



## Herpesviral Protein Networks and Their Interaction with the Human Proteome

Peter Uetz, *et al.*  
*Science* **311**, 239 (2006);  
DOI: 10.1126/science.11116804

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of February 21, 2009):**

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

<http://www.sciencemag.org/cgi/content/full/311/5758/239>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/cgi/content/full/1116804/DC1>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/cgi/content/full/311/5758/239#related-content>

This article **cites 13 articles**, 8 of which can be accessed for free:

<http://www.sciencemag.org/cgi/content/full/311/5758/239#otherarticles>

This article has been **cited by** 48 article(s) on the ISI Web of Science.

This article has been **cited by** 16 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/cgi/content/full/311/5758/239#otherarticles>

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

Cell Biology

[http://www.sciencemag.org/cgi/collection/cell\\_biol](http://www.sciencemag.org/cgi/collection/cell_biol)

Information about obtaining **reprints** of this article or about obtaining **permission to reproduce this article** in whole or in part can be found at:

<http://www.sciencemag.org/about/permissions.dtl>

# Herpesviral Protein Networks and Their Interaction with the Human Proteome

Peter Uetz,<sup>1\*</sup> Yu-An Dong,<sup>2\*</sup> Christine Zeretzke,<sup>3</sup> Christine Atzler,<sup>3</sup> Armin Baiker,<sup>3</sup> Bonnie Berger,<sup>2</sup> Seesandra V. Rajagopala,<sup>1</sup> Maria Roupelieva,<sup>3</sup> Dietlind Rose,<sup>3</sup> Even Fossum,<sup>3</sup> Jürgen Haas<sup>3†</sup>

The comprehensive yeast two-hybrid analysis of intraviral protein interactions in two members of the herpesvirus family, Kaposi sarcoma-associated herpesvirus (KSHV) and varicella-zoster virus (VZV), revealed 123 and 173 interactions, respectively. Viral protein interaction networks resemble single, highly coupled modules, whereas cellular networks are organized in separate functional submodules. Predicted and experimentally verified interactions between KSHV and human proteins were used to connect the viral interactome into a prototypical human interactome and to simulate infection. The analysis of the combined system showed that the viral network adopts cellular network features and that protein networks of herpesviruses and possibly other intracellular pathogens have distinguishing topologies.

Herpesviruses are widely spread throughout vertebrates and have large, double-stranded DNA genomes encoding between about 70 and 170 viral proteins, which is only one order of magnitude less than the number of proteins encoded by bacterial genomes. Although studies have revealed a large number of interactions between herpesviral and host proteins, relatively little is known about interactions among herpesviral proteins, particularly for those herpesviruses that replicate poorly in cell culture, e.g., Kaposi sarcoma-associated herpesvirus (KSHV) (1). We thus generated genomewide protein interaction maps for two human pathogens: KSHV, which is a member of the  $\gamma$ -herpesvirus subfamily associated with Kaposi sarcoma and B cell lymphomas (2), and varicella-zoster virus (VZV), which is a member of the  $\alpha$ -herpesvirus subfamily that causes chickenpox and shingles (3).

We cloned the open reading frames (ORFs) of both viruses (currently there are 89 ORFs identified in KSHV and 69 ORFs in VZV) by recombinatorial cloning and generated yeast two-hybrid (Y2H) bait and prey arrays. To circumvent the limitation of the Y2H system for transmembrane proteins, we cloned full-length proteins as well as extra- and intracellular domains separately. In KSHV, we tested more than 12,000 viral protein interactions involving both full-length proteins and protein fragments and identified 123 nonredundant interacting protein pairs (fig. S1 and table S1). To date,

only a small number of intraviral protein interactions have been reported for KSHV, of which 71% were captured by our screen (table S2). To further confirm the quality of our Y2H results and generate a set of high-confidence interactions, we tested all positive Y2H interactions in parallel by both  $\beta$ -galactosidase and coimmunoprecipitation (CoIP) assays (fig. S2). About 50% of the protein interactions could be confirmed by CoIP (table S1), and many of the remaining ones have orthologous interactions in other herpesviruses that could be confirmed (table S3). Although our array-based two-hybrid system is internally controlled, some of the interactions not confirmed by CoIP may be nonphysiological and, for example, caused by autoactivation. A comparison between protein interactions and expression profiles of KSHV-infected cells indicated that protein interactions occurred predominantly between proteins expressed at the same time point after infection (fig. S3). In VZV, we detected 173 nonredundant intraviral protein interactions out of ~10,000 tested bait-prey pairs (table S4).

Although cellular protein interaction networks exist for several model organisms, none of the studies on viral protein interactions published so far produced a large enough data set to constitute a protein interaction network. In many cellular protein interaction networks, most nodes have few neighbors, although some have many interaction partners (so-called hubs). The degree distributions of cellular protein networks were reported to follow a power-law decay, and they have been classified as scale-free (4). Like their cellular counterparts, KSHV and VZV had relatively many hubs, a key characteristic of scale-free networks (Fig. 1A; fig. S4). However, in contrast to known cellular protein interaction networks, in which nodes with a single interaction partner are most abundant, the viral networks had relatively few such “peripheral” nodes lying on

the “edge” of the network. In KSHV, for example, the degree distribution peaks at nodes with three neighbors. This unusual characteristic at low node degrees is one of the reasons that viral networks appear as single, highly coupled modules and presumably reflects their incompleteness as stand-alone networks. Whereas cellular protein networks have been shown or assumed to be scale-free, the degree distribution of viral networks does not present such a clear-cut picture (fig. S5). If viral networks are approximated by a power-law distribution, they have unusually small power coefficients distinctive from those of known cellular networks and defying current dynamic network evolution models (Fig. 1B and Table 1).

Another important characteristic of complex networks is the so-called small-world property (5). In a small-world network, the average distance between any two nodes is short (short characteristic path length, or the six degrees of separation phenomenon) and local neighborhoods are more densely connected (high clustering coefficient). Both viruses exhibited a short characteristic path length and a short network diameter (the maximum distance between any two nodes), which also suggests their coupling as single modules (Table 1). To assess the viral levels of local clustering, we generated random networks of the same size and degree distribution. The level of local clustering was low in KSHV and VZV, indeed, comparable to the level in equivalent random networks, and thus these viral networks cannot be classified as small-world. In contrast, most cellular protein interaction networks are unambiguously small-world, even after the effect on local clustering due to degree distribution and network size, which is substantial, is filtered out. For example, the clustering coefficient of the *Saccharomyces cerevisiae* protein interaction network is increased ~30-fold over simulated networks (Table 1). In *S. cerevisiae*, Maslov and Sneppen demonstrated that hubs tend to avoid each other while preferring low-connectivity nodes (6). As a result, the yeast network has well-separated modules, and errors in one module do not easily propagate to other modules. In the viral networks, there was no such declining degree correlation, and hubs did not tend to avoid each other, which offers additional evidence that these viral networks could be viewed as single, highly coupled modules (fig. S6). Because of these unusual topological features, viral networks should be more resistant to deliberate attacks than cellular networks, as both network size and characteristic path length remain more stable after the most highly connected nodes are removed (Fig. 1, C and D).

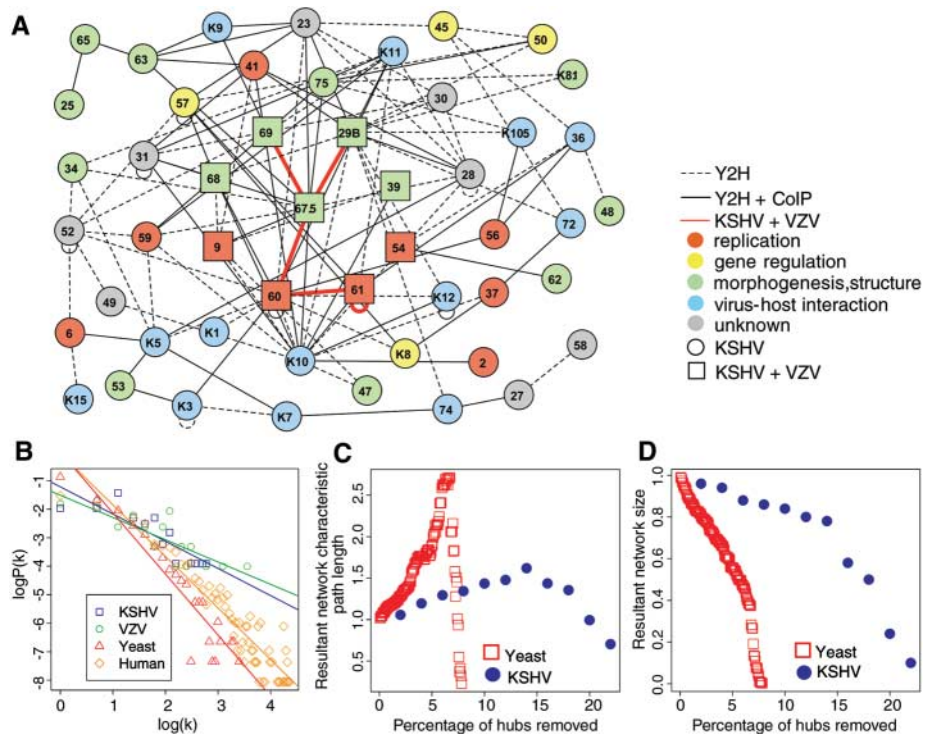
Whereas sequence and phylogenetic analyses identify a core set of genes conserved in all herpesviruses, the KSHV interactome allowed us to determine a core set of interactions conserved in all herpesviruses. Using the recip-

<sup>1</sup>Institut für Genetik, Forschungszentrum Karlsruhe, Postfach 3640, Karlsruhe, D-76021 Germany. <sup>2</sup>Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA. <sup>3</sup>Max-von-Pettenkofer Institut, Ludwig-Maximilians-Universität München, Pettenkoferstrasse 9a, 80336 München, Germany.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: haas@lmb.uni-muenchen.de

**Fig. 1.** Topology of the KSHV protein interaction network. **(A)** Protein interaction map of KSHV. KSHV proteins are indicated as nodes, and protein interactions either as hatched (found only by Y2H) or solid (confirmed by CoIP) edges. Orthologous proteins interactors in KSHV and VZV are marked by squares, and orthologous interactions detected in both viruses by red edges. KSHV ORFs were assigned to five functional classes depicted in different colors on the basis of GenBank annotations for the corresponding ORF or its orthologs. **(B)** Comparison of the approximated power-law degree distribution of two herpesviral (KSHV and VZV) networks and two cellular (yeast and human) networks. The yeast data set is derived from Schwikowski *et al.* (9), and the *Homo sapiens* (predicted) data set from Lehner and Fraser (12). For each network, node degrees  $k$  and their relative frequency (i.e., probability) are plotted on a bilogarithmic scale and fitted by linear regression. **(C and D)** Simulations of deliberate attack on KSHV and yeast networks by removing their most highly connected nodes (in decreasing order). After each node is removed, the new network characteristic path length (average distance between any two nodes) and size (number of nodes) of the remaining single largest connected component (SLCC) are computed and plotted as a multiple or fraction of the original parameters. KSHV exhibits much higher attack tolerance, because the increase in path length and the decrease in network size are considerably smaller.



**Table 1.** Comparison of network parameters of cellular and viral protein interaction networks. The table indicates key parameters of eight viral and cellular networks (all analyses were done using the SLCC of the respective network). Among 123 nonredundant KSHV protein interactions, 8 are self-interactions and the SLCC consists of 115 edges; among 173 nonredundant VZV protein interactions, 13 are self-interactions and one edge is isolated, leaving the SLCC with 159 edges. Only 1 of the 123 interactions in KSHV and 10 of the 173 in VZV were detected bidirectionally, and 1 interaction in KSHV and 13 in VZV were redundantly detected by distinct fragments of the same proteins. Interactions detected only in one direction are a common phenomenon in two-hybrid assays and most likely are due to steric hindrance of either bait or prey fusion proteins. The data sets were derived from the following sources: vaccinia virus, from McCraith *et al.* (13); *S. cerevisiae I*, from Schwikowski *et al.* (9); *S. cerevisiae II*, from the Database of Interacting Proteins (DIP; October 2004 release); *H. sapiens I*, from Rual *et al.* (8); *H. sapiens II*, from Stelzl *et al.* (7); and *H. sapiens* (predicted), from Lehner and Fraser (12). The table includes the number of nodes and edges in the SLCC; the average node degree (i.e.,

number of neighbors); the power coefficient  $\gamma$  and its  $P$  value (the slope and its significance under linear regression) as fitted by a power-law degree distribution ("scale-free" property); the characteristic path length and diameter, as well as the clustering coefficient and its fold enrichment over comparable random networks ("small-world" property). For each real network, a corresponding ER (Erdős-Rényi) randomization has the same number of nodes and edges, and an ES randomization, generated through an edge-swapping algorithm, also has the same degree distribution. The fold enrichments shown are over the theoretical clustering coefficient under the ER model and the median clustering coefficient of 1000 ES randomizations, respectively (see supporting online material). Note that the network parameters of the two yeast networks are surprisingly stable, although a large number of interactions have been included additionally into the DIP database in comparison to the initial data set generated by Schwikowski and colleagues (9). The recently reported *H. sapiens* networks have rather low levels of local clustering, which was discussed as being caused by their incompleteness ( $\beta$ ). NA, not applicable owing to the low number of edges.

Network parameters	KSHV	VZV	Vaccinia virus	<i>S. cerevisiae I</i>	<i>S. cerevisiae II</i>	<i>H. sapiens I</i>	<i>H. sapiens II</i>	<i>H. sapiens</i> (predicted)
Nodes	50	55	7	1,548	2,397	1,307	1,598	3,169
Edges	115	159	6	2,358	6,101	2,483	3,072	10,636
Average degree	4.60	5.78	1.71	3.05	5.09	3.80	3.84	6.71
Power coefficient	0.95	0.78	NA	2.14	2.01	1.54	1.66	1.81
$P$ value	$1.2 \times 10^{-4}$	$1.1 \times 10^{-4}$	NA	$3.6 \times 10^{-11}$	$7.7 \times 10^{-23}$	$1.2 \times 10^{-20}$	$5.8 \times 10^{-25}$	$1.4 \times 10^{-30}$
Characteristic path length	2.84	2.34	NA	7.28	5.10	4.36	4.85	6.40
Diameter	7	5	NA	23	13	12	13	20
Clustering coefficient	0.146	0.393	NA	0.213	0.296	0.060	0.012	0.186
Enrichment over ER	1.55	3.67	NA	108.1	139.5	20.6	4.8	87.9
Enrichment over ES	0.76	1.01	NA	29.2	29.5	0.92	0.28	11.7

rocal best BLAST hit approach, we determined the orthology relationships among the viruses and predicted 114 orthologous protein interactions in herpes simplex virus 1 (HSV-1),

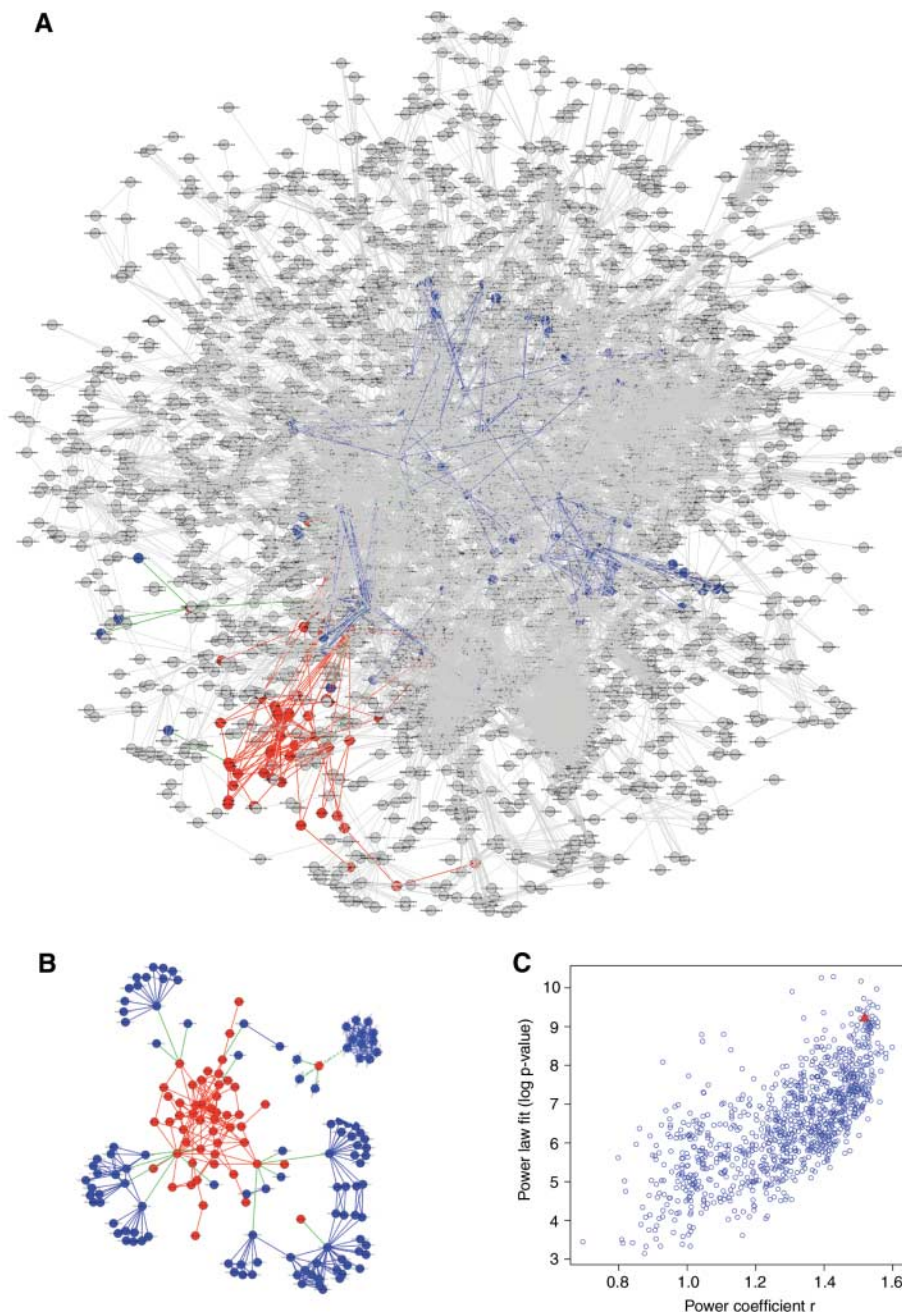
VZV, murine cytomegalovirus (mCMV), and Epstein-Barr virus (EBV) (table S1). Intriguingly, most of the predicted interactions (72 out of 112 tested, or 64%; table S3) could be

confirmed by CoIP despite the rather low level of sequence similarity between KSHV proteins and their orthologs (only in the 20 to 40% range). Although in general KSHV proteins

with viral interaction partners are not more conserved than are those without (fig. S7a), for KSHV proteins with interaction partners there was a significant correlation between the number of protein interactions and homology to the respective EBV ortholog (fig. S7b). When we compared the protein networks of KSHV and VZV, we found that only 9 of the 50 protein interactors in KSHV have orthologous interactors in VZV (Fig. 1A; table S1). Analysis of the KSHV network suggests that 19 interactions should exist between orthologous viral proteins in VZV, of which we could detect 5 in our Y2H screen (26%). This low result is not surprising, because the number of predicted interactions in HSV-1, mCMV, and EBV that were confirmed by Y2H was also considerably lower than the number confirmed by CoIP, indicating an inherent technical limitation of the Y2H system.

The network analyses of the KSHV and VZV interactomes revealed unique features of viral systems that also manifested themselves on the local level (fig. S8). Because we hypothesized that many of these features could be attributed to missing virus-host interactions, we modeled the interplay between viral and human protein networks. To date, only rather small subnets of the human interactome have been reported (7, 8). Unfortunately, only a few published human proteins targeted by KSHV lie within these reported human subnets and only a few virus-host interactions could be predicted on the basis of homology between KSHV and human proteins. However, using a prototypical human protein interaction network [derived from high-confidence interactions in *S. cerevisiae*, *Caenorhabditis elegans*, and *Drosophila melanogaster* (9–11)] that is considerably larger than the experimental subnets, we were able to efficiently connect the viral and the human networks by predicting interactions between KSHV and human proteins if both proteins have known interacting orthologs in either *S. cerevisiae*, *C. elegans*, or *D. melanogaster* (12). By this approach, we found 20 predicted interactions between 8 KSHV and 20 human proteins that are connected within the human network. Nineteen of these 20 virus-host interactions were tested by CoIP and an unexpectedly large percentage (13 out of 19, or 68%) could be confirmed (table S9). Although published viral-host interactions tend to involve genes or interactions specific to human or higher eukaryotes (because most human targets have no orthologs or orthologous interactions in the three lower-eukaryotic model organisms), our predicted viral-host interactions involve genes and interactions conserved from yeast to human and hence might reflect more general host-interacting mechanisms.

Using the predicted KSHV-human interactions, we were able to merge the two interactomes with each other (Fig. 2, A and B). The topology of the KSHV network changed



**Fig. 2.** Interplay between the KSHV and a predicted high-confidence human network. **(A)** Global view of the interplay between the KSHV and a predicted high-confidence human interaction network consisting of 10,636 edges among 3169 nodes. Viral proteins are depicted as red nodes, cellular interacting proteins (level 1 and 2) as blue nodes, and cellular proteins (level >2) as gray nodes. Interactions between viral proteins are depicted as red edges, those between viral and cellular proteins as green edges, and those between cellular level 1 and 2 proteins as blue edges. **(B)** Local view of the combined KSHV and human interaction networks (level  $\leq 2$ ). **(C)** The combined virus-host network has a bigger and statistically more significant power coefficient and thus adopts scale-free features. The combined virus-host network was compared to 1000 random networks, which were generated by rewiring fixed virus interactors to swapped cellular proteins with the same degree as the actual target. The power coefficient and the power-law fit of the predicted (red triangle) and 1000 random (blue circles) KSHV-human networks are indicated on the x axis and the y axis, respectively. Among 1000 random networks, 65 have a higher power coefficient, 19 a more significant power-law fit, and only 8 a higher power coefficient that is more significant (empirical  $P$  value  $<0.01$ ).

completely from a highly coupled module to a more typical scale-free network of interacting submodules once it was connected to its host. To rigorously assess the impact of the two systems upon each other, we performed a combined viral-host network analysis. Starting with the KSHV network (level 0), we first added in their direct human targets (level 1), then added in those human targets' own cellular interactions partners (level 2), and so on, until the viral network was completely assimilated into the host network. To evaluate the topology of the combined virus-host network, we reasoned that a correctly combined system should be different from randomly combined networks. To generate an ensemble of equivalent random viral-host networks, we adopted the following simulation strategy: the identity and degree of KSHV interactors are fixed, while their human targets are randomly chosen from the host network so that each random target has the same degree as the predicted target that it replaces. Because degree distribution reflects global topology, with the KSHV and human protein networks differing substantially in this respect, it offers an ideal measure to assay the coupled system. Indeed, at level 2, the predicted viral-host network not only exhibited a better power-law fit, but the power coefficient was also bigger, both crucial features of the human network (empirical  $P$  value < 0.01 in 1000 simulations) (Fig. 2C).

Thus, at the level carrying most biological relevance (KSHV's human targets and their direct interaction partners) and subjected to minimal noise (level 3 already includes a sizable fraction of the human network and many of the interactions are conceivably no longer relevant to the viral-host context), the combined virus-host network assimilated human network properties to a large extent (fig. S9). When we predicted VZV-human interactions and modeled the combined system, we saw very similar results, demonstrating the general utility of our approach (fig. S10; table S10).

Although we have shown that virus and host interactomes possess distinct network topologies, their interplay may lead to emergent new system properties that represent specific features of the viral pathogenesis. Obviously, numerous biological hypotheses resulting from our study remain to be investigated in detail. The availability of protein interaction networks in other herpesviruses and large-scale virus-host interaction data in the near future will boost our knowledge of the function of many still poorly characterized viral proteins and the phylogeny of herpesviruses. It may eventually lead to a considerably improved understanding of viral pathogenesis and evoke new therapeutic strategies.

#### References and Notes

1. P. Uetz, S. V. Rajagopala, Y. Dong, J. Haas, *Genome Res.* **14**, 2029 (2004).

2. Y. Chang *et al.*, *Science* **266**, 1865 (1994).
3. A. J. Davison, J. E. Scott, *J. Gen. Virol.* **67**, 1759 (1986).
4. A. L. Barabasi, R. Albert, *Science* **286**, 509 (1999).
5. D. J. Watts, S. H. Strogatz, *Nature* **393**, 440 (1998).
6. S. Maslov, K. Sneppen, *Science* **296**, 910 (2002).
7. U. Stelzl *et al.*, *Cell* **122**, 957 (2005).
8. J. F. Rual *et al.*, *Nature* **437**, 1173 (2005).
9. B. Schwikowski *et al.*, *Nat. Biotechnol.* **18**, 1257 (2000).
10. L. Giot *et al.*, *Science* **302**, 1727 (2003).
11. S. Li *et al.*, *Science* **303**, 540 (2004).
12. B. Lehner, A. G. Fraser, *Genome Biol.* **5**, R63 (2004).
13. S. McCraith, T. Holtzmann, B. Moss, S. Fields, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4879 (2000).
14. This work was supported by grants provided by Bayerisches Staatsministerium für Wissenschaft, Forschung und Kunst (BayGene), Deutsche Forschungsgemeinschaft (SFB 576), Sander-Stiftung (1999.096.1), Helmholtz Gemeinschaft (P.U. and Y.-A.D.), and LMU München (FoeFoLe). We thank A. Arvin, Y. Chang, D. Ganem, S.-J. Gao, W. Hammerschmidt, Y. Kawaguchi, U. Koszinowski, P. Moore, B. Moss, F. Neipel, P. Pitha-Rowe, T. Schulz, B. Sodeik, C. Wood, and Y. Yuan for plasmids, cosmids, bacterial artificial chromosomes, and viruses. This manuscript is dedicated to C.A..

#### Supporting Online Material

www.sciencemag.org/cgi/content/full/1116804/DC1  
Materials and Methods  
Figs. S1 to S10  
Tables S1 to S10  
References and Notes

30 June 2005; accepted 21 November 2005

Published online 8 December 2005;

10.1126/science.1116804

Include this information when citing this paper.

## Magnetosomes Are Cell Membrane Invaginations Organized by the Actin-Like Protein MamK

Arash Komeili,<sup>1\*</sup> †‡ Zhuo Li,<sup>2\*</sup> Dianne K. Newman,<sup>1,3</sup> Grant J. Jensen<sup>2</sup>

Magnetosomes are membranous bacterial organelles sharing many features of eukaryotic organelles. Using electron cryotomography, we found that magnetosomes are invaginations of the cell membrane flanked by a network of cytoskeletal filaments. The filaments appeared to be composed of MamK, a homolog of the bacterial actin-like protein MreB, which formed filaments *in vivo*. In a *mamK* deletion strain, the magnetosome-associated cytoskeleton was absent and individual magnetosomes were no longer organized into chains. Thus, it seems that prokaryotes can use cytoskeletal filaments to position organelles within the cell.

**P**rokaryotes are highly organized cells with many ultrastructural similarities to eukaryotic cells (1). In addition to a dynamic cytoskeleton composed of homologs of actin, tubulin, and intermediate filaments (2), some prokaryotes also possess intracellular membranous organelles (3–6). The magnetosome chains of magnetotactic bacteria are among the best studied examples of membranous prokaryotic organelles. Magnetosome chains contain 15 to 20 ~50-nm magnetite crystals that together act like the needle of a compass to orient magnetotactic bacteria in

geomagnetic fields, thereby simplifying their search for their preferred microaerophilic environments (3). The distinct properties of magnetosomal magnetite crystals have drawn attention to their potential use in biotechnology, bioremediation, and geobiology (3). Furthermore, the cell biological characteristics of magnetosomes make them an ideal model system for the study of organelle biology in prokaryotes. Each magnetite crystal within a magnetosome is surrounded by a lipid bilayer, and specific soluble and transmembrane proteins are sorted to the magnetosome membrane (7, 8).

Electron cryotomography (ECT) is an emerging technology that can reveal the three-dimensional (3D) ultrastructure of small, intact prokaryotic cells in a near-native, “frozen-hydrated” state (9, 10). Here, we used ECT to define the precise physical characteristics of the magnetosome chain and to identify the molecular factors controlling its subcellular positioning and organization.

To image magnetotactic bacteria by ECT, exponential-phase cultures of *Magnetospirillum magneticum* sp. AMB-1 were plunge-frozen and imaged in a transmission electron cryomicroscope. Dual-axis image tilt series were recorded (11), from which 3D reconstructions of individual cells were calculated (Movie S1). A slice through the middle of one typical reconstruction (Fig. 1A and Movie S2) clearly showed the inner and outer membranes and the peptidoglycan layer. The cytoplasm was full of ribosomes, recognized as the numerous, dense, ~25-nm particles. Large, approximately

<sup>1</sup>Division of Geological and Planetary Sciences, <sup>2</sup>Division of Biology, <sup>3</sup>Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA 91125, USA.

\*These authors contributed equally to this work.

†Present address: Department of Plant and Microbial Biology, University of California, Berkeley, Mail Code 3102, Berkeley, CA 94720, USA.

‡To whom correspondence should be addressed. E-mail: komeili@nature.berkeley.edu