Title: Sulfur oxidation genes in diverse deep-sea viruses

Authors: Karthik Anantharaman[1], Melissa B. Duhaime[2], John A. Breier[3], Kathleen Wendt[4], Brandy M. Toner[4], and Gregory J. Dick[1,2,5*]

Affiliations:

[1]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA

[2]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

[3]Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

[4]Department of Soil, Water and Climate, University of Minnesota – Twin Cities, St. Paul, MN, USA

[5]Center for Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

[*]To whom correspondence should be addressed. E-mail: gdick@umich.edu

One sentence summary: Viruses of ubiquitous sulfur-oxidizing bacteria in the dark oceans possess genes for oxidation of elemental sulfur.

Abstract :

Viruses are the most abundant biological entities in the oceans and a pervasive cause of mortality of microorganisms that drive biogeochemical cycles. Although the ecological and evolutionary impacts of viruses on marine phototrophs are well-recognized, little is known about their impact on ubiquitous marine lithotrophs. Here we report 18 genome sequences of double-stranded DNA viruses that putatively infect widespread sulfur-oxidizing bacteria. Fifteen of these viral genomes contain auxiliary metabolic genes for the alpha and gamma subunits of reverse dissimilatory sulfite reductase *(rdsr)*. This enzyme oxidizes elemental sulfur, which is abundant in the hydrothermal plumes studied here. Our findings implicate viruses as a key agent in the sulfur cycle and as a reservoir of genetic diversity for bacterial enzymes that underpin chemosynthesis in the deep oceans.

Main text:

Chemolithoautotrophic bacteria are ubiquitous in the dark oceans (*1*), where they serve as a sink for CO_2 (*2*) through primary production that equals up to 53% of the particulate organic carbon exported from the photic zone (*3*). Uncultured sulfur-oxidizing bacteria of the SUP05 clade are among the most abundant and widespread marine chemolithoautotrophs, fixing carbon and oxidizing reduced sulfur species and hydrogen in diverse marine environments such as hydrothermal vent plumes (*4*), hydrothermal vent-associated animals (*5, 6*), and oxygen minimum zones (*7*), where they underpin cryptic links between the sulfur and nitrogen cycles (*8*). Although viruses are abundant in these deep-sea ecosystems (*9*), little is known about viruses that infect lithotrophic primary producers.

We conducted shotgun metagenomic sequencing on samples from five different hydrothermal vent plumes and associated deep ocean waters at the Eastern Lau Spreading Center (Lau Basin) in the Western Pacific Ocean and one plume at Guaymas Basin in the Gulf of California (*10, 11*) (Table S1). *De novo* assembly of sequence reads and binning by tetranucleotide signatures (*12*) revealed discrete genomic 'bins' (Fig. S1). Five bins (henceforth Lau77, Lau85, Lau87, Lau218 and Lau220) contained 18 viral genome sequences of putative SUP05 viruses. Phylogeny of the viral large terminase gene (*terL*) (Fig. S2) (which reflects phage DNA packaging mechanisms (*13*)), synteny with well-characterized phage of known taxonomy (Fig. S3), and results of protein sequence similarity searches against public sequence databases (Fig. S4) indicated that the five viruses belonged to three marine viral families of the orders *Caudovirales* (dsDNA viruses, no RNA stage), *Podoviridae*, *Siphoviridae* and *Myoviridae* (Table S2).

Fifteen of the 18 viral genomes (from four of the five SUP05 viral genomic bins) contained genes encoding the alpha (*rdsrA*) and gamma (*rdsrC*) subunits of the reverse-acting dissimilatory sulfite reductase (Rdsr) complex for elemental sulfur oxidation (Fig. 1). No other *rdsr* genes or other sulfur oxidation genes were present on the viral genomes. Analysis of bacterial genome bins recovered from Lau and Guaymas metagenomes revealed co-localized *rdsr* genes in the order *rdsrABEFHCMKLJOPN* in the Gammaproteobacteria Lau10 (SUP05), Lau62 (EC-01-9C-26), and Lau60 (unclassified). The deltaproteobacterium Lau20 (Sar324) (*14*) possessed only *rdsrAB*. Regions flanking the bacterial *rdsr* genes showed no similarity to the viral genome sequences, suggesting that viral *rdsr* genes were derived from selective retention rather than recent homologous recombination with bacterial genomic DNA.

Phylogenetic analyses indicated that all viral *rdsrA* genes recovered affiliated with SUP05 Gammaproteobacteria (74-96% amino acid identity, Fig. S5) and were distinct from *rdsrA* genes of other bacteria (Fig. 2). We identified two distinct groups of *rdsrA* sequences that each included both viral and bacterial sequences. All viral *rdsrA* genes fall into Group one except for Lau85, which contained two copies of *rdsrA* with one representative in each group. Bacterial representatives of Group one included the SUP05 GB-1 and GB-2 from Guaymas as well as *Bathymodiolus* mussel symbionts (*6*). Group two was populated by SUP05 from oxygen minimum zones (*7*) and symbionts of deep-sea clams (*5*). The tight phylogenetic clustering of *rdsrA* gene sequences of three distinct phage families with SUP05 bacteria in two separate lineages suggests that the phage *rdsrA* genes originated from SUP05 and were transferred to viruses. These observations are analogous to those of core photosynthesis genes in cyanobacterial phages and other microbe-derived auxiliary metabolic genes (*15, 16*) (e.g. *psbA, psbD, mazG*) that are similar but not identical to known hosts, forming sub-clusters distinct from host proteins (*17, 18*).

The amino acid sequences deduced from the viral *rdsrA* and *rdsrC* genes indicated the capacity to serve as functional sulfur-oxidizing enzymes. Phage RdsrA contained all conserved sulfite reductase residues and secondary structure elements for α-helix and β-sheets (Fig. S6). Similarly, a multiple alignment of RdsrC amino acid sequences indicated highly conserved residues across two distinct groups (Fig. S7). We also identified other genes in the viral genomes with high sequence identity to SUP05, including iron-sulfur cluster proteins, cytochromes, and sulfur relay proteins (Table S3). The existence of these additional SUP05-like genes on viral genomes supports their specificity to SUP05 bacteria and suggests a role for viral genes in supplementing host metabolism.

Elemental sulfur is a key and stable intermediate of bacterial oxidation of reduced sulfur compounds and constitutes a 'bottleneck' in the sulfur cycle (*19*). Sulfur-oxidizing bacteria with an incomplete Sox pathway, such as SUP05, form intracellular or extracellular globules of elemental sulfur that serve as a store of electron donor for energy metabolism (Fig. 3) (*4*, *19*, *20, 21*). The presence of *rdsrA* and *rdsrC* on viral genomes may offer selective advantages to the viruses by supplementing host pathways for oxidation of this sulfur during infection. First, enhanced expression of *rdsrA* could replenish proteins involved in a rate limiting reaction in the host, as previously demonstrated with cyanobacterial phage D1 proteins involved in photosynthesis (*22*). Second, phage *rdsrC* could maintain or increase high transcription levels to ensure efficient delivery of sulfur-substrate to the RdsrAB complex during infection. Thus, phage auxiliary metabolic genes that can supplement or sustain sulfur oxidation metabolism in their hosts may ensure continued viral infection and replication.

A coupled bioenergetic-thermodynamic reaction path model (*23*) indicated that aerobic elemental sulfur oxidation potentially accounted for 80-93% of the total lithotrophic energy available in the ABE and Mariner hydrothermal plumes at Lau Basin at two temperatures that were representative of the plumes studied here (Fig. 4A, Fig. S8). The model predicted a minor role for 19 other lithotrophic microbial metabolisms considered, including oxidation of other reduced sulfur species such as hydrogen sulfide, sulfite and thiosulfate. These results were consistent with a previous model of hydrothermal plumes that found elemental sulfur to be the most abundant source of chemosynthetic energy (*24*), and with our previous work (*4*, *25*, *26*), which was based on the same approach. Our model predicted greater relative energy yields for more energetically favored metabolisms such as sulfur oxidation because we assumed that metabolic reactions occurred in sequence, from most to least energetically favored, and enabled

reactions to explicitly modify the product and reactant pools. Although there is some uncertainty regarding the speciation of sulfur present in the plume, elemental sulfur is a central intermediate in the oxidation of other reduced forms of sulfur used by SUP05 (Fig. 3). X-ray fluorescence maps showed that sulfur was abundant in the plumes (Fig. 4B, Fig. S8), while microprobe X-ray diffraction showed elemental sulfur was widely present in particle aggregates with other crystalline phases, such as pyrite (Fig. 4C, Fig. S8). Although we did not conclusively identify intracellular elemental sulfur, our results showed that elemental sulfur presented an abundant source of energy for SUP05 bacteria in hydrothermal plumes and deep ocean waters of Guaymas and Lau basins.

The abundance and diversity of viruses infecting SUP05 bacteria in hydrothermal plumes suggests that chemolithoautotrophs in the deep sea face viral predation pressures similar to photosynthetic microbes in the surface waters (*27*). The remarkable synteny and conservation of the four viruses studied here (95-99% genome nucleotide identity) across hydrothermal vent environments and ocean basins suggests that these viruses are ubiquitous in marine environments dominated by SUP05 bacteria. Analyses of the Pacific Ocean Virome dataset (*28*) (Fig. S9), which notably contains viral communities from oxygen minimum zones dominated by SUP05 (*7*), revealed the presence of *rdsrA* and *rdsrC* genes (Table S5), consistent with the prevalence of phage-encoded sulfur oxidation beyond hydrothermal plumes and in the wider pelagic oceans.

To date, deep-sea SUP05 has evaded growth in laboratory cultures, thus direct host-phage manipulations and validation of the underlying mechanisms of phage-influenced sulfur oxidation remain a challenge. Yet, this study demonstrates the sequence-based elucidation of microbial community dynamics through the discovery of phages that infect a widespread deep-sea bacterium. The existence of *rdsr* genes in viral genomes reveals a mechanism for horizontal

transfer of genes associated with sulfur cycling (*29*) and implicates viruses in the evolutionary dynamics of a central step in the planetary cycling of sulfur.

Supplementary Materials

www.sciencemag.org

Materials and Methods

Supplementary Text

Figs. S1 to S10

Tables S1 to S6

References (30-80)

Database S1

References and Notes

1. B. K. Swan *et al.*, *Science* **333**, 1296 (2011).
2. J. Aristegui, J. M. Gasol, C. M. Duarte, G. J. Herndl, *Limnol. Oceanogr.* **54**, 1501 (2009).
3. T. Reinthaler, H. M. van Aken, G. J. Herndl, *Deep Sea Res. (II Top. Stud. Oceanogr.)* **57**, 1572 (2010).
4. K. Anantharaman, J. A. Breier, C. S. Sheik, G. J. Dick, *Proc. Natl. Acad. Sci. USA* **110**, 330 (2013).
5. I. L. Newton *et al.*, *Science* **315**, 998 (2007).
6. J. M. Petersen *et al.*, *Nature* **476**, 176 (2011).
7. D. A. Walsh *et al.*, *Science* **326**, 578 (2009).
8. D. E. Canfield *et al.*, *Science* **330**, 1375 (2010).
9. S. Hara *et al.*, *Mar. Ecol. Prog. Ser.* **145**, 269 (1996).
10. G. J. Dick, B. M. Tebo, *Environ. Microbiol.* **12**, 1334 (2010).
11. R. A. Lesniewski *et al.*, *ISME J.* **6**, 2257 (2012).
12. G. J. Dick *et al.*, *Genome Biol.* **10**, R85 (2009).

13. S. R. Casjens *et al.*, *J. Bacteriol.* **187**, 1091 (2005).
14. C. S. Sheik, S. Jain, G. J. Dick, *Environ. Microbiol.* **16**, 304 (2014).
15. M. Breitbart, L.R. Thompson, C.A. Suttle, and M.B. Sullivan, *Oceanography* **20**, 135 (2007).
16. M. Breitbart, *Ann Rev Mar Sci.* **4**, 425 (2012).
17. J. C. Ignacio-Espinoza, M. B. Sullivan, *Environ. Microbiol.* **14**, 2113 (2012).
18. D. Lindell *et al.*, *Proc. Natl. Acad. Sci. USA* **101**, 11013 (2004).
19. W. Ghosh, B. Dam, *FEMS Microbiol. Rev.* **33**, 999 (2009).
20. D. Hensen *et al.*, *Mol. Microbiol.* **62**, 794 (2006).
21. K. T. Marshall, R. M. Morris, *ISME J.* **7**, 452 (2012).
22. D. Lindell *et al.*, *Nature* **438**, 86 (2005).
23. Materials and methods are available as Supplementary Materials on Science Online.
24. T. M. McCollom, *Deep Sea Res. (I Oceanogr. Res. Pap.)* **47**, 85 (2000).
25. J. A. Breier *et al.*, *Geochim. Cosmochim. Acta* **88**, 216 (2012).
26. M. Li *et al.*, *Nat. Commun.* **5**, 3192 (2014).
27. S. Avrani *et al.*, *Nature* **474**, 604 (2011).
28. B. L. Hurwitz, M. B. Sullivan, *PLoS One* **8**, e57355 (2013).
29. M. Klein *et al.*, *J. Bacteriol.* **183**, 6028 (2001).

Acknowledgements: This project is funded in part by the Gordon and Betty Moore Foundation Grant GBMF2609 and National Science Foundation Grant OCE1038006. We thank the University of Michigan Rackham Graduate School Faculty Research Fellowship Program for their support and Meng Li and Sunit Jain for their assistance. DNA sequencing was conducted at the University of Michigan DNA Sequencing Core. We acknowledge the captain and crew of R/V *Thomas G. Thompson*, Chief Scientist Anna-Louise Reysenbach, and Jason Sylvan for sampling and logistical support. For assistance with synchrotron measurements at BL10.3.2, we thank K.J. Edwards, J.V. Sorensen, M.A. Marcus, and S.C. Fakra. The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The nucleotide sequences are available from DOE JGI-IMG/MER - Taxon Object IDs (Kilo Moana: 3300001680, ABE: 3300001681, Mariner: 3300001678, Tahi Moana: 3300001679, Tui Malila: 3300001676 and

192 Guaymas: 3300001683) and NCBI (BioProject: PRJNA234377). The authors declare no
193 competing financial interests.

194

195 **FIGURE CAPTIONS**

196 **Fig. 1**

197 Gene content of 15 phage genomes from three viral families retrieved from Lau and Guaymas
198 basins. Colored nested circles represent syntenous viral genomes/contiguous genomic fragments
199 from locations indicated in the legend. Grey boxes on outermost circle indicate predicted genes
200 of the four viruses. *rdsrA* and *rdsrC* genes are highlighted in yellow.

201

202 **Fig. 2**

203 **A.** Phylogenetic tree of *rdsrA* genes inferred by Maximum Likelihood. **B.** Detailed view of the
204 SUP05 *rdsrA* clade. Group one and Group two sub-clades are shown on the right. Sequences are
205 colored by geographical origin; Blue – Guaymas Basin; Red – Kilo Moana (Lau Basin); Green –
206 Tahi Moana (Lau Basin); Purple – ABE (Lau Basin); Brown – Tui Malila (Lau Basin); Orange –
207 Mariner (Lau Basin).

208

209 **Fig. 3.**

210 Schematic of the sulfur oxidation pathway in SUP05 bacteria. Grey box indicates the reaction
211 impacted by SUP05 viruses. Key genes are shown in italics: *sqr* (sulfide quinone reductase), *sox*

(sulfur oxidation), *rdsr* (reverse dissimilatory sulfite reductase), *apr* (adenosine 5'-phosphosulfate reductase), and *sat* (sulfate adenylyltransferase).

Fig. 4

A. Modeled free energies of catabolic reactions as a percentage of total available free energy in the ABE hydrothermal plume at 2.5 and 5^{o}C. Total available free energy in the plume is normalized per kilogram plume fluid (p.f.) and per kilogram vent fluid (v.f.). **B.** Distribution of iron (displayed in red) and sulfur (displayed in green) in particles collected at 0.5 m above the ABE vent. Locations where elemental sulfur was detected by micro-probe X-ray diffraction measurements are indicated as spots 0, 1, 2, and 19. **C.** Radially integrated diffractograms with elemental sulfur peaks annotated (*****) at 22.0, 25.7, and 34.1 degrees 2-theta.