly predicted 16 of 121 (recall of 13%) with 29% precision for the up-regulated genes. The recall and precision were 5.3 times and 11.5 times greater than the ME model for *E. coli* that did not account for ROS damage and repair processes.

OxidizeME helps investigate the scope of protein damage.

We tested whether the set of 42 protein complexes that we assumed to be damaged significantly affected the accuracy of our predictions. Accordingly, we generated 1000 random models, each perturbing a random set of 42 enzymes with the same damage rate constants as the OxidizeME model. We then compared the measured DEGs to those predicted by these random models under oxidative stress. Of the 1000 random models, 567 showed non-zero growth rate under the same superoxide stress as the OxidizeME model. Of these random models, 26/567 (4.6%) models had equal or greater F values for down-regulated genes, while all random models had lower F values for up-regulated genes, where $F = 2 \cdot (\text{Precision} \cdot \text{Recall})/(\text{Precision} + \text{Recall})$. These results show that despite the broad-reaching effects of ROS stress, accurately predicting stress response relies significantly on knowing which proteins are damaged.

Based on this finding, we further refined the scope of damaged proteins by incorporating a structural systems biology approach to understanding the stress response. Specifically, we computed protein structural properties relevant for determining whether ROS can reach and damage a cofactor and its accompanying enzyme. We used the computed solvent accessibilities and depths of cofactor coordinating residues within the enzyme, along with the presence of nearby cysteine residues coordinating the cofactor (≤ 5 Å). These properties were utilized as inputs into a Bayesian network, whose conditional probability tables were estimated from prior data and expert knowledge (see SI Methods). This Bayesian network took in the above pre-computed properties for all known metal-binding enzymes within our model, and as output predicted a probability for ROS damage, allowing us to refine damage rate constants for each enzyme.

Altogether, OxidizeME accurately models cofactor damage by ROS at the molecular level for over 40 enzymes. These enzymes are involved in diverse biological processes including amino acid biosynthesis, TCA cycle, pentose phosphate pathway, and translation. When ROS decreases activity of these enzymes, complex cellular responses were observed. Accordingly, simulated and measured cellular responses to ROS damage involves hundreds of differentially expressed genes and global metabolic shifts. The predicted responses were categorized, in increasing complexity as follows: compensatory up-regulation, direct cofactor repair, alternate pathway usage, and RpoS-mediated metabolic and proteome adjustment. These predicted responses were largely consistent with measured transcriptome changes.

Investigating how environmental context shapes ROS stress response.

We used OxidizeME to investigate how environmental context shapes ROS stress response. To this end, we simulated growth on 140 carbon sources with and without ROS stress. The relative growth rate, μ_{rel}, the growth rate under ROS stress over that without ROS, ranged between 0.40 to 0.80 (Fig. 3A). To understand the network basis of the diverse μ_{rel} values, we computed N_{toggle}, the number of metabolic fluxes that toggled on or off with ROS stress. We observed a significant negative correlation between μ_{rel} and N_{toggle} (Fig. 3A). Thus, the growth rate was generally affected less by ROS stress in conditions requiring fewer changes in metabolic pathway usage.

An exception was glycolate, which was the least ROS-sensitive carbon source. In contrast to most ROS-tolerant conditions, the optimal ROS response on glycolate involved 70 toggled fluxes, which was above the 96th percentile of N_{toggle}. Response to ROS involved increased flux in anaplerotic reactions and pyruvate metabolism including malic enzyme (NADP) and pyruvate dehydrogenase, and decreased flux in the TCA cycle and pentose phosphate pathway (Fig. 3B).

The least tolerant condition was D-galactose, in which a 10-fold increase in superoxide concentration decreased growth rate by 92%. Despite decrease in most fluxes with lowered growth rate, reactions involving folate metabolism increased under ROS stress (Fig. 3C). We traced this increase to changing sources of NADPH and NADH under ROS stress. Under ROS stress, the model predicted increased NADPH production from methylenetetrahydrofolate dehydrogenase (FolD) to supplement NADPH production by the pentose phosphate pathway (PPP). Meanwhile, NADH production relied greatly on the glycine cleavage system under ROS stress, whereas glyceraldehyde-3-phosphate dehydrogenase was the primary source of NADH without stress. Thus, growth on D-galactose in ROS stress required increased tetrahydrofolate and its derivates to drive NADPH and NADH production.

We next investigated the rates of protein damage by ROS between glycolate (most robust) and D-galactose (most fragile) growth conditions (Fig. 3D). The model predicted that biosynthesis of aromatic amino acids was strongly impacted by ROS in D-galactose but not in glycolate due to a bottleneck in shikimate metabolism. This prediction is consistent with experiments showing that shikimate supplementation alleviates aromatic amino acid (AA) auxotrophy imposed by superoxide in *E. coli* [5]. The model suggested a decrease in Mg^{2+} uptake under ROS stress in galactose but not in glycolate. Mg^{2+} is needed for shikimate kinase 1 and 2, and decreased uptake was associated with reduced shikimate biosynthesis and downstream aromatic AA biosynthesis. The model also predicted a difference in the metabolic flexibility of NADPH production pathways. On D-galactose, the predicted primary source of NADPH was the oxidative PPP (Gnd and Zwf) with and without ROS. On glycolate, the primary source of NADPH switched from the TCA cycle under no ROS stress to malic enzyme

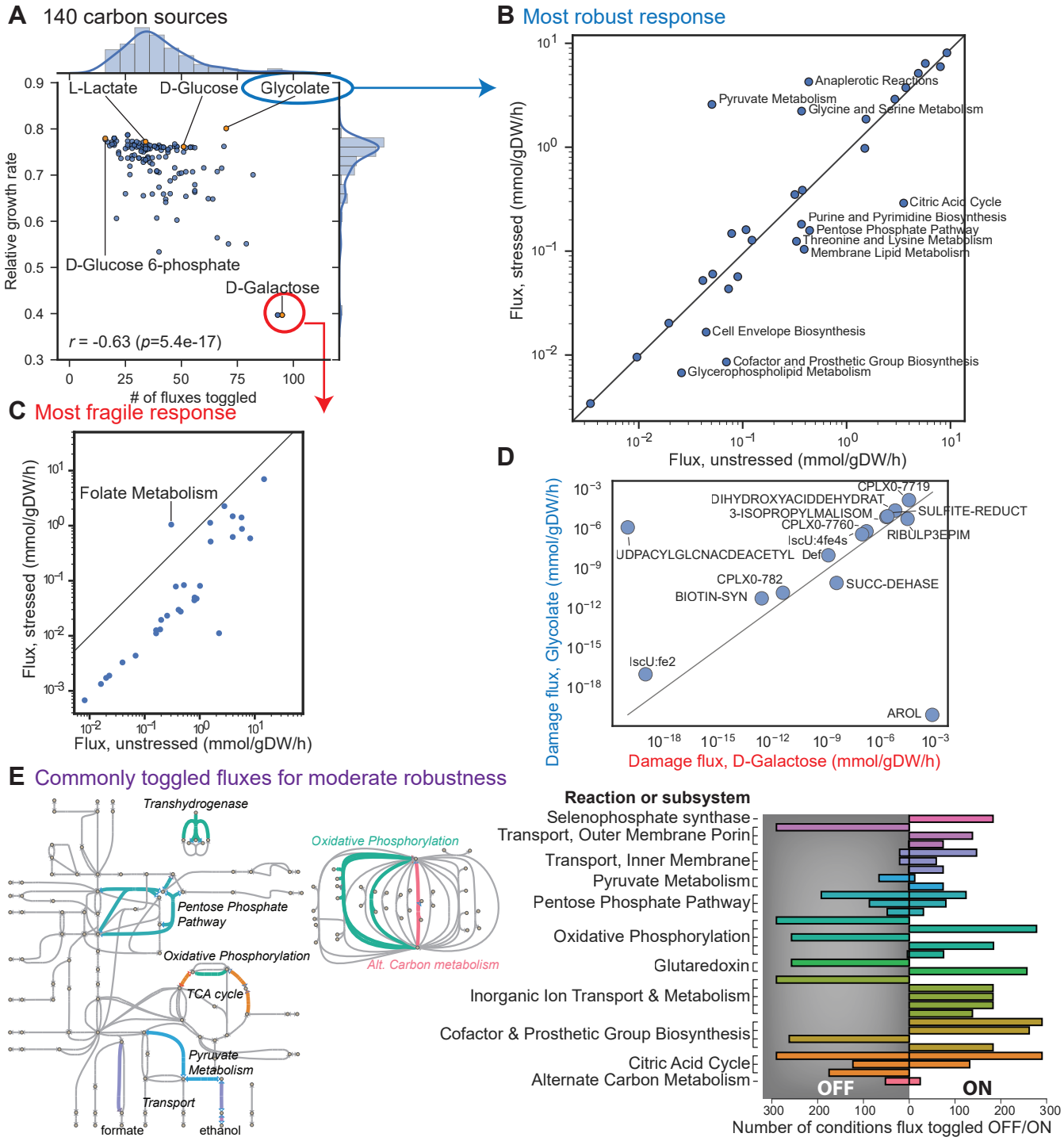

Figure 3: Stress response capacity across 333 carbon, nitrogen, sulfur, and phosphorous sources. (A) Simulated metabolic shifts (number of fluxes toggled on or off) versus relative growth rate (stressed over unstressed growth rates). (B) Simulated growth on glycolate. Flux for stressed versus unstressed models (fluxes averaged by subsystem). (C) Simulated growth on D-galactose. Flux for stressed versus unstressed models (fluxes averaged by subsystem). (D) Simulated damage fluxes for growth on glycolate (most robust) versus D-galactose (most fragile). (E) 37 reactions were commonly toggled on or off across the 140 conditions. This reaction set comprises a core metabolic response against ROS.

under stress. Fragility to ROS thus appears to be associated with metabolic rigidity in a given environment. Finally, we identified a core metabolic response module consisting of 37 metabolic reactions spanning 11 processes (Fig. 3E). This module contained reactions that most frequently toggled on or off in response to oxidative stress.

DISCUSSION

We extended the ME modeling framework to account for protein damage by ROS. This OxidizeME model includes three new major mechanisms: damage to metalloproteins by ROS, repair of these metalloproteins by dedicated biochemical pathways, and re-allocation of metabolic and macromolecular resources as a systems-level response to ROS. OxidizeME recapitulated measured consequences and responses to ROS, including inactivation of enzymes by superoxide, amino acid auxotrophies under ROS stress, and differential expression under PQ treatment.

We further used OxidizeME to investigate how environmental context shapes stress response. Overall, OxidizeME suggested that ROS tolerance was negatively correlated with the amount of change in metabolic state due to ROS stress. However, a number of environments including growth on glycolate supported high ROS tolerance because the metabolic state could be flexibly reconfigured. OxidizeME predicted that ROS tolerance was lowest when grown on D-galactose. Under ROS stress, the model predicted substantially higher dependence on folate metabolism to drive NADPH and NADH production under stress. A number of drugs (e.g., trimethoprim) target folate metabolism. Our computational analysis suggests that folate metabolism may be an effective antimicrobial target by when coupled with oxidative stress by disrupting NAD(P)H production; however, this efficacy depends on environmental context.

An ongoing goal of computational systems biology is to model the whole-cell response to environmental and genetic perturbations. ME models, in particular, can be extended to model major physicochemical stresses, including thermal, oxidative, pH, and osmotic stress. To this end, the recent FoldME model [6] enabled simulation of thermal stress response for *E. coli*, and reconstructed the cytoplasmic proteostasis network. OxidizeME represents another step in this direction. There is increasing recognition that microbial stress response plays an important role in infectious disease and antibiotic tolerance [7, 8]. Stress-ME models (FoldME, OxidizeME, etc.) will thus facilitate model-driven investigation of antimicrobial strategies.

METHODS

Methods are described fully in the Supporting Information (SI) Methods.

ACKNOWLEDGMENTS

We thank Daniel Zielinski for valuable discussions. This work was funded by the National Institute of General Medical Sciences of the National Institutes of Health (awards U01GM102098 and R01GM057089), the US Department of Energy (DE-SC0008701), and the Novo Nordisk Foundation Grant Number NNF10CC1016517. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the US Department of Energy (DE-AC02-05CH11231).

AUTHOR CONTRIBUTIONS

COMPETING FINANCIAL INTERESTS

The authors declare that they have no conflict of interest.

References

1. Bringer, M.-A., Barnich, N., Glasser, A.-L., Bardot, O. & Darfeuille-Michaud, A. *Infection and immunity* **73**, 712–721 (2005).
2. Lloyd, C. J. *et al.* *bioRxiv* doi:10.1101/106559 (2017).
3. Gardner, P. & Fridovich, I. *Journal of Biological Chemistry* **267**, 8757–8763 (1992).
4. Imlay, J. A. *Nat Rev Microbiol* **11**, 443–454 (2013).
5. Benov, L. & Fridovich, I. *Journal of Biological Chemistry* **274**, 4202–4206 (1999).
6. Chen, K. *et al.* *Proc Natl Acad Sci USA* **114**, 11548–11553 (2017).
7. Radzikowski, J. L. *et al.* *Molecular systems biology* **12**, 882 (2016).
8. Radzikowski, J. L., Schramke, H. & Heinemann, M. *Current Opinion in Biotechnology* **46**, 98–105 (2017).
9. Brynildsen, M. P., Winkler, J. A., Spina, C. S., MacDonald, I. C. & Collins, J. J. *Nat Biotechnol* **31**, 160–165 (2013).
10. Imlay, J. A. *J Biol Chem* **289**, 28121–28128 (2014).
11. Sobota, J. M. & Imlay, J. A. *Proc Natl Acad Sci USA* **108**, 5402–5407 (2011).
12. Sobota, J. M., Gu, M. & Imlay, J. A. *J Bacteriol* **196**, 1980–1991 (2014).
13. Keseler, I. M. *et al.* *Nucleic Acids Research* **41**, D605–D612 (2013).
14. van der Heijden, J. *et al.* *mBio* **7**, e01238–16 (2016).
15. Seaver, L. C. & Imlay, J. A. *J Bacteriol* **183**, 7182–7189 (2001).
16. Hassan, H. M. & Fridovich, I. *J Biol Chem* **254**, 10846–10852 (1979).
17. Liochev, S. I., Hausladen, A., Beyer, W. F. & Fridovich, I. *Proc Natl Acad Sci USA* **91**, 1328–1331 (1994).
18. Bianchi, V., Haggård-Ljungquist, E., Pontis, E. & Reichard, P. *J Bacteriol* **177**, 4528–4531 (1995).
19. Brunk, E. *et al.* *BMC systems biology* **10**, 26 (2016).
20. https://github.com/SBRG/ssbio (2017).
21. Rose, P. W. *et al.* *Nucleic acids research* **39**, D392–D401 (2010).
22. Andreini, C., Cavallaro, G., Lorenzini, S. & Rosato, A. *Nucleic acids research* **41**, D312–D319 (2012).
23. Xu, D. & Zhang, Y. *Scientific reports* **3** (2013).
24. Mitternacht, S. *F1000Research* **5** (2016).
25. Sanner, M. F., Olson, A. J. & Spehner, J.-C. *Biopolymers* **38**, 305–320 (1996).
26. Hamelryck, T. & Manderick, B. *Bioinformatics* **19**, 2308–2310 (2003).
27. https://github.com/jmschrei/pomegranate (2016).
28. Anjem, A. & Imlay, J. A. *Journal of Biological Chemistry* **287**, 15544–15556 (2012).
29. Outten, F. W. & Theil, E. C. *Antioxidants & redox signaling* **11**, 1029–1046 (2009).
30. *Andrade, S. L.* et al. (Walter de Gruyter GmbH & Co KG, 2014).
31. Love, M., Anders, S. & Huber, W. *Genome Biol* **15**, 550 (2014).

SUPPORTING INFORMATION METHODS

Differentiating stress–specific from growth-associated responses. We simulated transcriptomes over lowered growth rates using two models—one accounting for stress and the other not. The growth rate was determined by the stress model over increasing superoxide concentrations, leading to lowered growth rate, and this growth rate was used for the non-stress model. DEGs that were correctly predicted by the stressed model but not the unstressed model were considered to be stress-associated. DEGs that were correctly predicted by the unstressed model, and were not considerably more differentially expressed by the stressed model were considered growth-associated.

Damage to iron-sulfur clusters We note that one of the spontaneous H_2O_2-generation reaction classes by [9] was the cycling of Fe-S cluster damage by O_2^- and repair resulting in net generation of H_2O_2. Here, we have separate damage reactions of Fe-S clusters by H_2O_2 and O_2^-.

Demetallation and mismetallation of mononuclear iron enzymes Free metal ions compete to bind to enzymes according to different stability constants. In the absence of dedicated metallochaperones, the relative metalloenzyme composition will depend on the relative size of metal concentrations.Under oxidative stress, Fe(II)-complexes are demetallated, increasing mismetallation by other metals [4]. The mismetallated complexes are also demetallated, possibly by cysteine in *E. coli* [10]. This observation is important because without demetallation of mismetallated complexes, the high stability of these complexes would make mismetallation considerably more deleterious for cells.

We thus model the direct proteomic consequence of ROS damage to mononuclear Fe(II) metalloenzymes: increased mismetallation by other divalent metal ions, specifically Mn(II) and Zn(II). Mn(II) and Zn(II) metalloenzymes have reduced turnover (by about 60% and 99%, respectively) and are not damaged by ROS [10]. Therefore, increasing Mn(II) pools is a potential response to oxidative stress, while mismetallation by Zn(II) is a deleterious consequence of Fe(II) enzyme demetallation.

In our model, we assume that metallation occurs rapidly and is close to equilibrium [10]. Therefore, the concentrations of metal and complexes will depend on stability constants.

Demetallation of mononuclear iron enzymes occurs with rate constants ($k_{\mathrm{cat}}/K_{\mathrm{M}}$) of 10^3 to 10^4 $\mathrm{M^{-1}s^{-1}}$ *in vitro* [4]. We added reactions for mismetallation of mononuclear iron enzymes by alternate metals. The k^{eff} was scaled relative to Fe^{2+} for enzymes mismetallated with Mn^{2+}, Co^{2+}, and Zn^{2+}, based on published data [11, 12]. We know the total metal concentrations and some free concentrations (Imlay, 2014). We then fix the metal concentrations within these ranges. We then use proteomics to constrain the total enzyme concentrations. We assume that metallation proceeds rapidly and sufficiently close to equilibrium [10] that apoenzyme concentrations and dilution are negligible. We use varyME to explore a range of feasible metallation states. From the ME computations, we substitute in the fluxes into the equations above, and calculate the binding constants. Due to alternate optima, we end up with a range of feasible constants, from which we generate a random ensemble.

Association constants in the ensemble were constrained according to the Irving-Williams series [10]: Mn(II) < Fe(II) < Co(II) < Ni(II) < Cu(II) > Zn(II).

Limits of ROS damage rates. The K_{M} of catalase for H_2O_2 is 8.65×10^7 nM [13], which represents a lower limit K_{M} for reaction of H_2O_2 with the mononuclear metal or Fe-S cluster enzymes. Under glucose aerobic conditions, H_2O_2 concentration is around 50 nM. At 200 nM, OxyR is activated, and at 400 nM, growth defects are observed [4]. Therefore, for a viable cell, the H_2O_2 concentration is far below K_{M} and we can approximate the damage rate as $v^{\mathrm{dmg}} \approx k_{\mathrm{cat}}/K_{\mathrm{M}}ES$, where E and S are enzyme and H_2O_2 concentrations, respectively. Values for $k_{\mathrm{cat}}/K_{\mathrm{M}}$ for dehydratase and mononuclear iron enzymes range from 10^3 to 10^4 $\mathrm{M^{-1}s^{-1}}$ [4]. We thus used a lower limit of 10^3 $\mathrm{M^{-1}s^{-1}}$ and created an ensemble of damage scenarios where damage rate constants varied within this range. Thus, at 50 nM H_2O_2, effective rate constants (i.e., $k^{\mathrm{eff}} = k_{\mathrm{cat}}/K_{\mathrm{M}}\mathrm{S}$) were between 5×10^{-5}—5×10^{-4} $\mathrm{s^{-1}}$ and at 400 nM H_2O_2, k^{eff} was between 4×10^{-4}—4×10^{-3} $\mathrm{s^{-1}}$.

Similarly, rate constants for damage by O_2^- exceed 10^6 $\mathrm{M^{-1}s^{-1}}$ [4]. Thus, at a basal O_2^- concentration of 0.2 nM, k^{eff} of damage by O_2^- is about 2×10^{-4} $\mathrm{s^{-1}}$.

ROS influx as a function of outer membrane permeability Hydrogen peroxide influx rate is a function of outer membrane permeability (P), surface area (A), extracellular (C_{out}) and intracellular (C_{in}) ROS concentrations for passive diffusion [14] :

$$v_{in} = P \cdot A \cdot (C_{\mathrm{out}} - C_{\mathrm{in}}).$$

For *E. coli*, basal values are $P = 1.6\times10^{-3}$ cm/s, $A = 1.41\times10^{-7}$ $\mathrm{cm^2}$ [15]. We can simulate various ROS stress environments by changing C_{out}. We then simulate with various C_{in} to explore the range of intracellular ROS concentrations that *E. coli* can maintain. The solution to this problem is not simple: maintaining lower intracellular ROS concentrations creates a larger gradient leading to higher influx and subsequently higher detoxification requirements. In contrast, maintaining higher C_{in} lowers ROS influx but drives higher

rates of ROS damage. This non-trivial trade-off is captured by the model through ROS-damage coupling constraints that force damage proportional to intracellular ROS concentrations. Also, outer membrane permeability can be altered in Gram-negative bacteria: *Salmonella* can decrease permeability to increase tolerance to ROS stress, while increased permeability can potentiate antibiotics [14]. Outer membrane permeability is a function of the composition of membrane lipid and channel proteins. Thus, $P = 1.6 \times 10^{-3}$ cm/s represents a basal reference value, which we used to calibrate the permeability in the model. Note that overall, H_2O_2 diffusion is rate-limited by outer membrane permeability, since due to lipid composition, diffusion over the inner membrane is orders of magnitude faster than the OM [14].

Modeling damage rates as functions of ROS concentration. We simulate ROS damage and repair by specifying intracellular ROS concentrations, which affect effective rate constants of damage and detoxification enzymes. Assuming ROS concentrations are much smaller than K_{M} of damage reactions, individual damage fluxes are coupled to enzyme abundance as follows:

$$\begin{aligned} v^{\mathrm{dmg}} &= k_{\mathrm{cat}}/K_{\mathrm{M}}[ROS]E \\ &= k_{\mathrm{eff}}^{\mathrm{dmg}} \frac{v^{\mathrm{dil}}}{\mu}. \end{aligned}$$

Thus, damage fluxes are proportional to ROS concentration and abundance of the damageable protein. Generated ROS is either detoxified or damages macromolecules:

$$\sum_{j \in \mathrm{Damaged}} s_j v_j + \sum_{j \in \mathrm{Detox}} s_j v_j = \sum_{j \in Generated} s_j v_j$$

$$\begin{aligned} [ROS] &= [ROS]^{\mathrm{detox}} + [ROS]^{\mathrm{damage}} \\ &= \sum_{j \in \mathrm{Detox}} \frac{v_j \mu}{k_{\mathrm{eff},j} v_j^{\mathrm{dil}}} + \sum_{j \in \mathrm{Damage}} \frac{v_j \mu}{k_{\mathrm{eff},j} v_j^{\mathrm{dil}}} \end{aligned}$$

Thus, increased detoxification decreases the ROS left to damage macromolecules. The concentration of ROS that is not detoxified determines the rate of macromolecule damage. For a fixed [ROS], OxidizeME determines the optimal distribution of ROS between detoxification and damage. Typically, detoxification will be maximized until its capacity is exceeded. We also assume [ROS] concentrations do not change at steady-state:

$$\frac{d[ROS]}{dt} = \sum_{j \in \mathrm{Production}} s_j v_j - \sum_{j \in \mathrm{Consumption}} s_j v_j \quad = 0$$

ROS generation rate from paraquat concentrations. Superoxide generation from paraquat was assumed to proceed by the following two-step mechanism [16]:

1. PQ2 + NADPH → PQ1 (catalyzed by Fpr[17, 18], TrxB, CysJ)
2. PQ1 + O_2 → O_2^- + PQ2 (spontaneous)

We validated the computed concentrations of intracellular superoxide given extracellular paraquat using data from Gardner and Fridovich [3]

Solving the OxidizeME model as an optimization problem. Combining all new variables and constraints, the final OxidizeME optimization problem for maximizing growth rate subject to oxidative stress is the following:

Cellular objective	$\max_{\mu,v}$	μ	(1)
(Macro)molecule mass balance	s.t.	$Sv = 0$	(2)
ROS influx		$v_{\mathrm{in}} = P \cdot A \cdot \Delta C$	(3)
ROS-enzyme damage coupling		$v^{\mathrm{dmg}} = k_{\mathrm{cat}}/K_{\mathrm{M}} \cdot [ROS] \cdot v^{\mathrm{dil}}/\mu$	(4)
Damaged complex mass balance		$v_j^{\mathrm{dmg}} - v_j^{\mathrm{repair}} - v_j^{\mathrm{dil}} = 0$	(5)
		$v_j^{\mathrm{dil}} = \frac{\mu}{k_{\mathrm{repair},j}} \cdot v_j^{\mathrm{repair}}, \quad j \in \mathrm{DamagedComplex}$	(6)
Mismetallation coupling		$v_i^{\mathrm{met}} = \frac{\beta^i [Metal\ i]}{\beta^{Fe}[Fe(II)]} \cdot \left(\sum_j v_{E:Fe}^{\mathrm{demet}} + v_{E:Fe}^{\mathrm{dil}} \right), \quad i \in \mathrm{AltMetals}$	(7)

where W is the cell specific weight (gDW/L)= $\frac{278 \times 10^{-15}\ \mathrm{gDW}}{6.8 \times 10^{-16}\ \mathrm{L}}$, μ is the specific growth rate (1/h), $[E]$, $[E:i]$, $[i]$ are concentrations (M) of apoenzyme, holoenzyme, and metal ions, respectively, $AltMetals = \{Zn(II), Mn(II), \ldots\}$, and β^k are metal-protein stability constants (M).

Protein structural property computation To enable a structural systems biology approach of computing ROS damage probabilities for each metal-binding enzyme in the OxidizeME model, we utilized the GEM-PRO pipeline [19] implemented in the ssbio Python package [20] to gather all available protein structures for model enzymes from the Protein Data Bank [21]. Metal-binding residues were taken as annotated in each protein's UniProt entry, and mapped to the correct residue numbering scheme in the structure file. We set a single representative structure for each enzyme based on the following four factors: 1) sequence identity and coverage of the wild-type amino acid sequence; 2) presence of the annotated metal binding site 3) presence of the model-annotated metal ion; and 4) presence or absence of any model-annotated cofactors other than the metal ion. The MetalPDB database [22] was used to expedite this analysis. Per protein, these factors contributed to a weighted score enabling a custom rank-ordering of all available structures. Finally, for those proteins with no experimental structure available, previously generated homology models from the I-TASSER pipeline were set as representative [23].

Next, all heteroatoms (excluding non-metal model-annotated cofactors) were deleted from experimental protein structures. This was done so calculations were consistent between experimental and homology models, as metal ions were not present in select models. Solvent accessibilities for each metal-binding residue were then calculated using FreeSASA [24], residue depths using the MSMS program [25], and the presence of cysteine residues (either directly contributing to metal binding, or within a 5 Å radius around the binding site) with the Biopython PDB module [26] and information from MetalPDB. This information was used as input to a Bayesian network created with the pomegranate Python package [27]. Conditional probability tables for ROS damage based on these parameters were generated from manual curation of data and comments from [11], [28], [29], and [30]; however, it is noted that pure quantitative data is unavailable for most of these enzymes and represents further experimental work to be done.

The conditional nodes of the Bayesian network are: A) the probability of ROS reaching the metal-binding site which depends on 1) the maximum calculated solvent accessibility of each of the metal-binding residues as well as 2) its associated depth calculation; and B) the probability of ROS damaging the metal-binding site which depends on A), as well as if the site uses a cysteine residue for binding or is in the proximity. A final probability of activity decrease for this enzyme is then calculated depending on the damage probability, and used as input into the OxidizeME model. These above methods are available as two separate notebooks (for representative structure selection and Bayesian network construction) in the iStress platform.

RNA sequencing RNA-sequencing data were generated under conditions of exponential-phase, aerobic growth in glucose M9 minimal media with appropriate concentration paraquat. Cells were washed with Qiagen RNA-protect Bacteria Reagent and pelleted for storage at -80 °C prior to RNA extraction. Cell pellets were thawed and incubated with Readylyse Lysozyme, SuperaseIn, Protease K, and 20 % sodium dodecyl sulfate for 20min at 37 °C. Total RNA was isolated and purified using the Qiagen RNeasy Mini Kit columns, following vendor procedures. An on-column DNase-treatment was performed for 30 min at room temperature. RNA was quantified using a Nano drop and quality assessed by running an RNA-nano chip on a bioanalyzer. Paired-end, strand-specific RNA-seq was performed following a modified dUTP method (Latif et al. 2013). The rRNA was isolated using Epicentre's Ribo-Zero rRNA removal kit for Gram Negative Bacteria. Sequences were run on an Illumina HiSeq using a KAPA Stranded RNA-seq kit. Reads were mapped to the *E. coli* K12 MG1655 Genome (NC_000913.2) with bowtie2 (Langmead and Salzberg 2012). Expression levels in fragments per kilobase per million mapped (FPKM), and differentially expressed genes (DEGs) were found using DESeq2 [31]. A fold-change of 1.5 and FDR-adjusted p-value cutoff of 0.05 were used to call significant differential expression.