
This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of April 8, 2011):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/319/5866/1083.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2008/02/21/319.5866.1083.DC1.html>

This article has been **cited by** 31 article(s) on the ISI Web of Science

This article has been **cited by** 17 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/319/5866/1083.full.html#related-urls>

This article appears in the following **subject collections**:

Biochemistry

<http://www.sciencemag.org/cgi/collection/biochem>

26. B. D. Santer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15248 (2007).
27. A. Wood *et al.*, *Clim. Dyn.* **16**, 755 (2004).
28. E. P. Maurer, H. G. Hidalgo, *Hydrol. Earth Syst. Sci. Discuss.* **4**, 3413 (2007).
29. X. Liang, D. Lettermaier, A. Wood, S. Burges, *J. Geophys. Res.* **99**, 14415 (1994).
30. T. P. Barnett *et al.*, *Science* **309**, 284 (2005); published online 2 June 2005 (10.1126/science.1112418).
31. X. Zhang *et al.*, *Nature* **448**, 461 (2007).
32. The choices were CCSM3-FV noise for significance testing, PCM fingerprint, and statistical downscaling with the CA method. In the multivariable case, PCM noise was used for normalization.
33. The MIROC data were generously supplied by the National Institute for Environmental Studies, Onogawa, Tsukuba, Ibaraki, Japan. The PCM simulation had previously been made available to the Scripps Institution of Oceanography (SIO) by the National Center for Atmospheric Research for the Accelerated Climate Prediction Initiative project. Supported by the Lawrence Livermore National Laboratory (LLNL) through a Laboratory-Directed Research and Development grant to SIO via the San Diego Super Computer Center for the LUCSID project; the U.S. Department of Energy and NOAA through the International Detection and Attribution Group (T.P.B.); Program of Climate Model Diagnoses and Intercomparison grant DOE-W-7405-ENG-48 (C.B., B.D.S., G.B., A.A.M.); the U.S. Geological Survey and SIO (D.R.C., M.D.D.); and the California Energy Commission (D.W.P., H.G.H.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1152538/DC1

SOM Text

Figs. S1 to S3

References

2 November 2007; accepted 23 January 2008

Published online 31 January 2008;

10.1126/science.1152538

Include this information when citing this paper.

Atomic-Level Models of the Bacterial Carboxysome Shell

Shiho Tanaka,¹ Cheryl A. Kerfeld,^{3,4} Michael R. Sawaya,² Fei Cai,⁵ Sabine Heinhorst,⁵ Gordon C. Cannon,⁵ Todd O. Yeates^{1,2,*}

The carboxysome is a bacterial microcompartment that functions as a simple organelle by sequestering enzymes involved in carbon fixation. The carboxysome shell is roughly 800 to 1400 angstroms in diameter and is assembled from several thousand protein subunits. Previous studies have revealed the three-dimensional structures of hexameric carboxysome shell proteins, which self-assemble into molecular layers that most likely constitute the facets of the polyhedral shell. Here, we report the three-dimensional structures of two proteins of previously unknown function, CcmL and OrfA (or CsoS4A), from the two known classes of carboxysomes, at resolutions of 2.4 and 2.15 angstroms. Both proteins assemble to form pentameric structures whose size and shape are compatible with formation of vertices in an icosahedral shell. Combining these pentamers with the hexamers previously elucidated gives two plausible, preliminary atomic models for the carboxysome shell.

The carboxysome enhances CO₂ fixation inside many photosynthetic and chemoautotrophic bacterial cells by encapsulating the key enzymes ribulose-1,5-bisphosphate

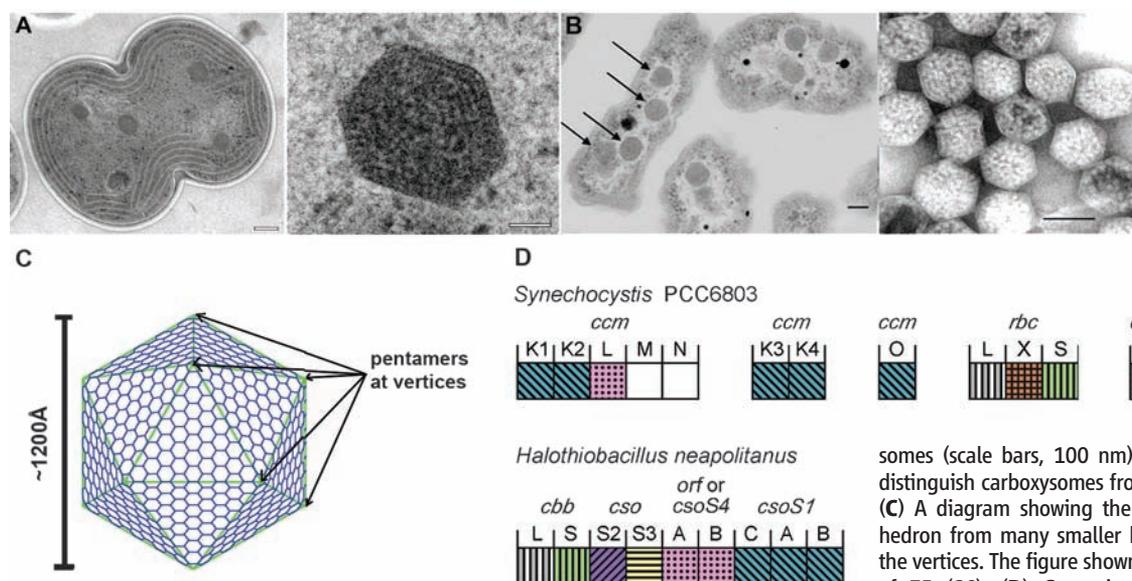
carboxylase-oxygenase (RuBisCO) and carbonic anhydrase (1–3). In contrast to membrane-bound eukaryotic organelles, carboxysomes and related bacterial microcompartments (4, 5) have a pro-

teinaceous outer shell, which is roughly polyhedral in shape (Fig. 1). The earliest observations of carboxysomes by electron microscopy date back more than 40 years (6, 7). Subsequent genetic and biochemical studies have provided essential information about their function and composition (3, 8–10); structural studies (11–14) have begun to illuminate the functional mechanisms and architectural details of the carboxysome.

Carboxysomes are found in all cyanobacteria and in some chemoautotrophic bacteria. Two types of carboxysomes have been defined by patterns of genomic organization and by protein sequence comparisons (15). They are represented in two model organisms, *Halothiobacillus neapolitanus*

¹Department of Chemistry and Biochemistry, University of California at Los Angeles (UCLA), Los Angeles, CA 90095, USA. ²UCLA–Department of Energy Institute for Genomics and Proteomics, Los Angeles, CA 90095, USA. ³Department of Energy–Joint Genome Institute, Walnut Creek, CA 94598, USA. ⁴Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA. ⁵Department of Chemistry and Biochemistry, University of Southern Mississippi, Hattiesburg, MS 39406, USA.

*To whom correspondence should be addressed. E-mail: yeates@mbi.ucla.edu



associated proteins. Homologous proteins from the two model organisms are shaded and colored similarly. The CcmL and OrfA (or CsoS4A) proteins under study here are colored in pink. The *rbc* and *ccb* genes code for the large and small subunit of RuBisCO. (A) is adapted from ref. (11); (B) is adapted from ref. (12).

(a chemoautotroph containing α -carboxysomes) and *Synechocystis* sp. strain PCC (Pasteur Culture Collection) 6803 (*Syn.* 6803, a cyanobacterium containing β -carboxysomes); potential functional differences between the two types are not clear. The carboxysome shell is composed predominantly of CsoS1 proteins (homologous proteins CsoS1A, B, and C) in *H. neapolitanus* (16, 17) and of CcmK proteins (homologous proteins CcmK1 to 4) in *Syn.* 6803 (Fig. 1D). We have previously determined the structures of four of these shell proteins and found that they are all hexamers and that they tend to form extended, tightly packed molecular layers hypothesized to represent the flat facets of the polyhedral carboxysome shell. The structures suggest that diffusion of metabolites into and out of the carboxysome is limited by the small hexameric pores and gaps between the hexamers (11, 12).

Two recent electron microscopy studies have confirmed that carboxysomes are approximately icosahedral in shape (13, 14). The construction of large icosahedral structures typically requires a combination of hexameric and pentameric units (18, 19). Pentamers generate curvature in an otherwise flat hexagonal sheet and occupy

the vertices of an icosahedral shell (Fig. 1C). Candidates for such pentameric components of the shell had not been identified. Functions have been elucidated for most of the proteins encoded within well-characterized carboxysome operons (3). However, functions had not been assigned to the gene products of two open reading frames in *H. neapolitanus*, designated *orfA* and *orfB*, or their homolog in *Syn.* 6803, CcmL (Fig. 1D).

The CcmL protein from *Syn.* 6803 was expressed in *Escherichia coli* cells, and its crystal structure was determined by selenomethionine multiwavelength anomalous phasing methods (20). An atomic model was built and refined at a resolution of 2.4 Å, with a final model consisting of 96 of the 100 residues of the complete CcmL protein (20). OrfA was expressed in *E. coli* cells, and its crystal structure was determined by molecular replacement using the CcmL structure as a search model; the amino acid sequence identity between the two proteins is 37%. The OrfA structure was refined at a resolution of 2.15 Å, yielding a final model that contains 82 of the 83 amino acid residues in that protein (20). The structures of CcmL and OrfA are highly similar;

their backbones superimpose with a root mean square deviation (RMSD) of only 1.0 Å (Fig. 2). CcmL consists of seven β strands and one α helix, with five of the β strands forming a β barrel (Fig. 2). OrfA contains the same central five-stranded β barrel and the α helix, but it lacks the two C-terminal β strands present in CcmL, owing to its smaller size.

Both CcmL and OrfA formed symmetric pentamers in their respective crystal structures (Fig. 2). The individual subunits of the pentamers are held tightly together by their protruding C-terminal regions, which form extensive interactions with neighboring subunits. The large subunit interfaces and the structural agreement between the two proteins argue that the observed pentameric structures represent their natural biological forms. The CcmL pentamer is shaped roughly like a pentagonal disk 30 to 35 Å thick. One side of the disk, where both the N and C termini reside, is broader than the other, giving the pentamer the appearance of a truncated pyramid. Its base has a pentagonal edge of ~42 Å (Fig. 2), which narrows to ~35 Å in the middle of the pentamer. The subunits are packed tightly around the axis of symmetry, leaving only a narrow pore through the center, with a diameter of ~5 Å in CcmL and ~3.5 Å in OrfA. This tight packing around a narrow pore is reminiscent of the hexameric carboxysome shell proteins. The tight packing of those shell subunits has been proposed to limit movement of substrates and products into and out of the carboxysome (11, 12).

The database of known protein structures was searched for proteins similar to CcmL and OrfA. A single protein, EutN from *E. coli*, was identified as having a similar three-dimensional fold. The protein backbone of EutN superimposes on that of CcmL with an RMSD of only 1.3 Å (Fig. 2C). The structure of EutN had been determined as part of a structural genomics program (21), so detailed biological interpretation was not provided. It is known, however, that the *eut* operon, which has been studied in *Salmonella typhimurium* (22), encodes several proteins involved in ethanolamine utilization, which takes place inside the *eut* microcompartment. Remarkably, however, the reported EutN structure is hexameric rather than pentameric (Fig. 2). The difference between the oligomeric state of EutN, compared with those of CcmL and OrfA, presumably reflects structural differences between the *eut* microcompartment and the carboxysome, as well as a different functional role for EutN. Aside from EutN, no other homolog of CcmL or OrfA could be identified in the genome of *E. coli*, which suggests that some other potentially unrelated protein might serve as a pentameric shell protein in the *eut* microcompartment. Alternatively, the *eut* microcompartment could lack pentamers, which might explain why the microcompartments of enteric bacteria tend to have less-regular icosahedral shapes than carboxysomes.

The pentameric organization of CcmL and OrfA suggests that these proteins serve as

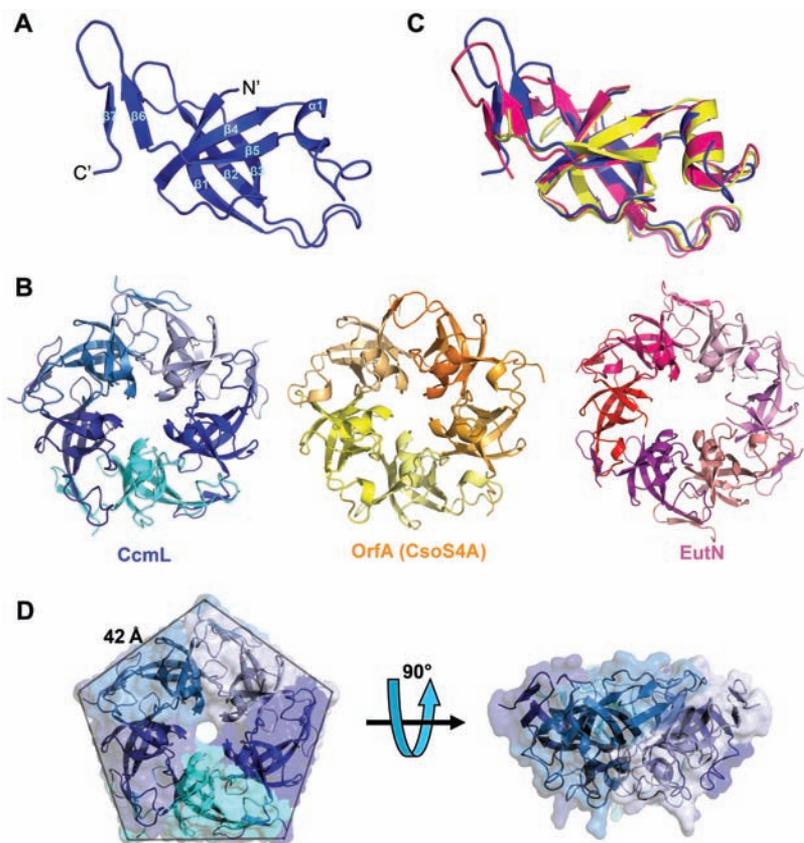


Fig. 2. Crystal structures of the carboxysome proteins CcmL and OrfA revealing pentagonal symmetry. (A) Structure of the CcmL monomer from *Syn.* 6803. (B) A comparison of similar structures: CcmL (blue), OrfA (or CsoS4A) from *H. neapolitanus* (yellow), and EutN from *E. coli* (pink) (PDB 2Z9H). The RMSD between the protein backbones of CcmL and OrfA is 1.0 Å, and 1.3 Å between CcmL and EutN. (C) CcmL and OrfA assemble as natural pentamers. EutN, which is part of the *eut* operon that encodes proteins presumed to comprise the distinct *eut* microcompartment in *E. coli*, is instead hexameric. (D) Top and side views of the CcmL pentamer showing a pentagonal disk with slanted sides.

vertices in their respective carboxysome shells. Whether there are significant functional differences between OrfA and its adjacent paralog in *H. neapolitanus*, OrfB, is unknown. The role proposed here for CcmL and OrfA is consistent with the earlier observation that deleting the *ccmL*

gene leads to elongated, rod-shaped carboxysomes (9), as would be expected if the component required for vertex formation is lost. In large icosahedral shells, 12 pentamers are present among a much greater number of hexamers. In the carboxysome, only 60 copies of the CcmL or OrfA

(and possibly OrfB) subunits would be expected to be present among about four or five thousand hexameric shell subunits (3, 12, 23) and a similar number of total RuBisCO subunits (14). This explains the inability in earlier experiments to identify OrfA or OrfB in preparations of carboxysomes.

Knowing the structures of both the hexameric and pentameric components of the carboxysome, it is possible to model the architecture of the complete shell. The size and shape of the CcmL and OrfA pentamers make them suitable for insertion into the hexagonally packed molecular layers previously elucidated for the carboxysome shell. The hexameric units (i.e., of CcmK subunits) in the *Syn.* 6803 shell are packed together with an edge length of 40 Å (20) [supporting online material (SOM) text]. This size is consonant with the edge length of the CcmL pentamers. The suggestion that the CcmL and OrfA pentamers generate curvature by being inserted into an otherwise flat hexagonal molecular layer is consistent with recent electron microscopy investigations in which carboxysomes exhibit a relatively simple polyhedral shape at their vertices (13, 14). Some large viral capsids have defining features, such as major protrusions, at their pentameric vertices (24, 25). The lack of such protrusions at the vertices of the carboxysome places useful constraints on atomic models of the carboxysome that can be built from the known hexameric and pentameric components.

CcmL (or OrfA) pentamers can be fit into vacancies created by folding up hexameric molecular layers of CcmK (or CsoS1) proteins in four distinct ways (Fig. 3). It has not yet been possible either by biophysical methods or by electron microscopy to define which side of the hexagonally packed protein layers represents the inside of the carboxysome and which side faces outward into the bacterial cytosol (12). The hexagonal layers, therefore, could be folded up with either side facing outward (fig. S1). Likewise, which side of the CcmL or OrfA pentamer faces inward and which faces outward is unknown. Computational attempts were made to fit the CcmL and OrfA pentamers into their respective hexagonal layers, evaluating all four distinct (inside versus outside) orientations of the pentamers and hexamers. In addition, rotation and translation of the pentamer about its central axis was allowed, but given the tightness of the fit, only a narrow range of rotations and translations was feasible. The best-fitting solutions were identified by evaluating geometric fit (26) and empirical energy functions (20, 27). Among the four possible orientations, one was a very poor fit and could be rejected. The other three led to potential packing solutions whose fits could be compared (fig. S2). One of these provided a much lower amount of buried surface area between the pentamer and its surrounding hexamers, and left larger gaps in the resulting shell and was, therefore, judged unlikely to be correct. Two plausible solutions for building complete models of the shell remained (Fig. 3). Both candidate

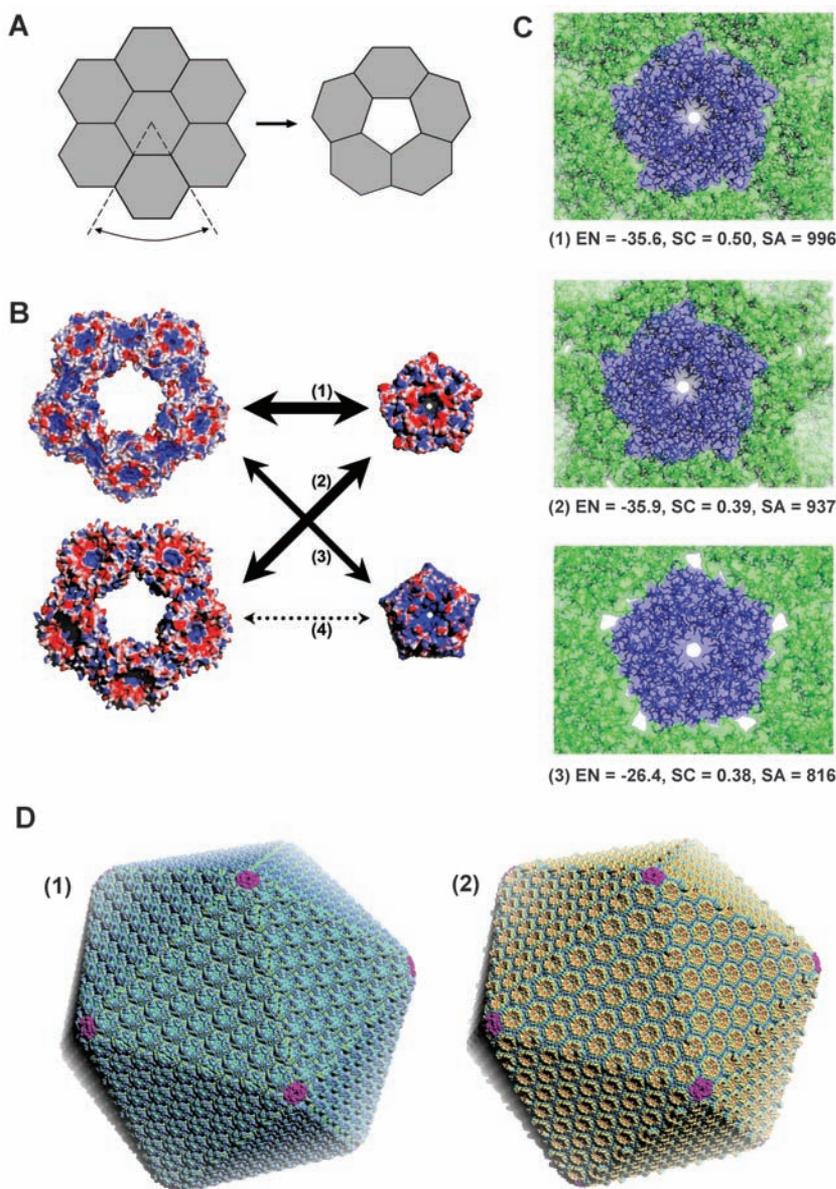


Fig. 3. Models of the carboxysome shell based on pentamer and hexamer components. **(A)** A flat layer of hexagons can be folded to give pentagonal vertices by removing one sector at each vertex. Twelve such vertices are present in an icosahedral shell. **(B)** Taken in combination, alternate choices for the curvature of the hexagonal layer and the orientation of the pentamer lead to four possible constructions, numbered 1 to 4 according to the quality of fit. Combination 4 led to impossible steric collisions. The structures are colored according to calculated electrostatic potential, from negative (red) to positive (blue). **(C)** Illustration of the best packing solutions for constructions 1 to 3. EN, calculated packing energies (27) (with more negative values being favorable); SC, surface complementarity (26); and SA, buried surface area between a pentamer and a single neighboring hexamer (with higher values of these parameters being favorable). **(D)** Two alternate models for the complete carboxysome shell, based on the two constructions, 1 and 2, judged to be most plausible. There are 740 hexamers and 12 pentamers in a $T = 75$ arrangement. The packing of hexamers is derived from multiple consistent crystal structures. The two models differ with respect to the orientation of the hexameric layer. The hexagonal layer is colored according to hydrophobicity, with increases showing as blue to orange. The CcmL pentamers are shown in magenta. The diameter from vertex to vertex is 1150 Å.

solutions orient the pentamer with its broader base facing outward. This is consistent with the role the pentamer appears to play in introducing curvature into the hexagonal layer. A notable feature of the hexagonal layer of the carboxysome shell is the presence of a bowl-shaped depression or concavity on one side of the hexameric building block (figs. S3 and S4). This depression is situated on the same side of the hexamer as both the N and C termini of the protein chain, and the prominence of the depression is affected by the disposition of the C terminus, which tends to vary in conformation between different homologs of the carboxysome hexamer (11, 12). The narrow pore through the carboxysome hexamer tends to be positively charged, but the bowl-shaped depression surrounding the pore has a partial hydrophobic character, giving one side on the hexameric layer a distinctive appearance (Fig. 3). The model building does not resolve whether this side of the layer faces inward or outward. In one of the plausible models constructed here, it faces inward and could interact with RuBisCO, carbonic anhydrase, or possibly other protein components.

The present findings clarify the roles of previously uncharacterized proteins in the carboxysome; accordingly, we propose that the genes *orfA* and *orfB* be renamed *csoS4A* and *csoS4B*, consistent with other known shell protein genes in α -type carboxysomes. The results also lead to atomic models for the carboxysome shell, but these are incomplete in numerous respects, and considerably more work will be required to complete our understanding of this structure. For instance, the details of how the pentamers and hexamers fit together are only approximated

here, and how the edges of the icosahedral shell are formed where a hexagonal layer bends is unknown. In addition, interactions that are almost certain to exist between the shell and the other enzymatic components of the carboxysome are only beginning to be elucidated (28).

The emerging structure of the carboxysome emphasizes common principles under which bacterial microcompartments and certain previously characterized viral capsids are constructed and illustrates the ability of evolution to produce a diversity of highly complex molecular structures from common principles and a small repertoire of basic building blocks. The elucidation of the component structures of the carboxysome shell also opens up strategies for designing new kinds of molecular containers or reaction chambers on the mid-nanometer scale.

References and Notes

1. J. M. Shively, R. S. English, S. H. Baker, G. C. Cannon, *Curr. Opin. Microbiol.* **4**, 301 (2001).
2. M. R. Badger, G. D. Price, *J. Exp. Bot.* **54**, 609 (2003).
3. S. Heinhorst, G. C. Cannon, J. M. Shively, in *Complex Intracellular Structures in Prokaryotes*, J. M. Shively, Ed. (Springer, Berlin, 2006), vol. 2, pp. 141–165.
4. T. A. Bobik, *Appl. Microbiol. Biotechnol.* **70**, 517 (2006).
5. J. T. Penrod, J. R. Roth, *J. Bacteriol.* **188**, 2865 (2006).
6. J. M. Shively, F. L. Ball, B. W. Kline, *J. Bacteriol.* **116**, 1405 (1973).
7. E. Gantt, S. F. Conti, *J. Bacteriol.* **97**, 1486 (1969).
8. R. S. English, S. C. Lorbach, X. Qin, J. M. Shively, *Mol. Microbiol.* **12**, 647 (1994).
9. G. D. Price, S. M. Howitt, K. Harrison, M. R. Badger, *J. Bacteriol.* **175**, 2871 (1993).
10. A. K. So *et al.*, *J. Bacteriol.* **186**, 623 (2004).
11. C. A. Kerfeld *et al.*, *Science* **309**, 936 (2005).
12. Y. Tsai *et al.*, *PLoS Biol.* **5**, e144 (2007).
13. M. F. Schmid *et al.*, *J. Mol. Biol.* **364**, 526 (2006).
14. C. V. Iancu *et al.*, *J. Mol. Biol.* **372**, 764 (2007).

15. M. R. Badger, D. T. Hanson, G. D. Price, *Funct. Plant Biol.* **29**, 161 (2002).
16. J. M. Shively *et al.*, *Can. J. Bot.* **76**, 906 (1998).
17. G. C. Cannon *et al.*, *Appl. Environ. Microbiol.* **67**, 5351 (2001).
18. J. E. Johnson, J. A. Speir, *J. Mol. Biol.* **269**, 665 (1997).
19. D. L. Caspar, A. Klug, *Cold Spring Harbor Symp. Quant. Biol.* **27**, 1 (1962).
20. Materials and methods are available as supporting material on Science Online.
21. Z. Wunderlich *et al.*, *Proteins* **56**, 181 (2004).
22. E. Kofoid, C. Rappleye, I. Stojiljkovic, J. Roth, *J. Bacteriol.* **181**, 5317 (1999).
23. J. M. Shively, R. S. English, *Can. J. Bot.* **69**, 957 (1991).
24. J. T. Huiskonen, V. Manole, S. J. Butcher, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 6666 (2007).
25. P. L. Stewart, R. M. Burnett, M. Cyrklaff, S. D. Fuller, *Cell* **67**, 145 (1991).
26. M. C. Lawrence, P. M. Colman, *J. Mol. Biol.* **234**, 946 (1993).
27. J. J. Gray *et al.*, *J. Mol. Biol.* **331**, 281 (2003).
28. B. M. Long, M. R. Badger, S. M. Whitney, G. D. Price, *J. Biol. Chem.* **282**, 29323 (2007).
29. See the Virus Particle Explorer, <http://viprdb.scripps.edu>.
30. The authors thank D. Cascio for crystallographic advice and the staff at the Advanced Light Source beamline 8.2.2 for technical assistance. This work was supported by a grant from the Biological and Environmental Research program of the Department of Energy Office of Science. Support was also provided by NSF (grants MCB-0444568 and DMR-0213883 to G.C.C. and S.H., respectively). Protein structure coordinates and structure factors have been deposited in the Protein Data Bank (PDB); accession numbers are 2QW7 (CcmL), 2RCF (OrfA or CsoS4A), and 3BN4 (CcmK1).

Supporting Online Material

www.sciencemag.org/cgi/content/full/319/5866/1083/DC1

Materials and Methods

SOM Text

Figs. S1 to S4

Tables S1 and S2

References

9 October 2007; accepted 3 January 2008

10.1126/science.1151458

Differential Regulation of Dynein and Kinesin Motor Proteins by Tau

Ram Dixit, Jennifer L. Ross,* Yale E. Goldman, Erika L. F. Holzbaur†

Dynein and kinesin motor proteins transport cellular cargoes toward opposite ends of microtubule tracks. In neurons, microtubules are abundantly decorated with microtubule-associated proteins (MAPs) such as tau. Motor proteins thus encounter MAPs frequently along their path. To determine the effects of tau on dynein and kinesin motility, we conducted single-molecule studies of motor proteins moving along tau-decorated microtubules. Dynein tended to reverse direction, whereas kinesin tended to detach at patches of bound tau. Kinesin was inhibited at about a tenth of the tau concentration that inhibited dynein, and the microtubule-binding domain of tau was sufficient to inhibit motor activity. The differential modulation of dynein and kinesin motility suggests that MAPs can spatially regulate the balance of microtubule-dependent axonal transport.

Active transport of cytoplasmic material along microtubules is critical for cell organization and function, and defects in this process are associated with dysfunction and disease (1). Much of the active transport in cells depends on the molecular motor proteins cytoplasmic dynein and kinesin-1, which transport

cargo toward the minus-end (toward the cell center) and plus-end of microtubules (toward the cell periphery), respectively. Dynein and kinesin have very different structures and translocation mechanisms (2). Kinesin has a compact motor domain and walks unidirectionally along single protofilaments with 8-nm steps (2). In contrast,

dynein has a larger, more complex motor domain and is capable of variable step sizes, lateral steps across the microtubule surface, and processive runs toward both the minus- and plus-end of the microtubule (3–5). Cytoplasmic dynein function in vivo also requires an accessory complex, dynein. This large, multiprotein complex is thought to facilitate dynein processivity (6) and may also regulate dynein activity (5). Within the cell, the balance between oppositely directed transport determines the steady-state distribution of organelles and biomolecules.

In the crowded cell environment, dynein and kinesin compete with nonmotile microtubule-associated proteins (MAPs) for binding to the microtubule surface. MAPs bound to microtubules might also block the path of motor proteins. Thus, MAPs can provide spatio-

Department of Physiology and Pennsylvania Muscle Institute, University of Pennsylvania, Philadelphia, PA 19104, USA.

*Present address: 302 Hasbrouck Laboratory, Department of Physics, University of Massachusetts at Amherst, Amherst, MA 01003, USA.

†To whom correspondence should be addressed. E-mail: holzbaur@mail.med.upenn.edu