changes with bound cross-bridges requires energy to disrupt attachments, resulting in negative power production. However, intermediate temperatures permit some detachment to accommodate length changes in addition to some attachment at the extrema of the length cycle. At these intermediate temperatures, cross-bridges that remain bound at the very end of lengthening or shortening can store energy in their axial or radial extension, respectively. This stored energy may return energy into the lattice when the cross-bridges detach at the start of the subsequent phase. In doing so, the deformed cross-bridges could assist antagonistic muscles. Prior studies have shown that elastic energy storage is indeed crucial for meeting the high inertial power costs of flight (*3*, *4*). If even a portion of these cross-bridges facilitate elastic energy savings via a temperature gradient, they would contribute to the overall energy savings in locomotion. Because temperature gradients are an inevitable consequence of internal energy generation and heat dissipation in both vertebrates and invertebrates, this mechanism of energy storage could be a general phenomenon in locomotor systems (*11*, *12*).

References and Notes

1. M. Jensen, T. Weis-Fogh, *Philos. Trans. R. Soc. London Ser. B* **245**, 137 (1962).
2. R. M. Alexander, *Am. Zool.* **24**, 85 (1984).
3. R. M. Alexander, H. C. Bennet-Clark, *Nature* **265**, 114 (1977).
4. C. P. Ellington, *J. Exp. Biol.* **115**, 293 (1985).
5. H. E. Huxley, A. Stewart, H. Sosa, T. Irving, *Biophys. J.* **67**, 2411 (1994).
6. M. Dickinson *et al*., *Nature* **433**, 330 (2005).
7. J. G. Tidball, T. L. Daniel, *Am. J. Physiol.* **250**, R56 (1986).
8. I. Dobbie *et al*., *Nature* **396**, 383 (1998).
9. A. F. Huxley, R. M. Simmons, *Nature* **233**, 533 (1971).
10. K. C. Nishikawa *et al*., *Proc. R. Soc. B* **279**, 981 (2012).
11. N. T. George, T. L. Daniel, *J. Exp. Biol.* **214**, 894 (2011).
12. F. G. Carey, J. M. Teal, *Proc. Natl. Acad. Sci. U.S.A.* **56**, 1464 (1966).
13. N. T. George, S. Sponberg, T. L. Daniel, *J. Exp. Biol.* **215**, 471 (2012).
14. R. K. Josephson, *J. Exp. Biol.* **108**, 77 (1984).
15. A. F. Bennett, *J. Exp. Biol.* **115**, 333 (1985).
16. C. D. Williams, M. Regnier, T. L. Daniel, *PLOS Comput. Biol.* **8**, e1002770 (2012).
17. A. Miller, R. T. Tregear, *Nature* **226**, 1060 (1970).
18. Information on materials and methods is available on *Science* Online.
19. M. S. Tu, T. L. Daniel, *J. Exp. Biol.* **207**, 2455 (2004).
20. T. C. Irving, in *Nature's Versatile Engine: Insect Flight Muscle Inside and Out*, J. O. Vigoreaux, Ed. (Springer, New York, 2006), pp. 197–213.
21. G. Wang, M. Kawai, *J. Physiol.* **531**, 219 (2001).

Acknowledgments: We thank H.-M. Hsu, E. Carrington, and R. Huey for their contributions. This research was supported by NSF Graduate Research Fellowship no. DGE-0718124 to N.T.G., NSF grant IOS-1022471 to T.L.D and T.C.I., NSF grant EEC 1028725 to T.L.D., and the University of Washington Komen Endowed Chair to T.L.D. Use of the Advanced Photon Source was supported by the U.S. Department of Energy under contract no. DE-AC02-06CH11357. Use of BioCAT was supported by grants from the NIH (2P41RR008630-17 and 9 P41 GM103622-17). The data reported in this paper are deposited in the Dryad Repository (http://dx.doi.org/10.5061/dryad.cg873). For inquiries about materials, please contact N.T.G. (ntgeorge@uw.edu).

Supplementary Materials

www.sciencemag.org/cgi/content/full/science.1229573/DC1
Materials and Methods
Figs. S1 and S2
References (*22*, *23*)
Movie S1

31 August 2012; accepted 28 March 2013
Published online 25 April 2013;
10.1126/science.1229573

Structural Systems Biology Evaluation of Metabolic Thermotolerance in *Escherichia coli*

Roger L. Chang,[1] Kathleen Andrews,[2] Donghyuk Kim,[2] Zhanwen Li,[3,4] Adam Godzik,[3,4] Bernhard O. Palsson[1,2,5]*

Genome-scale network reconstruction has enabled predictive modeling of metabolism for many systems. Traditionally, protein structural information has not been represented in such reconstructions. Expansion of a genome-scale model of *Escherichia coli* metabolism by including experimental and predicted protein structures enabled the analysis of protein thermostability in a network context. This analysis allowed the prediction of protein activities that limit network function at superoptimal temperatures and mechanistic interpretations of mutations found in strains adapted to heat. Predicted growth-limiting factors for thermotolerance were validated through nutrient supplementation experiments and defined metabolic sensitivities to heat stress, providing evidence that metabolic enzyme thermostability is rate-limiting at superoptimal temperatures. Inclusion of structural information expanded the content and predictive capability of genome-scale metabolic networks that enable structural systems biology of metabolism.

Cellular thermosensitivity depends on proteome stability. Chaperones and proteases are well-characterized heat shock proteins (HSPs), and chaperones improve survival at superoptimal temperatures (*1*). Protein folding and structural stability required for function are disrupted at high temperatures. Many individual proteins and their mutant variants have been studied to identify structural loci within a protein that are destabilized at high temperatures, leading to denaturation. Replacing heat-sensitive loci with more stabilizing residues has allowed engineering of thermostable proteins (*2*). By analogy, identifying the proteins that confer susceptibility to heat within the cellular system is critical to uncovering mechanisms for cellular thermosensitivity. Strategies for increasing thermotolerance have included introduction of chemical chaperones, overexpression of HSPs, pretreatment with moderate heat, or random mutagenesis to evolve stress tolerance (*3*). Instead, we sought to directly identify the particular proteins that confer thermosensitivity in the system.

The emerging discipline of structural systems biology (*4*) has enabled new insights into topics that include the structure-function relations in metabolism in a hyperthermophile (*5*), identification of causal off-target actions of drugs that cause adverse side effects (*6*), identification of protein-protein interactions (*7*, *8*), and determination of causal mutations for disease susceptibility (*8*, *9*). We used a structural systems biology approach to discover points of thermosensitivity in the mesophilic bacterium *Escherichia coli* K-12 MG1655. Metabolic thermosensitivity, affected by enzyme activity in a genome-scale model (GEM), was assessed as a function of protein thermostability, providing mechanistic explanations for effects of mutations in evolved thermotolerant strains (*10*, *11*) and leading to the discovery of metabolic limitations to thermotolerance.

To assess protein thermostability, we integrated a genome-scale model of *E. coli* metabolism (*i*JO1366) (*12*) with protein structures (GEM-PRO) by associating metabolic reactions with structures of their catalytic enzymes (database S1), thereby enabling parameterization of the network model on the basis of protein structural properties. The main objectives of this reconstruction (Fig. 1A) were to (i) maximally cover amino acid sequence (Fig. 1B), (ii) represent the native structure of each wild-type (WT) protein (Fig. 1C), (iii) map existing amino acid functional annotations to structures (*13*–*16*) (Fig. 1C), and (iv) represent changes in functional conformation or induced fit caused by protein-substrate

[1]Bioinformatics and Systems Biology Graduate Program, University of California San Diego, La Jolla, CA 92093–0412, USA. [2]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093–0412, USA. [3]Bioinformatics and Systems Biology Program, Sanford-Burnham Medical Research Institute, La Jolla, CA 92037, USA. [4]Joint Center for Structural Genomics (JCSG), La Jolla, CA 92037, USA. [5]The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK-2970 Hørsholm, Denmark.

*Corresponding author. E-mail: palsson@ucsd.edu

binding (Fig. 1D). Thus, in this model a protein may be represented by zero, one, or multiple separate structures. Experimentally determined structures (*17*) and structures from homology modeling (fig. S1) were used to achieve 93% structural coverage of proteins in the network (Fig. 1B) and between 24 and 33% coverage of protein-substrate binding conformations. The majority of coverage was enabled by structure-modeling techniques (*5*), without which such a reconstruction would not currently be possible.

Experimentally measured critical temperatures (*18*, *19*) accounting for optimal, half-maximum, and total loss of protein activity were supplemented with bioinformatic predictions (*20*–*23*), based on protein three-dimensional structures, of protein melting temperatures (see table S1 for a full list and fig. S2 for a comparison of experimental and predicted melting temperatures). We used these critical temperatures to define protein activity functions that imposed temperature-dependent constraints on the metabolic model (fig. S3). In this way, temperature, by affecting protein function, became a parameter for genome-scale metabolic simulation. Simulated temperature-dependent growth showed good qualitative agreement with experimental growth data in three different nutrient media (Fig. 2A) in the range from 32° to 43°C, where growth is above 50% of maximum. Thus, thermostability of metabolic proteins appears to explain much of the thermosensitivity within this temperature range.

Our modeling framework also enabled precise prediction of points of thermosensitivity in the metabolic network (fig. S4 and table S2). The most temperature-limited protein activities cluster in cofactor synthesis pathways, identifying them as most growth limiting (Fig. 2B). We tested how temperature sensitivity of the model would change with the introduction of thermostable proteins into the network by alleviating temperature-dependent activity constraints on predicted growth-limiting proteins. For instance, optimally increasing critical temperatures of all proteins predicted to be limiting at 42.2°C produced a twofold increase in maximum growth rate and shifted the optimal temperature to 42.2°C, but narrowed the range of growth temperatures because of incompatibility of these more thermophilic activity functions with lower growth temperatures (Fig. 2B). Adjusting

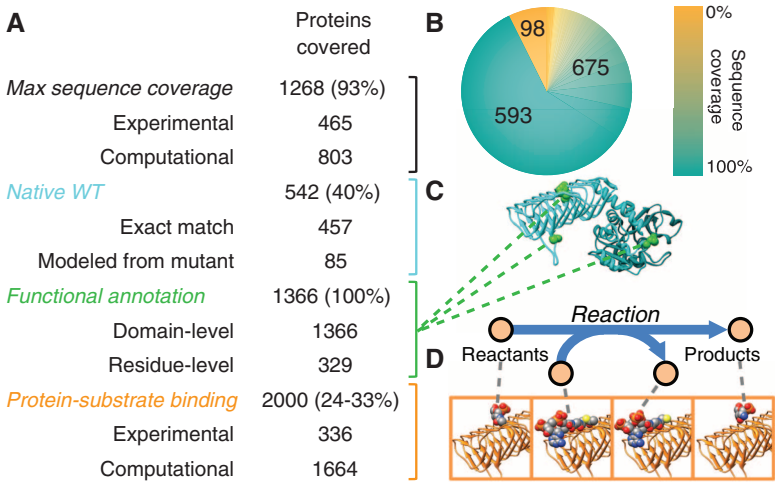

Fig. 1. Properties of the *E. coli* metabolic model integrated with protein structures. (**A**) The *E. coli* GEM-PRO model provides maximal amino acid sequence coverage, native WT structures, functional annotation, and protein-substrate binding for proteins included in *i*JO1366. Percentages are out of 1366 total proteins, except for the percentage of protein-substrate binding pairs, which is out of an estimated total between 6144 and 8448 such pairs. (**B**) The distribution of maximum amino acid sequence coverage of proteins by structures included in the GEM-PRO. Numbered wedges indicate the number of proteins with 0%, 100%, or partial sequence coverage. (**C**) Example of a native WT structure included in the GEM-PRO. Green highlighted residues denote annotated functional sites. (**D**) Protein-substrate binding is structurally represented as the pairwise interactions between each protein and the reactants or products of the catalyzed metabolic reaction.

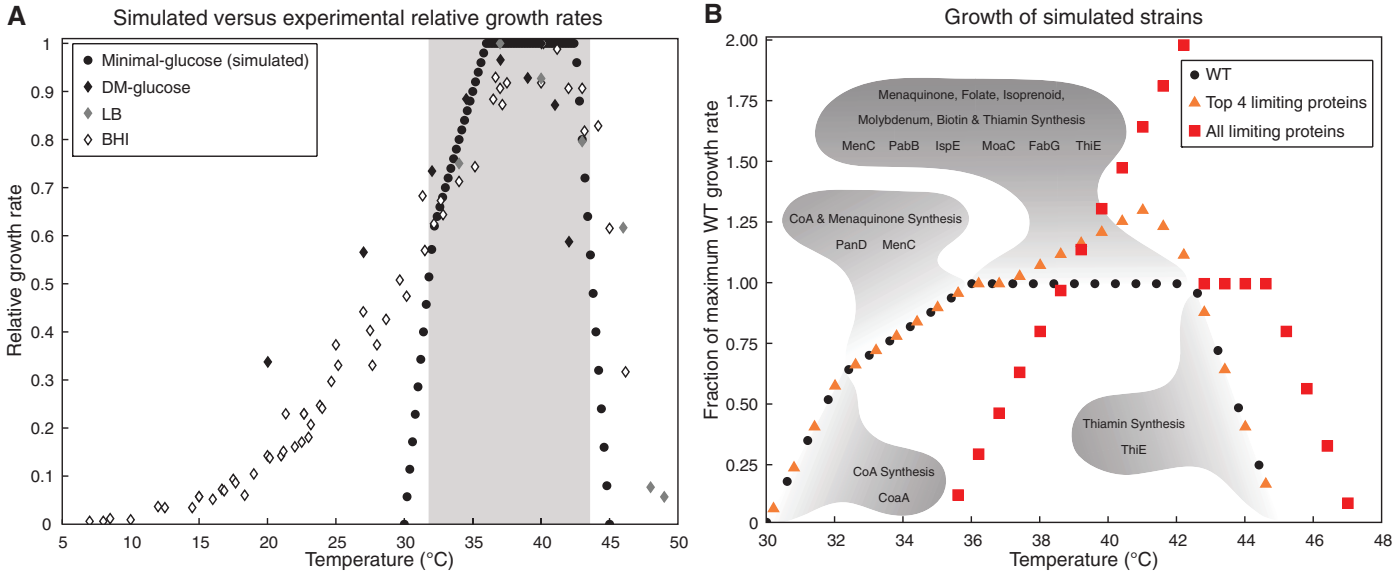

Fig. 2. Simulated and experimental growth rates as a function of temperature. (**A**) Growth rates relative to the maximum are depicted under each condition. Growth was simulated on minimal medium with glucose (circles) or experimentally measured (diamonds) on Davis minimal medium (DM) with glucose (*29*), lysogeny broth (LB) (*11*), or brain heart infusion (BHI) broth (*30*). The shaded region highlights the temperature range in which the model best predicts relative growth rates. (**B**) Simulated growth rates relative to maximum WT growth rate are shown for the WT strain, a strain with critical temperatures of the four most growth-limiting proteins at 42.2°C adjusted for maximum activity at that temperature, and a strain with critical temperatures of all growth-limiting proteins at 42.2°C adjusted for maximum activity at that temperature. The gray regions indicate the predicted most growth-limiting proteins and pathways for each phase of WT growth.

activity functions for just the four most-growth-limiting proteins had similar but smaller effects on temperature-dependent growth (Fig. 2B).

Adaptive laboratory evolution experiments have yielded 119 thermotolerant *E. coli* mutants (*10*, *11*). Investigating mutations in metabolic genes and their regulators (*24*) with our modeling framework yielded classification of potential causal mutations for 51 strains (supplementary text and table S3) and possible mechanistic explanations for their functionality in thermotolerance through compensating for heat-limited growth factors (Fig. 3), often consisting of cofactors. Statistical analysis (table S3) established that predicted causal mutant gene combinations conferring thermotolerance have a low probability of being identified by chance, signifying the predictive accuracy of heat-affected metabolic activities.

Mutations decreasing thermosensitivity of metabolic activities could stabilize or otherwise increase protein activity at high temperatures, for example, through increased gene expression. We thus profiled gene-expression of WT *E. coli* at 37° and 42°C (table S4) to identify genes with heat-induced transcription. Such genes participate in native heat-shock response and offer possible mechanisms for adaptive evolution of thermotolerance.

Specific susceptible proteins may be directly characterized by replacement with more thermostable proteins, by increasing gene expression to compensate for decreased activity (*25*), or by bypassing their function through nutrient supplementation. We chose metabolites produced immediately downstream of model-predicted growth-limiting proteins (fig. S4) and for which transport mechanisms are known in *E. coli* as supplements and found a set of six compounds that supplemented heat-limited growth factors (table S5). Each individual compound and a mixture combining all compounds were tested for effects on growth rate at 42° and 37°C. The mixture increased log-phase growth rate at 42°C by about 13% but yielded no benefit at 37°C (Fig. 4).

Triplicate experiments for components of the supplement mixture (table S6) were prioritized for the two compounds resulting in the highest growth rates at 42°C in single experiments (fig. S5). Pantothenate and biotin both provided heat-dependent supplementation, although to a lesser degree than the mixture (Fig. 4). Production of the coenzyme A (CoA) precursor pantothenate in WT grown at 37°C has been measured in excess of the minimum requirement for growth by as much as 15-fold, leading to excretion (*26*). A pathway with such excess activity at 37°C being successfully supplemented at 42°C indicates a substantial heat-dependent loss of function. This experimental result supports our predictions that this pathway has lowered activity because of thermal deactivation of PanB, PanC, PanD, and IlvC proteins. The heat dependency of this supplementation indicates that supplements do not simply alleviate the burden of synthesizing cofactors from the nutrient carbon source; it confirms the accuracy of thermosensitive metabolic activities predicted by our modeling framework and supports the precise proteins predicted to be limiting at 42°C. The supplement mixture elicited a combined benefit beyond that observed for individual compounds. The combination of pantothenate and biotin accounts for part but not all of the benefit of the full mixture (Fig. 4). Additions to pantothenate and biotin within the mixture appear to compensate for less rate-limiting growth factor deficits.

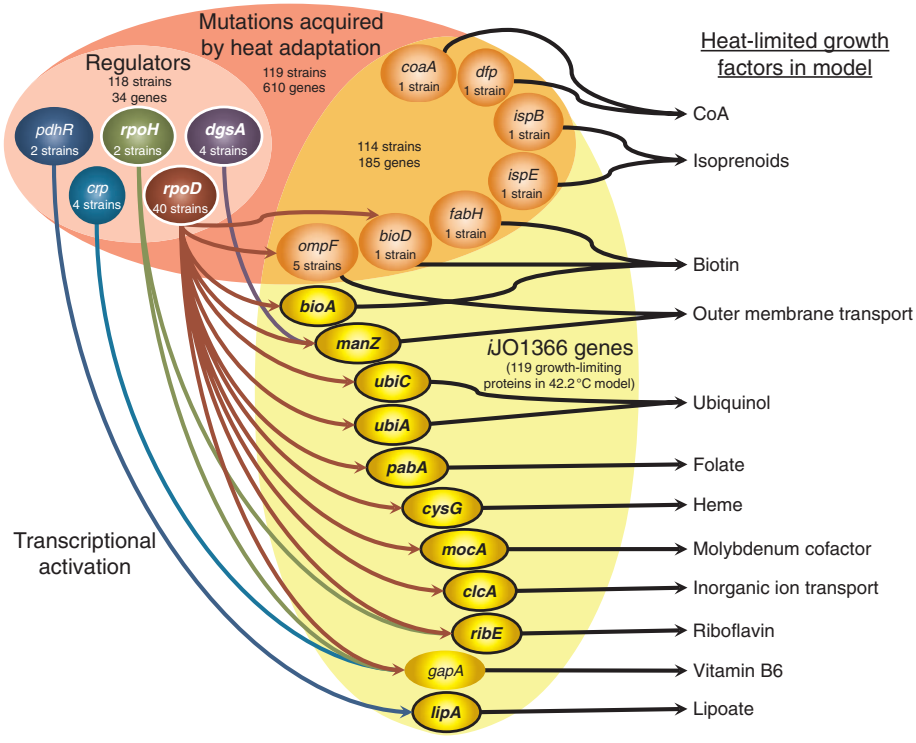

Fig. 3. Mechanisms predicted to confer thermotolerance are summarized for heat-adapted *E. coli* strains. The total number of heat-adapted strains and mutated genes is given and also noted for the regulatory and metabolic subsets of mutated genes. Only regulators acting on metabolic genes both predicted to lead to thermosensitivity and with heat-induced transcription in WT are depicted, except for *crp*. Only metabolic genes predicted to lead to thermosensitivity and either mutated in the set of evolved strains or both activated by depicted regulators and with heat-induced transcription in WT are depicted, except for *gapA*. Encircled, bolded genes show heat-induced transcription in WT. Predicted growth factors limited by heat-dependent decreases in protein activity are indicated at right.

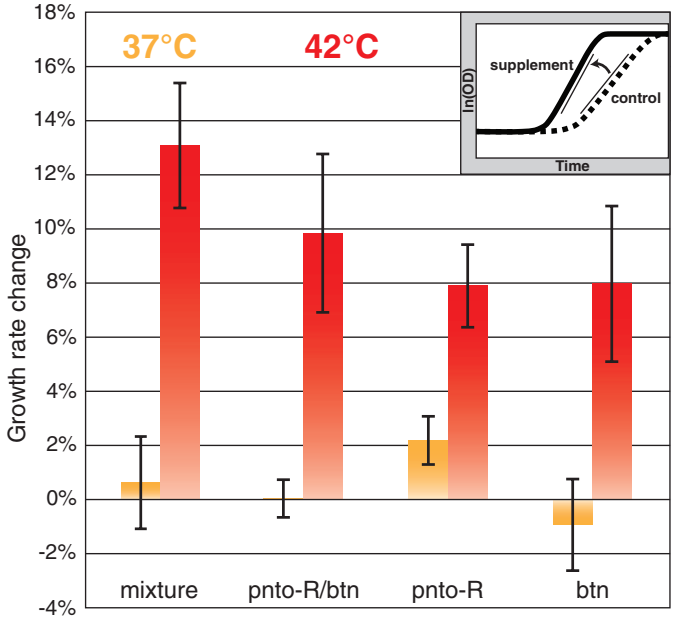

Fig. 4. Changes in specific growth rate upon supplementation relative to a no-supplement control are depicted in orange for 37°C and red for 42°C. Error bars show standard deviations with $n = 3$ for each condition. (**Inset**) How growth rate changes were computed by comparing the maximum slopes of growth curves between the control and supplement condition. Mixture, combination of all six supplements; pnto-R, pantothenate; btn, biotin; OD, optical density.

The *E. coli* GEM-PRO platform reconciled disparate data types to explain fundamental properties of thermosensitivity, providing evidence that metabolic processes are growth-limiting under heat stress and providing mechanistic interpretations of complex thermotolerant mutation data. Our model suggests that these dependencies arise directly from the systemic constraints that proteome thermostability imposes upon growth, which could be relieved through exogenous supplementation of the most-limiting processes, among them CoA and biotin synthesis. Understanding thermotolerance in microbes has important implications for developing industrial microbial biocatalysts (*27*), probiotics, and bacterial vaccines (*28*). Most efficient producers of compounds of interest are not naturally thermotolerant, but the absence of a genetic system often limits the utility of native thermophiles in industrial processes. Therefore, strategies for increasing thermotolerance of production strains may be useful. Our result supports the necessity of systems biology in understanding complex stress responses. Furthermore, these findings would not have been possible with use of either protein structure data or the metabolic network in isolation, illustrating the potential of advancing the field of structural systems biology.

References and Notes

1. E. Van Derlinden, K. Bernaerts, J. F. Van Impe, *J. Appl. Microbiol.* **104**, 438 (2008).
2. A. Korkegian, M. E. Black, D. Baker, B. L. Stoddard, *Science* **308**, 857 (2005).
3. C. R. Fischer, D. Klein-Marcuschamer, G. Stephanopoulos, *Metab. Eng.* **10**, 295 (2008).
4. P. Beltrao, C. Kiel, L. Serrano, *Curr. Opin. Struct. Biol.* **17**, 378 (2007).
5. Y. Zhang *et al.*, *Science* **325**, 1544 (2009).
6. R. L. Chang, L. Xie, L. Xie, P. E. Bourne, B. Ø. Palsson, *PLOS Comput. Biol.* **6**, e1000938 (2010).
7. Q. C. Zhang *et al.*, *Nature* **490**, 556 (2012).
8. X. Wang *et al.*, *Nat. Biotechnol.* **30**, 159 (2012).
9. T. M. Cheng *et al.*, *PLOS Comput. Biol.* **8**, e1002738 (2012).
10. O. Tenaillon *et al.*, *Science* **335**, 457 (2012).
11. I. K. Blaby *et al.*, *Appl. Environ. Microbiol.* **78**, 144 (2012).
12. J. D. Orth *et al.*, *Mol. Syst. Biol.* **7**, 535 (2011).
13. C. T. Porter, G. J. Bartlett, J. M. Thornton, *Nucleic Acids Res.* **32**, D129 (2004).
14. I. M. Keseler *et al.*, *Nucleic Acids Res.* **39**, D583 (2011).
15. Z. Huang *et al.*, *Nucleic Acids Res.* **39**, D663 (2011).
16. UniProt Consortium, *Nucleic Acids Res.* **40**, D71 (2012).
17. H. M. Berman *et al.*, The Protein Data Bank, *Nucleic Acids Res.* **28**, 235 (2000).
18. M. Scheer *et al.*, *Nucleic Acids Res.* **39**, D670 (2011).
19. M. D. Kumar *et al.*, *Nucleic Acids Res.* **34**, D204 (2006).
20. T. Ku *et al.*, *Comput. Biol. Chem.* **33**, 445 (2009).
21. M. Oobatake, T. Ooi, *Prog. Biophys. Mol. Biol.* **59**, 237 (1993).
22. K. P. Murphy, E. Freire, *Pure Appl. Chem.* **65**, 1939 (1993).
23. K. A. Dill, K. Ghosh, J. D. Schmit, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 17876 (2011).
24. S. Gama-Castro *et al.*, *Nucleic Acids Res.* **39**, D98 (2011).
25. T. S. Gunasekera, L. N. Csonka, O. Paliy, *J. Bacteriol.* **190**, 3712 (2008).
26. S. Jackowski, C. O. Rock, *J. Bacteriol.* **148**, 926 (1981).
27. P. Turner, G. Mamo, E. N. Karlsson, *Microb. Cell Fact.* **6**, 9 (2007).
28. B. N. Duplantis *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 13456 (2010).
29. V. S. Cooper, A. F. Bennett, R. E. Lenski, *Evolution* **55**, 889 (2001).
30. E. Van Derlinden, J. F. Van Impe, *Int. J. Food Microbiol.* **158**, 73 (2012).

Acknowledgments: We thank H. Nagarajan, A. M. Feist, P. Charusanti, P. E. Bourne, J. A. Lerman, E. T. O'Brien, D. Zielinski, Y. Zhang, and L. Jaroszewski for insightful discussion on this work and J. Orth for providing the *i*JO1366 network map. This work was supported by NSF grant GK-12 742551 to R.L.C.; U.S. Department of Energy grant DE-SC0004917 to R.L.C., K.A., D.K., and B.O.P.; NIH grants R01GM101457, R01 GM057089, and U54GM094586 to Z.L. and A.G.; and the Novo Nordisk Foundation to B.O.P. Data reported in this report are available in the supplementary materials and the National Center for Biotechnology Information Gene Expression Omnibus (GEO) database (GSE42675). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The University of California San Diego has filed a provisional patent for the method to predict proteins responsible for cellular thermosensitivity.

Supplementary Materials

www.sciencemag.org/cgi/content/full/340/6137/1220/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S5
Tables S1 to S6
Database S1
References (*31*–*53*)

12 December 2012; accepted 5 April 2013
10.1126/science.1234012

Widespread Production of Extracellular Superoxide by Heterotrophic Bacteria

Julia M. Diaz,[1]* Colleen M. Hansel,[1,2]†‡ Bettina M. Voelker,[3] Chantal M. Mendes,[1] Peter F. Andeer,[2] Tong Zhang[2]

Superoxide and other reactive oxygen species (ROS) originate from several natural sources and profoundly influence numerous elemental cycles, including carbon and trace metals. In the deep ocean, the permanent absence of light precludes currently known ROS sources, yet ROS production mysteriously occurs. Here, we show that taxonomically and ecologically diverse heterotrophic bacteria from aquatic and terrestrial environments are a vast, unrecognized, and light-independent source of superoxide, and perhaps other ROS derived from superoxide. Superoxide production by a model bacterium within the ubiquitous *Roseobacter* clade involves an extracellular oxidoreductase that is stimulated by the reduced form of nicotinamide adenine dinucleotide (NADH), suggesting a surprising homology with eukaryotic organisms. The consequences of ROS cycling in immense aphotic zones representing key sites of nutrient regeneration and carbon export must now be considered, including potential control of carbon remineralization and metal bioavailability.

Heterotrophic bacteria are ubiquitous and abundant components of natural ecosystems. These metabolically versatile organisms alter the global environment by mediating the redox transformation of many elements, including carbon, iron, and mercury. Biologically produced reactive oxygen species (ROS) may play an unrecognized role in these pathways because of the susceptibility of these elements to redox reactions with ROS (*1*–*4*), but the ability of common heterotrophic bacteria to produce extracellular ROS has remained largely unexplored (*2*). Nevertheless, the recent discovery of extracellular superoxide (O_2^-) production by a manganese-oxidizing marine bacterium belonging to the prolific *Roseobacter* clade (*5*) suggests that this capability may be present in environmentally relevant heterotrophs.

ROS production in natural waters (fig. S1) has long been linked to abiotic photooxidation of organic compounds (*6*). Yet, biological ROS production has been historically recognized (*7*), and recent field evidence indicates that biological production is likely the dominant ROS source in many marine and freshwater systems (*8*–*10*). Although all aerobic organisms produce intracellular ROS as a metabolic byproduct (*11*), ROS are generally toxic to living cells. Intracellular concentrations are therefore kept very low by enzymes that scavenge ROS, such as superoxide dismutase and catalase. This strict regulation of intracellular ROS and the inability of charged ROS to pass the lipid bilayer (*2*) point to biologically directed processes rather than adventitious cell rupture or leakage as the primary source of biogenic ROS in the environment. In fact, a number of cultivated eukaryotic phytoplankton and cyanobacteria produce extracellular

[1]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA. [2]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA. [3]Department of Chemistry and Geochemistry, Colorado School of Mines, Golden, CO 80401, USA.

*Present address: Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA.
†Present address: Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA.
‡Corresponding author. E-mail: chansel@whoi.edu