UNIVERSITY OF CALIFORNIA
Los Angeles

Using Advances in Electron Microscopy to Study Microbial Interactions

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctorate of Philosophy
in Microbiology, Immunology, and Molecular Genetics

by

Nicole Lynn Poweleit

2016

ABSTRACT OF THE DISSERTATION

Using Advances in Electron Microscopy to Study Microbial Interactions

by

Nicole Lynn Poweleit

Doctorate of Philosophy in Microbiology, Immunology and Molecular Genetics

University of California, Los Angeles, 2016

Professor Robert P. Gunsalus, Co-Chair

Professor Hong Zhou, Co-Chair

Microbes interact with their surroundings through a variety of mechanisms, ranging from extracellular machineries like flagella, pili, and surface layer proteins to protein complexes embedded in the cell membrane. In this dissertation, we used a variety of techniques to characterize these mechanisms of interaction, with a focus on exploiting recent advances in electron microscopy to better understand these systems. The dissertation opens with an overview of the history and development of electron microscopy (EM) with a focus on its historical contributions to microbial interactions in the literature, its recent technical developments in electron microscopy, and helical reconstruction of protein filaments by cryo electron microscopy (cryoEM).

EM has developed into an indispensable tool for studying all aspects of microbial interactions from the gross cellular level to protein structures at atomic resolution. The following chapter of the thesis focuses on characterizing the interaction between two bacteria in the oral microbiome by utilizing scanning electron microscopy (SEM) in conjunction with light microscopy and genetic experiments. The relationship between the newly described obligately parasitic bacterium TM7x (Candidatus Saccharimonas formerly Candidate Phylum TM7) and its host *Actinomyces odontolyticus* species XH001 is described. Evidence from qPCR experiments, light microscopy, and SEM show that TM7x causes stress to its host XH001 and that this stress is additive with other stress factors in the environment. This demonstrates that the relationship between the two is actively harmful to XH001 whereas it was previously unclearly if XH001 host was impacted by the growth of TM7x. Light microscopy and SEM were also used to demonstrate that TM7x divides by budding and that no flagella or pili are visible on either cell. This data suggests that TM7x adheres to the host cell in a directional manner using cell surface or membrane proteins.

In chapter three, the cell envelope of the bacterium *Syntrophamonus wolfei* was characterized using biochemical assays and transmission electron microscopy (TEM). S. wolfei is a syntroph which must live in symbiotic relationships or consortium with other prokaryotes that consume the syntrophic metabolic products. A method was devised to separate the membrane portion of the cells from the soluble cell contents to produce cell ghosts. These cell ghosts were analyzed via mass spectroscopy to identify the three major protein components. One of these proteins, Swol_0141 has domains which identify it as a potential surface (S) layer protein domains. Transmission electron microscopy (TEM) of the cell ghosts revealed paracrystalline array of P4 symmetry, consistent with the production of a

proteinaceous S-layer. CryoEM of the cells shows the protein arrangement of the cell envelope.

Finally, I present an atomic structure of the archaeal flagellum from *Methanospirillum hungatei* strain JF-1, obtained with cryo electron microscopy, helical reconstruction, and *de novo* model building. The archaeal flagellum is a nanomachine which rotates to drive cell motility and adheres to other cells and surfaces. The thin filament of flagellum is only 10 nm in diameter, but can extend to be several times longer than the cell length. This structure is the first complete atomic resolution model of an archaeal flagellin, and it describes the intermolecular interactions which allow for the stability of the flagellar filament under rotational stress. The cryoEM structure of the native protein also revealed eight sites of post-translational modification. To conclude, a comparison with the bacterial flagella and type IV pili shows that the archaeal flagellum is a structurally distinct cell motility and adhesion apparatus.

In summary, these thesis projects demonstrate the breadth of utility electron microscopy has for studying microbes and their environmental interactions. These include the cell to cell interactions of an oral parasitic bacteria and host bacteria, the definitions of an undescribed single cell envelope, and the atomic resolution protein structure of a flagellar nanomachine. The depth of information which can be explored using electron microscopy to solve complex microbial cell and protein structures continues to expand.

The dissertation of Nicole Lynn Poweleit is approved.

Jeffery F. Miller,

Joseph A. Loo,

Robert P. Gunsalus, Committee Co-Chair,

Hong Zhou, Committee Co-Chair

University of California – Los Angeles

2016

This work is dedicated to my husband Ivan, my grandmother Margaret Poweleit and my grandfather Nelson Jenquin.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

ACKNOWLEDGEMENTS

Chapter 2 is a reprint of “Phenotypic and Physiological Characterization of the Epibiotic Interaction Between TM7x and Its Basibiont Actinomyces” reprinted here with permissions. Chapter 3 is currently in preparation for publication. Chapter 4 is the uncorrected proofs for an article which has been accepted for publication by Nature Microbiology under the title: “CryoEM structure of the *Methanospirillum hungatei* archaellum reveals structural features distinct from the bacterial flagellum and type IV pili” and will appear in the print shortly after this thesis is completed. Appendix 1 is a reprint of the previously published work “Complete genome sequence of *Methanospirillum hungatei* type strain JF1” and is reprinted here with permissions.

VITA
Nicole Poweleit

Education

University of California – Los Angeles, Microbiology Expected Dec 2016
Ph.D. in Microbiology, Immunology, and Molecular Genetics
Los Angeles, California
Dissertation Title: *Using Advances in Electron Microscopy to Study Microbial Interactions*

University of Wisconsin – Milwaukee May 2010
B.S. double major in Microbiology and Chemistry Milwaukee, Wisconsin

Honors and Awards

- Warsaw Grant Aug 2015 – July 2016
- NIH Biotechnology Training Grant Oct 2011 – Oct 2013
- Dean's List Fall 2006, Spring 2010

Professional Experience

Graduate Student Researcher
Graduate Student Researcher July 2010 – Nov 2016
Thesis Advisors: Dr. Hong Zhou and Dr. Robert Gunsalus
- Cryo electron microscopy of helical filaments
- Sample preparation and imaging using TEM and SEM
- Cloning in *E. coli* and protein purification

Undergraduate Student Researcher
Lab of Dr. Gyaneshwar Prasad, Microbiology Sep 2008 – Dec 2009
Project: Interactions between rhizobia microbes and plants
University of Wisconsin - Milwaukee

Lab of Dr. Andrew Pacheco, Chemistry Jan 2010 – May 2010
Project: Purification and crystallization of *Shewanella oneidensis* heme-related proteins
University of Wisconsin - Milwaukee

Teaching Experience

Teaching assistant, MIMG 105: Biological Microscopy Fall 2016
Mentored three undergraduates and one high school student 2011 – 2015
Teaching assistant, MIMG 101: Introduction to Microbiology Spring 2012, Winter 2013

Professional Presentations

The structure and post-translational modification of an archaeal flagellum. *Southern California CryoEM Symposium*, San Diego, CA, 2016. **Platform**

The Structure of the *Methanospirillum hungatei* flagellum as determined by Cryo Electron Microscopy. *Annual Meeting of the Biophysical Society,* Los Angeles, CA, 2016. **Platform**

Biochemical and Structural Characterization of an Archaeal Flagellum. *Annual Meeting of the Biophysical Society*, San Francisco, CA, 2014. **Poster**

Isolation and Biochemical Characterization of the *Methanospirillum hungatei* strain JF-1 Flagellum. Annual Meeting of the American Society for Microbiology, Denver, CO, 2013. **Poster**

Publications

Poweleit, N., Ge, P., Nguyen, H.H., Ogorzalek Loo, R., Gunsalus, R.P., and Zhou, Z.H. CryoEM structure of the *Methanospirillum hungatei* archaellum reveals structural features distinct from the bacterial flagellum and type IV pili. *Nature Microbiology*. Accepted.

Bor, B., **Poweleit**, **N.**, Bois, J., Cen. L., Bedree, J., Zhou, Z.H., Gunsalus, R., Lux, R., McLean, J., He, X. and Shi, W. Phenotypic and Physiological Characterization of the Epibiotic Interaction Between TM7x and Its Basibiont Actinomyces. *Microbial Ecology* 71(1): 243-255. Jan 6, 2016.

Gunsalus, R. P., Cook, L. E., Crable, B., Rohlin, L., McDonald, E., Mouttaki, H., Sieber, R. S., **Poweleit, N.**, … McInerney, M. J. Complete genome sequence of *Methanospirillum hungatei* type strain JF1. *Standards in Genomic Sciences*, *11*, 2. Jan 2016.

Gehlot, H. S., Tak, N., Kaushik, M., Mitra, S., Chen, W.-M., **Poweleit, N.**, … Gyaneshwar, P. An invasive *Mimosa* in India does not adopt the symbionts of its native relatives. *Annals of Botany*, *112*(1), 179–196. May 26, 2013.

References

Dr. Hong Zhou, Professor, Department of Microbiology, Immunology, and Molecular Genetics, University of California – Los Angeles, Los Angeles, CA, 90024.
Tel: (310) 983-1033, **Email:** Hong.Zhou@ucla.edu

Dr. Robert Gunsalus, Professor, Department of Microbiology, Immunology, and Molecular Genetics, University of California – Los Angeles, Los Angeles, CA, 90024.
Tel: (310) 206-8201, **Email:** RobG@microbio.ucla.edu

Dr. Peter Bradly, Professor, Department of Microbiology, Immunology, and Molecular Genetics, University of California – Los Angeles, Los Angeles, CA, 90024.
Tel: 310-825-8386, **Email:** pbradley@ucla.edu

CHAPTER 1: Introduction

Overview

Traditionally, the resolution of light microscopy has been limited by the wavelength of the light being used to form the image. Most conventional light microscopes are unable to achieve better than 200nm resolution or about half the wavelength of visible light, although in the past ten years, advances in super-resolution light microscopy have changed this paradigm. Electron microscopes were developed in the 1930's and can achieve superior resolution, due to the smaller wavelength of the electron relative to the photons from visible light. Over the past ninety years, electron microscopy has developed into an essential tool in microbiology. This chapter gives an overview of a variety of sample preparation and imaging techniques used in electron microscopy with a focus on those techniques which are most important for studying microbial interactions.

Scanning electron microscopy

Scanning electron microscopes (SEM) focus an electron beam to a small point and raster scan the focused beam across a sample (Fig. 1.1a)[1]. The electron beam in an SEM has a low accelerating voltage and does not penetrate deeply into the sample, only mapping the surface. Either secondary or back scatter electrons are emitted from the atoms at the sample surface and detected to produce an image of the sample (Fig. 1.1b).[2]

Typically, samples are fixed before being imaged by SEM as samples must be imaged in vacuum. Although, many modern instruments offer the option to switch into

an Environmental SEM mode (ESEM) and image wet samples at micro-pressure which obviates the need for fixation. However, imaging at micro-pressure often results in poorer data quality as the back scattered electrons leaving the sample will interact with a thin layer of water vapor surrounding the sample. This interaction dampens the signal detected as fewer back scattered electrons reach the detector resulting in a decrease in signal and increase in noise and therefor, a lower signal to noise ratio[3]. The exact sample fixation procedure will differ somewhat from sample to sample, but for most biological samples, the following is the basic method followed: 1) fixation with a primary fixative such as glutaraldehyde or formaldehyde, 2) optionally, fix with a second fixative such as osmium tetroxide, 3) gradual dehydration into a volatile solution such as ethanol or propanol, 4) critical point drying to remove the volatile solution, 5) coating with a heavy metal to improve contrast. The sample preparation process can result in fixation artifacts for biological samples. However, SEM is still widely used in microbiology as it is effective at giving information about the gross organization of microbes and is not limited by sample thickness.

SEM has been used to study microbial interactions in a variety of ways such as in the study of biofilms and the formation of microbial consortia. *Myxoccocus xanthus* is a bacterium which exhibits multicellular behavior. *M. xanthus* cells coordinate to form multicellular fruiting bodies and spores when the cells become stressed, and this behavior has been imaged extensively by SEM.[4-6] Multispecies biofilms such as those in the human microbiome at wound sites[7] and in the oral microbiome[8], and bacterial biofilms which are useful for bioremediation[9], are also frequently imaged by SEM. In both cases, the number of prokaryotic cells involved in the interaction results in a

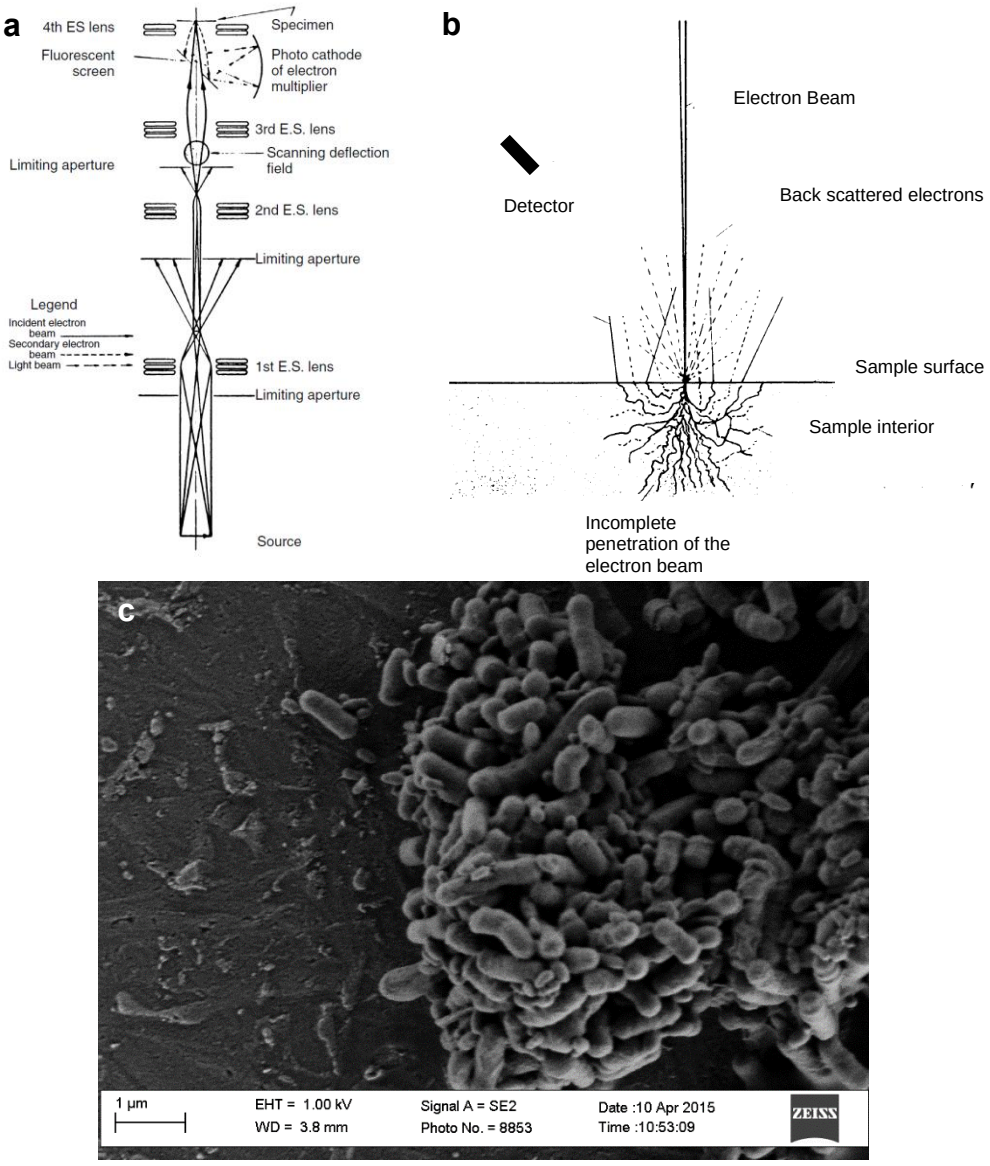


Figure 1.1 | Overview of scanning electron microscopes. a, the arrangement of lenses in the first SEM adapted from Zworykin *et al.,* 1942a **b,** How the electron beam in an SEM interacts with the sample. The relatively weak beam does not penetrate deeply into the sample and the back scattered electrons are collected to form an image of the sample surface adapted from a sketch by von Ardenne, 1940 which appeared in McMullan 1995. **c,** An example micrograph taken with SEM. This is micrograph is of the bacteria XH001 described in chapter 2.

sample which is too thick to be effectively imaged by transmission electron microscopy (TEM). However, a recent development in SEM sample preparation technology, known as focused ion beam (FIB) milling[10], allows for some of these thick samples to be cut down in an SEM by a focused ion beam and then imaged in a TEM. An example of this technique is also present in the *Myxoccocus xanthus* literature[11].

For the work involving dental microbes in chapter two, we chose to use SEM because very little was known about the samples being studied. The two microbes being studied represent a unique type of epibiotic relationship and one of the microbes, TM7x, had been successfully cultured for the first time only a few months previous. As such, the gross morphological information provided by SEM complemented other lines of inquiry and helped characterize this novel microbe and its interaction with the host, *Actinomyces odontolyticu*s species XH001 (Fig. 1.1c). SEM is a useful tool for characterizing novel microbes and microbial interactions, particularly when samples are too thick to be imaged by transmission electron microscopy.

Transmission electron microscopy

The optical path of a transmission electron microscope (TEM) is similar to that of a compound light microscope. An electron source is located at the top of the instrument and emits electrons into a column kept under vacuum. Early electron microscopes used tungsten filaments, like those used in incandescent bulbs, as an electron source. These filaments lack the coherence required for high resolution imaging, and the best electron microscopes now use field emission guns[12]. The electrons are then focused into a coherent beam using a series of electromagnetic lenses and apertures (Fig. 1.2a)[13].

Samples are inserted near the focal point of the beam and moved to an under focused position to provide contrast. Electrons which pass through the sample are recorded by a camera or detector. The original detectors used were film based cameras, in the 1990's charged couple device (CCD) cameras became the standard digital cameras used for data collection[14], and in the past five years, a new wave of direct electron detectors has greatly enhanced the imaging capabilities of TEMs (direct electron detectors are discussed in detail later in this chapter).

In traditional TEM imaging, the sample is prepared by drying or fixing and staining with a heavy atom before insertion into the column under vacuum. Thin samples are deposited onto an electron microscopy grid and negatively or positively stained before being dehydrated. Thick samples are fixed *en bloc* and sectioned to create a thin enough sample that the electron beam may pass completely through the sample. These sample fixation procedures can cause artifacts which lead to the development of different sample preservation procedures to preserve proteins in near natural environments.

Transmission electron microscopes are the workhorses of any electron microscopy lab and as such, they have been used extensively to study microbial interactions. For example, bacterial surface (S)-layer proteins have been studied extensively using traditional TEM sample preparation techniques[15-18]. S-layer proteins self-assemble to from paracrystalline sheets which completely cover the cell surface. These proteins form a symmetrical mesh which is easily observed by TEM imaging. As the outermost layer of the cell envelope, S-layer proteins are important for prokaryotes to interact with their environment and are important in biofilm formation and

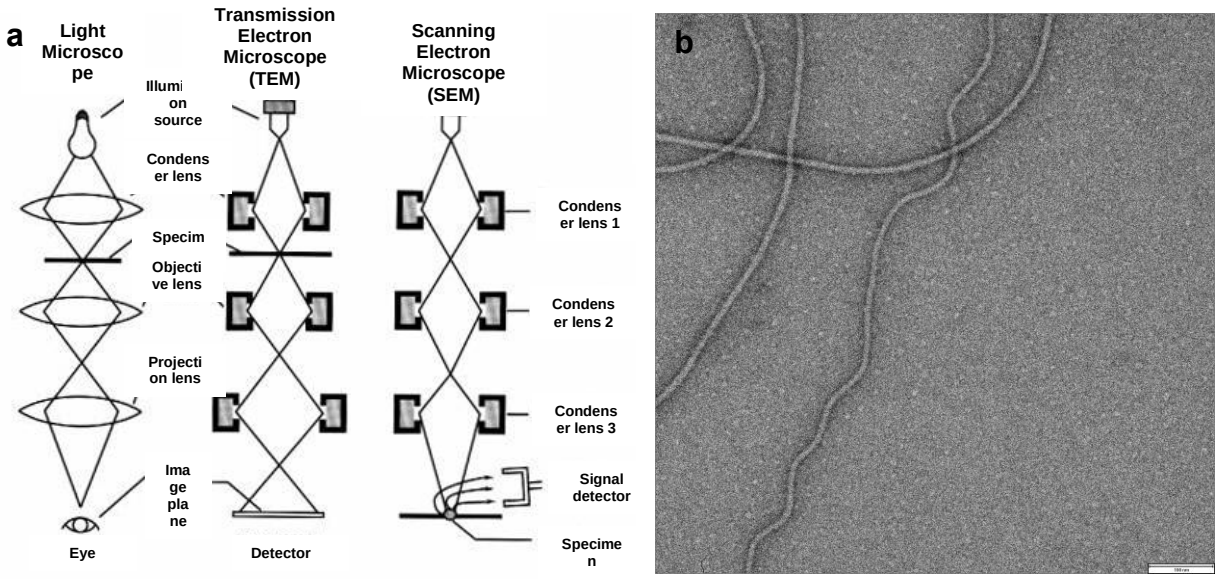


Figure 1.2 | An overview of transmission electron microscopes. a, Comparison between the optical paths of a light microscope, transmission electron microscope (TEM), and a scanning electron microscope (SEM) Figure adapted from the Central Microscopy Research Facility. **b,** An example of a TEM micrograph with a negatively stained sample, scale bar is 100nm. The sample is the *Methanospirillum hungatei* flagellum described in detail in chapter 4.

pathogenesis[19]. Traditional TEM sample preparation has also been used to study bacterial pili such as the *Geobacter sulfurreducens* type IV pili which allows the bacterium to break down inorganic material extracellularly and harvest the electrons for energy and was first visualized via traditional TEM[20]. Several other types of bacterial pili have been described via traditional TEM such as the conjugation pili[21,22] and type IV pili used in adhesion[23] and motility[24,25].

In this thesis, every project described used a TEM at least once, even if the data did not make it into the final manuscript. All of the samples used had an initial quality assay performed by TEM negative staining (Fig. 1.2b). Negative staining is by far the simplest and least time consuming of any electron microscopy sample preparation technique and as such, it is used routinely to assess the quality of a new samples. For whole cells and purified proteins, negative staining can determine if the samples are intact and in good condition and what their approximate concentration is in solution. Negatively stained samples also have higher contrast relative to cryoEM sample preparations which can help to distinguish certain features, although the staining process does result in a loss of some high-resolution information and can introduce artifacts.

For example, both chapter three where the *S. wolfei* S-layer proteins are characterized and chapter four where the *M. hungatei* flagella structure is characterized began with visualization of whole cells by negative stain TEM. As a syntrophic bacterium and methanogenic archaea, these two organisms have been isolated together and have a symbiotic relationship where *M. hungatei* cells consume the metabolic waste products of *S. wolfei* cells, thereby allowing both organisms to grow

effectively. The initial goal of this project was to use microscopy to characterize this interaction. However, after visualizing both cells by TEM and reviewing the literature, it became apparent that not enough was known about the extracellular assemblies, such as the archaeal flagella and bacterial S-layer proteins, to characterize this interaction.

Cryo electron microscopy

To overcome artifacts introduced by fixation, a new form of sample preparation was developed, cryo electron microscopy (cryoEM)[26]. Samples are flash frozen in liquid nitrogen cooled liquid ethane or propane. By rapidly freezing at very low temperatures, the water surrounding the sample does not have time to form ice crystals and is instead frozen in a liquid water like state called vitreous ice[27]. Vitreous ice does not diffract the electron beam as ice crystals do, and allows for samples to be imaged using a TEM microscope under near-native protein conditions (Fig. 1.3a)[28].

Unfortunately, cryoEM has a major drawback. Biological samples frozen in vitreous ice without any staining have very little contrast and can be difficult to see. Typically, the only contrast provided is phase contrast produced by under focusing the sample. Imaging samples at an under focused position introduces a contrast transfer function which must be corrected for to achieve high resolution protein reconstructions. Additionally, the high energy electron beam damages samples frozen in vitreous ice more rapidly than fixed samples and results in more pronounced beam induced drifting[29]. As the electron beam impacts the sample, some volatile atoms are lost. Water molecules are unstable enough to have their bonds broken by the electron beam. As water molecules are stripped off, the water matrix distorts and causes

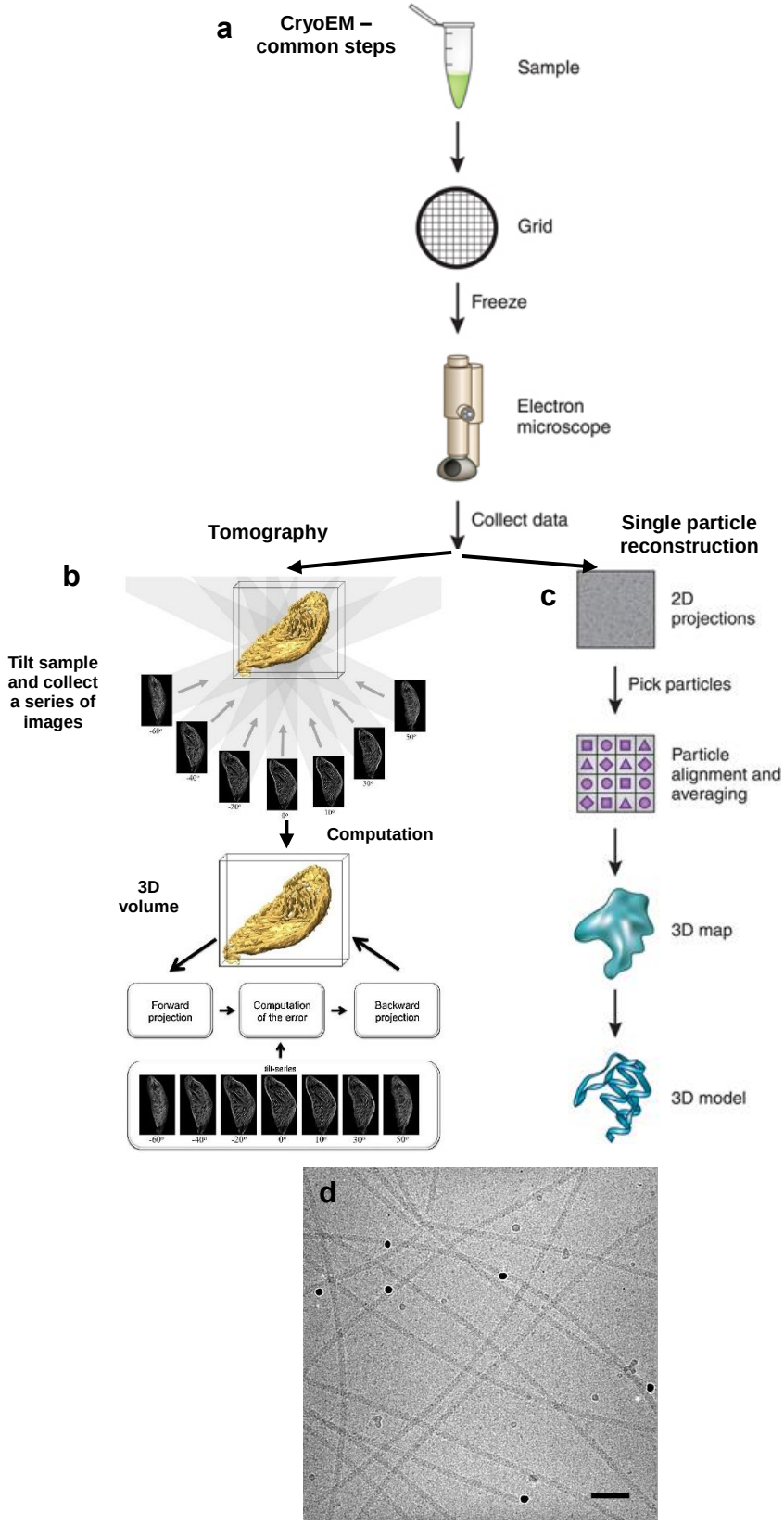


Figure 1.3 | An overview of the workflow for cryoEM data collection and processing. a, CryoEM sample preparation steps which are the same for tomography and single particle reconstruction, adapted from Doerr 2016 **b,** CryoEM tomography workflow from freezing to final result, adapted from Fernandez 2012 **c,** CryoEM single particle reconstruction work flow from sample freezing to *de novo* atomic model, adapted from Doerr 2016. **d,** An example cryoEM micrograph, scale bar is 50nm. The sample is the *Methanospirillum hungatei* flagellum described in chapter 4.

samples embed in the ice to move resulting in sample drift[30-32]. Eventually, water around the sample will be completely lost and the sample will be directly exposed to the beam resulting in boiling and other damage to the sample itself. There are two major imaging techniques used for cryoEM, which moderate the drawbacks of cryoEM and gain useful biological information about the sample in different ways, tomography and single particle reconstructions.

In tomography imaging, a single specimen of interest, such as a whole prokaryotic cell[33], very thin eukaryotic cell such as neurons[34], or large protein complex such as the nuclear pore[35], is positioned at the eucentric height in the microscope. At the eucentric height, the stage can be rotated in the z direction without changing the height or location of the particle of interest. The microscope stage is tilted over a wide range of heights to create a series of images. These images are then compiled to form a 3D representation of a single particle of interest (Fig. 1.3b)[36]. By collecting a series of low electron dose images of a single specimen, tomography allows electron microscopists to mitigate the issue of sample damage while producing useful three-dimensional information about the species of interest. Tomography is extremely useful for studying non-identical or large specimens which are not suitable to single particle imaging and can be used with samples prepared for cryoEM or traditional TEM sample preparation methods. Electron tomography allows for fascinating insights into cellular machinery.

Unfortunately, tomography has two main limitations that hold it back from producing data with high enough resolution information to give rise to *de novo* protein structures. The first is the decreased signal to noise ratio in each individual image within

the tomography tilt series. To prevent sample damage during the tilt series, each individual image only receives a low dose from the electron beam. This issue is overcome by sub-tomographic averaging. In subtomographic averaging, an area of interest within several tomograms is identified. These areas are then isolated and averaged together computationally to give a higher resolution structure for the area of interest. An example of subtomographic averaging which was used to study microbial interactions is the work done on the basal body of the bacterial type IV pili[37,38].

The other factor which limits resolution of tomographic tilt series is the difficulty of properly correcting for the contrast transfer function (CTF) in tilted images[39]. In electron microscopy, images are typically taken at an under-focused position to give contrast to the specimen. This results in the CTF which needs to be corrected for during image processing. The CTF is dependent on the defocus value of an image. In images taken at high tilt, such as those taken in a tomography tilt series, the defocus value is different in different regions of the image, as one side of the grid is tilted closer to the electron beam and the other side is tilted further away. This results in a gradient of different CTFs across the image which need to corrected for. Currently, there is no one program which can completely deconvolute and correct for this effect, although several groups are working to create a solution[40,41].

Single particle reconstructions are accomplished by taking images of several thousand particles. Ideally, these particles will be random in orientation and the data set will contain images (i.e. two-dimensional projections) of every view of the sample. Using the central projection theorem, these two-dimensional projections of the sample are then Fourier transformed. The Fourier transforms of the two-dimensional particles are

combined computationally in three-dimensional Fourier space. The combined Fourier transform data is then inversely Fourier transformed back into real space in three-dimensions to produce an electron density map of the particles that were imaged. In general, the workflow for single particle cryo electron microscopy is as follows: sample optimization and freezing, data collection, CTF correction, two-dimensional class averaging, three-dimensional class averaging, final reconstruction of the electron density map, and, if the electron density map is of sufficiently high resolution (usually 4 Å resolution or better) building an atomic model and model refinement (Fig. 1.3c)[28].

There is one major disadvantage to using single particle cryoEM which has traditionally held the technique back from routinely solving structures to atomic resolution, the low signal to noise ratio of cryoEM images. This has given rise to several data processing issues, for example, because cryoEM images have a low signal to noise ratio and are subject to extensive computational procedures, an issue with model bias was discovered[42,43]. In the early 2000's it was demonstrated that using some existing data processing tools for cryoEM, any input reference model could adversely affect the result. Due to the noise in cryoEM images and the large size of the datasets being used to reconstruct electron density maps, an input model for refinement can bias the result and give rise to a structure which is closer to the original model than the data supports. This resulted in a number of controls being developed, such as the so-called "gold standard" FSC method, where the data set is divided in half and the two halves are reconstructed independently, and spurred a series of modifications to existing data processing programs and the development of new programs which dealt with this

problem[44,45]. However, the gold standard FSC method is not fool-proof as systematic errors in data processing can mask model bias. The development of the direct electron detector discussed in the next section has dramatically increased the number of cryoEM single particle reconstructions which achieve sufficient resolution to solve a protein structure *de novo*. At atomic resolution, using the sequence of the protein as an internal control is a more robust demonstration that the cryoEM density map is accurate. The FSC method remains an important tool for assessing the quality of cryoEM density maps and reducing model bias.

Single particle cryoEM has been used to study microbial interactions such as bacterial secretion systems and toxin export pores. The bacterial type II, type III[46], and type IV[47,48] secretion system complexes have been described at intermediate resolution using single particle cryoEM. The atomic structure of the anthrax pore complex shows how a pathogenic bacterium is able to export its toxin[49]. This structure also highlights one of the major advantages of single particle cryoEM over X-ray crystallography and NMR, the ability to routinely solve membrane protein structures.

In this thesis, we used cryoEM for two separate capacities. To study the *S. wolfei* surface layer protein in chapter three, we imaged whole cells by cryoEM. This allowed us to confirm the existence of the S-layer and to characterize the density with which the protein covers the outside of the cell. To determine the structure of the *M. hungatei* flagellum in chapter four, we used single particle imaging (Fig. 1.3d) and helical reconstruction techniques, described in greater detail later in this chapter.

Direct electron detectors

In 2005, a paper was published showing that a material used to detect single electron strikes in nuclear reactors, known as a high-sensitivity active pixel sensory, could be used to build a better detector for transmission electron microscopes (Fig. 1.4a)[50]. Over the next ten years, three companies developed and released direct electron detectors based on this technology for commercial use. As a whole, these new detectors offer significant improvements in data quality by increasing the signal to noise ratio of individual micrographs and by making it possible to correct for beam induced sample damage and drifting. In contrast to CCDs or film cameras which collect a single image developed over the exposure period, direct electron detectors collect a short movie over the exposure period. Due to the increased sensitivity to electron strikes of the detector, a series of short, usually around 100-300 microsecond long, frames can be collected during each exposure period. These frames can then be aligned to form one image of the full exposure period. Frames from latter in the exposure process when the sample is damaged can be discarded, and beam induced motion of the sample can be corrected[51]. The invention of these new detectors coupled with ongoing improvements to data processing software created what has been termed the “resolution revolution” in cryoEM. In the past four years, the number of cryoEM structures resolved to 5 Å or better has increased exponentially (Fig. 1.4b)[52]. Direct electron detectors have improved resolution of cryoEM experiments across the board. This has included improvements to samples important for microbial interactions, including all of the samples referenced in the previous section about cryo electron microscopy.

Chapter four of this thesis is an excellent example of how the resolution revolution has helped provide new insights into microbial interactions. This chapter

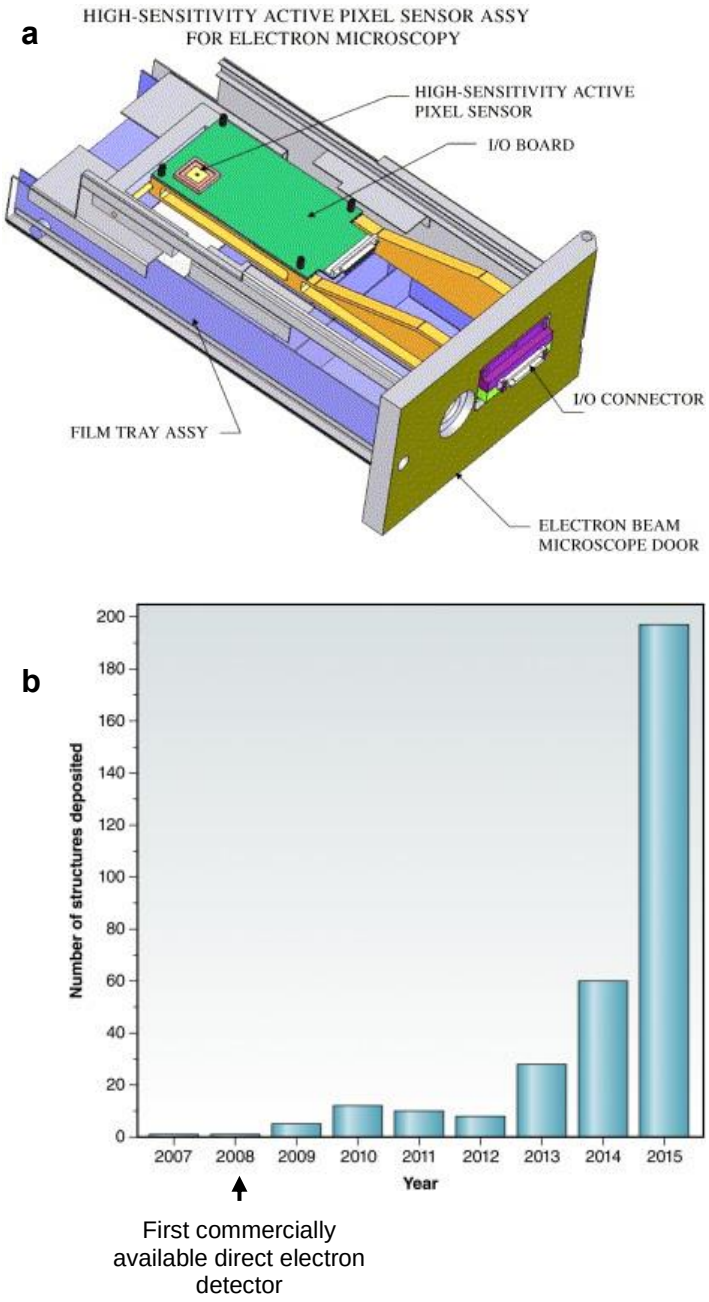


Figure 1.4 | The direct electron detector and the "resolution revolution". a, Initial concept for the direct electron detector in a modified film tray, adapted from Milazzo *et al* 2005 **b,** The number of 5 Å or better cryo EM maps deposited in the EMDB per year since 2007, adapted from Egelman 2016.

details the structure of an archaeal flagellum. This project began with single particle cryoEM data collection on photographic film. Issues with the sample, such as the thin diameter of the filaments and their lack of apparent high resolution features made structure determination with the film dataset impossible. However, new direct electron detectors and the incorporation of Baysian statistics into helical reconstruction of filaments through the RELION[53] software package made it possible to solve this challenging structure to 3.4 Å resolution. This resolution was sufficient to allow for *de novo* backbone tracing of the flagellin monomer and was the first complete structure of an archaeal flagellin to be solved at atomic resolution. Previously, only the structure of a non-functional bacterial flagellum mutant had been solved to atomic detail using cryoEM helical reconstruction[54]. Motility filaments such as bacterial and archaeal flagella and pili were extremely difficult to solve by cryoEM due to their small diameter, curvature, and lack of high resolution features; however, by demonstrating that this is now possible by solving the archaeal flagella structure, we hope to open the structure of bacterial and archaeal motility filaments up to further study by cryoEM.

Helical reconstruction

Solving atomic resolution structures of helical filaments is a unique challenge. Many of these filaments autopolymerize in high concentrations creating problems for X-ray crystallography and NMR studies. Single particle cryo electron microscopy is ideal for structural determination of large complexes, and the first proof of concept article to demonstrate the ability of electron microscopy to solve protein structures used a helical sample, the tail of the T4 bacteriophage[55]. However, the nature of helical filaments results in a unique set of challenges when attempting to solve these structures. The

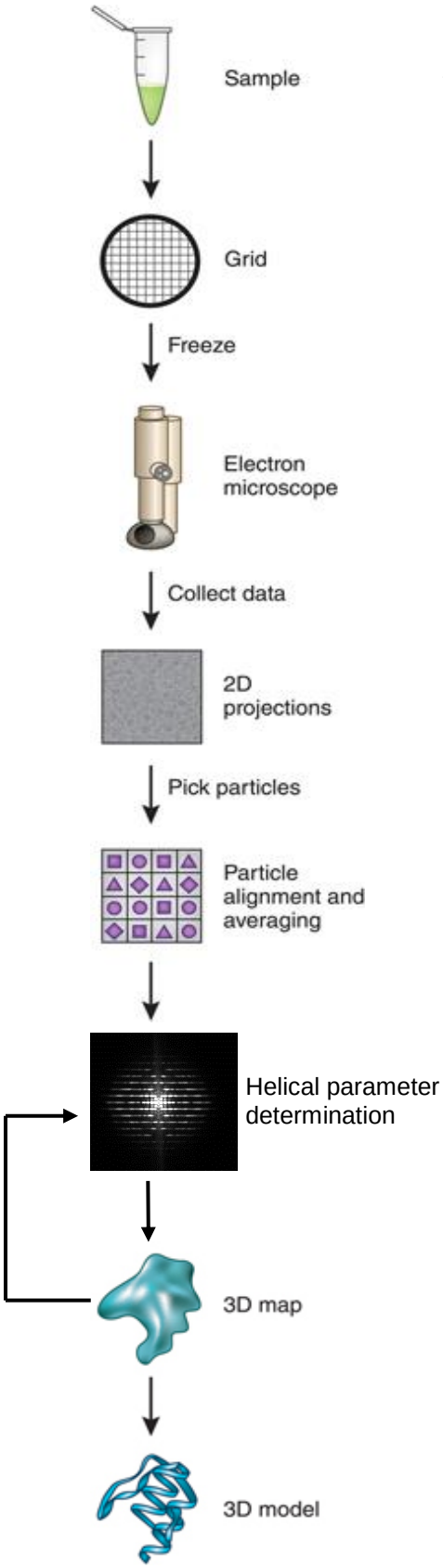


Figure 1.5 | An overview of the workflow for helical data processing. While the overall workflow is very similar to single particle reconstruction, note the addition of the helical parameter determination step, adapted from Doerr 2016.

overall workflow for helical reconstruction is similar to single particle reconstruction with sample optimization, data collection, CTF correction, 2D class averaging, 3D class averaging, final refinement, and model building and refinement. However, the helicial organization results in the modification of some of these steps and inserts a new step, helical parameter determination, in between 2D class averaging and 3D class averaging (Fig. 1.5).

Molecular complexes with helical symmetry are a unique challenge for cryoEM reconstructions. The filament does not provide every view of the protein complex as is desirable for single particle reconstructions (top views of the complex are completely absent). However, the helical symmetry itself can be used to align particles and back project a model for refinement, if the symmetry can be correctly identified. Yet applying incorrect helical parameters to the reconstruction will result in model bias, as discussed above for single particle reconstruction, and an incorrect structure. For this reason, correctly determining the helical parameters is the most crucial step when using cryoEM to reconstruct helical filaments.

When the projection (i.e. the electron micrograph) of a helical filament is converted into Fourier space, the result is a Fourier transform with a characteristic X shape (Fig. 1.6a)[56]. This is due to the arrangement of subunits in the helical polymer. Every continuous helix in the real space projection of a helical polymer can be defined by a characteristic cylindrical harmonic function. A cylindrical harmonic function is a Bessel function where the Bessel order (n) is an integer value. When the solutions to a cylindrical harmonic function are plotted in three dimensions, the result is a wave. The arrangement of subunits in a helical polymer can be described by a series of

characteristic cylindrical harmonic functions which are unique to that filament. Transforming the two-dimensional projection (electron micrograph) of a helical polymer into Fourier space decomposes the polymer's characteristic Bessel functions into their frequencies. This results in only the Bessel orders with an integer value being plotted in Fourier space. These are known as layer lines and the layer lines form a characteristic X shape. A sample's layer lines can be indexed to identify the cylindrical coordinates of each Bessel function using a helical net (Fig. 1.6b)[57] and the selection rule:

$$l = tn + um \tag{1}$$

Where l, n, and m are cylindrical coordinates which can be determined from the helical net, t is the number of turns around the filament needed to return to a point identical to the helical starting point, and u is the number of subunits in the filament needed to return to a point identical to the helical starting point. The selection rule for any one cylindrical coordinate is undetermined and impossible to solve, however by defining the helical net in terms of l,n, and m for two characteristic lines we are able to solve an algebraic system of two equations which provide the basic information needed to define the arrangement of subunits in a helical polymer by a cylindrical harmonic function. Due to the undetermined nature of the selection rule and the difficulty of indexing the layer lines from noisy data, correctly determining helical parameters is very difficult and time consuming. By imposing these helical parameters on a cryoEM dataset of helical filaments, we are able to reconstruct the filament in three dimensions and solve its structure using single particle cryoEM methods as described above[58].

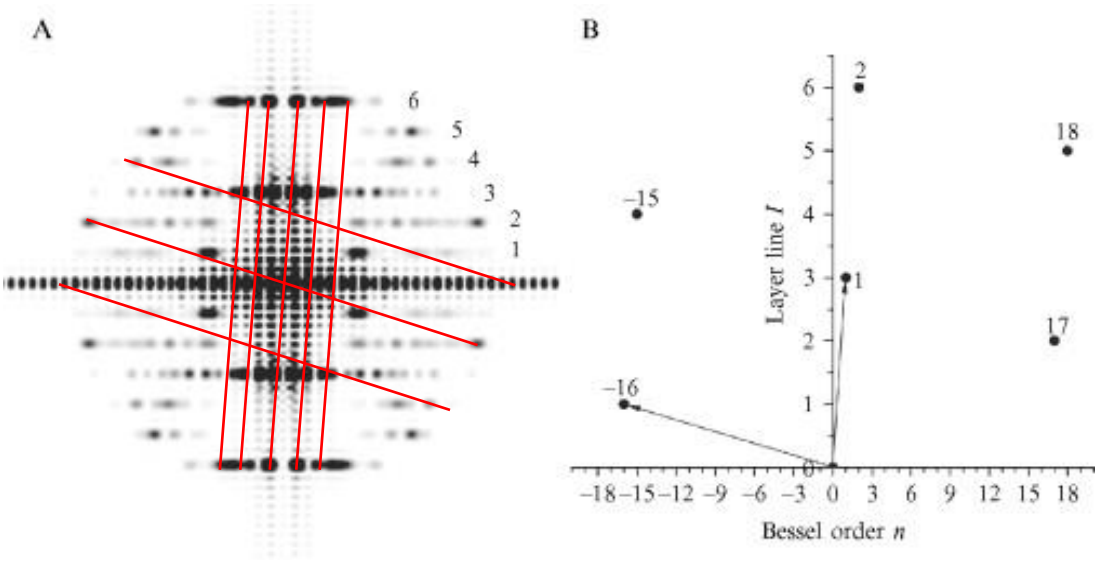


Figure 1.6 | An overview of helical parameter determination. a, The Fourier transform of Tabaco Mosaic Virus (TMV) with a helical net overlaid in red. **b,** A plot of the layer line number and Bessel order of the points in the helical net. Both panels adapted from Egelman 2000.

Helical reconstruction has been used to study a variety of bacterial and archaeal flagella, pili, and adhesion filaments as well as the bacterial type VI secretion system. These extracellular filaments are important players in microbial interactions and studying their structures offers insight into how each filament is able to assemble and function. In 2003 the first atomic resolution bacterial flagellin structure was solved using helicial reconstruction of a straight, non-functional mutant of the *Salmonella enterica* flagella[54]. Several intermediate resolution cryoEM structures have been published for bacterial pili[59], archaeal flagella[60,61], and archaeal pili[62], and a partial atomic model was published for an archaeal adhesion filament[63]. These intermediate resolution structures were often combined with partial or full length atomic resolution x-ray crystallography and NMR protein structures to give an idea of how these extracellular filaments are assembled and function. The trunk portion of the bacterial type VI secretion system was solved to atomic resolution via helical reconstruction[64,65].

Knowing how the structure of bacterial pilin and flagellins as well as how the subunits are arranged to from a filament has helped answer many questions about the function of these filaments, yet no comparable structures existed for any of the archeal motility and adhesion filaments. Chapter four of this thesis is the final draft of an article in which we solved the first complete atomic model of an archaeal flagellin using cryoEM and helical reconstruction. This model allowed us to compare the archaeal flagellum to the bacterial type IV pili and flagellum and demonstrates that the archaeal flagellum is structurally distinct from the two bacterial filaments.

References

1 Zworykin, V. K. The Scanning Electron Microscope. *Scientific American* **167**, 111-113 (1942).

2 McMullan, D. Scanning electron microscopy 1928–1965. *Scanning* **17**, 175-185, (1995). Muscariello, L. *et al.* A critical overview of ESEM applications in the biological field. *Journal of Cellular Physiology* **205**, 328-334, (2005).

3 Bozzola, J. J. & Russell, L. D. *Electron microscopy: principles and techniques for biologists*. (Jones & Bartlett Learning, 1999).

4 O'Connor, K. A. & Zusman, D. R. Patterns of cellular interactions during fruiting-body formation in *Myxococcus xanthus*. *Journal of Bacteriology* **171**, 6013-6024 (1989).

5 Behmlander, R. M. & Dworkin, M. Extracellular fibrils and contact-mediated cell interactions in *Myxococcus xanthus. Journal of Bacteriology* **173**, 7810-7820 (1991).

6 Licking, E., Gorski, L. & Kaiser, D. A Common Step for Changing Cell Shape in Fruiting Body and Starvation-Independent Sporulation of *Myxococcus xanthus. Journal of Bacteriology* **182**, 3553-3558, (2000).

7 Sun, Y., Dowd, S. E., Smith, E., Rhoads, D. D. & Wolcott, R. D. In vitro multispecies Lubbock chronic wound biofilm model. *Wound Repair and Regeneration* **16**, 805-813, (2008).

8 Shen, Y., Qian, W., Chung, C., Olsen, I. & Haapasalo, M. Evaluation of the Effect of Two Chlorhexidine Preparations on Biofilm Bacteria In Vitro: A Three-Dimensional Quantitative Analysis. *Journal of Endodontics* **35**, 981-985, (2009).

9 Massol-Deyá, A. A., Whallon, J., Hickey, R. F. & Tiedje, J. M. Channel structures in aerobic biofilms of fixed-film reactors treating contaminated groundwater. *Applied and Environmental Microbiology* **61**, 769-777 (1995).

10 Heymann, J. A. W. *et al.* Site-specific 3D imaging of cells and tissues with a dual beam microscope. *Journal of Structural Biology* **155**, 63-73, (2006).

11 Remis, J. P. *et al.* Bacterial social networks: structure and composition of *Myxococcus xanthus* outer membrane vesicle chains *Environmental Microbiology* **16**, 598-610 (2014).

12 Crewe, A. V., Eggenberger, D. N., Wall, J. & Welter, L. M. Electron Gun Using a Field Emission Source. *Review of Scientific Instruments* **39**, 576-583, (1968).

13 Facility, C. M. R. *Transmission Electron Microscopy*, https://cmrf.research.uiowa.edu/transmission-electron-microscopy

14 Tsay, T.-T., Inman, R., Wray, B., Herman, B. & Jacobson, K. Characterization of low-light-level cameras for digitized video microscopy. *Journal of Microscopy* **160**, 141-159, (1990).

15 Waśko, A., Polak-Berecka, M., Kuzdraliński, A. & Skrzypek, T. Variability of S-layer proteins in *Lactobacillus helveticus* strains. *Anaerobe* **25**, 53-60, (2014).

16 Poppinga, L. *et al.* Identification and Functional Analysis of the S-Layer Protein SplA of *Paenibacillus larvae*, the Causative Agent of American Foulbrood of Honey Bees. *PLoS Pathog* **8**, e1002716, (2012).

17 Sleytr, U. B., Sára, M., Küpcü, Z. & Messner, P. Structural and chemical characterization of S-layers of selected strains of *Bacillus stearothermophilus* and *Desulfotomaculum nigrificans*. *Arch. Microbiol.* **146**, 19-24, (1986).

18 Farchaus, J. W., Ribot, W. J., Downs, M. B. & Ezzell, J. W. Purification and characterization of the major surface array protein from the avirulent *Bacillus anthracis* Delta Sterne-1. *Journal of Bacteriology* **177**, 2481-2489 (1995).

19 Gerbino, E., Carasi, P., Mobili, P., Serradell, M. A. & Gómez-Zavaglia, A. Role of S-layer proteins in bacteria. *World Journal of Microbiology and Biotechnology* **31**, 1877-1887, (2015).

20 Reguera, G. *et al.* Extracellular electron transfer via microbial nanowires. *Nature* **435**, 1098-1101, (2005).

21 Chumakov, M. I. & Kurbanova, I. V. Localization of the protein VirB1 involved in contact formation during conjugation among *Agrobacterium* cells. *FEMS Microbiology Letters* **168**, 297-301, (1998).

22 Maher, D., Sherburne, R. & Taylor, D. E. H-pilus assembly kinetics determined by electron microscopy. *Journal of Bacteriology* **175**, 2175-2183 (1993).

23 Tomich, M. & Mohr, C. D. Adherence and autoaggregation phenotypes of a *Burkholderia cenocepacia* cable pilus mutant. *FEMS Microbiology Letters* **228**, 287-297, (2003).

24 Bradley, D. E. Evidence for the retraction of *Pseudomonas aeruginosa* RNA phage pili. *Biochemical and Biophysical Research Communications* **47**, 142-149, (1972).

25 Melville, S. & Craig, L. Type IV Pili in Gram-Positive Bacteria. *Microbiology and Molecular Biology Reviews : MMBR* **77**, 323-341, (2013).

26 Dubochet, J. *et al.* Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics* **21**, 129-228, (1988).

27 Bruggeller, P. & Mayer, E. Complete vitrification in pure liquid water and dilute aqueous solutions. *Nature* **288**, 569-571 (1980).

28 Doerr, A. Single-particle cryo-electron microscopy. *Nat Meth* **13**, 23-23, (2016).

29 Henderson, R. Overview and future of single particle electron cryomicroscopy. *Archives of Biochemistry and Biophysics* **581**, 19-24, (2015).

30 Henderson, R. *et al.* Tilt-Pair Analysis of Images from a Range of Different Specimens in Single-Particle Electron Cryomicroscopy. *J Mol Bio* **413**, 1028-1046, (2011).

31 Wright, E. R., Iancu, C. V., Tivol, W. F. & Jensen, G. J. Observations on the behavior of vitreous ice at ~82 and~12 K. *Journal of Structural Biology* **153**, 241-252, (2006).

32 Brilot, A. F. *et al.* Beam-induced motion of vitrified specimen on holey carbon film. *Journal of Structural Biology* **177**, 630-637, (2012).

33 Farley, M. M., Tu, J., Kearns, D. B., Molineux, I. J. & Liu, J. Ultrastructural analysis of bacteriophage Φ29 during infection of *Bacillus subtilis*. *Journal of Structural Biology*.

34 Perkins, G. A., Jackson, D. R. & Spirou, G. A. Resolving Presynaptic Structure by Electron Tomography. *Synapse (New York, N.Y.)* **69**, 268-282, (2015).

35 Kosinski, J. *et al.* Molecular architecture of the inner ring scaffold of the human nuclear pore complex. *Science* **352**, 363-365, (2016).

36 Fernandez, J.-J. Computational methods for electron tomography. *Micron* **43**, 1010-1030, (2012).

37 Chang, Y.-W. *et al.* Architecture of the type IVa pilus machine. *Science* **351**, (2016).

38 Gold, V. A. M., Salzer, R., Averhoff, B. & Kühlbrandt, W. Structure of a type IV pilus machinery in the open and closed state. *eLife* **4**, e07380, (2015).

39 Hoenger, A. High-resolution cryo-electron microscopy on macromolecular complexes and cell organelles. *Protoplasma* **251**, 417-427, (2014).

40 Xiong, Q., Morphew, M. K., Schwartz, C. L., Hoenger, A. H. & Mastronarde, D. N. CTF Determination and Correction for Low Dose Tomographic Tilt Series. *Journal of structural biology* **168**, 378-387, (2009).

41 Galaz-Montoya, J. G. *et al.* Alignment algorithms and per-particle CTF correction for single particle cryo-electron tomography *Journal of structural biology* **194** 383-394 (2016)

42 Yang, S., Yu, X., Galkin, V. E. & Egelman, E. H. Issues of resolution and polymorphism in single-particle reconstruction. *Journal of Structural Biology* **144**, 162-171, (2003).

43 Stewart, A. & Grigorieff, N. Noise bias in the refinement of structures derived from single particles. *Ultramicroscopy* **102**, 67-84, (2004).

44 Scheres, S. H. W. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat Meth* **9**, 853-854, (2012).

45 Henderson, R. *et al.* Outcome of the First Electron Microscopy Validation Task Force Meeting. *Structure* **20**, 205-214, (2010).

46 Tosi, T. *et al.* Structural Similarity of Secretins from Type II and Type III Secretion Systems. *Structure* **22**, 1348-1355, (2014).

47 Rivera-Calzada, A. *et al.* Structure of a bacterial type IV secretion core complex at subnanometre resolution. *The EMBO Journal* **32**, 1195 (2013).

48 Low, H. H. *et al.* Structure of a type IV secretion system. *Nature* **508**, 550-553, (2014).

49 Jiang, J., Pentelute, B. L., Collier, R. J. & Zhou, Z. H. Atomic structure of anthrax protective antigen pore elucidates toxin translocation. *Nature* **521**, 545-549, (2015).

50 Milazzo, A.-C. *et al.* Active pixel sensor array as a detector for electron microscopy. *Ultramicroscopy* **104**, 152-159, (2005).

51 Campbell, Melody G. *et al.* Movies of Ice-Embedded Particles Enhance Resolution in Electron Cryo-Microscopy. *Structure* **20**, 1823-1828, (2012).

52 Egelman, Edward H. The Current Revolution in Cryo-EM. *Biophysical Journal* **110**, 1008-1012, (2016).

53 Scheres, S. H. W. A Bayesian View on Cryo-EM Structure Determination. *Journal of Molecular Biology* **415**, 406-418, (2012).

54 Yonekura, K., Maki-Yonekura, S. & Namba, K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643-650, (2003).

55 De Rosier, D. & Klug, A. Reconstruction of three dimensional structures from electron micrographs. *Nature* **217**, 130-134 (1968).

56 Cochran, W., Crick, F. H. & Vand, V. The structure of synthetic polypeptides. I. The transform of atoms on a helix. *Acta Crystallographica* **5**, (1952).

57 Klug, A., Crick, F. H. C. & Wyckoff, H. W. Diffraction by helical structures. *Acta Crystallographica* **11**, 199-213, (1958).

58 Egelman, E. H. A robust algorithm for the reconstruction of helical filaments using single-particle methods. *Ultramicroscopy* **85**, 225-234, (2000).

59 Craig, L. *et al.* Type IV Pilus Structure by Cryo-Electron Microscopy and Crystallography: Implications for Pilus Assembly and Functions. *Molecular Cell* **23**, 651-662, (2006).

60 Trachtenberg, S., Galkin, V. E. & Egelman, E. H. Refining the Structure of the Halobacterium salinarum Flagellar Filament Using the Iterative Helical Real Space Reconstruction Method: Insights into Polymorphism. *Journal of Molecular Biology* **346**, (2005).

61 Cohen-Krausz, S. & Trachtenberg, S. The Structure of the Archeabacterial Flagellar Filament of the Extreme Halophile *Halobacterium salinarum* R1M1 and Its Relation to Eubacterial Flagellar Filaments and Type IV Pili. *Journal of Molecular Biology* **321**, 383-395, (2002).

62 Henche, A.-L. *et al.* Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environmental microbiology* **14**, 3188-3202, (2012).

63 Braun, T. *et al.* Archaeal flagellin combines a bacterial type IV pilin domain with an Ig-like domain. *PNAS* **113**, 10352-10357, (2016).

64 Clemens, Daniel L., Ge, P., Lee, B.-Y., Horwitz, Marcus A. & Zhou, Z. H. Atomic Structure of T6SS Reveals Interlaced Array Essential to Function. *Cell* **160**, 940-951, (2015).

65 Kudryashev, M. *et al.* Structure of the Type VI Secretion System Contractile Sheath. *Cell* **160**, 952-962, (2015).

CHAPTER 2:

Phenotypic and Physiological Characterization of the Epibiotic Interaction Between TM7x and Its Basibiont *Actinomyces*

Batbileg Bor[1] · Nicole Poweleit[2] · Justin S. Bois[3] · Lujia Cen[1] · Joseph K. Bedree[1] · Z. Hong Zhou[2,4] · Robert P. Gunsalus[2] · Renate Lux[1] · Jeffrey S. McLean[5] · Xuesong He[1] · Wenyuan Shi[1]

Received: 6 August 2015 / Accepted: 11 November 2015 / Published online: 23 November 2015

Abstract Despite many examples of obligate epibiotic symbiosis (one organism living on the surface of another) in nature, such an interaction has rarely been observed between two bacteria. Here, we further characterize a newly reported interaction between a human oral obligate parasitic bacterium TM7x (cultivated member of Candidatus Saccharimonas formerly Candidate Phylum TM7), and its basibiont *Actinomyces odontolyticus* species (XH001), providing a model system to study epiparasitic symbiosis in the domain Bacteria. Detailed microscopic studies indicate that both partners display extensive morphological changes during symbiotic growth. XH001 cells manifested as short rods in monoculture, but displayed elongated and hyphal morphology when physically associated with TM7x. Interestingly, these dramatic morphological changes in XH001 were also induced in oxygen-depleted conditions, even in the absence of TM7x. Targeted quantitative real-time PCR (qRT-PCR) analyses revealed that both the physical association with TM7x as well as oxygen depletion triggered up-regulation of key stress response genes in XH001, and in combination, these conditions act in an additive manner. TM7x and XH001 co-exist with relatively uniform cell morphologies under nutrient-replete conditions. However, upon nutrient depletion, TM7x-associated XH001 displayed a variety of cell morphologies, including swollen cell body, clubbed-ends, and even cell lysis, and a large portion of TM7x cells transformed from ultrasmall cocci into elongated cells. Our study demonstrates a highly dynamic interaction between epibiont TM7x and its basibiont XH001 in response to physical association or environmental cues such as oxygen level and nutritional status, as reflected by their morphological and physiological changes during symbiotic growth.

Keywords Obligate · Epibiont · Symbiosis · Bacterial interaction · TM7 · Actinomyces

Electronic supplementary material The online version of this article (doi:10.1007/s00248-015-0711-7) contains supplementary material, which is available to authorized users.

✉ Xuesong He
xhe@g.ucla.edu

✉ Wenyuan Shi
wenyuan@ucla.edu

[1] Section of Oral Biology, School of Dentistry, University of California, Los Angeles, CA 90095, USA

[2] Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA 90095, USA

[3] Division of Biology and Biological Engineering, California Institute of Technology, MC 114-96, Pasadena, CA 91125, USA

[4] California Nanosystems Institute, University of California, Los Angeles, CA 90095, USA

[5] Department of Periodontics, University of Washington, Seattle, WA 98195, USA

Introduction

A recurring theme in ecology is that organisms of different species interact with each other in a variety of ways. However, intimate relationships, such as obligate epibiont symbiosis, where an organism lives exclusively on the surface of other organisms, are mainly observed between interacting pairs with at least one eukaryote [1]. Only a few examples of prokaryote-to-prokaryote obligate epibiotic interactions have been reported [2–4]. This is surprising considering the large diversity of prokaryotic organisms that have been detected by DNA sequence-based techniques [5–8]. The infrequent reporting of obligate epibiont interactions between prokaryotes can be largely explained by our inability to cultivate and study many of the

microorganisms from the environment and the human body. Culturing these bacterial species in order to study their intra- and inter-species interactions is one of the major challenges in current microbiology. The ability to cultivate these species will allow for better understanding of their roles in microbial ecology, as well as microbial community-based pathogenesis.

There are numerous examples of epibiotic interactions between different bacterial species, such as phototrophic microbial consortia from lakes, corn-cob forming consortia from the oral cavity, and methane-oxidizing consortia from deep-sea sediments [9–11]. However, these symbiotic interactions are not obligate; bacteria involved in these consortia can also grow independently. To our knowledge, the only known obligate epibiotic relationships between bacterial species are predatory bacteria such as *Vampirococcus* and *Micavibrio*, and their respective bacterial prey, which include a variety of species [2, 4, 12, 13]. For this type of obligate epibiotic interaction, the predatory bacteria live on the surface of the prey and feed by leeching externally, which often results in killing of the prey in a short period of time.

Recently, we described a unique and intimate relationship between two bacterial species, TM7x (TM7 phylotype) and *Actinomyces odontolyticus* strain XH001 which were co-isolated from the human oral cavity [14]. TM7x is characterized as an obligate epibiont parasite (epiparasite) that lives on the surface of its host XH001 (referred to as the basibiont). Unlike predatory bacteria, TM7x shows a high degree of specificity for its basibiont. Furthermore, TM7x does not immediately induce cell death like most predatory bacteria, but rather maintains a stable parasitic relationship with XH001 under nutrient-replete conditions, suggesting a novel class of bacteria-to-bacteria interaction.

TM7x belongs to the TM7 phylum which, among other uncultivated phyla [15], has long been referred to as "microbial dark matter," due to its cosmopolitan existence without a cultivable representative [14]. This phylum was recently renamed Saccharibacteria [16], after an additional TM7 genome was reconstructed from metagenomic reads and combined with the existing partial single-cell TM7 genomes that were available [17, 18]. The cultivated oral TM7 described here is currently designated as Saccharibacteria oral taxon TM7x (NCBI Taxonomy ID: 1476577), and a complete genome has been deposited (ACCESSION CP007496) [14]. Our recent study revealed that TM7x has an ultrasmall cell size (200–300 nm), a reduced genome, and is devoid of many biosynthetic pathways including de novo biosynthesis of all essential amino acids [14]. Moreover, the TM7 phylum was recently included in the proposed candidate phyla radiation (CPR), a subdivision of the domain Bacteria, due to these shared genomic characteristics with other novel genomes discovered in the bacterial domain [19]. Ultrasmall cells in groundwater have also been subsequently reported and suspected to belong to candidate phyla [20]. Although it is a typical constituent of the human microbiome, TM7 has been implicated in multiple human mucosal diseases such as vaginosis, inflammatory bowel disease, and periodontitis [16, 17, 21–34]. An increase in abundance of TM7 members (as high as 21 % of the whole oral bacteria population in some cases) was detected in patients with various types of periodontitis [35–39]. However, little is known about the role of TM7 in the pathogenesis of these mucosal diseases.

Although it serves as a basibiont for TM7x, *A. odontolyticus* strain XH001 can also be cultured independently of TM7x [14]. Similar to many other oral *Actinomyces* spp, *A. odontolyticus* can be found among the healthy human oral flora. However, it is also considered an opportunistic pathogen and has been implicated in many diseases, such as childhood caries, periodontitis, human oral carcinomas [40–45], and most notably, Actinomycosis, the formation of painful abscesses in the mouth, lungs, or gastrointestinal tract [46, 47].

The identification of the unique TM7x/XH001 interaction provides a new model system to study obligate epiparasitic symbiosis in the domain Bacteria. In an effort to further characterize the intimate interaction between epibiont TM7x and its basibiont, we performed a detailed phenotypic and physiological analysis. Microscopic analysis revealed reciprocal morphological changes in the two interacting partners during their symbiotic growth under varying nutritional conditions. Furthermore, we demonstrate that the TM7x-induced morphological changes in XH001 can be partially recapitulated by depleting oxygen in the absence of TM7x. Most intriguingly, we present evidence that suggests that TM7x-induced morphological changes in XH001 are a result of a cellular stress response and negatively affect XH001 cell growth.

Materials and Methods

Bacterial Strains and Growth Conditions

XH001 monoculture and XH001/TM7x co-culture were isolated from the oral cavity as described in our previous study [14]. Strains were cultured in brain heart infusion (BHI) at 37 °C under different oxygen conditions as specified in the main text: anaerobic (0 % O_2, 5 % CO_2, 5 % H_2, balanced with N_2), microaerophilic (2.6 % O_2, 5 % CO_2, balanced with N_2), high oxygen (19.7 % O_2, 5 % CO_2, balanced by N_2), and atmospheric conditions (~21 % O_2, 0.04 % CO_2, 0.9 % Ar, 78 % N_2). To acquire growth kinetics and phase contrast images, three independent cell cultures were grown under the specified oxygen condition for two passages (1 ml culture inoculated into 10 ml BHI and incubated 24 h each) before being re-inoculated into 20 ml fresh BHI with a starting OD_{600}

of 0.03. Cells were grown for two passages to insure homogeneity. Subsequently, cell density was measured using Spectronic Genesys 5 spectrophotometer.

Phase Contrast Imaging and Quantification

Phase contrast images were acquired during growth at each time point as indicated in the growth curve. At early time points, due to low cell density, we concentrated the cells by centrifugation and resuspended the cell pellets in a smaller volume of fresh medium. Cells were directly observed and imaged using Nikon Eclipse E400 microscope equipped with a Nikon Plan Fluor ×100/1.30 oil immersion objective.

For cell length and branch point analysis, XH001 monoculture and XH001/TM7x co-culture were grown in microaerophilic condition for 24 h (after two passages) before samples were collected and observed under a phase contrast microscope. Due to the formation of micro-aggregates in XH001/TM7x co-culture, samples were treated with mild sonication (see supplementary methods) before images were acquired as described above. Images of XH001 were processed using tools available in SciPy version 0.10.0 [48] and scikit-image version 0.14.0 [49] to obtain the length and degree of branching of each bacterium.

For each image, we performed the following operations:

1. We eliminated large-scale artifacts such as uneven illumination by performing a strong Gaussian blur of the image and then subtracting this from the original image.
2. We used Otsu's method to threshold the image. Henceforth, as "region" is defined as a collection of connected pixels in a thresholded image.
3. We cleared regions that were either touching the border of the image or below a size of 50 pixels.
4. We skeletonized the image using the skimage.morphology.skeletonize function and labeled the non-adjacent skeletonized regions. These regions are only one pixel thick and may be branched.
5. For each region in the skeletonized image, we converted the pixels into a connectivity matrix, *A*, where *$A_{ij}>0$* if pixels *i* and *j* are nearest neighbors. This connectivity matrix describes a graph.
6. We checked the graph for cycles. If there is a cycle, the graph is neglected. Thus, we only analyze trees. XH001 had 17/811 cycles and XH001/TM7x had 41/912.
7. Using the connectivity matrix, we computed the shortest path between each pair of pixels using the scipy.sparse.csgraph.shorestpath function.
8. Of all these shortest paths, we report the longest as the cell length.
9. Again using the connectivity matrix, we computed the number of leaves (vertices of degree one), *n_leaves*, of the branched structure. We used the number of leaves as a metric for the degree of branching. We assume we have only vertices of degree one or three, so the number of branch points (vertices of degree three) is *$n_\mathrm{leaves}2$*.

The source code is available upon request.

Attachment of TM7x to XH001

Phase contrast images revealed that TM7x associated with XH001 at the highest abundance ratio when cultured under microaerophilic condition, thus this condition was used for establishing symbiotic growth. After 24 h of growth at 37 °C, 2 ml of the co-culture was re-inoculated into 20 ml fresh BHI medium, incubated for 24 h, then followed by another tenfold dilution in 200 ml fresh medium, and incubated for an additional 24 h. Cells were collected by centrifugation at 13000×*g* for 10 min, and the resulting cell pellet was resuspended in 1 ml fresh BHI. Cell suspension was passed through 1 cm^3 insulin syringe (U-100 28 g ½, Becton Dickinson 329410) for 10 min under sterile conditions to physically separate TM7x from XH001. The disassociated TM7x cells were collected by filtering the solution through a 0.22 μm Millex GP filter unit (SLGP033RB).

To re-attach TM7x to XH001, 1 ml XH001 monoculture at OD_{600} of 0.2 was centrifuged at 13000×*g* for 5 min, and the cell pellet was resuspended in 200 μl BHI medium containing the isolated TM7x collected as described above. Mixed cells were grown under microaerophilic conditions for 8 h at 37 °C to allow the establishment of symbiosis followed by addition of 800 μl of fresh BHI. Cultures were further incubated for 24 h before 50 μl of the culture was re-inoculated into 2 ml fresh BHI. Cells were similarly passaged every 24 h thereafter.

Isolation of mRNA and Construction of cDNA

For messenger RNA (mRNA) isolation, cells were grown in a 100 ml culture in BHI under the appropriate oxygen condition for 24 h. Cells were collected by centrifugation at 13000×*g* for 5 min. Total RNA was isolated using High Pure RNA Isolation Kit (Ref 11828665001), and RNA Clean Concentrator Kit (ZYMO RESEARCH RNA Clean Concentrator: Cat#R1015) was used to clean up the RNA. Further purification of the RNA required the use of a modified version of Ambion TURBO DNA-FREE Kit protocol (Cat# AM 1907). Briefly, 1 μl of 10× TURBO DNase Buffer and DNase were added to the 6 μl of concentrated RNA. The samples were mixed gently and incubated at 37 °C for 1 h. Samples were then spiked with 1 μl of additional TURBO DNase and incubated at 37 °C for 1 h. To inactivate the DNase, 1 μl of DNase Inactivation Reagent was added and incubated for

5 min at room temperature. Sample was centrifuged at 10000×*g* for 1.5 min and the supernatant containing the RNA was transferred to a new tube. The RNA samples were tested for DNA contamination by PCR using universal 16S bacterial primers (015: 5′-ACTACGTGCCAGCAGCC-3′ and 016: 5′-GGACTACCAGGGTATCTAATC-3′). Once the RNA was free of DNA, it was reverse-transcribed into complementary DNA (cDNA) using Takara PrimeScript 1st strand cDNA Synthesis Kit (Cat# 6110A). The cDNA library was confirmed by PCR with universal bacterial primers (see above), *A. odontolyticus*-specific 16S primers (966F: 5′-ACGGCGGCACTGCAGAGATGTG-3′ and 1410R: 5′-CCACAAACGCGGTTAGGC-3′) and TM7x-specific 16S primers (400F: 5′-TATGAGTGAAGAATATGAC-3′ and 1110R: 5′-CAGTCCAAGTAGAAAAATAC-3′).

Quantitative Real-Time PCR

Quantitative real-time PCR (qRT-PCR) was performed using the iQ SYBR Green supermix (Bio-Rad Laboratories, Cat# 170-8882) on a Bio-Rad iQ5 real-time PCR detection system (Bio-Rad Laboratories, Inc., CA). The final qRT-PCR mixture (20 μl) contained 1× iQ SYBR Green supermix, 1 μg cDNA, and 4.25 μM of the appropriate forward and reverse qRT-PCR primers (Table S1) designed for selected stress response genes. The reactions were incubated at an initial denaturation at 95 °C for 3 min, followed by a 40-cycle amplification consisting of denaturation at 95 °C for 10 s, annealing at 60 °C for 30 s, and extension at 72 °C for 15 s. All primers pairs were checked for primer-dimer formation by using the dissociation curve analysis. The critical threshold cycle (C_t) was defined as the cycle in which fluorescence became detectable above the background fluorescence, and is inversely proportional to the logarithm of the initial number of template molecules. To calculate the fold change in mRNA expression, we used the equation developed by Michael W. Pfaffl [50]. To determine primer efficiency, a standard curve was plotted for each primer set with C_t values obtained from amplification of known quantities of XH001 gDNA. We also crosschecked the sequences of the five stress genes in the TM7x genome to ensure that TM7x did not contain these sequences. Each assay was performed with at least two independent RNA samples in triplicates.

Fluorescence In Situ Hybridization Imaging

Fluorescence in situ hybridization (FISH) was carried out as previously described [14] with modifications. The cells were fixed using 4 % formaldehyde for 3 h and permeabilized by 2 mg/ml lysozyme in 20 mM Tris at pH 7.0 for 9 min at 37 °C. Fixed cells were resuspended in 500 μl of hybridization buffer (20 mM Tris•Cl, pH 8.0, 0.9 M NaCl, 0.01 % SDS, 30 % deionized formamide) and incubated at 37 °C for 30 min. TM7x-specific (TM7-567: Cy5-5′-CCTACGCAACTCTTTACGCC-3′) and *A. odontolyticus*-specific (M33910: Cy3-5′-CAGTGTCGCCGTGCAT-3′) probes were used to stain the cells for 3 h at 42 °C. In Figs. 5, 6, and S3, we assigned white color to Cy5-labeled TM7x to provide the best contrast. Cells were then washed three times with 0.1× saline-sodium citrate buffer, 15 min each, and mounted on the cover slip with SlowFade Gold anti-fade reagent (Invitrogen, Ref: S36937). Cells were visualized with a Leica SPE I inverted confocal microscope equipped with an ACS APO 100×/1.15 oil CS immersion objective. We repeated each experiment and FISH analysis at least three times and showed only representative images in the figures.

Scanning Electron Microscope Sample Preparation and Imaging

Cells were grown overnight under the indicated gas conditions (main text) and washed with 1 ml PBS. Isolation of TM7x was conducted as described above. Samples were pelleted by centrifugation at 13000×*g* for 5 min and resuspended in 1 ml of fixative solution (3 % glutaraldehyde buffered by 0.1 M phosphate) and incubated at room temperature for 2–4 h. Samples were pelleted by centrifugation and washed with 0.1 M phosphate buffer four times. Samples were dehydrated with ethanol using a gradient until the water was replaced with 100 % ethanol. Samples were transferred to plastic-coated copper quantifoil grids by incubation with the samples for 15 min. The grids were transferred to the chamber of a Tousimis Autosamdri-810 Critical Point Dryer, and critical point drying was performed as per the manufacturer's instructions. Grids were imaged directly using the appropriate scanning electron microscope (SEM) holder or were adhered to SEM stubs via silver paste and coated with iridium using the South Bay Technology Ion Beam Sputtering/Etching System. Imaging was performed using a ZEISS Supra 40VP SEM operating at 1 kV under high vacuum.

Results

Growth Kinetics and Morphology of XH001 as Mono-Species or as Basibiont to TM7x

The oral cavity is a semi-oxygenated environment where saliva, supragingival tooth surfaces, and subgingival pockets harbor 40–114 (~5–15 %), 60 (~8 %), and 5–27 (~0.6–3.7 %) mmHg oxygen tension, respectively (in comparison, atmosphere has 155 mmHg or 21 % oxygen) [51, 52]. Therefore, we monitored the growth and morphology of XH001 as mono-species, as well as when forming symbiosis with TM7x, under microaerophilic conditions (2.6 % O_2 and 5 % CO_2). XH001 monoculture had a short lag phase (~2 h) and an approximately 16-h exponential phase with doubling time of

~2.5 h (Fig. 1a). The cell density plateaued at ~20 h (stationary phase) and slowly decreased thereafter (death phase). The XH001/TM7x co-culture exhibited similar growth kinetics and plateaued around ~22 h and had a doubling time of ~3 h (Fig. 1a). XH001 cells from monoculture displayed short rod morphology during growth (Fig. 1b(a–d)). However, multiple condensed spotty structures were observed throughout the cell body at later growth stages (Fig. 1b(d)) such as death phase (~48 h).

Interestingly, a more dynamic morphological change was observed in XH001 cells physically associated with TM7x. In early lag phase, TM7x-associated XH001 cells exhibited elongated and hyphal (branched) morphology (Fig. 1b(e)), whereas during exponential phase, the majority of the cells exhibited a short rod shape (Fig. 1b(f)). When entering stationary phase, the co-culture was again dominated by long and hyphal XH001 cells that were heavily decorated with TM7x (Fig. 1b(g)). In addition, as TM7x-associated XH001 increased in cell length during stationary phase, we also observed micro-aggregate formation (Fig. S1a). During death phase, we frequently observed XH001 with large clubbed-ends, swollen cell shape, and cell lysis, which were not observed in XH001 monoculture (Fig. 1b(h)).

In order to validate the TM7x-induced morphological changes in XH001, we performed a re-attachment experiment. Individually isolated TM7x cells were collected by repeatedly passing the co-culture through a 28-gauge needle followed by filtering the mixture through 0.22-μm filter. The isolated TM7x cells were then co-incubated with a pure culture of XH001. Light microscopy showed that XH001 in pure culture

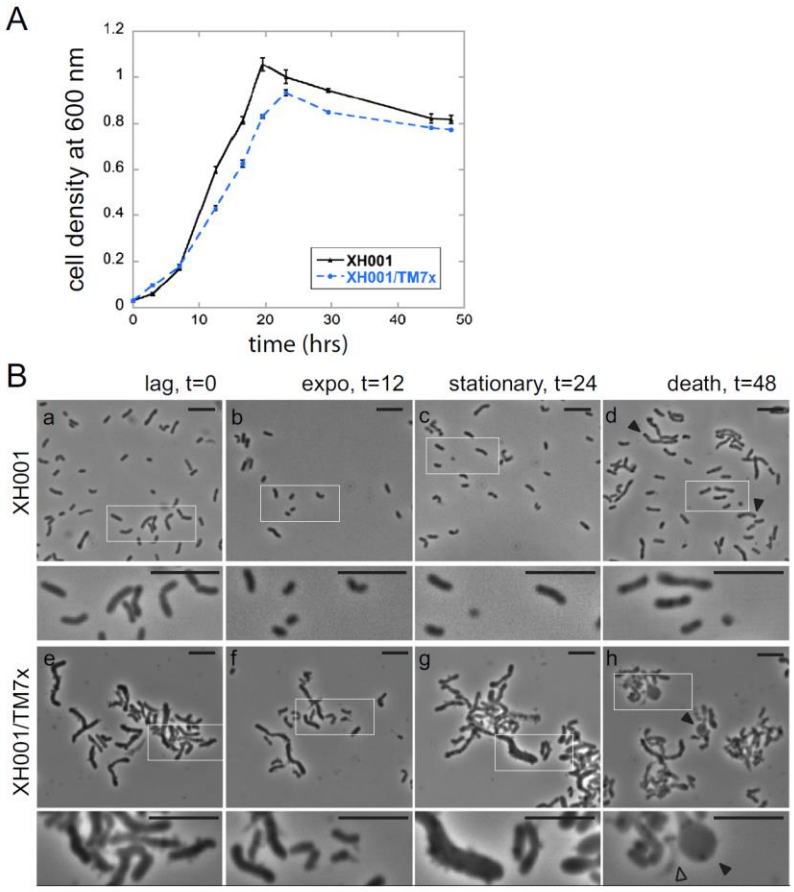


Fig. 1 Growth and morphology of XH001 monoculture and XH001/TM7x co-culture. **a** Mono (*triangle*, *black solid line*) and co-culture (*circle*, *blue broken line*) cell densities were determined by measuring the optical density at 600 nm (OD_{600}). Each point represents the average of three independent cultures (*error bars*, SD). The time points were connected by straight line to guide the eye. **b**, *a–d* XH001 monoculture grown in microaerophilic conditions has short rod morphology from early to late growth phases. The lower images are a higher magnification of the boxed region in the upper image. During death phase (**b**, *d*), many of the cells had condensed dot structures within the cell body (*arrow heads*). **b**, *e–h* TM7x-associated XH001 have elongated and hyphal morphology from early to late growth phases. During exponential phase (**b**, *f*), many of the XH001 cells are short rods similar to XH001 alone. At death phase (**b**, *h*), many of the XH001 cells display large clubbed-ends (*arrow heads*), swollen cells, and cell lyses (*open arrow head*). During all growth phases, XH001 was heavily decorated with TM7x. All *scale bars* indicate 5 μm

initially manifests as short rod cells, but becomes long and hyphal after establishing the physical association with TM7x (Fig. S1b–d). Our data suggests that the physical attachment with TM7x induces elongation and hyphal formation in XH001. Moreover, this association slightly decreases the apparent doubling time of XH001 compared to that of monoculture. Note that the cell density measurements of the growth curve account for both XH001 and TM7x and therefore, the decrease in the doubling time of XH001 we observe in co-culture is likely greater than represented in the graph. Additionally, the exact cell enumeration in the co-culture is difficult to obtain due to the micro-aggregation facilitated by the elongated XH001 cells.

Quantification of XH001 Elongation and Branching

To further characterize the XH001 morphological changes that occur in the presence of TM7x, we evaluated the cell length and the number of branch points. We developed and utilized a python-based software that segments the phase contrast images followed by skeletonizing each bacterium to determine cell length and number of branch points (see Materials and Methods). The results revealed a striking difference between XH001 cells grown in monoculture and those forming a physical association with TM7x under microaerophilic condition for 24 h. XH001 from monoculture had an average cell length of ~1.7 μm, whereas TM7x-associated XH001 averaged ~3.7 μm (Fig. 2a). Some of the XH001 cells in co-culture were as long as ~20 μm in length, while the longest XH001 cell length observed in monoculture was only ~8 μm. Furthermore, ~37 % of the XH001 cells analyzed had one or more branch points during co-culture growth compared to only ~5 % in monoculture. Furthermore, ~13 % of the co-culture and ~1 % of the monoculture XH001 cells had two or more branch points. Our quantitative data clearly indicate that TM7x induced significant elongation and branching in XH001.

The Effect of Oxygen Availability on the Morphology of XH001 as Monospecies or as Basibiont to TM7x

Although the majority of *Actinomyces* spp., including oral species, are classified as facultative anaerobes with an obligate carbon dioxide requirement, they demonstrate a wide range of oxygen preferences and tolerances during in vitro culture [53]. To determine the optimal oxygen condition, XH001 monoculture and XH001/TM7x co-culture were incubated under conditions with varying oxygen and carbon dioxide concentrations. In high oxygen (19.7 % O_2, 5 % CO_2) conditions, both monoculture and co-culture displayed retarded growth compared to microaerophilic conditions (both yielding a doubling time of ~4 h, Fig. 3a, b). Under the same conditions, XH001 monoculture demonstrated short rod morphology akin to cells

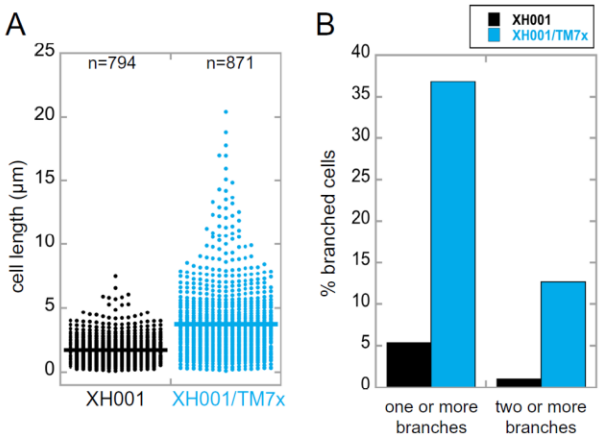


Fig. 2 Quantification of cell length and branch points in TM7x associated XH001. Mild sonication-dispersed co-culture cells were quantified by python-based image analysis software (see Materials and Methods) that automatically quantifies the longest length of the XH001 cells, and number of branch points. **a** *n* is the total number of cells measured from multiple microscope slides prepared from three independent mono or co-cultures incubated under microaerophilic conditions. *Error bars* are not shown as we combined the data to show the cell length distribution. Student's *t* test was used to determine significance ($p<0.001$). **b** Each *bar* shows the percent of cells in **a** (from *n* cells) that had one or more, or two or more branch points. TM7x-associated XH001 is ~7 times more likely to have one or more branch points or ~13 times more likely to have two or more branch points compare to XH001 alone

grown in microaerophilic conditions (Fig. S2a), while in the co-culture, XH001 displayed similar elongated and hyphal morphology, although with a reduced number of associated TM7x (Fig. 3c(a–d)). This observation suggests that an increase in oxygen negatively impacts the growth of both XH001 and TM7x. Under normal atmospheric conditions, where oxygen is high (20.9 %) and carbon dioxide concentrations are low (0.04 %), XH001 in both monoculture and co-culture completely ceased growth (Fig. 3a(b)). Morphology of the XH001 in monoculture and co-culture under this condition assumed short branching-filaments, branching rods, and diphtheroid forms after 24 h of growth (Fig. S2b, c). Therefore, carbon dioxide is absolutely necessary for XH001, and therefore enables TM7x growth.

In anaerobic growth conditions (0 % O_2 and 5 % CO_2), monoculture or co-culture growth was decreased relative to microaerophilic conditions and plateaued at a much lower cell density (Fig. 3a, b). Likewise, the co-culture had decreased growth when compared to monoculture (doubling time of ~5 and ~4 h, respectively, Fig. 3a, b). Most interestingly, an XH001 monoculture incubated in anaerobic conditions exhibited elongated and hyphal morphology, including clubbed-ends, swollen cells, and cell lysis throughout all stages of growth comparable to the observed morphologies of XH001/TM7x co-culture experiencing nutrient-depleted environments under microaerophilic condition (Fig. 3d(a–d)).

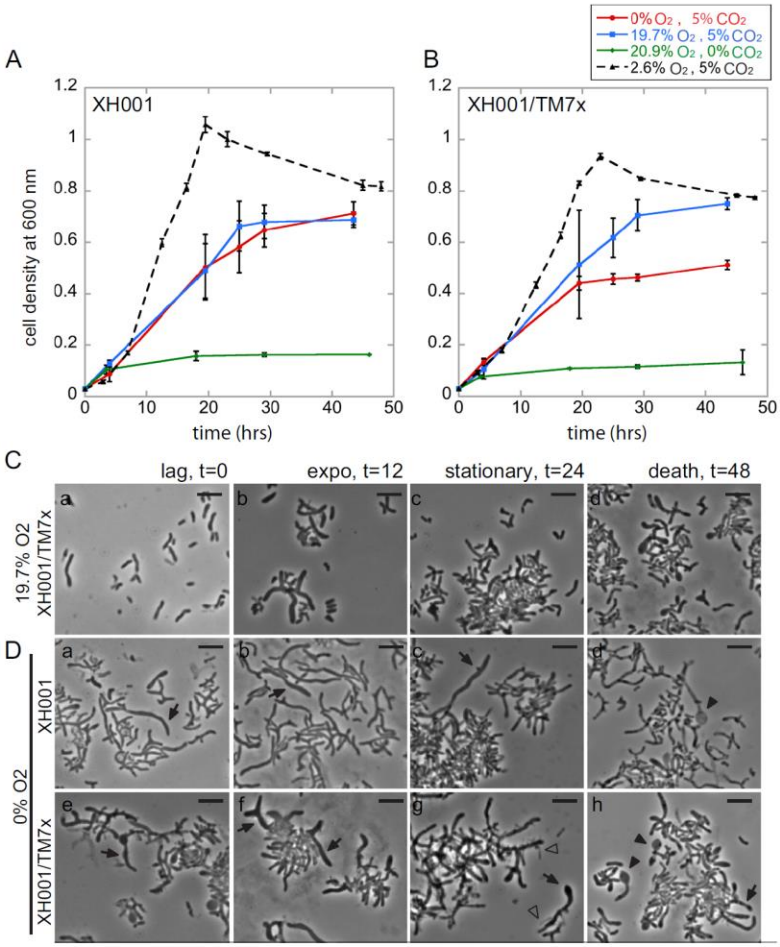


Fig. 3 Oxygen depletion-induced morphological changes in XH001 monoculture and co-cultures. **a**, **b** Cell density measurements of XH001 mono (**a**) or co-culture (**b**) under microaerophilic (2.6 % O_2, 5 % CO_2, *triangles*, *black broken line*), high oxygen (19.7 % O_2, 5 % CO_2, *squires*, *blue solid line*), atmospheric (20.9 % O_2, 0.04 % CO_2, *diamonds*, *green solid line*), or anaerobic (0 % O_2, 5 % CO_2, *circles*, *red solid lines*) conditions. The *dashed lines* are the same data points from Fig. 1. Each point represents the average of three independent cultures (*error bars*, SD). **c**, **d** Phase contrast images of cells grown under indicated conditions (see also Fig. S2). **c**, *a–d* Co-culture grown under high oxygen condition for indicated amount of time. XH001 alone (**d**, *a–d*) or with TM7x (**d**, *e–h*) grown under anaerobic condition for indicated amount of time. In these images, we can clearly see clubbed-ended (*filled arrow heads*) and swollen (*arrows*) XH001 cells as well as elongated TM7x cells (*open arrow heads*). All *scale bars* indicate 5 μm

However, these morphological alterations were reversible, switching from short rod XH001 cells to long and branched, then back to short rod morphology when oxygen levels switched from 2.6 to 0 % and back to 2.6 %, respectively (Fig. S2d–f). The TM7x-associated XH001 cells incubated in anaerobic condition also had very similar morphology to XH001 monoculture (Fig. 3d(e–h)). These results indicate that the depletion of oxygen induces a cellular response in XH001 similar to the response observed when associated with TM7x.

Association with TM7x and Depletion of Oxygen Induces Stress Response in XH001 in an Additive Manner

Several lines of evidence suggest that the observed morphological changes in XH001 during its symbiotic growth with TM7x is likely due to a stress response: (1) increase in cell length is generally correlated with a sub-lethal stress response [54]; (2) our previous transcriptomic study and live/dead staining suggest that TM7x is parasitic to XH001, and may cause cell stress and a decrease in viability [14]; and (3) there is a correlation between cell elongation/branching and decreased growth rate. To test this, we performed targeted qRT-PCR to monitor the mRNA levels of five stress-related genes that were up-regulated in the previous transcriptomic data (see Materials and Methods) [14]. qRT-PCR analysis revealed increased expression of *kefA*, which encodes part of the potassium efflux system, heat shock protein 60 [*groES*], and universal stress protein family in XH001 during symbiotic growth with TM7x under microaerophilic conditions. The expression of genes encoding Chaperone protein DnaJ and putative NADPH-quinone reductase was not changed compared to XH001 monoculture (Fig. 4). Meanwhile, the XH001 monoculture under anaerobic conditions showed even higher mRNA expression levels of all genes except NADPH-quinone reductase compared to co-culture, suggesting that stress induced by oxygen depletion is more severe than that caused by establishing an epiparasitic relationship with TM7x under microaerophilic conditions. Most intriguingly, when serving as basibiont for TM7x and incubated in an anaerobic condition, XH001 displayed the most drastic increase in all stress response genes evaluated, with several genes increasing 1000-fold or higher, indicating an additive effect of these

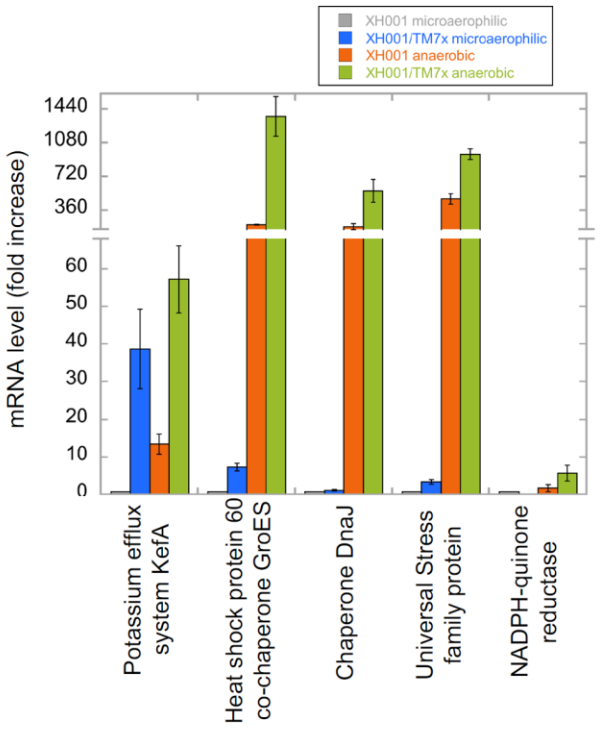


Fig. 4 Stress response of XH001 under TM7x-associated and/or anaerobic conditions. We tested mRNA levels of five stress genes using qRT-PCR in XH001 grown with TM7x under microaerophilic (*blue*) or anaerobic conditions (*green*), or XH001 alone grown under anaerobic conditions (*orange*), and compared that to XH001 monoculture grown under microaerophilic conditions (*gray*). Each *bar* represents the average of three independent experiments (*error bars*, SD)

conditions. These results suggest that the XH001 morphological changes due to TM7x binding or depletion of oxygen are likely a result of a cellular stress response.

Morphological Changes in TM7x During Epibiotic Growth with XH001

The low resolution of the phase contrast microscope allowed us to characterize the morphology of XH001, but not TM7x. To monitor the possible morphological changes in TM7x, which are ultrasmall cocci with a diameter of 200–300 nm (measured previously by transmission electron microscope (TEM)) during early symbiotic growth with XH001 [14], we employed fluorescence in situ hybridization (FISH) staining followed by fluorescence confocal microscopy. To distinguish between XH001 and TM7x, we developed fluorescently tagged XH001- and TM7x-specific FISH probes targeting the 16S rRNA gene. Using these probes, we visualized the morphology of TM7x associated with XH001 over different growth phases under microaerophilic conditions. As expected, XH001 monoculture cells exhibited short rod morphology when stained with an XH001-specific probe throughout all stages of growth (Fig. S3a, only 24 h shown). We did not observe non-specific staining of XH001 monoculture with TM7x-specific probes (Fig. S3a). During lag, exponential, and stationary phase of the co-culture, TM7x mostly appeared as cocci with a diameter of 0.5–1 μm, although we also observed a slightly elongated form of TM7x (1–2 μm) that is much smaller than XH001 (Fig. 5a–c). Cocci cell diameter, however, is different from our previous measurements as well as measurements through SEM imaging (200–300 nm, see following section). The observed discrepancy is likely due to the low-quality resolution of the FISH images. Consistent with our phase contrast images, TM7x-associated XH001 frequently displayed elongated and hyphal morphology except during exponential phase, where short rod morphology dominated (Fig. 5b). During death phase, in addition to cocci and short rods, we also observed elongated TM7x cells that were 2–5 μm in length (Fig. 5d). These elongated forms of TM7x were also enriched in the XH001/TM7x co-culture grown in anaerobic conditions (Figs. 3d(g, h), 5e, and S3c). Our FISH imaging does not clearly indicate whether these are chains of multiple cells or a single, elongated TM7x cell. Similar to the phase contrast images, XH001 in monoculture and co-culture under anaerobic conditions demonstrated elongated and hyphal morphology, as well as clubbed-ends and swollen cell shape (Fig. S3b, c). Moreover, FISH imaging showed non-uniform staining of XH001, possibly due to leakage of cytoplasmic content, which supports the hypothesis that XH001 is highly stressed in an anaerobic environment (Fig. S3b, c).

Fluorescent images also revealed that individual TM7x cells decorating XH001 could be classified into distinct subsets based on their morphology. Specifically, TM7x morphology ranged from cocci (Fig. 6a), cocci with various tail lengths (Fig. 6b–d), two connected cocci (Fig. 6e), or two slightly separated cocci (Fig. 6f). These distinct morphologies can be observed during all growth phases and suggest that TM7x may undergo bud formation and growth.

Morphology of XH001 and TM7x Under Scanning Electron Microscopy

To acquire more detailed knowledge of the morphological dynamics of XH001 and TM7x, SEM was used to visualize XH001 and TM7x. SEM images confirmed the short rod appearance of XH001 in monoculture as well as the more elongated and branched morphology of XH001 in co-culture that is heavily decorated with TM7x (Fig. 7a, b). XH001 in monoculture and co-culture in a microaerophilic and nutrient-rich condition displayed smooth cell surfaces without any appendages (pili and/or flagella). Similar to our FISH images, SEM revealed TM7x cells with different morphologies, such as short rods as well as elongated cells. Notably, the TM7x morphology resembled budding yeast or budding bacteria (Fig. 7b inset) [55, 56], supporting our speculation that TM7x growth occurs by budding. Furthermore, when TM7x cells were isolated from co-culture using a 0.22-μm filter, SEM images

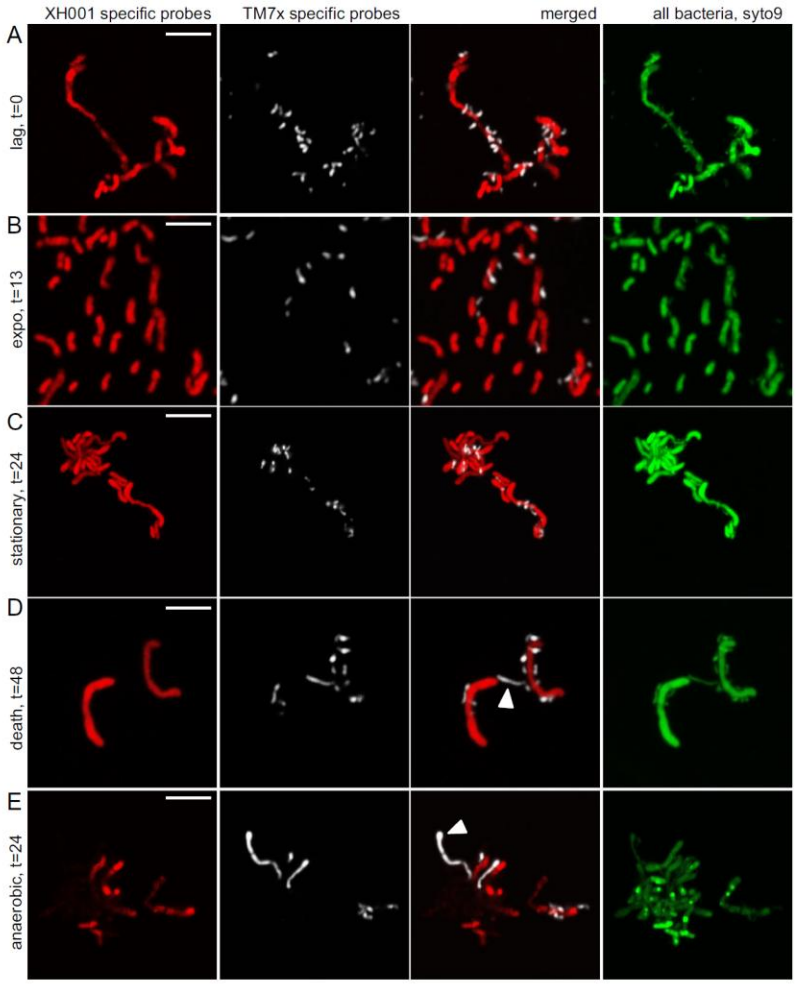


Fig. 5 Morphology of TM7x during growth. Co-culture grown under microaerophilic conditions was FISH stained using TM7x-specific (*white*, TM7567) and XH001-specific (*red*, M33910) DNA-probes that target the 16S rRNA gene (see Materials and Methods). *Green* is a universal DNA dye, syto9, which stains both XH001 and TM7x DNA and RNA. XH001 monoculture staining is shown in Fig. S3. At time points 0 (**a**), 24 (**c**), and 48 h (**d**), XH001 had long and hyphal morphology. At 13 h (**b**), many of the XH001 cells were short rods. TM7x also assumed different morphologies. At time points 0–24 h (**a**–**c**), TM7x appeared as small cocci or short rods. At 48 h (**d**), the elongated form of TM7x was observed more frequently (*white arrow head*). **e** Interestingly, TM7x also assumed an elongated morphology when the co-culture was incubated under anaerobic conditions (*white arrowhead*). Long TM7x can be observed throughout all time points of the growth but only *t*= 24 h is shown in (**e**). All *scale bars* indicate 5 μm

revealed individual, non-XH001 attached TM7x with a coccoid morphology with a diameter of ~200 nm (Fig. 7c), confirming our previous measurements with TEM sections [14]. Cells grown in anaerobic conditions exhibited aggregation under SEM, consistent with our phase contrast and FISH images (Fig. 7d). Nevertheless, we were still able to visualize the elongated morphologies of XH001 and TM7x. Some of the XH001 cells appeared contracted with pitted surfaces (data not shown), reflecting a potential stressed state. Similar "stress-related" cell surface morphologies have been reported in other bacteria, but with less severe changes in the membrane surface [57, 58]. Overall, the SEM images validated our observation of elongated and branched XH001, and further revealed that stressed XH001 has uneven surfaces.

Discussion

Using novel cultivation methods, a member of the phylum Candidatus Saccharibacteria (formerly TM7) was recently isolated from the human oral cavity (designated Candidatus Saccharibacteria oral taxon TM7x) as an obligate epibiont parasite growing on the surface of its host bacterial species *A. odontolyticus* (XH001) [14]. In this study, we further reveal the dynamic interaction between the two partners, as reflected by their reciprocal morphological changes and induction of stress genes during co-culture growth. Elongation and cell branching were the most substantial morphological changes in XH001 induced by the symbiotic association of TM7x. Previous studies have indicated that *Actinomyces* spp. assumes different morphologies ranging from short curved rods to filamentous hyphal cells. Other morphologies include, but are not limited to, clubbed-ends, diphtheroid arrangements, short chains, small clusters such as spider colonies, and coccoid shape [53]. These cell morphologies are thought to be dynamic and determined by different environmental conditions such as growth media and temperature [53]. Our findings are consistent with these studies to the extent that XH001 can take different cell morphologies when experiencing different growth conditions. More importantly, we determined that the observed

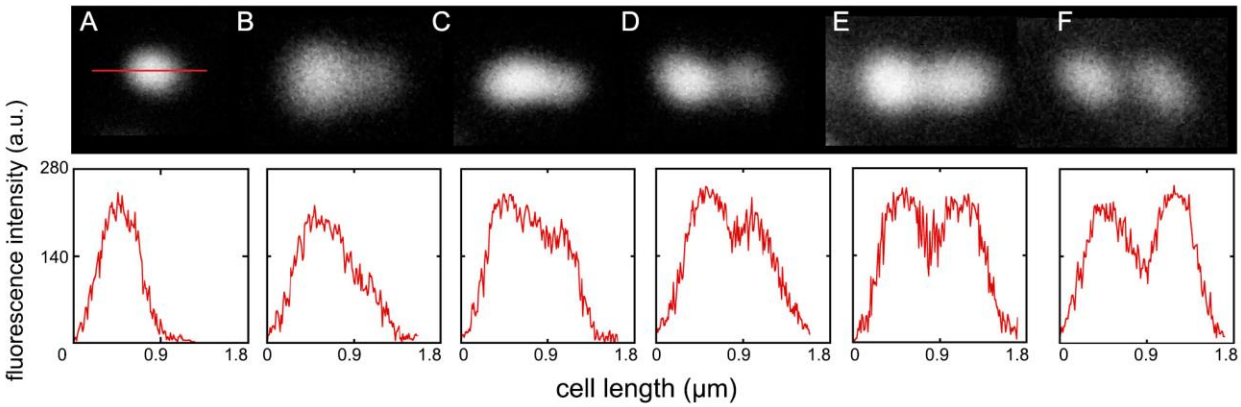


Fig. 6 Bud-like morphology of TM7x. **a**–**f** Magnified images of TM7x from the co-culture stained with a TM7x-specific FISH probe. Representative images were selected from FISH staining of co-culture grown under microaerophilic conditions. These images are not a time series, but rather a collection of different cells from the FISH imaging. XH001 staining is not shown. Different stages of bud growth can be seen throughout all time points of the co-culture growth. These images were further analyzed by graphing the length (x-axes) versus fluorescent intensity of Cy-5-TM7567 (y-axes) along the *red line* drawn across the budding cells (see *red line* in **a**). **a** Representative image of cocci TM7x is shown. **b** Very beginning of bud formation is shown by one large coccus with a small tail-like formation. **c** Further bud development is evidenced by a clearer tail-like formation. **d** Two connected cells are represented. One cell is smaller than the other, possibly formed from the tail-like structure. **e** Two cells that have clear separation but still connected by weak fluorescent signal. **f** Finally, two cells that are separated and have clear boundaries. The *scale bar* in the images are not displayed since the graph is accurately showing the cell length

transformation from short rods to elongated and branched morphology in XH001 is likely a result of a cellular stress response induced by TM7x, as revealed by our targeted qRT-PCR analysis (Figs. 4 and 8a). To our knowledge, this is the first evidence that cell stress is linked to elongated and hyphal morphology of *Actinomyces* species. Intriguingly, these morphological changes, as well as up-regulation of stress-related gene expression, can be recapitulated in a heightened manner by incubating XH001 monoculture in anaerobic conditions (Fig. 8b). This is an indication that TM7x and anaerobic environments may induce similar morphological changes in XH001 as a result of the same cellular stress response. Supporting this hypothesis, association with TM7x and cultivation under anaerobic conditions resulted in an additive effect on the up-regulation of stress response genes in XH001 (Fig. 8c). However, the molecular mechanisms by which TM7x association or anaerobic environment induces the stress response have yet to be elucidated.

In many bacterial species, one of the main cellular responses to external insults or internal stress is elongation of the cell body [54]. In the current study, we reveal a novel culprit for induction of the stress response, XH001's obligate epibiotic

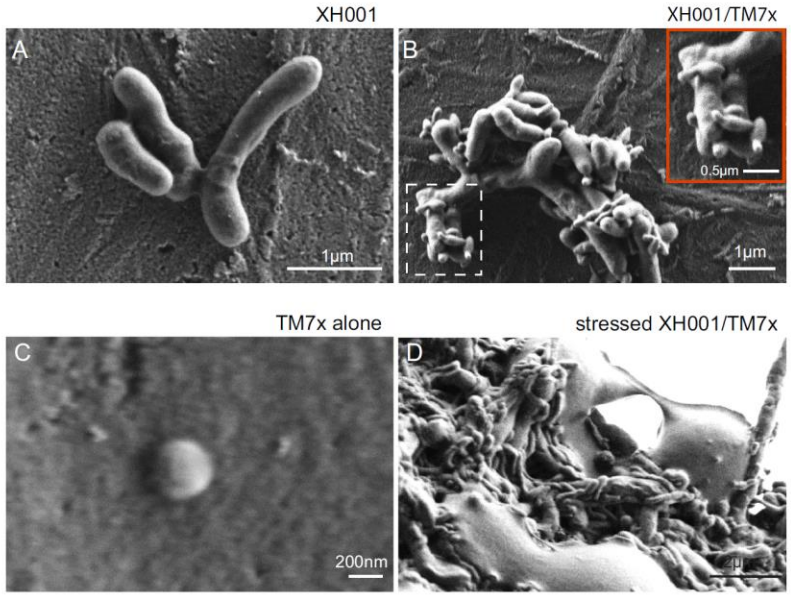


Fig. 7 Morphology of XH001 and TM7x under SEM. **a** XH001 monoculture grown under microaerophilic conditions shows short rod morphology with smooth surfaces. **b** Co-culture grown under microaerophilic conditions reveals elongated and branched XH001 cells decorated with smaller cells, presumably TM7x. Inset shows a zoomed-in image of the smaller TM7x (*box with dashed line*) that seems to be budding. **c** Isolated TM7x using 0.22-μm filter shows cocci TM7x with diameter of ~200 nm. **d** Co-culture grown under anaerobic conditions shows elongated XH001 decorated with many elongated TM7x

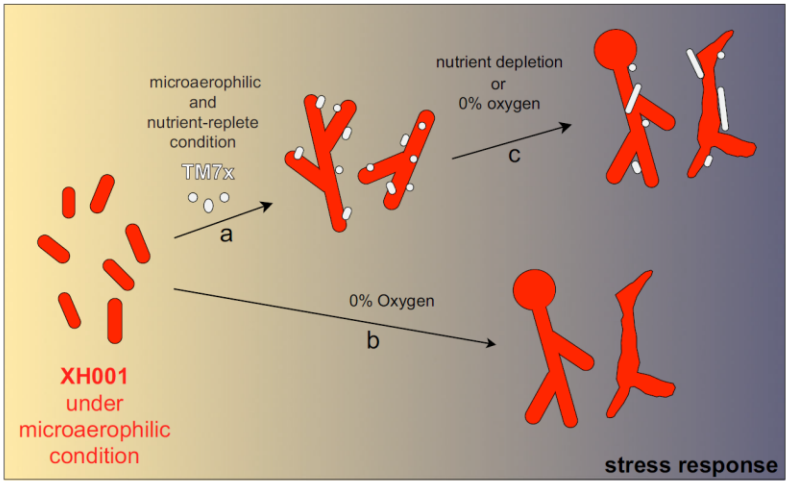


Fig. 8 Diagram depicting morphological changes of XH001 and TM7x under different conditions. XH001 and TM7x bacteria are shown in *red* and *gray* respectively. *Arrows* represent the changes in condition of the culture environment: **a** attachment of TM7x to XH001 under microaerophilic nutrient-replete environment, **b** incubation of XH001 alone under oxygen deplete environment, and **c** co-culture under nutrient or oxygen deplete condition. Increasing dark color in the background represents increase in stress response genes

partner, TM7x. Since TM7x is an epibiotic parasite that exists on the surface of its basibiont, we speculate that the induction of the stress response in XH001 that results in cell elongation allows TM7x greater surface area in which to grow. Under nutrition-replete conditions, TM7x causes moderate elongation and branching in XH001 presumably by keeping its basibiont relatively healthy (Fig. 8a). The balance between achieving greater surface area for growth and maintaining basibiont viability plays an important role in determining the success of the epibiotic relationship. Under nutrition-depleted conditions, however, the balance shifts, with TM7x causing severe cell morphology and cell lysis in XH001, similar to that of a predatory bacteria (Fig. 8c). In plant symbiosis, the terms biotroph (a parasite that depends on the host's survival), necrotroph (a parasite that kills the host), and hemi-biotrophs (a parasite that displays characteristics of both) are used regularly [59]. These terms can be applied to bacterial symbiosis, as we see the relationship between TM7x and XH001 shift from biotrophic (early nutrient-replete condition) to necrotrophic (late nutrient deplete condition), which is in contrast to predatory bacteria in which the interaction is mainly dominated by a necrotrophic relationship. It is fascinating that biotrophic and necrotrophic symbiosis not only exists in higher organisms but also among the most basic life forms, bacteria. Future studies are needed to explore this dynamic interaction between TM7x and XH001 at the molecular and genetic level.

TM7x grows dependently on the surface of basibiont XH001 and has different morphologies, such as cocci and a filamentous cell body. More intriguingly, our data suggest that TM7x might divide by budding. The FISH and SEM images clearly reveal TM7x morphologies that may reflect different budding stages (Figs. 6 and 7). Without analyzing large number of cells, however, we could not determine if the different morphologies are correlated with specific growth phases. Previous studies show that alpha-proteobacteria *Hyphomonas*, *Pedomicrobium*, and *Ancalomicrobium*, as well as *Planctomyces maris* grow by budding [56], although they are all free-living bacteria. In contrast, TM7x seems to bud while attached to XH001. Future studies are needed to understand the mechanism of budding and the physical association of TM7x with XH001. In addition, we also observed an elongated form of TM7x. This morphology was observed most frequently when the co-culture was incubated under microaerophilic conditions for long periods of time (nutrient-depleted) or incubated under anaerobic conditions (Fig. 8c), suggesting that elongation of TM7x may also be a result of a stress response.

Both *A. odontolyticus* and TM7 are known members of the healthy bacterial flora of the oral cavity with increased abundance in certain disease conditions. The intimate interaction between TM7x and XH001 demonstrated in our study offers a glimpse into additional ways bacteria may interact with each other within the oral cavity and possibly in other body sites where TM7 members are detected. This unique interaction provides a great model system for understanding the intricacies of bacteria-to-bacteria hemi-biotrophic symbiosis and the impact of this interaction relative to bacterial ecologically induced human diseases.

Acknowledgments We thank the members of the Shi and Lux laboratories for their feedback and invaluable discussion. We also thank Melissa Agnello for providing extensive editing of the manuscript. We thank the Chemistry and Biochemistry instrumentation facility at UCLA for providing access to the confocal microscope. This work was supported in part by grants from the National Institutes of Health (1R01DE023810-01) and Oral Health-Research Postdoctoral Training Program (B.B., UCLA School of Dentistry T90 award).

Compliance with Ethical Standards

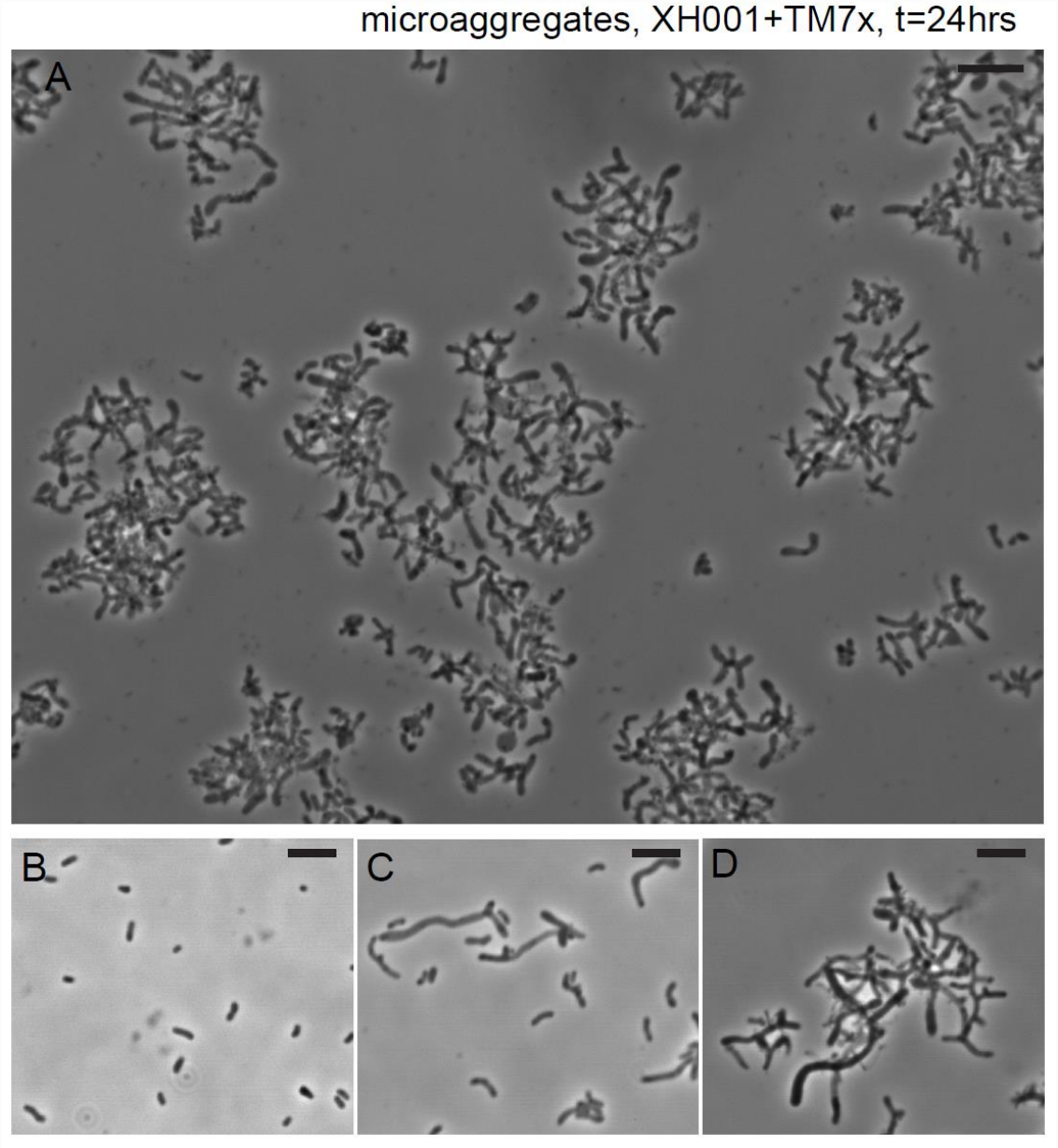
Conflict of Interest The authors declare that they have no conflict of interests.

References

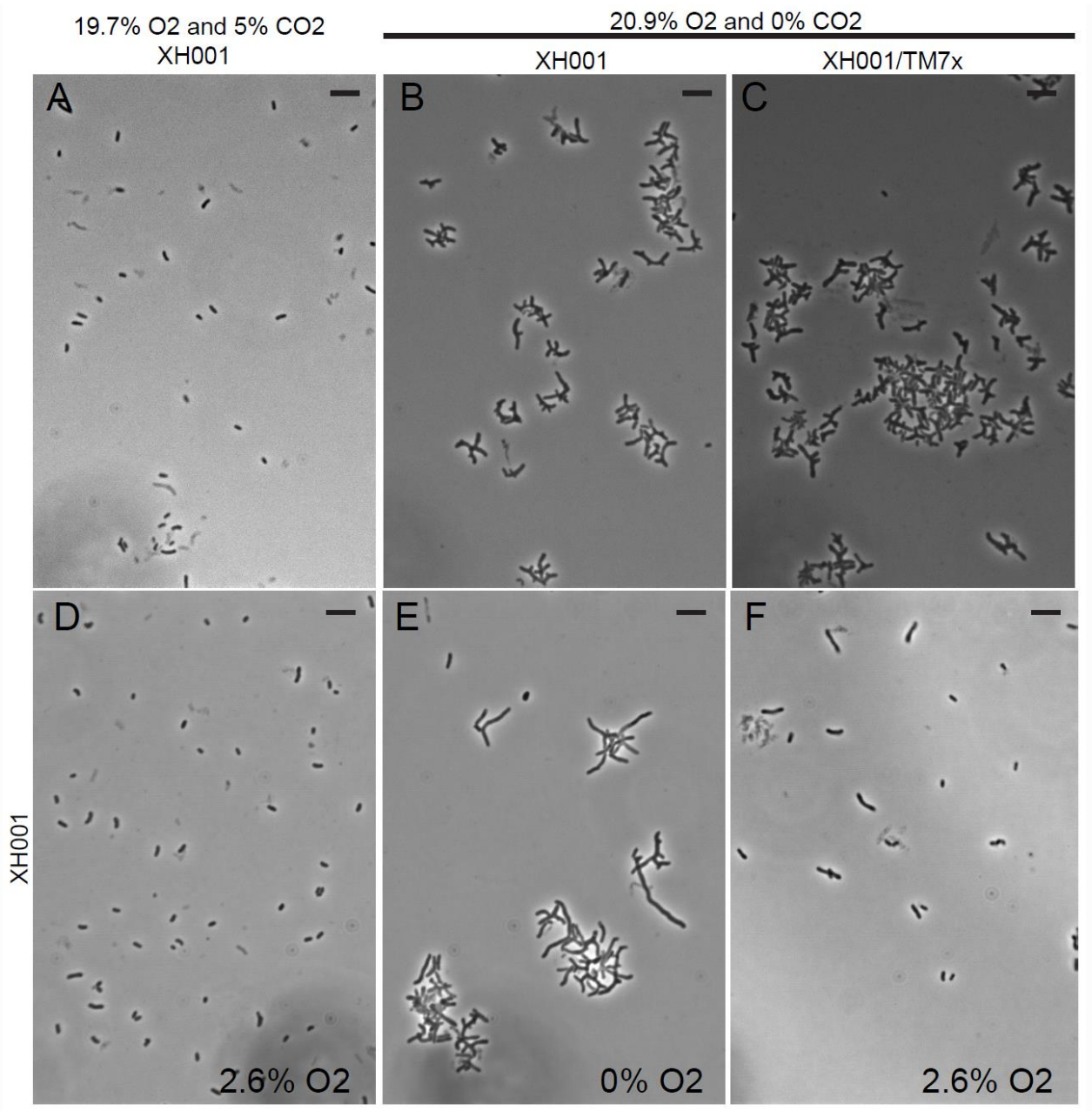
1. Joseph S (2002) Symbiosis: mechanisms and model systems. Kluwer Academic Publisher, New York
2. Guerrero R, Pedros-Alio C, Esteve I et al (1986) Predatory prokaryotes: predation and primary consumption evolved in bacteria. Proc Natl Acad Sci 83:2138–2142. doi:10.1073/pnas.83.7.2138
3. Huber H, Hohn MJ, Rachel R et al (2002) A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature 417:63–67. doi:10.1038/417063a
4. Lambina VA, Afinogenova AV, Romaĭ Penabad S et al (1982) *Micavibrio admirandus* gen. et sp. nov. Mikrobiologiia 51:114–117
5. Grice EA, Segre JA (2012) The human microbiome: our second genome. Annu Rev Genomics Hum Genet 13:151–170. doi:10.1146/annurev-genom-090711-163814
6. Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. Nature 486:207–214. doi:10.1038/nature11234
7. Roesch LFW, Fulthorpe RR, Riva A et al (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. ISME J 1:283–290. doi:10.1038/ismej.2007.53
8. Schauer R, Bienhold C, Ramette A, Harder J (2010) Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. ISME J 4:159–170. doi:10.1038/ismej.2009.106
9. Bryant MP, Wolin EA, Wolin MJ, Wolfe RS (1967) *Methanobacillus omelianskii*, a symbiotic association of two species of bacteria. Arch Für Mikrobiol 59:20–31
10. Lancy P, Dirienzo JM, Appelbaum B et al (1983) Corncob formation between *Fusobacterium nucleatum* and *Streptococcus sanguis*. Infect Immun 40:303–309
11. Overmann J (2010) The phototrophic consortium "*Chlorochromatium aggregatum*"—a model for bacterial heterologous multicellularity. Adv Exp Med Biol 675:15–29. doi:10.1007/978-1-4419-1528-3_2
12. Davidov Y, Huchon D, Koval SF, Jurkevitch E (2006) A new *Alpha-proteobacterial* clade of *Bdellovibrio*-like predators: implications for the mitochondrial endosymbiotic theory. Environ Microbiol 8:2179–2188. doi:10.1111/j.1462-2920.2006.01101.x
13. Dashiff A, Junka RA, Libera M, Kadouri DE (2011) Predation of human pathogens by the predatory bacteria *Micavibrio aeruginosavorus* and *Bdellovibrio bacteriovorus*. J Appl Microbiol 110:431–444. doi:10.1111/j.1365-2672.2010.04900.x
14. He X, McLean JS, Edlund A et al (2015) Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. Proc Natl Acad Sci U S A 112:244–249. doi:10.1073/pnas.1419038112
15. Lasken RS, McLean JS (2014) Recent advances in genomic DNA sequencing of microbial species from single cells. Nat Rev Genet 15:577–584. doi:10.1038/nrg3785
16. Albertsen M, Hugenholtz P, Skarshewski A et al (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat Biotechnol 31:533–538. doi:10.1038/nbt.2579
17. Marcy Y, Ouverney C, Bik EM et al (2007) Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. Proc Natl Acad Sci 104:11889–11894. doi:10.1073/pnas.0704662104
18. Podar M, Abulencia CB, Walcher M et al (2007) Targeted access to the genomes of low-abundance organisms in complex microbial communities. Appl Environ Microbiol 73:3205–3214. doi:10.1128/AEM.02985-06
19. Brown CT, Hug LA, Thomas BC et al (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. Nature. doi:10.1038/nature14486
20. Luef B, Frischkorn KR, Wrighton KC et al (2015) Diverse uncultivated ultra-small bacterial cells in groundwater. Nat Commun 6:6372. doi:10.1038/ncomms7372
21. Bik EM, Eckburg PB, Gill SR et al (2006) Molecular analysis of the bacterial microbiota in the human stomach. Proc Natl Acad Sci 103:732–737. doi:10.1073/pnas.0506655103
22. Dewhirst FE, Chen T, Izard J et al (2010) The human oral microbiome. J Bacteriol 192:5002–5017. doi:10.1128/JB.00542-10
23. Dinis JM, Barton DE, Ghadiri J et al (2011) In search of an uncultured human-associated TM7 bacterium in the environment. PLoS One 6:e21280. doi:10.1371/journal.pone.0021280
24. Eckburg PB, Bik EM, Bernstein CN et al (2005) Diversity of the human intestinal microbial flora. Science 308:1635–1638. doi:10.1126/science.1110591
25. Ferrari BC, Binnerup SJ, Gillings M (2005) Microcolony cultivation on a soil substrate membrane system selects for previously uncultured soil bacteria. Appl Environ Microbiol 71:8714–8720. doi:10.1128/AEM.71.12.8714-8720.2005
26. Fredricks DN, Fiedler TL, Marrazzo JM (2005) Molecular identification of bacteria associated with bacterial vaginosis. N Engl J Med 353:1899–1911. doi:10.1056/NEJMoa043802
27. Gao Z, Tseng C, Pei Z, Blaser MJ (2007) Molecular analysis of human forearm superficial skin bacterial biota. Proc Natl Acad Sci 104:2927–2932. doi:10.1073/pnas.0607077104
28. Hanada A, Kurogi T, Giang NM et al (2014) Bacteria of the candidate phylum TM7 are prevalent in acidophilic nitrifying sequencing-batch reactors. Microbes Environ JSME 29:353–362. doi:10.1264/jsme2.ME14052
29. Hugenholtz P, Tyson GW, Webb RI et al (2001) Investigation of candidate division TM7, a recently recognized major lineage of the domain Bacteria with no known pure-culture representatives. Appl Environ Microbiol 67:411–419. doi:10.1128/AEM.67.1.411-419.2001
30. Kianoush N, Adler CJ, Nguyen K-AT et al (2014) Bacterial profile of dentine caries and the impact of pH on bacterial population diversity. PLoS One 9:e92940. doi:10.1371/journal.pone.0092940
31. Kuehbacher T, Rehman A, Lepage P et al (2008) Intestinal TM7 bacterial phylogenies in active inflammatory bowel disease. J Med Microbiol 57:1569–1576. doi:10.1099/jmm.0.47719-0
32. Paster BJ, Boches SK, Galvin JL et al (2001) Bacterial diversity in human subgingival plaque. J Bacteriol 183:3770–3783. doi:10.1128/JB.183.12.3770-3783.2001
33. Rheims H, Spröer C, Rainey FA, Stackebrandt E (1996) Molecular biological evidence for the occurrence of uncultured members of the *Actinomycete* line of descent in different environments and geographical locations. Microbiol Read Engl 142(Pt 10):2863–2870. doi:10.1099/13500872-142-10-2863
34. Soro V, Dutton LC, Sprague SV et al (2014) Axenic culture of a *Candidate* division TM7 bacterium from the human oral cavity and biofilm interactions with other oral bacteria. Appl Environ Microbiol 80:6480–6489. doi:10.1128/AEM.01827-14
35. Brinig MM, Lepp PW, Ouverney CC et al (2003) Prevalence of bacteria of division TM7 in human subgingival plaque and their association with disease. Appl Environ Microbiol 69:1687–1694. doi:10.1128/AEM.69.3.1687-1694.2003
36. Kumar PS, Griffen AL, Barton JA et al (2003) New bacterial species associated with chronic periodontitis. J Dent Res 82:338–344

37. Liu B, Faller LL, Klitgord N et al (2012) Deep sequencing of the oral microbiome reveals signatures of periodontal disease. PLoS One 7:e37919. doi:10.1371/journal.pone.0037919
38. Paster BJ, Russell MK, Alpagot T et al (2002) Bacterial diversity in necrotizing ulcerative periodontitis in HIV-positive subjects. Ann Periodontol Am Acad Periodontol 7:8–16. doi:10.1902/annals.2002.7.1.8
39. Rylev M, Bek-Thomsen M, Reinholdt J et al (2011) Microbiological and immunological characteristics of young Moroccan patients with aggressive periodontitis with and without detectable *Aggregatibacter actinomycetemcomitans* JP2 infection. Mol Oral Microbiol 26:35–51. doi:10.1111/j.2041-1014.2010.00593.x
40. Becker MR, Paster BJ, Leys EJ et al (2002) Molecular analysis of bacterial species associated with childhood caries. J Clin Microbiol 40:1001–1009
41. Colombo AV, Silva CM, Haffajee A, Colombo APV (2006) Identification of oral bacteria associated with crevicular epithelial cells from chronic periodontitis lesions. J Med Microbiol 55:609–615. doi:10.1099/jmm.0.46417-0
42. Kanasi E, Dewhirst FE, Chalmers NI et al (2010) Clonal analysis of the microbiota of severe early childhood caries. Caries Res 44:485–497. doi:10.1159/000320158
43. Ling Z, Kong J, Jia P et al (2010) Analysis of oral microbiota in children with dental caries by PCR-DGGE and barcoded pyrosequencing. Microb Ecol 60:677–690. doi:10.1007/s00248-010-9712-8
44. Nagy KN, Sonkodi I, Szöke I et al (1998) The microflora associated with human oral carcinomas. Oral Oncol 34:304–308
45. Sato T, Watanabe K, Kumada H et al (2012) Peptidoglycan of *Actinomyces naeslundii* induces inflammatory cytokine production and stimulates osteoclastogenesis in alveolar bone resorption. Arch Oral Biol 57:1522–1528. doi:10.1016/j.archoralbio.2012.07.012
46. Aas JA, Paster BJ, Stokes LN et al (2005) Defining the normal bacterial flora of the oral cavity. J Clin Microbiol 43:5721–5732. doi:10.1128/JCM.43.11.5721-5732.2005
47. Paul D, Reddy D, Mukherjee D et al (2011) Actinomyces. In: Liu D (ed) Mol Detect Hum Bact. Pathog. CRC Press, p 23–30
48. Jones E, Oliphant E, Peterson P et al (2001) SciPy: open source scientific tools for python. http://www.scipy.org/
49. Walt S, Schonberger JL, Nunez-Iglesias J et al (2014) Scikit-image: image processing in pythong. PeerJ 2:e453. doi:10.7717/peerj.453
50. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29:e45
51. MJ Hill, PD Marsh (1990) Factors controlling the microflora of the huealthy mouth. In: Hum Microb Ecol. CRC Press. Inc., p 34–37
52. Loesche WJ, Gusberti F, Mettraux G et al (1983) Relationship between oxygen tension and subgingival bacterial flora in untreated human periodontal pockets. Infect Immun 42:659–667
53. Schaal KP, Yassin AA (2012) Genus I. *Actinomyces*. In: Bergeys Man Syst Bacteriol. Springer, p 42–109
54. Jones TH, Vail KM, McMullen LM (2013) Filament formation by foodborne bacteria under sublethal stress. Int J Food Microbiol 165:97–110. doi:10.1016/j.ijfoodmicro.2013.05.001
55. Herskowitz I (1988) Life cycle of the budding yeast *Saccharomyces cerevisiae*. Microbiol Rev 52:536–553
56. Hirsch P (1974) Budding bacteria. Annu Rev Microbiol 28:391–440
57. Greif D, Wesner D, Regtmeier J, Anselmetti D (2010) High resolution imaging of surface patterns of single bacterial cells. Ultramicroscopy 110:1290–1296. doi:10.1016/j.ultramic.2010.06.004
58. Patil S, Valdramidis VP, Karatzas KAG et al (2011) Assessing the microbial oxidative stress mechanism of ozone treatment through the responses of *Escherichia coli* mutants. J Appl Microbiol 111:136–144. doi:10.1111/j.1365-2672.2011.05021.x
59. Glazebrook J (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu Rev Phytopathol 43:205–227. doi:10.1146/annurev.phyto.43.040204.135923

Supplementary Information



Supplementary Figure 1 | TM7x induces morphological changes in XH001. a XH001/TM7x co-culture grown under microaerophilic condition for 24 h showing clear micro-aggregation. Scale bar is 10 µm. **b** Monoculture of XH001. **c, d** Establishment of physical association between XH001 and TM7x via attachment assay (see supplementary methods). The re-attached XH001 cells were passaged two times (c) and four times (d) respectively in fresh medium. Scale bars are 5 µm (PDF 4683 kb)



Supplementary Figure 2 | Morphology of XH001 under different oxygen conditions. Phase contrast image of XH001 alone a under high oxygen condition (19.7 % O2, 5 % CO2). Phase contrast images of XH001 alone **b** and with TM7x **c** under normal atmospheric condition (20.9 % O2, 0.04 % CO2) after 24 h. d–f XH001 alone cells grown in microaerophilic condition **d** were shifted to anaerobic condition **e** and then back to the microaerophilic condition **f** before taking the phase contrast images. All scale bars indicate 5 μm (PDF 8569 kb)

Supplementary Figure 3 | FISH staining of XH001 alone and with TM7x. FISH probes specific to TM7x (*white*) and XH001 (*red*) were used to stain the fixed samples. *Green* represents syto9 staining of all bacteria. **a** XH001 monoculture grown under a microaerophilic condition for 24 h shows short rod morphology with XH001-specific probe (*red*) and universal DNA stain syto9 (*green*), but no staining with TM7x-specific probe, confirming our probe specificity. **b**, **c**XH001 alone (**b**) and with TM7x (**c**) grown under anaerobic condition for 24 h. Similar to **a**, we do not see any staining of TM7x probe in the XH001 alone cells, whereas in the co-culture, we saw elongated TM7x. Under anaerobic condition, XH001-specific probe stained the cells non-uniformly, suggesting that these cells were stressed and probably lost their cell content. All *scale bars* indicate 10 μm. (PDF 26435 kb

Table S1. qPCR primers targeting 16S and stress genes

Primer name	Primer target	sequence
F5	XH001 16S	5’-GCGGAGCATGCGGATTA-3’
R3	XH001 16S	5’-AACGTGCTGGCAACATAGGG-3’
STRS1F	Potassium efflux system KefA protein	5’-AAGACGTACGCCGTGCTCGTCATC-3’
STRS1R	Potassium efflux system KefA protein	5’-GTGCCGAGTTGAGTCGTCGTTAGC-3’
STRS7F	Potassium uptake protein	5’-CTGATCCTTGCATTCGTGG-3’
STRS7R	Potassium uptake protein	5’-GGACGAGCGCGAGTTAACG-3’
STRS2F	Heat shock protein 60, GroES	5’-AAGGAGAAGCCGCAGGAAG-3’
STRS2R	Heat shock protein 60, GroES	5’-GTCGTACTTGACCTCGGTGC-3’
STRS3F	Chaperone protein DnaJ	5’-ACGGTAAGAAGACCGTGACG-3’
STRS3R	Chaperone protein DnaJ	5’-GAATCTGAACGTCCACGTGC-3’
0867F	Universal stress family protein	5’-GAGGGAATGGATACTGCGAT-3’
0867R	Universal stress family protein	5’-GTGGAGAACTCGATGAGCAG-3’
0069F	NADPH-quinone reductase	5’-ATGTCCAACGTGCTGATTGT-3’
0069R	NADPH-quinone reductase	5’-ACTCCGGATAGAGGTCATCG-3’

Supplementary Methods. This section contains detailed procedure regarding disaggregation of coculture, XH001 cell quantification, re-attachement of TM7x to XH001, isolation of mRNA, construction of cDNA, and qPCR of key stress genes.

Disaggregation of micro-aggregated XH001/TM7x. XH001/TM7x coculture tends to form micro-aggregates and in order to accurately quantify the cell length and branch points, we attempted to disperse the micro-aggregates using different reagents, including complexing agents: EDTA, EGTA, sodium pyrophosphate, sodium citrate and sodium potassium tartrate, sugars: D-glucose, L-arabinose, L-fucose, D-galactose, D-mannose and N-acetyl-glucoseamine, denaturant: urea and formamide, detergents: SDS and Triton x-100, hydrolytic enzymes: proteinase K and lysozyme, and reducing agent: L-cystein. We also tried physical separation including vigorously pipetting or mild sonication. Sonication was carried out by using 60 Sonic Dismembrator (Fisher Scientific) with power output setting between 2-3. Our data demonstrated that only mild sonication resulted in dispersed cells, although we observed that a small amount of cells also lysed during the sonication, particularly the longer and swollen cells (data not shown). Therefore, quantification in Figure 2 is an underestimation of cell length and branching.

CHAPTER 3: Identification and characterization of the major cell envelope proteins in the syntrophic bacterium, *Syntrophomonas wolfei subsp. wolfei str. Goettingen G311*

Nicole Poweleit[1,2], Rachel R. Ogorzalek Loo[3], Joseph A. Loo[3], Z. Hong Zhou[1,2], Robert Gunsalus[1]

[1]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA

[2]Electron Imaging Center for Nanomachines, California Nano Systems Institute, UCLA, Los Angeles (UCLA), Los Angeles, CA 90095, USA

[3]Department of Chemistry and Biochemistry, UCLA, Los Angeles 90095, UCLA, Los Angeles (UCLA), Los Angeles, CA 90095, USA

Abstract

Syntrophomonas wolfei subsp. wolfei str. Goettingen G311 stains as a Gram-negative bacterium; yet it is phylogenetically a member of the Gram-positive Firmicutes and has a complex cell wall ultrastructure which may prevent them from staining Gram-positive. A tryptic digest of the whole cell revealed a number of highly expressed proteins of unknown function with possible transmembrane domains. These candidates were analyzed using bioinformatic tools to determine proteins of related function in other organisms as well as in *S. wolfei* itself. To resolve which of these hypothetical proteins were present in the cell envelope, cell protein fractionations were prepared and analyzed by SDS-page and LC-MS MS mass spectrometry methods. Products of the putative operon of Swol_0141, Swol_0142, and Swol_0143 as well as the protein coded by Swol_0133 were identified as the major cell envelope proteins. By glycostaining, conventional and cryo electron microscopy, we show that these cell envelop proteins form ordered arrays, are not glycosylated and form a multi-layered structure.

Introduction

Syntrophomonas wolfei is a syntroph which grows in association with other microorganisms that use hydrogen and formate in an anaerobic environment[1]. Syntrophy is an essential step in the degradation of natural polymers and accounts for much of the carbon interconversion in methanogenic ecosystems[2,3]. However, under standard biochemical conditions it is an endothermic process and the symbiotic partners of syntrophic organisms are required to consume the reaction end products in the surroundings so as to make this process energetically favorable[4]. In the environment,

this is accomplished by growth alongside a partner, for *S. wolfei* the specific partner is the methanogic archaea *Methanospirillum hungatei*[5], which consumes H_2 and/or formate in its own energetic processes[6].

A metabolic specialist, *S. wolfei* inhabits a narrow niche where it metabolizes four to eight-carbon saturated and unsaturated fatty acids by the β-oxidation pathway[5,7]. The means of acquiring these starting materials as well as the removal of end products from the cell are unknown. *S. wolfei* interacts with its symbiotic partner *M. hungatei* through hydrogen and formate transfer[8] and has been shown to have adhesive properties critical for the formation of syntrophic consortium[9,10]. One might expect the *S. wolfei* cell envelope, of which the surface layer proteins are a part, and its partner to be involved in this relationship. Indeed, several recent studies have shown that external cellular machinery such as flagella[11], pili[12-15], and the cell envelope[16] can be directly involved in electron transport[17,18]. However, *S. wolfei* lacks any of the genes known to be required for the formation of bionanowires, although a flagellum synthesis operon is present in the genome[19]. Studying the cell envelope of *S. wolfei* helps elucidate how the bacteria establishes functional syntrophic relationships.

While the genome study of *S. wolfei* place it in a Gram positive family[19], it was observed to stain Gram negatively[1]. The anomaly between staining and phylogeny, as well as the high abundance of putative S-layer protein (SLP) detected by quantitative proteomics[20] indicate the possibility of a previously undescribed S-layer involved in syntrophic growth. Many archaea and bacteria have S-layer proteins which form a paracrystalline array completely covering the cell[21]. S-layers are shown to have an important role in adhesion[22-25] for symbiosis and pathogeneses and as barriers to entry

to the cell. *S. wolfei* is related to the class Clostridia, including some human pathogens. The S-layer proteins of the Clostridia species are the best studied S-layer proteins amongst bacteria as S-layer proteins play a critical role in host infection[26-30]. Yet relatively little is known about S-layer assembly and function when compared to other cell envelope structures such as flagella, pili, or the peptidoglycan cell wall[21].

The mutualistic interactions between *S. wolfei* and its archaeal partner may provide insight into the pathogenic interactions between Clostridium species and humans. With little known about the surface layer proteins of *S. wolfei*, a general study to determine which genes encode them and general physical properties. To this end, cells were fractionated and subjected to SDS-PAGE, LC MS-MS analysis, and glycostaining. Electron microscopy was employed to image the outer surface of whole cells and cell ghosts and a bioinformatic study was undertaken to determine relative abundance and phylogeny of the S-layer proteins.

Materials and Methods

Cell Culture. *Syntrophomonas wolfei* ssp. wolfei strain Göttingen DSM 2245B was grown anaerobically in basal medium[31] as previously described with 20mM sodium crotonate. 50mL of culture was grown in a 200mL serum bottle, and the headspace was N2/CO2 (80:20 v/v). Cultures were incubated at 37˚C with gentle shaking until an OD 600nm of at least 0.26 was achieved, typically at least three weeks.

SDS-PAGE and glycostain. A 2mL aliquot of the culture was collected by centrifuging at 13,000g for 10 minutes at 4˚C, and the pellet was resuspended in dionized water for a final optical density of 8. The remaining culture (about 48mL) was centrifuged for 10

minutes at 7,500g at 4˚C. The cell pellet was washed in 1mL of TBS (100 mM Tris 50mM NaCl at pH 7.2) buffer 3 times and the washes were pooled and saved. The cell pellet was resuspended in TBS buffer and passed through a French Press 3 times at a pressure of 800psi at room temperature. The lysate was then centrifuged at 13,000g for 10 minutes at 4˚C, the supernatant was saved, and the pellet which contained the cell membrane fraction was resuspended in 1mL TBS. The fractions which were collected were then run on 4-12% NuPage gradient gels (Invitrogen) using a MES buffer per manufactures recommendation. Gels were stained for protein with the Coomassie based Imperial Stain (Thermo Scientific).

For glycostaining, samples were run on 4-12% NuPage gradient gels (Invitrogen) using a MES buffer per manufactures recommendation and stained with SYPRO Ruby protein stain (Invitrogen) and for glycoprotein with Pro-Q Emerald 300 glycoprotein stain (Invitrogen) per the manufactures protocols.

Protease time course. A 50mL *S. wolfei* cell culture was concentrated as described above for SDS-PAGE electrophoresis. Pelleted cells were resuspended in 1mL of TBS and aliquoted out in ten reaction tubes each with 100µL of cell suspension. 2.0 ng/mL of proteinase were added to each tube and the tubes were incubated at 60˚C. Tubes were removed from heat and immediately put on ice at 10 minute intervals. Samples were run on an SDS-PAGE gel as described above and the gel was stained with SYPRO Ruby total protein stain (Invitrogen) as per the manufactures protocol.

In-gel tryptic digestion and mass spectroscopy. Protein bands from Coomassie stained SDS-PAGE gels were digested by trypsin in-gel as previously described[32]. The

proteolytic peptides were injected for analysis by LC-MS/MS on a Thermo-Fisher Q-Exactive.

Informatics analysis and data visualization. Protein super families were determined using BLAST[33,34], and sequence comparisons were done using Clustal Omega[35].

Electron Microscopy. Electron microscopy was performed using an FEI Tecnai 20 microscope at an accelerating voltage of 200kV with a Gatan 2kx2k CCD camera. Whole cell samples were concentrated to an optical density of 8. Ghosts were prepared by 10 freeze/thaw cycles of whole cells in ethanol and dry ice. The lysate was centrifuged at 7,500g for 20 minutes at 4˚C. The pellet was washed once in 1 M NaCl, then 3 more times in deionized water to remove salt.

The ghosts and whole cells were negatively stained with 0.5% uranyl acetate, 4µL of sample was deposited on a carbon coated copper grid and allowed to settle for 1 minute. Excess sample was blotted away and 4µL of stain was added for 1 minute. Excess stain was blotted away and the grids were washed with stain 2 more times to remove any debris.

For cryo electron microscopy (CryoEM) of whole cells, 4µL of sample was applied to a holey carbon quantifoil grid. Grids were manually blotted and plunge frozen into liquid nitrogen cooled liquid ethane. Samples were loaded into a Gatan 626 cryo holder, and imaged on the same FEI Tecnai 20 described above using low dose imaging conditions.

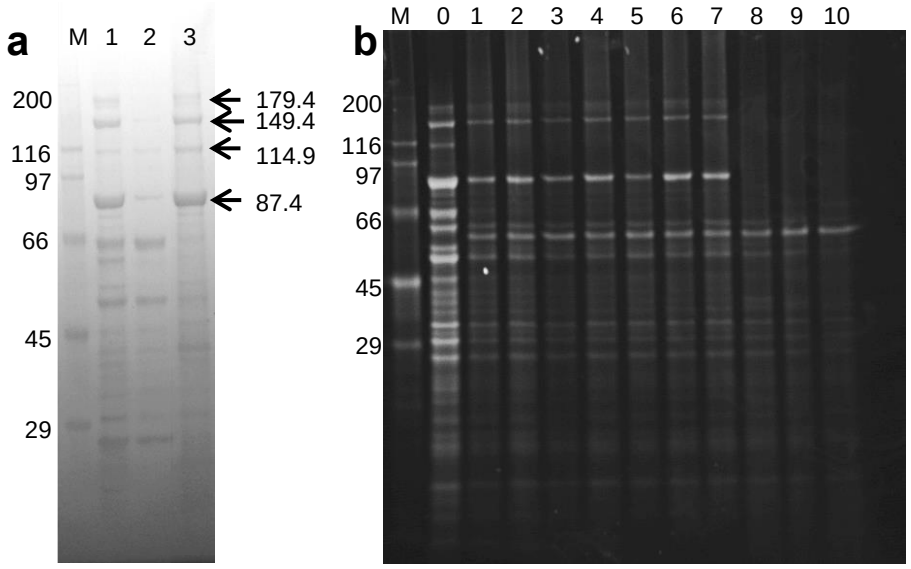


Figure 1 | SDS-PAGE analysis of the *S. wolfei* S-layer. a, SDS-PAGE gel of *S. wolfei* cell fractionation. Lane M is the ladder, Lane 1 is whole cells, lane 2 is the soluble fraction of lysed cells, and lane 3 is the membrane fraction of lysed cells. **b,** SDS-PAGE gel of S. wolfei whole cells treated with proteinase over a time course. Lane M is ladder, lane 0 is the whole cells which were not treated with proteinase, and lanes 1-10 are whole cells treated with proteinase at 60˚C over an increasing period from 10 to 100 minutes.

Results and Discussion

Identification of the major cell surface protein in *S. wolfei*. SDS-PAGE of cell extracts revealed five major bands in the membrane fraction of approximate size of 87.4kDa, 114.9 kDa, 149.4 kDa and a doublet centered at 179.4kDa (Fig. 1a). In order to confirm the localization of these proteins to the outside of the cell, whole cells were exposed to protease over a time course of 100 minutes. The bands corresponding to the membrane fraction disappeared from the SDS-PAGE gel before any of the bands recovered in the soluble fraction, suggesting they are surface exposed (Fig. 1b).

The major protein bands in Figure 1a were excised from the SDS-PAGE gel, in-gel tryptic digested, and identified by tandem mass spectrometry analysis. One standalone gene, Swol_0133 and three genes in a cluster, Swol_0141-0143 were identified to encode the major S-layer components (Fig. 2a). This analysis revealed the four bands at 114.9 kDa, 149.4 kDa, and the 179.4 kDa doublet were a mixture of Swol_0141 and Swol_0143, both genes were annotated in the genome as encoding surface layer proteins based on the presence of signal peptide (SP) and S-layer protein homology domains (Fig. 2bc). Swol_0143 has a predicted size of 141 kDa after removal of the signal peptide, while the Swol_0141 has a predicted size of 75.5 kDa after removal of the signal peptide, making the apparent molecular weight of both proteins larger on the SDS-PAGE gel. The other gene in this cluster, Swol_0142 is only annotated as encoding a 129 amino acid long gene product and does not produce any unique peptides by tryptic digest, making it impossible to confirm the presence of this protein. The major component of the 87.4 kDa band was Swol_0133 which has a predicted size of 78 kDa, and no SP sequence (Fig. 2d). However, analysis of the

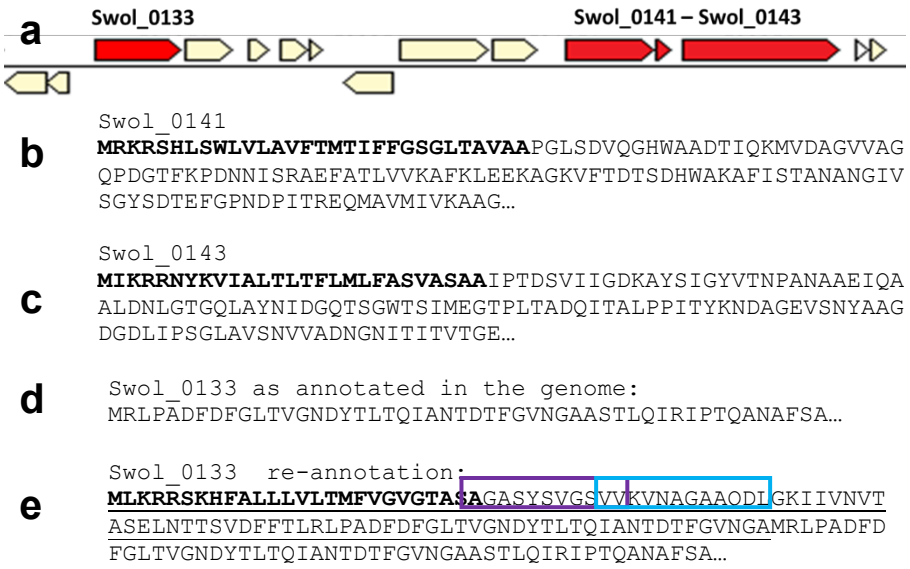


Figure 2 | Putative S-layer proteins. a, Context of the putative *S. wolfei* S-layer protein encoding genes, Swol_0133 and Swol_0141-Swol_0143. **b,** First 100 amino acids of Swol_0141 with signal peptide indicated in bold. **c,** First 100 amino acids of Swol_0143 with signal peptide indicated in bold. **d,** First forty amino acids of Swol_0133 as originally annotated in the genome, **e,** Frist 100 amino acids of Swol_0133 as re-annotated based on mass spectroscopy data. Signal peptide in bold, underlined region was not previously annotated as being part of this protein, and the boxed regions were recovered in the tryptic digest of this protein.

genome indicated a possible mis-call of the open reading frame encoding Swol_0133. Two peptide fragments from region upstream of the annotated gene were recovered during the mass spectroscopy experiment, leading us to conclude the open reading frame begins at an earlier methionine (Fig. 2e). Based on the re-assigned reading frame, Swol_0133 is predicted to have a signal peptide, and the protein's predicted molecular weight is 87 kDa, consistent with the observed molecular weight. Analysis of the whole cell tryptic digest of *S. wolfei* identified Swol_0143, Swol_0141, and Swol_0134 as the 24th, 28th, and 3rd most abundant proteins in the cell respectively[20]. S-layer proteins are typically some of the most abundant proteins in the cell as they must cover the entire cell surface.

Bioinformatic analysis of the putative S-layer genes. Bioinformatic analysis of the putative S-layer proteins encoded by Swol_0133, Swol_0141, Swol_0142, and Swol_0143 creates a more complex picture of the putative S-layer proteins we identified. BLAST analysis of Swol_0133 reveals a single protein domain super family at the C-terminus of the protein corresponding to a copper amine oxidase superfamily (Fig. 3a). Swol_0141 has four protein domain super family regions (Fig. 3b). Three regions belong to the SLH, S-layer homology domain a protein domain shown to be involved in non-covalently linking S-layer proteins to the peptidoglycan cell wall[36]. The final domain identified by BLAST of Swol_0141 is part of the E "early" set superfamily, a group of protein with carbohydrate binding regions. Phylogenetic comparisons between the *S. wolfei* putative S-layer proteins and other known bacterial S-layer proteins demonstrates the gene products of Swol_0142, and Swol_0143 group are closely related to each other. Swol_0133 and the Swol_0141-Swol_0143 group have limited similarity to each

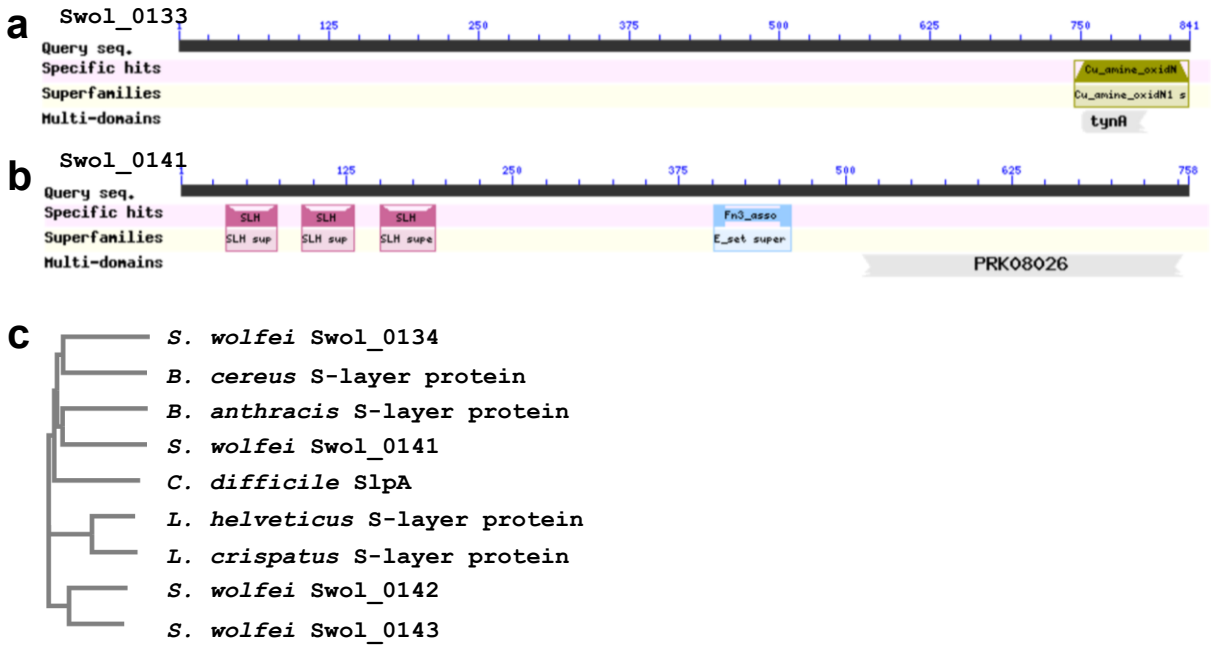


Figure 3 | Bioinformatic analysis of *S. wolfei* S-layer proteins. a, BLASTP results for Swol_0133 **b,** BLATP results for Swol_0141 **c,** phylogenetic tree of the *S. wolfei* S-layer proteins and other well studied bacterial S-layer proteins.

other and to the other bacterial S-layer proteins (Fig. 3c); however, bacterial S-layer proteins are known to be a very diverse group of proteins with limited sequence and functional homology.

Post-translational modification of the S-layer proteins. Due to the higher apparent molecular mass of the Swol_141-Swol_0143 bands, we suspect these proteins were post-translationally modified. However, a glycostain of the membrane fraction was negative (Fig. 4), ruling out the presence of the most common glycan moieties, N or O linked glycans, attached to the S-layer.

Electron microscopy of *S. wolfei*. Negative stain electron microscopy of the whole cells revealed a repeating pattern in the cell envelope (Fig. 5a). Unfortunately, the curvature of the whole cell distorts the pattern. We therefore prepared ghosts of the S-layer where the cells were lysed by freeze-thaw to release the cell contents, without destroying the S-layer (Fig. 5b). Close inspection of the negative stained cell ghosts reveals a 4-fold symmetrical pattern (Fig. 5c,d).

Cryo electron microscopy (cryoEM) of *S. wolfei* whole cells shows a cell architecture consisting of three layers (Fig. 6). A diffuse outer layer resembling other S-layer proteins is present, followed by a middle layer which could be an extremely thick peptidoglycan layer, and the inner layer which is the cell inner membrane. The outermost layer near the ends of the cell appear abraded or incomplete in both the cryoEM and negative stain micrographs (Fig. 5a and Fig. 6). Under the conditions imaged, *S. wolfei* was not observed to produce flagella, although the cells have previously been reported to be flagellated in co-culture[1].

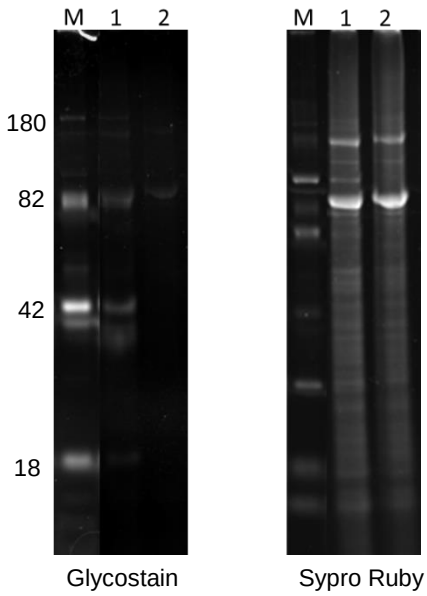


Figure 4 | Glycostain of *S. wolfei* cell envelope proteins. Lane M is the ladder, lane 1 is *S. wolfei* whole cells, and lane 2 is the *S. wolfei* envelope fraction. The bands in the envelop fraction did not react to the glycostain, suggesting these proteins are not glycosylated.

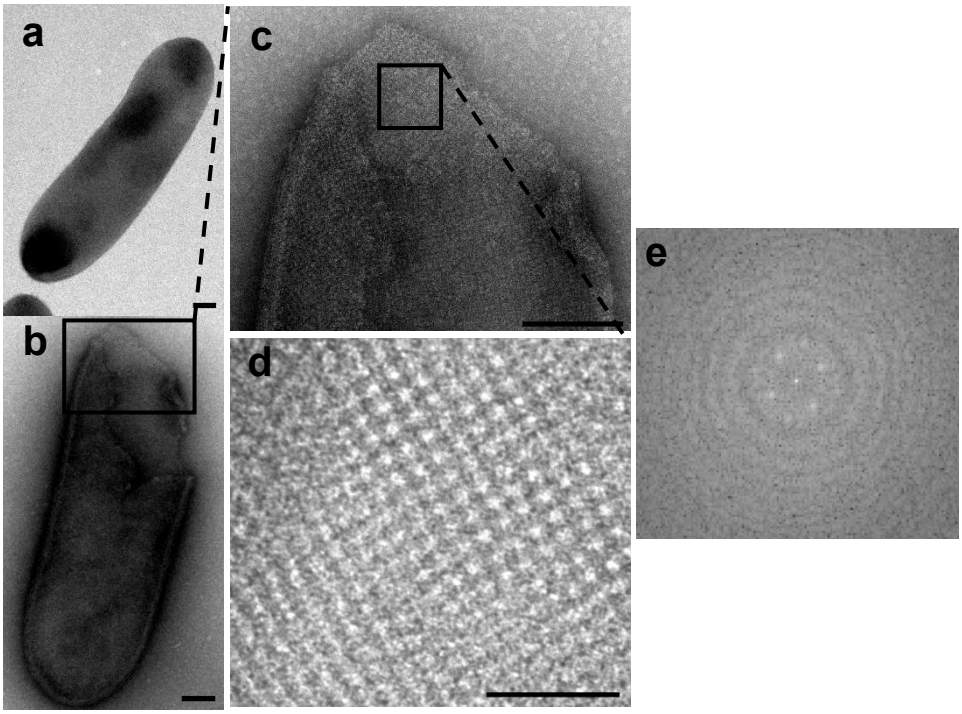


Figure 5 | Negative stain electron microscopy of *S. wolfei* S-layer.
a, Negative stain electron micrograph of an S. wolfei whole cell **b,** Negative stain electron micrograph of an S. wolfei cell ghost **c,d,** enlarged views of the S. wolfei cell ghost to show P4 protein symmetry, **e,** Fourier transform of d clearly showing P4 symmetry. All scale bars 50nm.

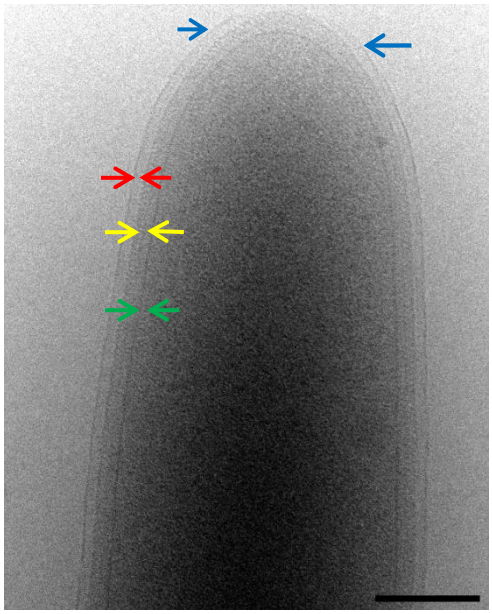


Figure 6 | Cryo electron microscopy of *S. wolfei* whole cells. The *S. wolfei* cell envelope has three layers. The inner membrane (green), peptidoglycan (yellow), and S-layer (red). At the cell ends, the outermost S-layer is abraded (blue). Scale bar 100nm.

The S-layer of *S. wolfei* is more complex than the typical bacterial S-layer with as many as four proteins comprising the paracrystalline array. Three of these proteins, Swol_0134, Swol_0141, and Swol_143 were directly identified by in-gel tryptic digest. The open reading frame for Swol_0133 was incorrectly assigned in the genome and translation of the protein begins at a methionine in the same frame, but upstream of the original genome annotation. With the change in start site for Swol_0133, all three of the directly observed S-layer proteins contain signal peptides. While Swol_0133 runs close to the predicted size of the protein on SDS-PAGE gels, the four bands from which Swol_0141 and Swol_0143 were recovered run anomalously high. None of the S-layer bands reacted positively to a glycostain, which suggests the Swol_0141-Swol_0143 undergoes some other type of modification to produce four bands of anomalously high molecular weight. Bioinformatics show Swol_0141 has three S-layer homology domains and is distantly related to Swol_0133 and other bacterial S-layer proteins. Electron microscopy of *S. wolfei* whole cells and ghosts reveals a paracrystalline protein array with P4 symmetry on the cell surface and a three-layered cell envelope.

The expression of multiple S-layer proteins and their presence in multiple highly abundant bands on SDS-PAGE gels suggests the *S. wolfei* S-layer proteins have some type of alternative splicing as observed in the *C. difficile* slpA protein system. In *C. difficle*, alternative splicing is used to facilitate infection and it could conceivably be used to facilitate the formation of symbiotic consortium in *S. wolfei*. Unfortunately, the lack of genetics tools available for *S. wolfei* and the long growth time of this organism makes experiments to confirm this hypothesis difficult to perform.

Acknowledgements

This work was supported by Department of Energy contract DE-FG02-08ER64689 (to RPG). This project received support from National Institutes of Health grants GM071940 which supports the Electron Imaging for NanoMachines core facility at the University of California – Los Angeles.

References

1 McInerney, M. J., Bryant, M. P., Hespell, R. B. & Costerton, J. W. *Syntrophomonas wolfei* gen. nov. sp. nov., an Anaerobic, Syntrophic, Fatty Acid-Oxidizing Bacterium. *Applied and Environmental Microbiology* **41**, 1029-1039 (1981).

2 Sieber, J. R., McInerney, M. J. & Gunsalus, R. P. Genomic Insights into Syntrophy: The Paradigm for Anaerobic Metabolic Cooperation. *Annual Review of Microbiology* **66**, 429-452, (2012).

3 McInerney, M. J., Sieber, J. R. & Gunsalus, R. P. Syntrophy in anaerobic global carbon cycles. *Current Opinion in Biotechnology* **20**, 623-632, (2009).

4 Schink, B. Energetics of syntrophic cooperation in methanogenic degradation. *Microbiology and Molecular Biology Reviews* **61**, 262-280 (1997).

5 Wofford, N. Q., Beaty, P. S. & McInerney, M. J. Preparation of cell-free extracts and the enzymes involved in fatty acid metabolism in *Syntrophomonas wolfei*. *Journal of Bacteriology* **167**, 179-185 (1986).

6 Stams, A. J. M. & Plugge, C. M. Electron transfer in syntrophic communities of anaerobic bacteria and archaea. *Nat Rev Micro* **7**, 568-577 (2009).

7 Hansen, K. H., Ahring, B. K. & Raskin, L. Quantification of Syntrophic Fatty Acid-β-Oxidizing Bacteria in a Mesophilic Biogas Reactor by Oligonucleotide Probe Hybridization. *Applied and Environmental Microbiology* **65**, 4767-4774 (1999).

8 Sieber, J. R., Le, H. M. & McInerney, M. J. The importance of hydrogen and formate transfer for syntrophic fatty, aromatic and alicyclic metabolism. Environmental Microbiology 16, 177-188, (2014).

9 van Schie, P. M. & Fletcher, M. Adhesion of Biodegradative Anaerobic Bacteria to Solid Surfaces. *Applied and Environmental Microbiology* **65**, 5082-5088 (1999).

10 Cutter, L. A., Van Schie, P. M. & Fletcher, M. Adhesion of anaerobic microorganisms to solid surfaces and the effect of sequential attachment on adhesion characteristics. *Biofouling* **19**, 9-18, (2003).

11 Shimoyama, T., Kato, S., Ishii, S. i. & Watanabe, K. Flagellum Mediates Symbiosis. *Science* **323**, 1574-1574, (2009).

12 Summers, Z. M. *et al.* Direct Exchange of Electrons Within Aggregates of an Evolved Syntrophic Coculture of Anaerobic Bacteria. *Science* **330**, 1413-1415, (2010).

13 Morita, M. *et al.* Potential for Direct Interspecies Electron Transfer in Methanogenic Wastewater Digester Aggregates. *mBio* **2**, (2011).

14 Shrestha, P. M. *et al.* Syntrophic growth with direct interspecies electron transfer as the primary mechanism for energy exchange. *Environmental Microbiology Reports* **5**, 904-910, (2013).

15 Nagarajan, H. *et al.* Characterization and modelling of interspecies electron transfer mechanisms and microbial community dynamics of a syntrophic association. *Nat Commun* **4**, (2013).

16 Rotaru, A.-E. *et al.* A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to *Methanosaeta* for the reduction of carbon dioxide to methane. *Energy & Environmental Science* **7**, 408-415, (2014).

17 Shrestha, P. M. & Rotaru, A.-E. Plugging in or going wireless: strategies for interspecies electron transfer. *Frontiers in Microbiology* **5**, (2014).

18 Kouzuma, A., Kato, S. & Watanabe, K. Microbial interspecies interactions: recent findings in syntrophic consortia. *Frontiers in Microbiology* 6, (2015).

19 Sieber, J. R. et al. The genome of *Syntrophomonas wolfei:* new insights into syntrophic metabolism and biohydrogen production. *Environmental Microbiology* 12, 2289-2301, (2010).

20 Sieber, J. R. et al. Proteomic analysis reveals metabolic and regulatory systems involved in the syntrophic and axenic lifestyle of *Syntrophomonas wolfei*. *Frontiers in Microbiology* 6, (2015).

21 Gerbino, E., Carasi, P., Mobili, P., Serradell, M. A. & Gómez-Zavaglia, A. Role of S-layer proteins in bacteria. *World Journal of Microbiology and Biotechnology* 31, 1877-1887, (2015).

22 Sakakibara, J. *et al.* Loss of adherence ability to human gingival epithelial cells in S-layer protein-deficient mutants of *Tannerella forsythensis*. *Microbiology* **153**, 866-876, (2007).

23 Hynönen, U. & Palva, A. *Lactobacillus* surface layer proteins: structure, function and applications. *Applied Microbiology and Biotechnology* **97**, 5225-5243, (2013).

24 Shimotahira, N. *et al.* The Surface Layer of *Tannerella forsythia* Contributes to Serum Resistance and Oral Bacterial Coaggregation. *Infection and Immunity* **81**, 1198-1206, (2013).

25 Zhang, W. *et al.* Adhesive ability means inhibition activities for *lactobacillus* against pathogens and S-layer protein plays an important role in adhesion. *Anaerobe* **22**, 97-103, (2013).

26 Ausiello, C. M. *et al.* Surface layer proteins from *Clostridium difficile* induce inflammatory and regulatory cytokines in human monocytes and dendritic cells. *Microbes and Infection* **8**, 2640-2646, (2006).

27 Kirby, J. M. *et al.* Cwp84, a Surface-associated Cysteine Protease, Plays a Role in the Maturation of the Surface Layer of *Clostridium difficile*. *Journal of Biological Chemistry* **284**, 34666-34673, (2009).

28 Ryan, A. *et al.* A Role for TLR4 in *Clostridium difficile* Infection and the Recognition of Surface Layer Proteins. *PLoS Pathog* **7**, e1002076, (2011).

29 Đapa, T. *et al.* Multiple Factors Modulate Biofilm Formation by the Anaerobic Pathogen Clostridium difficile. *Journal of Bacteriology* **195**, 545-555, (2013).

30 Pantaléon, V. *et al.* The *Clostridium difficile* Protease Cwp84 Modulates both Biofilm Formation and Cell-Surface Properties. *PLoS ONE* **10**, e0124971, (2015).

31 Beaty, P. S., Wofford, N. Q. & McInerney, M. J. Separation of *Syntrophomonas wolfei* from *Methanospirillum hungatii* in Syntrophic Cocultures by Using Percoll Gradients. *Applied and Environmental Microbiology* **53**, 1183-1185 (1987).

32 Huynh, M.-L., Russell, P. & Walsh, B. in *Two-Dimensional Electrophoresis Protocols* (eds Raymond Tyther & David Sheehan) 507-513 (Humana Press, 2009).

33 Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology* **292**, 195-202, (1999).

34 Altschul, S. F. *et al.* Protein database searches using compositionally adjusted substitution matrices. *FEBS Journal* **272**, 5101-5109, (2005).

35 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).

36 Mesnage, S. *et al.* Bacterial SLH domain proteins are non-covalently anchored to the cell surface via a conserved mechanism involving wall polysaccharide pyruvylation. *The EMBO Journal* **19**, 4473-4484, (2000).

CHAPTER 4:

nature microbiology

ARTICLES

PUBLISHED: XX XX 2016 | ARTICLE NUMBER: 16222 | DOI: 10.1038/NMICROBIOL.2016.222

CryoEM structure of the *Methanospirillum hungatei* archaellum reveals structural features distinct from the bacterial flagellum and type IV pili

Nicole Poweleit[1,2], Peng Ge[1,2], Hong H. Nguyen[3], Rachel R. Ogorzalek Loo[3], Robert P. Gunsalus[1,4] and Z. Hong Zhou[1,2]*

Archaea use flagella known as archaella—distinct both in protein composition and structure from bacterial flagella—to drive cell motility, but the structural basis of this function is unknown. Here, we report an atomic description of the archaella, based on the cryo-electron microscopy (cryoEM) structure of the *Methanospirillum hungatei* archaellum at 3.4 Å resolution. Each archaellum contains ~61,500 archaellin subunits organized into a curved helix with a diameter of 10 nm and average length of 10,000 nm. The tadpole-shaped archaellin monomer has two domains, a β-barrel domain and a long, mildly kinked α-helix tail. Our structure reveals multiple post-translational modifications to the archaella, including six O-linked glycans and an unusual N-linked modification. The extensive interactions among neighbouring archaellins explain how the long but thin archaellum maintains the structural integrity required for motility-driving rotation. These extensive inter-subunit interactions and the absence of a central pore in the archaellum distinguish it from both the bacterial flagellum and type IV pili.

Prokaryotes use a variety of extracellular protein complexes to interact with their surrounding environment. Both domains of prokaryotes archaea and bacteria express extracellular filaments (called archaella and flagella, respectively) that drive cellu lar motility in the form of 'swimming'. In addition to motility[1–3], the archaella[4] (also called archaeal flagella) have been shown to partici pate in cellular adhesion[5,6], biofilm formation[7] and other symbiotic and environmental activities[5]. Archaella have a significantly smaller diameter, on average, than bacterial flagella (10 12 nm and 20 26 nm, respectively), which suggested significant differences between the two analogous motility filaments[8]. Further studies of archaella gene sequences[9], archaellin post translational modifications[10], and archaellin export and assembly[11–14] have shown a greater similarity between bacterial type IV pili and archaella than between archaeal and bacterial flagella[1], although the structural basis was unknown. Additionally, the energy source archaea use in archaella driven swimming is ATP instead of a proton motive force[15]. However, archaella can grow to 10 20 times as long as the cell[8], much longer than bacterial pili, and the genes in archaella operons are unique to archaella, with a few proteins with homology to proteins in the bacterial pili operons and no proteins with homology to the bacterial flagella operons[16,17]. The results from these studies indicate that archaella exhibit some similarity to both bacterial flagella (swimming motility and length) and bacterial type IV pili (energy source and biogenesis), but how archaella use the combination of these traits to drive cellular motility is unknown.

Due to the tendency of purified flagellar and pili subunits to form filaments or aggregates in solution, efforts to determine the structure of the archaella by X ray crystallography and NMR have been unsuccessful. The best available structure of archaella is that from *Halobacterium salinarium*, which was determined at 8 Å resolution from images recorded on photographic films and is insufficient to resolve the two domains of monomers within the filament[18,19]. Because the domains appeared to be a central α helix and an exterior globular domain, it was hypothesized that the archaellin had a 'type IV pili like' structure. Here, we have used direct electron counting cryo electron microscopy (cryoEM) to (1) determine the structure of an endogenous archaellum at 3.4 Å resolution, (2) build an atomic model of the archaellin, (3) identify novel post translational modifications to the archaellin, and (4) define inter archaellin inter actions essential to the assembly of the archaellar filament. These findings allow us to establish differences between and similarities with the bacterial flagella and type IV pili.

Results and discussion

Atomic structure of the archaella. Our structural determination of purified archaella involved a complex strategy, beginning with helical parameter estimation in high contrast EM images of negatively stained sample, intermediate resolution cryoEM structure from film data and a refined structure from high resolution images by direct electron counting (for details see Methods). Typical frame aligned images from a Titan Krios with a Gatan K2 Summit direct electron counting camera (Fig. 1a and Supplementary Fig. 1) shows barely visible filaments, although their Fourier transforms indicate the existence of data better than 3.8 Å resolution (Fig. 1b and Supplementary Fig. 1). Indeed, characteristic high resolution features can be seen in two dimensional class averages (Fig. 1c) and Fourier transforms of such class averages show layer lines consistent with the determined helical symmetry parameters of an axial rise of 5.32 Å and a rotation of 108.02° per subunit (Fig. 1d). The filaments were reconstructed to an averaged resolution of 3.4 Å (Fig. 1e,

Supplementary Fig. 2 and Supplementary Table 1). This resolution is consistent with visible structural features such as amino acid side chains (Supplementary Fig. 3). The cryoEM map is characterized with an α helical core (yellow to green in Fig. 1e,f), surrounded by an all β sheet peripheral (blue to red in Fig. 1e,f and Supplementary Video 1). Both the N terminus and C terminus of a single polypeptide chain are clearly visible, permitting *de novo* tracing of the archaellin monomer without any ambiguity (Fig. 1g, h, Supplementary Fig. 3 and Supplementary Video 1). A cryoEM image (Fig. 1i) of intact cells reveals a distinct cell architecture of the cylindrical *M. hungatei* cell: an S layer enclosing the totality of the cytoplasm of the cell, an external ridged proteinaceous sheath, only permeable to gases and small solutes[20–22], and a characteristic plug for larger substrates to pass into and out of the cell at each end of the cell. Multiple long (10,000 nm on average) but thin (10 nm diameter) archaella attach at each end of the cell (Fig. 1i). Unlike its bacterial counterpart, the archaellum exhibits restricted filament flexibility without uniform super helicity (Fig. 1a,i and Supplementary Fig. 1). Based on the length and identified helical pitch of the *M. hungatei* archaella, we estimate that each archaellar 10 μm filament contains, on average, ~61,500 archaellin subunits.

The first six amino acid residues, MRKETA, predicted from all three *flaB* candidate genes from the *M. hungatei* genome[23], are not present in the cryoEM structure. All archaellins studied to date have type III signal peptides that must be cleaved to form the mature protein[24–27]. To identify which gene codes the observed archaellin and to determine how many amino acids are cleaved to form the mature archaellin or if the full length protein is present and simply not resolved in the structure due to flexibility, we performed a series of biochemical analyses of the purified archaella. First, mass spectroscopy analysis of the major SDS band (Supplementary Fig. 3a) confirms there is a single protein in the purified archaella, and it is the *flaB3* gene product of Mhun 3140 (Supplementary Table 2). Second, we performed Edman N terminal sequencing to establish that the functional FlaB3 protein's N terminal sequence is FSGLE, as seen in the atomic model (Supplementary Table 3). Third, we observed that the protein runs anomalously large on the SDS PAGE gel based on the predicted molecular weight of Mhun 3140 (17.6 kDa), a phenomenon that often indicates that the protein is post translationally modified. Taken together, our protein analyses identified FlaB3 protein as the *M. hungatei* archaellin subunit, which is cleaved following position six.

No structure is available for any protein of the archaeal motility and adhesion filament family, and we built the atomic model of the above identified FlaB3 protein *de novo* based on cryoEM data, without reference to any existing crystal or NMR structures (Fig. 1h and Supplementary Fig. 3b). The amino acid backbone can clearly be traced, and large amino acid side chains are readily identifiable (Fig. 1b, Supplementary Fig. 3c f and Supplementary Video 2).

Each archaellin monomer has a tadpole shape and is organized into two domains (Fig. 1h and Supplementary Fig. 3b): an N terminal domain (NTD) and a C terminal domain (CTD). The NTD forms a long, hydrophobic α helix (Fig. 1h and Supplementary Fig. 3c,d) and the CTD an eight stranded anti parallel β barrel (Fig. 1h and Supplementary Fig. 3d,e). The hydrophobic α helices of each subunit point into the centre of the filament and form a hydrophobic core down the helical axis of the filament (Fig. 1e h and Supplementary Fig. 3c,d). About two thirds of the helix corresponds to the conserved FlaB sequence[28] and one third extends into the variable domain of the protein, making the α helical domain significantly longer than predicted (Fig. 1j and Supplementary Fig. 4). The CTD β barrel decorates the outside of the filament (Fig. 1e h and Supplementary Fig. 3b f), and is composed entirely of the variable sequence region of the archaellin (Fig. 1j). As detailed below, this domain plays an important role in the archaella as the site of multiple post translational modifications and maintains ionic interactions between subunits that help to stabilize the long filament.

Post translational modification. The FlaB3 cryoEM map has eight sites of extra density per subunit (Fig. 2, Supplementary Fig. 5 and Supplementary Video 3) that form dead end protrusions off the peptide backbone that could not be accounted for by single amino acid side chains. All of the sites are located on the exterior surface of the β barrel domain and are therefore exposed to the environment. Seven of the extra densities are probably glycosylation sites based on their size, shape and location (Fig. 2a). Six sites are potential O linked glycans at positions Ser59, Ser61, Ser65, Thr100, Ser104 (Fig. 2b) and Thr108 (Supplementary Fig. 5a c). All previously described glycosylation sites for archaella are N linked glycans, making the primarily O linked glycans of the *M. hungatei* archaellin unique among archaellins thus far described. The seventh site is a potentially N linked glycan at position Asn159 (Fig. 2c and Supplementary Fig. 5e). There is one additional amino acid, Gln134, with an unknown modification (Fig. 2d and Supplementary Fig. 5d). Single or multiple sugar residues can be docked into, and fit well with, all of the extra densities except the Gln134 density, which is far too small to hold a glycan based on the previously described sugar moieties decorating the *Methanococcus maripaludis*[29] and *Methanococcus voltae*[30] archaellins. The N glycosidic linkage is particularly unusual because the protein sequence in this region does not contain the canonical N linked glycosylation sequence of N X T/S, nor does it occur anywhere in the protein. The amino acid sequence of the N linked glycan of the *M. hungatei* archaellin is N G, which is an atypical glycosylation site[31].

These post translational modifications were further confirmed biochemically. First, a mass ion spectrum (Fig. 2e) demonstrated that the mass of the archaellin is ~20 kDa, ~2.4 kDa more than the size predicted from the peptide sequence. The broadness of the matrix assisted laser desorption/ionization (MALDI) spectrum peak also suggests heterogeneity in the post translational modification. Second, a glycostain confirmed that the *M. hungatei* archaellin is glycosylated (Fig. 2f, lane 4). Third, deglycosylation with PNGaseF was performed (Fig. 2f, lane 5), and the *M. hungatei* archaellin showed sensitivity to PNGaseF. These SDS PAGE and mass spectroscopy analyses confirm the presence and heterogeneity of post translational modifications, as suggested by the cryoEM archaellum structure.

Interactions between archaellins within the archaellum. The cryoEM structure reveals intermolecular interactions between archaellin subunits that can shed light onto how the extremely long, yet thin, filament maintains physical integrity under stress caused by rotation driven motility. Individual archaellin subunits within the *M. hungatei* archaellum form a dense network of interactions, with each individual subunit forming intermolecular interactions with eight other subunits (Fig. 3a c and Supplementary Video 4). There are three types of intermolecular contact identifiable in the protofilament structure: a radial interaction between Phe1 residues at the archaellum helical axis, which create a spokes effect via a π stacking sandwich (Fig. 3d f and Supplementary Video 4), hydrophobic interactions between adjacent α helices, and ionic interactions between α helices and β sheets or between two β sheets of adjacent subunits (Fig. 3 and Supplementary Videos 5 and 6). Every subunit in the *M. hungatei* archaellum participates in a π stacking interaction between the Phe1 residues that creates a single, continuous, one start helix in

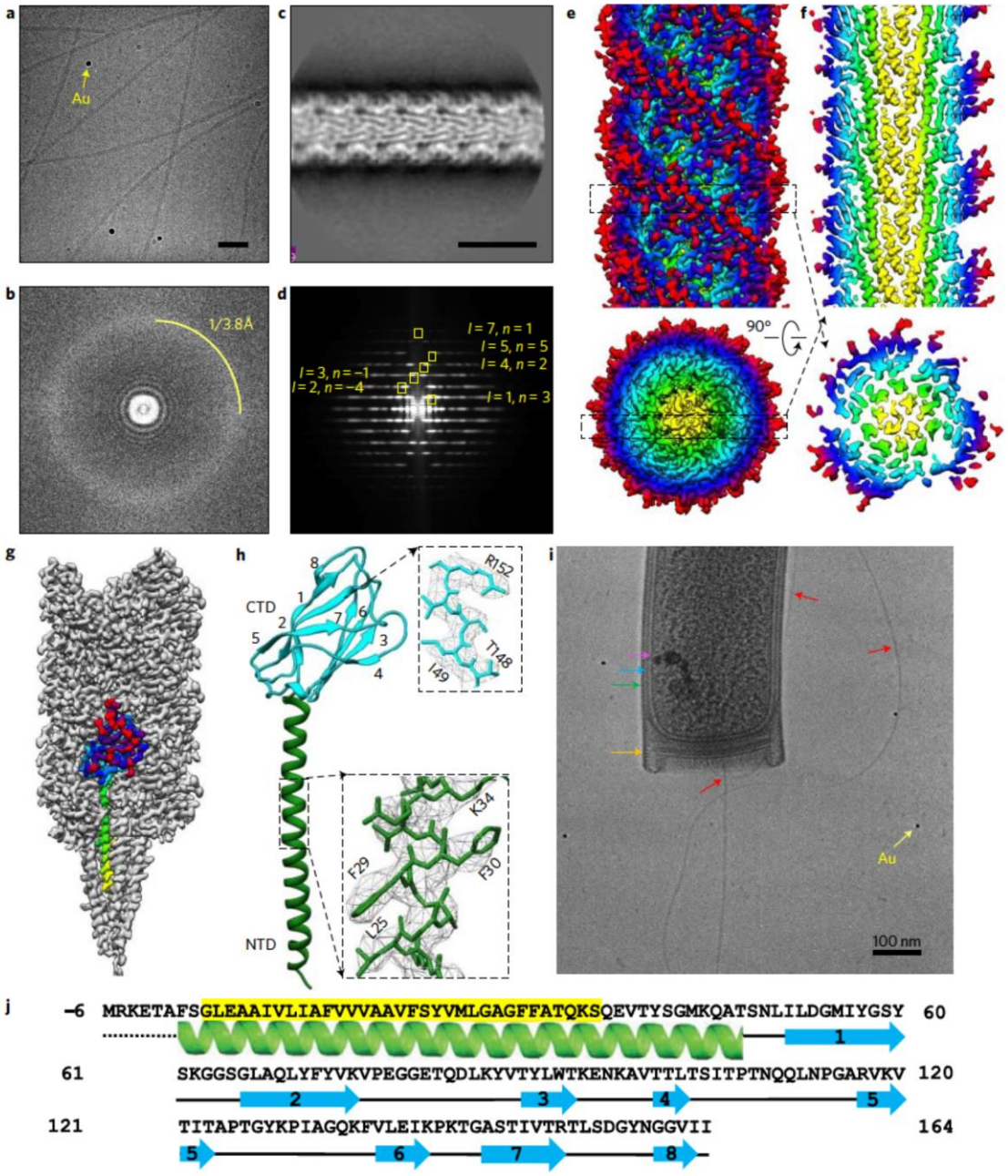


Figure 1 | CryoEM data and refinement. **a**, Representative cryoEM micrograph of *Methanospirillum hungatei* JF1 filaments with 10 nm gold fludicials (yellow arrow) added. Scale bar, 50 nm. Around 3,300 videos were collected in the data set, of which ~75% had filaments used in two-dimensional classification. **b**, Fourier transform of micrograph in **a** with 3.8 Å water ring indicated. **c**, A two-dimensional classification result of CTF corrected images for the *M. hungatei* archaella filaments. Eleven two-dimensional classes (of which the pictured class is one) were used in the final reconstruction. Scale bar, 10 nm. **d**, Fourier transform of the two-dimensional class average with helical indexing. **e**, Reconstruction of the *M. hungatei* archaella at 3.4 Å resolution. **f**, Slab view of the reconstruction showing central α-helices. **g**, Archaella reconstruction with a single subunit highlighted (rainbow). **h**, Reconstruction of an archaellin with two domains, a hydrophobic N-terminal α-helix (green) and the C-terminal eight-stranded β-barrel (cyan) with strands numbered. Insets: close-ups on the α-helix, and on strand 7 of the β-barrel. **i**, A cryoEM image of *M. hungatei* cell end with cytoplasmic membrane (pink), S-layer (blue), sheath (green), plug (orange), 10 nm gold fiducial (yellow) and protruding archaellum (red) indicated by arrows. Scale bar, 100 nm. **j**, Schematic of the primary and secondary sequences of the *M. hungatei* FlaB3 protein (Mhun_3140). The yellow highlighted portion of the sequence is conserved in archaellins.

the centre of the filament with the aromatic side chains projecting into the centre of the filament and stacking on top of one another (Fig. 3d f and Supplementary Video 4). Through this interaction, every subunit interacts radially with two other subunits, one aromatic side chain above and one below, to which it is not adjacent in any other area of the protein.

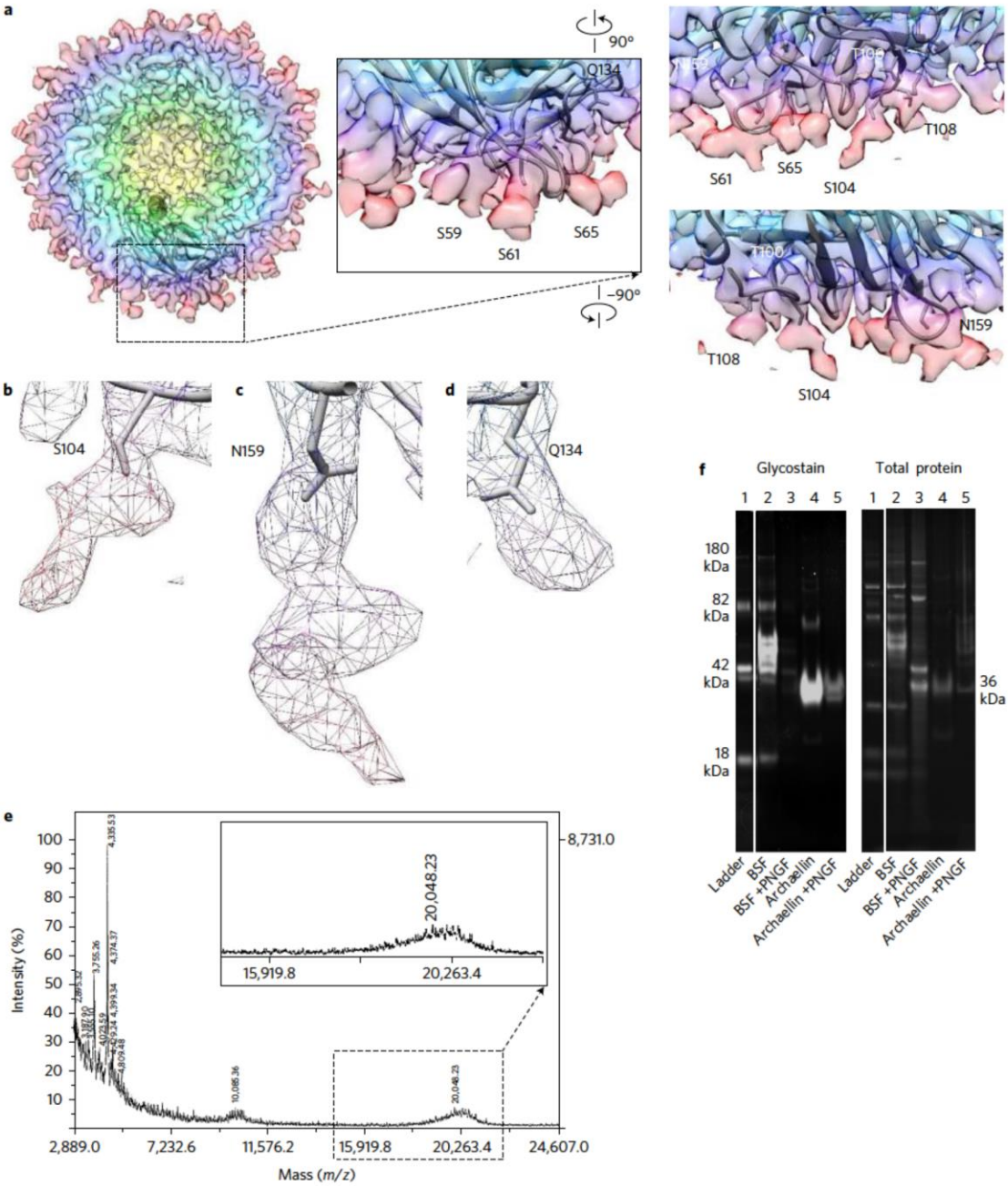


Figure 2 | Post-translational modification of the *M. hungatei* flagellin. **a**, CryoEM reconstruction of the *M. hungatei* archaella with a single archaellin subunit with surface exposed extra densities that do not correspond to the atomic model indicated. **b**, The extra density over Ser104. **c**, The extra density on Asp159. **d**, The extra density on Gln134. **e**, Mass ion spectra of *M. hungatei* archaellin. Inset: enlargement of the broad 20 kDa peak. **f**, Glycostain and total protein stain of *M. hungatei* archaella and treatment with PNGase F. Lane 1, Candy Cane ladder; lane 2, bovine serum fetuin not treated with PNGase F; lane 3, bovine serum fetuin treated with PNGase F; lane 4, *M. hungatei* archaella not treated with PNGase F; lane 5, *M. hungatei* archaella treated with PNGase F.

The hydrophobic and ionic intermolecular interactions occur between a subunit and the six subunits physically adjacent to it in the archaellar filament (Figs 3g, 4 and Supplementary Videos 5 and 6). First, the N terminal hydrophobic α helix is the defining characteristic of the archaellin and this region is highly conserved. Every subunit within the filament interacts hydrophobically through the conserved portion of the α helix with six adjacent sub units (Figs 3h,i, 4a d and Supplementary Video 5). In addition to the hydrophobic interactions, each subunit interacts with the same six adjacent subunits through ionic interactions at four positions around the β barrel (Fig. 4e), the leftmost (Fig. 4f), upper (Fig. 4g), lower (Fig. 4h) and middle interaction (Fig. 4i) areas (Supplementary Video 6). Second, the ionic interaction between adjacent subunits involves a total of eight amino acids Tyr22, Gln45, Asp53, Lys75, Glu78, Lys141, Lys143 and Asp156 and is confined to discrete pockets where four separate archaellins come

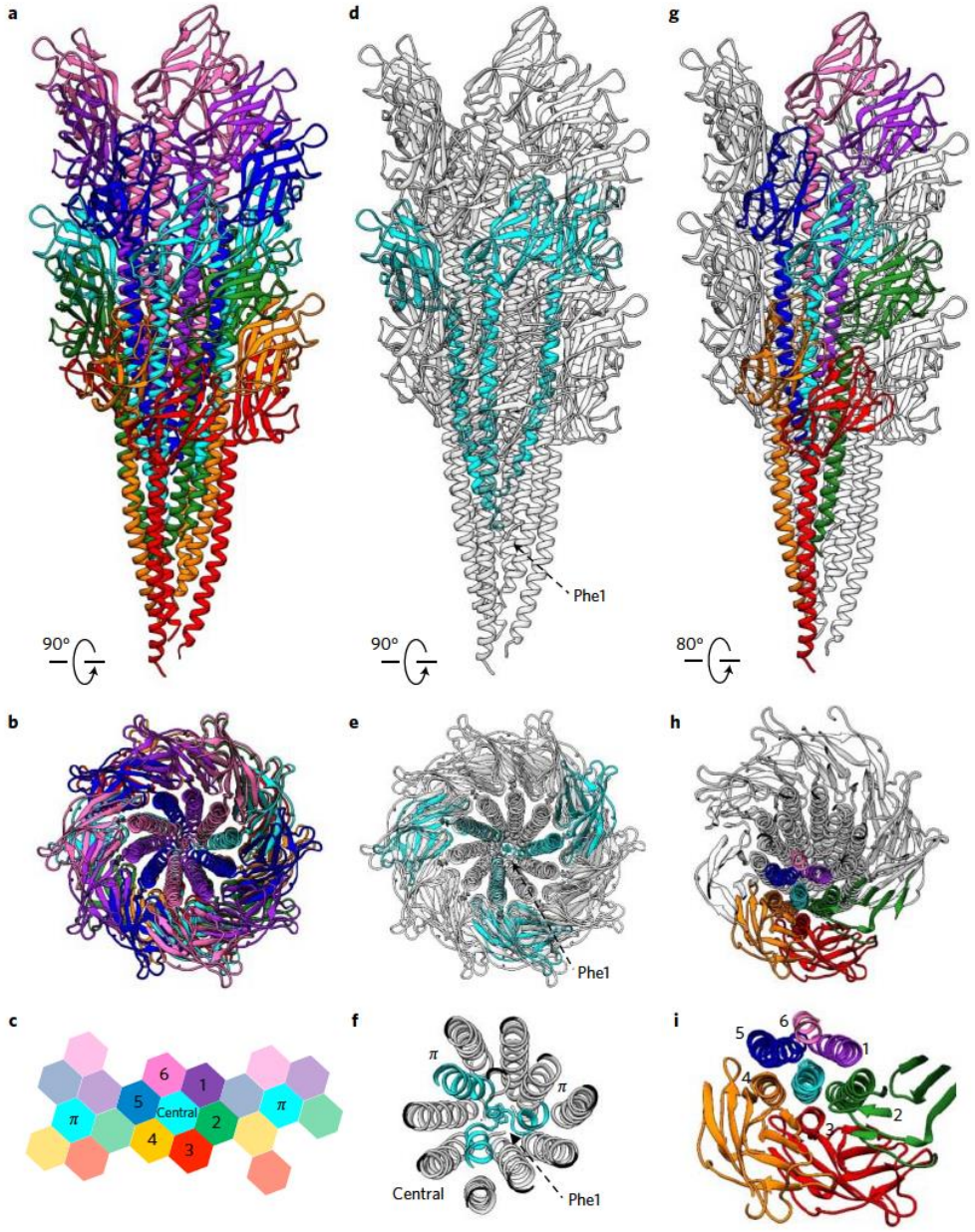


Figure 3 | Subunit–subunit interactions. a,**b**, Network of 21 subunits in the *M. hungatei* archaella filament **c**, Cartoon depicting the interaction unit for a single subunit marked 'central'. This subunit interacts radially with two subunits marked 'π' and hydrophobically and ionically with six subunits marked 1 to 6. **d**,**e**, The two archaellins (cyan) with which the central cyan archaellin interacts radially through π-stacking sandwiches formed by the N-terminal phenylalanine residue shown in the protofilament. **f**, Close-up on the Phe1 residues that form the π-stacking interactions. **g**,**h**, The six subunits with which the central cyan subunit has adjacent interactions through their hydrophobic α-helices and ionic interactions in the β-barrel domain.

together to participate in the ionic interaction network (Fig. 4c e and Supplementary Video 6). The complex series of interactions explains the intermolecular stability of the archaellar filament.

Although much is known about the structure, biogenesis and functional operation of bacterial flagella, practically nothing is known about the analogous processes in the archaea. Many funda mental questions remain such as how the archaella are anchored to the cell membrane[32], how the thin archaella are able to remain intact under the rotating stress of motility, whether subunits are added at the tip as they are in bacterial flagella or at the base of the filament as with bacterial pili[33], and the roles of glycosylation in filament bio genesis and function[34–37]. The high resolution structure presented here begins to answer some of these questions. The complex network of archaellin interactions allows the archaellum to remain stable. The lack of a central channel suggests new subunits are added at the base of the filament as they are in bacterial pili. The surface exposed glycans could be involved in adhesion and proper filament assembly. In several Euryarchaeota, mutations to glycosylation sites result in non archaellated cells, suggesting that glycan addition is required for filament biogenesis and/or stability[34,35].

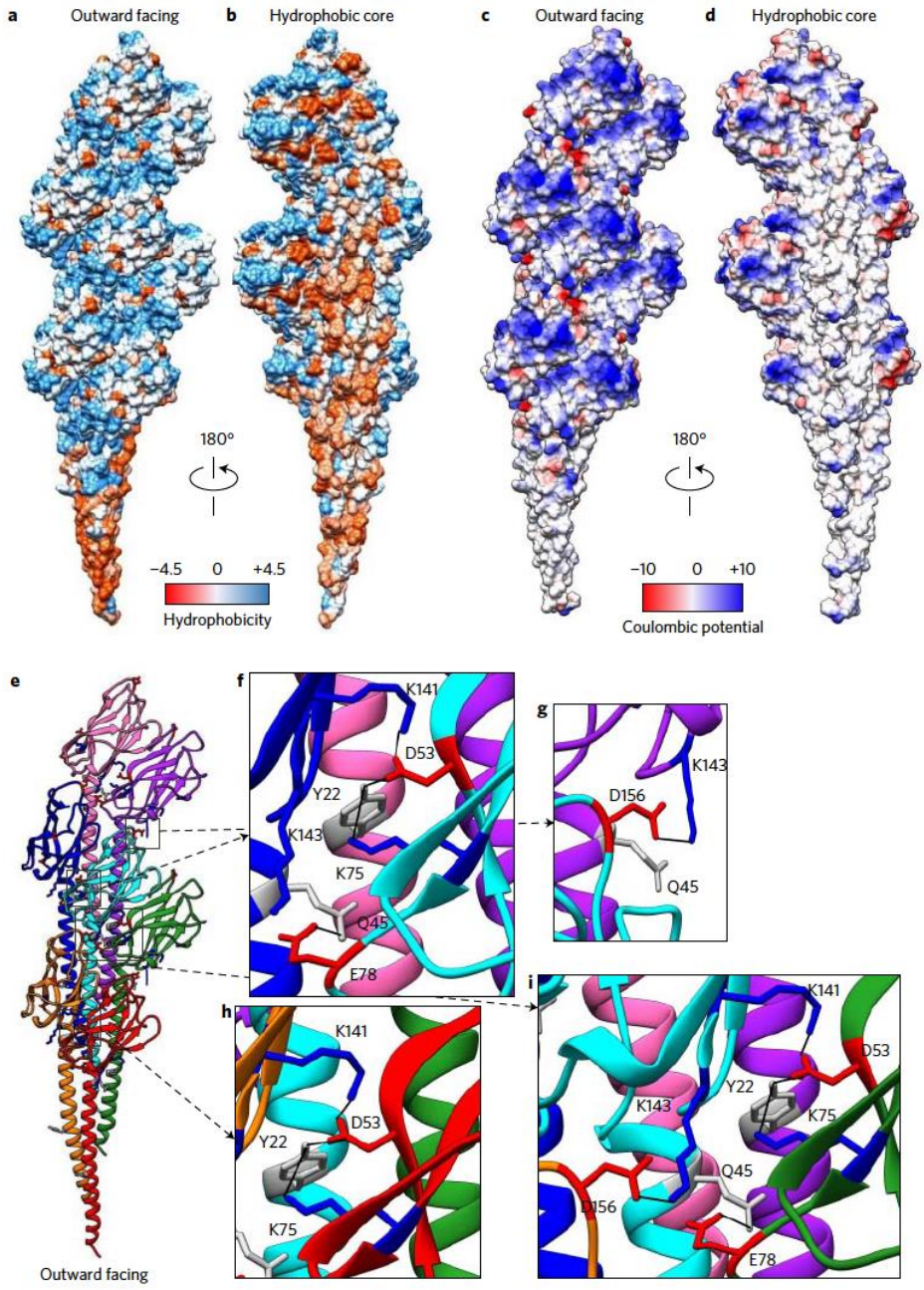


Figure 4 | Adjacent subunit interactions. **a**,**b**, Hydrophobicity plot of seven subunits involved in adjacent subunit interactions. The inward-facing α-helix domain is hydrophobic, while the outward-facing portion of the β-barrel is hydrophilic. **c**,**d**, Coulombic potential plot of seven subunits involved in adjacent subunit interactions. The inward-facing α-helix domain has very little charge, and the outside of the filament has patches of positive charge. **e**, An interaction unit of seven subunits. The central (cyan) archaellin interacts hydrophobically and ionically with the six adjacent archaellins pictured. The four areas of ionic interaction are indicated with dark grey boxes. The amino acids involved in ionic interactions are coloured in accordance with their charge potential. **f**, Leftmost interaction area of the central (cyan) subunit. The central subunit contributes three amino acids to this interaction: D53, K75 and E78. **g**, Top interaction area of the central (cyan) subunit. The central subunit contributes one amino acid to this interaction: D156. **h**, Lower interaction area of the central (cyan) subunit. The central subunit contributes one amino acid to this interaction: Y22. **i**, Middle interaction area of the central (cyan) subunit. The central subunit contributes three amino acids to this interaction: Q45, K141 and K143.

Comparison to bacterial motility filaments. In bacteria, two types of filamentous structure contribute to cell motility, bacterial flagella and type IV pili, both of which share some characteristics with archaella. At a global level, the bacterial flagella have a uniform super helical coiling that generates a waveform to drive motility[38]; however, our data (Fig. 1a,h and Supplementary Fig. 1) show the bacterial flagellar hallmark does not extend to archaellum that have a non uniform curvature. Before our current 3.4 Å resolution

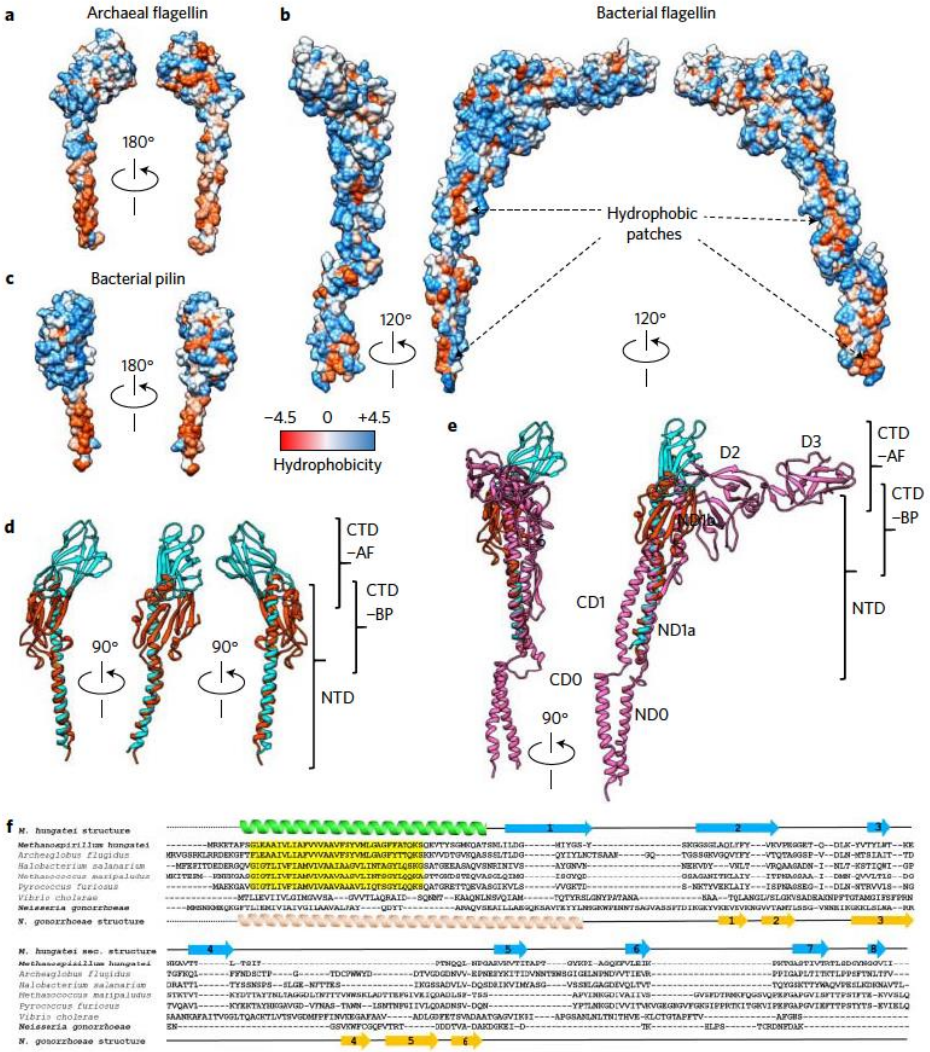


Figure 5 | Comparison between the *M. hungatei* archaellin and a bacterial type IV pilin and flagellin. a, Hydrophobicity plot of a single *M. hungatei* archaellin. **b**, Hydrophobicity plot of a single bacterial flagellin from *Salmonella enterica* (PDB: 1UCU). **c**, Hydrophobicity plot of a single bacterial type IV pilin from *Neisseria gonorrhoeae* (PDB: 2HIL). **d**, Comparison of domain organization between the atomic models of the *M. hungatei* archaellin (cyan) and the *N. gonorrhoeae* bacterial type IV pilin (orange). **e**, Comparison of domain organization between the atomic models of the *M. hungatei* archaellin (cyan), the *N. gonorrhoeae* bacterial type IV pilin (orange) and the *S. enterica* bacterial flagellin (pink). **f**, Sequence alignments of five archaellins (top) and two bacterial pilins, *V. cholerae* type IVa and *N. gonorrhoeae* type Ivb pilins (bottom). Sequences used are for the major expressed archaellins for each species. The secondary structure of the *M. hungatei* archaellin is shown above the alignments (green and blue), and the secondary structure of the *N. gonorrhoeae* pilin is shown below the alignments (orange). The sequence highlighted in yellow is the consensus sequence for archaeal FlaB proteins. Percent identity scores between the *M. hungatei* FlaB3 protein and the other sequences in the alignment are as follows: *A. flugidus* 39%, *H. salanarium* 27%, *M. maripaludus* 27%, *P. furiosus* 27%, *V. cholera* 20%, *N. gonorrhoeae* 21%. For comparison, the identity score between the two bacterial type IV pilins is 26%.

structure of the archaella, two other motility filament structures were known in atomic detail, the R type bacterial flagellum from *Salmonella enteracia*[39], solved to 4 Å resolution by cryoEM, and the type IV pili from *Neisseria gonorrhoeae*[40], solved to atomic resolution by X ray crystallography and docked into a 12 Å cryoEM map.

At the monomer level, the archaellin, bacterial pilin and bacterial flagellin structures contain two similar structural elements: a long helix domain in the core of the filament and a β sheet domain lining the periphery (Fig. 5 and Supplementary Video 7). The helix is hydrophobic in all three structures (Fig. 5a c), creating the non specific interactions necessary for slight flexibility during filament twisting and bending. The bulk of the β sheet domain consists of antiparallel strands and is decorated with glycans on the surface (Fig. 5d,e). Beyond these two common structural elements, the structures have a great deal of variation in monomer organiz ation. Within each type of motility filament, the helix domain has a characteristic conserved protein sequence, but the characteristic sequence is not conserved between motility filaments (Fig. 5f). The canonical helix domain of the bacterial pilin is a single helix with two sharp kinks, the archaellin is a single helix with a single very slight kink, and the bacteria flagellin contains four additional helices (ND0, ND1b, CD0 and CD1), which form a coiled coil domain (Fig. 5d,e). In bacterial pili, the β sheet domain contains

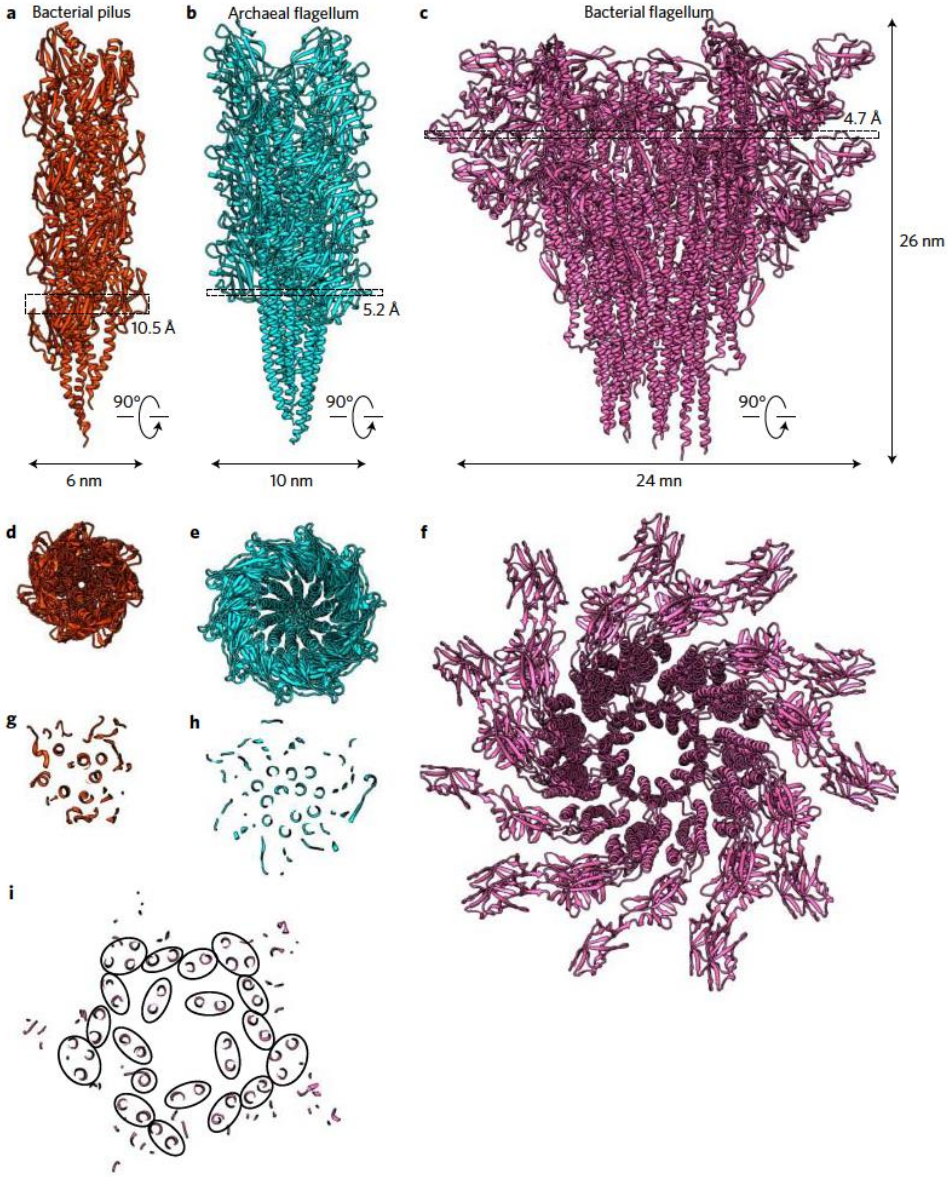


Figure 6 | Comparison between prokaryotic motility filaments. a, A 15-subunit protofilament of the *N. gonorrhoeae* bacterial type IV pilus (orange, PDB: 2HIL). **b**, A 26-subunit proto filament of the *M. hungatei* archaeal archaellum (cyan). **c**, A 20-subunit protofilament of the *S. enterica* bacterial flagellum (pink, PDB: 1UCU, protofilament generated based on published helical parameters). **d**, View down the helical axis of the *N. gonorrhoeae* bacterial type IV pilus protofilament. **e**, View down the helical axis of the *M. hungatei* archaellum. **f**, View down the helical axis of the *S. enterica* bacterial flagellum. **g**, Slice one subunit's axial rise (10.5 Å thick) of the *N. gonorrhoeae* bacterial pilus as viewed down the helical axis. **h**, Slice one subunit's axial rise (5.2 Å thick) of the *M. hungatei* archaellum as viewed down the helical axis. **i**, Slice one subunit's axial rise (4.7 Å thick) of the *S. enterica* bacterial flagellum as viewed down the helical axis. Ovals indicate helical bundles belonging to the same subunit.

only six β strands, lacks greater organization, and faces downward (Fig. 5d). The archaellin β domain forms an eight stranded β barrel projecting upwards, creating a taller monomer (Fig. 5d). The bacterial flagellin has two additional β sheet domains (D1 and D2) that are perpendicular to the helix domain (Fig. 5e). The increasingly complex monomer structures give rise to increasing filament diameter from bacterial pili (6 nm), to archaella (10 nm) to the bacteria flagella (24 nm) (Fig. 6 and Supplementary Video 8).

How these protomers assemble into a filament differs dramati cally between the three structures solved to atomic resolution. The helical parameters differ; for example, the number of subunits per turn is 3.6, 3.3 and 5.5 in the bacterial type IV pilus, archaellum and bacterial flagellum, respectively. To achieve a filament length of ~26 nm, 15 bacterial pilins (Fig. 6a), 26 archaellins (Fig. 6b) or 20 bacterial flagellins (Fig. 6c) are needed. Indeed, the axial views of these three filament types bear no similarity. The hollow central pore of the bacterial pilus measures 6 10 Å in diameter (Fig. 6d), the archaellum does not have a pore (Fig. 6e), and the bacterial flagellum has a 20 Å diameter pore (Fig. 6f). This difference is particularly notable for the two flagella, as the bacterial flagellar

pore functions as an export channel with subunits travelling through the channel to be assembled at the filament tip. The absence of such a pore in the archaellum supports the view that archaellins are added to the base of the filament as they are in bacterial pili. Central helices of the three filament types have different angles relative to the filament axis: the bacterial pilin helix is doubly kinked and has a variable offset (Fig. 6d); the archaellin is offset 20° (Fig. 6e); and the bacterial flagellin ND0 and CD0 domains are nearly parallel to the helical axis (Fig. 6f). The number of α helix domains visible in a slice equal to the axial rise of one subunit is different. In the bacterial pili seven helix domains are visible (Fig. 6g), in the archaella 12 helix domains are visible (Fig. 6h), and in the bacterial flagella 20 helix bundles from different subunits are visible (Fig. 6i). This high level of structural diversity between the three filaments indicates that the previous hypothesis, that archaella are similar to bacterial type IV pili, is over simplified. This picture of structural diversity is further complicated by other extracellular filaments such as the bacterial fimbria[41], conjugation pilus[42] and type II secretion system[43] and archaeal adhesion filaments such as the *Ignicoccus hospitalis* Iho670 filament[44] and the *Sulfolobus acidocaldarius* pili[45]. Indeed, we observed distinct pili at a low frequency in our *M. hungatei* archaella data set (Supplementary Fig. 6). The above identified canonical structural elements shared across bacteria and archaea provide a basis for the prediction of other extracellular filaments for which structural information is not yet available.

From a technical view point, cryoEM structures of archaella and bacterial pili have been hitherto limited in resolution to 8 Å due to the thinness and flexibility of the filaments. The recent technological advances in direct electron counting technology and helical recon struction incorporating a Bayesian statistics method have made it possible to solve *de novo* atomic structures by cryoEM of thin, curving filaments. Solving the *in situ* archaella structure has also led to the direct observation of post translational modifications of the *M. hungatei* archaellin. With this high resolution cryoEM tech nology and the canonical structure features reported here, micro biologists are now positioned to better understand the biogenesis and operational mechanism of a myriad of bacterial and archaeal filaments involved in diverse functions, such as 'sex' conjugation, charge conduction and toxin secretion.

Note added in proof: During the final review stage of this paper, we noted the publication of a paper by Braun *et al.*[46] that reports an ~4 Å structure of an adhesion filament from *Ignicoccus hospitalis*. This filament is archaeallum like in appearance, but *I. hospitalis* is not motile and has no archaellum operon. There are major differences between our structure reported here and the *I. hospitalis* structure. First, our atomic model is complete while the latter has 60 amino acids missing. Second, more post translational modifications have been resolved in our structure. Third, the helical parameters differ significantly (108.02° rotation per subunit and 5.32 Å axial rise per subunit in our structure versus 106.65° rotation per subunit and 5.45 Å axial rise per subunit in the latter).

Methods

***M. hungatei* growth and archaella isolation.** *Methanospirillum hungatei* strain JF1 (ATCC# 27890) was cultured in 1 l anaerobic bottles containing 500 ml of JMA medium modified to contain half the normal concentration of phosphate to elevate archaella abundance[47]. Bottles were inoculated (0.5% vol) with exponential phase cells and the vessel headspace was then pressurized (10 p.s.i.) with an 80:20 (vol/vol) mixture of H_2:CO_2. Cultures were incubated at 37 °C with gentle agitation for one week prior to cell harvest[48].

To isolate archaella, *M. hungatei* cells were collected by centrifugation at 5 K for 10 min. Cells were resuspended in 2 ml of TBS (100 mM NaCl, 50 mM Tris at pH 7.2) before passage through an 18 g syringe needle 30 times. Cells were pelleted at 5 K for 10 min and the supernatant fraction was saved while the cells were resuspended in TBS to be sheared via syringe twice more. All supernatant fractions were pooled and the archaella were collected by ultracentrifugation in a Beckman TLA100.2 rotor at 60 K for 20 min. The archaella were resuspended in 50 µl of TBS for a final 5 K spin for 10 min to remove unwanted cell debris. New cultures were grown, and samples were prepared fresh for each experiment.

SDS–PAGE and glycostaining. *M. hungatei* archaella were incubated for 1 h in 0.1% Triton X-100 (ref. 47), and subjected to SDS PAGE using 4 12% Bis-Tris gradient gels (Invitrogen) according to the manufacturer's instructions. Glycoprotein stains were performed using the Pro-Q Emerald 488 Glycoprotein Stain Kit from Thermo-Fischer (P21855) and incorporating CandyCane glycoprotein molecular weight standards (Molecular Probes) (Supplementary Fig. 7).

Mass spectroscopy and N-terminal identification. Bands excised from an SDS PAGE gel were digested with trypsin in-gel. The proteolytic peptides were injected for analysis by LC-MS/MS on a Thermo-Fisher Q-Exactive. For molecular mass determination, the *M. hungatei* archaella were first dissociated in Triton X-100 as described above. The dissociated archaella were mixed with a ferulic acid matrix and deposited on a sample stage for analysis by MALDI mass spectrometry (MS). MALDI MS was performed in positive ion mode on an Applied Biosystems Voyager-DE-STR MALDI-TOF with 337 nm irradiation. N-terminal identification was performed on archaellin using Edman sequencing as previously described[49].

Bioinformatics. Protein sequence alignments were performed with Clustal Omega[50] using default parameters. Percent identity scores were generated with BLASTP[51]. Parameters were changed to compensate for the short length of the protein sequence; specifically, the word size was reduced to 2 and the expected threshold was reduced to 8. Protein secondary structure was predicted *in silico* using psipred[52].

Negative stain electron microscopy. Purified archaella (4 µl) were pipetted onto carbon-coated copper grids and allowed to settle for 1 min before blotting to remove excess sample. A 2 µl volume of 1% uranyl acetate was added for 1 min and excess stain was blotted away. Data were collected using the FEI TF20 microscope with a 16-megapixel charge-coupled device (CCD) camera and the automated imaging software Leginon[53] (Supplementary Fig. 8a c). For cryoEM samples, the concentration of the *M. hungatei* archaella was first optimized by negatively staining the sample with 1% uranyl acetate and imaged in an FEI TF12 microscope with a Gatan 2k×2k CCD camera. Samples were diluted as needed to achieve the appropriate sample concentration.

CryoEM. Three successive cryoEM data sets were collected using the same FEI Titan Krios microscope, a data set on photographic film (Supplementary Fig. 8d f), a test data set collected with the FEI Falcon II direct electron detector (Supplementary Fig. 8g i), and the final data set used to reconstruct and build the *de novo* atomic model described in the paper collected with the Gatan K2 Summit direct electron detector (Supplementary Fig. 8j l). For all three data sets, an aliquot of 2 4 µl of the sample was applied to Quantifoil grids and an FEI Vitrobot was used to plunge-freeze the sample in liquid ethane cooled by liquid nitrogen.

For the first cryoEM data set collected on photographic films, an accelerating voltage of 120 kV was used at a magnification of ×59,000 and the defocus values of the images collected were between 2.5 and 4.0 µm. The automated imaging software Leginon was used to collect data. A total of 1,014 images were collected on film over 12 days and digitized using a Nikon Super Coolscan 9000, giving a pixel size of 1.02 Å per pixel on the specimen (Supplementary Fig. 8d). This data set is very noisy, and the thin filaments are difficult to see in the raw micrographs. However, at the time, we thought the flexibility and lack of contrast provided by the archaella limited the resolution of the refinement, not the use of film.

The cryoEM data set collected using the Falcon II electron detector in a pre-GIF configuration was a test data set of about 800 videos collected over a day and a half. Each video contains 24 frames with a total exposure of 8 s and a total dose of about 40 electrons per Å^2 per pixel. Direct electron detectors collect a series of very low-dose images that are combined into a video. The frames in the video are then aligned to correct for sample drift and beam-induced damage using whole frame alignment using the MotionCorr program[54]. The accelerating voltage was 300 kV, the magnification was 97.3k with a pixel size of 1.43 Å per pixel on the specimen, and the defocus values of the images collected were between 1.5 and 3 µm. The FEI automated imaging software EPU was used to collect data (Supplementary Fig. 8g). Imaging with the Falcon II provided a marked improvement in data quality over the previous film and negative stain data sets.

The final data set was recorded on a Gatan K2 Summit electron detector attached to a Gatan imaging filter (GIF Quanta) in counting mode with the automated imaging software Leginon (Supplementary Fig. 8j). To improve frame alignment, we also added gold fiducials (a 1:50 dilution of 10 nm gold fiducials was added to the optimized samples before freezing). Each video contains 30 frames with a total exposure of 7.5 s and an accumulated dosage of ~80 electrons per Å^2 per video, and frames were aligned using the MotionCorr program. A total of 3,300 videos were collected over four days with an accelerating voltage of 300 kV, with a pixel size of 1.3 Å per pixel on the specimen, and defocus between 1.0 and 2.5 µm.

Helical parameters. Identification of helical parameters went through multiple iterations as the quality of available data improved. First, we sought to generate a

high-resolution layer-line image using cryoEM images from the photographic film data set (Supplementary Fig. 8f). However, the low contrast and high noise in higher-resolution regions prevented sound alignment of filament segments.

Next, to improve the contrast, we used negatively stained micrographs to generate a layer-line image (Supplementary Fig. 8c). However, we were unable to determine the n numbers of many reflections, including the layer-lines $l = 1$ and $l = 2$, leading to major ambiguity of helical parameter assignment. Applying these parameters to the cryoEM data set on photographic films, we generated multiple possible structures, but we were unable to improve these structures beyond 7 Å resolution in any possible helical assignment.

The helical parameters were eventually established with direct electron detector data. Helical parameters were initially guessed from Fourier transforms of two-dimensional class averages (Supplementary Fig. 8i,l) by classical helical cryoEM methods[55]. Briefly, main reflections on the incoherently averaged Fourier transforms of two-dimensional class averages are aligned on a two-dimensional lattice (mesh) and their Bessel orders (n, l) are assigned empirically. The selection rule is therefore deduced from these n, l numbers and is converted to the format of turn and rise per subunit for processing with IHRSR (see section on image processing for details).

More importantly, the final structure shows amino acid side chains, providing a series of solid internal controls for the established helical parameters by matching the densities to the primary sequence.

Image processing packages used. As detailed in the reconstruction process below, we used a combination of different software packages for helical reconstructions. MotionCorr was used for whole frame alignment, EMAN[56] for image processing, particle boxing and two-dimensional classification, IHRSR (ref. 57) for low-resolution helical reconstruction, a modified EMAN package implemented with the IHRSR method for the 7 Å resolution reconstruction[58], and lastly a modified RELION[59] with IHRSR method for both the 4.8 Å resolution reconstruction and the 3.4 Å resolution reconstruction reported in this paper[60].

Image processing and structure refinement. For both the negatively stained and cryoEM on photographic film data sets, image processing was performed using EMAN 1.8 and IHRSR. CTF parameters were determined using ctffind3 and particles were boxed out with *helixboxer* in Eman 1.8 using a box size of 432 pixels. Two-dimensional classification was performed on both data sets in EMAN. The negative stain data set was processed first to estimate the helical parameters as described above. The helicity determined from the negative stain data and the initial model were used to refine the cryoEM data set collected on photographic films to ~7 Å resolution (Supplementary Fig. 9a).

Frames from the videos collected using the Falcon II electron detector were aligned with the first and last six frames discarded for a total exposure dose of about 30 electrons per Å^2. CTF correction was performed using ctffind3 and particles were boxed using *helixboxer* in EMAN 1.8 using a box size of 256 pixels. Reducing the box size mitigated the effects of filament curvature on the reconstruction. We use RELION for two-dimensional classification. We modified RELION to utilize the IHRSR method for three-dimensional classification and final refinement (Supplementary Fig. 9b).

Frames from the videos collected using the K2 Summit electron detector were first aligned. All frames were used to determine the ctf with ctffind3 and box particles in EMAN1.8 using a box size of 256 pixels. For the final reconstruction, the first two frames were discarded and frames with a total dose of 20 electrons per Å^2 were used. As described above, particles were boxed out with EMAN and three-dimensional structures were determined with the modified RELION (Fig. 1e and Supplementary Fig. 9c).

Once the final cryoEM density map was obtained, the protein backbone of the archaellin monomer was traced and amino acids assigned in Coot[61] using the baton building method. First, a poly-alanine backbone was built starting from the easily identifiable C terminus. Amino-acid side chains were then mutated based on the available sequence. Initial refinement was performed on a single subunit using pseudo crystallographic restraints with the PHENIX[62] and CNS software packages. Chi angles and rotomers were manual adjusted in Coot, and further refinement was performed in PHENIX using a protofilament of 26 subunits. Figures and videos were created in Chimera[63] and Filmora.

Data availability. The data that support the findings of this study are available from the corresponding author upon request. The cryoEM map and the atomic model have been deposited in the protein databank and EMDB under accession numbers 5TFY and EMD-8405, respectively.

Received 21 August 2016; accepted 7 October 2016; published xx xx 2016

References

1. Albers, S.-V. & Jarrell, K. F. The archaellum: how archaea swim. *Front. Microbiol.* **6,** 23 (2015).
2. Näther, D. J., Rachel, R., Wanner, G. & Wirth, R. Flagella of *Pyrococcus furiosus*: multifunctional organelles, made for swimming, adhesion to various surfaces, and cell cell contacts. *J. Bacteriol.* **188,** 6915 6923 (2006).
3. Szabó, Z. *et al.* Flagellar motility and structure in the hyperthermoacidophilic archaeon *Sulfolobus solfataricus*. *J. Bacteriol.* **189,** 4305 4309 (2007).
4. Jarrell, K. F. & Albers, S.-V. The archaellum: an old motility structure with a new name. *Trends Microbiol.* **20,** 307 312 (2012).
5. Bellack, A., Huber, H., Rachel, R., Wanner, G. & Wirth, R. *Methanocaldococcus villosus* sp. nov., a heavily flagellated archaeon that adheres to surfaces and forms cell cell contacts. *Int. J. Syst. Evol. Micr.* **61,** 1239 1245 (2011).
6. Jarrell, K. F., Stark, M., Nair, D. B. & Chong, J. P. J. Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. *FEMS Microbiol. Lett.* **319,** 44 50 (2011).
7. Schopf, S., Wanner, G., Rachel, R. & Wirth, R. An archaeal bi-species biofilm formed by *Pyrococcus furiosus* and *Methanopyrus kandleri*. *Arch. Microbiol.* **190,** 371 377 (2008).
8. Alam, M. & Oesterhelt, D. Morphology, function and isolation of halobacterial flagella. *J. Mol. Biol.* **176,** 459 475 (1984).
9. Faguy, D. M., Jarrell, K. F., Kuzio, J. & Kalmokoff, M. L. Molecular analysis of archaeal flagellins: similarity to the type IV pilin-transport superfamily widespread in bacteria. *Can. J. Microbiol.* **40,** 67 71 (1994).
10. Correia, J. D. & Jarrell, K. F. Posttranslational processing of *Methanococcus voltae* preflagellin by preflagellin peptidases of *M.* voltae and other methanogens. *J. Bacteriol.* **182,** 855 858 (2000).
11. Patenge, N., Berendes, A., Engelhardt, H., Schuster, S. C. & Oesterhelt, D. The *fla* gene cluster is involved in the biogenesis of flagella in *Halobacterium salinarum*. *Mol. Microbiol.* **41,** 653 663 (2001).
12. Thomas, N. A., Pawson, C. T. & Jarrell, K. F. Insertional inactivation of the *flaH* gene in the archaeon *Methanococcus voltae* results in non-flagellated cells. *Mol. Genet. Genomics* **265,** 596 603 (2001).
13. Chaban, B. *et al.* Systematic deletion analyses of the *fla* genes in the flagella operon identify several genes essential for proper assembly and function of flagella in the archaeon, *Methanococcus maripaludis*. *Mol. Microbiol.* **66,** 596 609 (2007).
14. Lassak, K. *et al.* Molecular analysis of the crenarchaeal flagellum. *Mol. Microbiol.* **83,** 110 124 (2012).
15. Streif, S., Staudinger, W. F., Marwan, W. & Oesterhelt, D. Flagellar rotation in the archaeon *Halobacterium salinarum* depends on ATP. *J. Mol. Biol.* **384,** 1 8 (2008).
16. Banerjee, A. *et al.* Flaf is a β-sandwich protein that anchors the archaellum in the archaeal cell envelope by binding the S-layer protein. *Structure.* **23,** 863 872 (2015).
17. Chaudhury, P. *et al.* The nucleotide-dependent interaction of FlaH and FlaI is essential for assembly and function of the archaellum motor. *Mol. Microbiol.* **99,** 674 685 (2016).
18. Trachtenberg, S., Galkin, V. E. & Egelman, E. H. Refining the structure of the *Halobacterium salinarum* flagellar filament using the iterative helical real space reconstruction method: insights into polymorphism. *J. Mol. Biol.* **346,** 665 676 (2005).
19. Trachtenberg, S. & Cohen-Krausz, S. The archaeabacterial flagellar filament: a bacterial propeller with a pilus-like structure. *J. Mol. Microb Biotech.* **11,** 208 220 (2006).
20. Beveridge, T. J., Sprott, G. D. & Whippey, P. Ultrastructure, inferred porosity, and Gram-staining character of *Methanospirillum hungatei* filament termini describe a unique cell permeability for this archaeobacterium. *J. Bacteriol.* **173,** 130 140 (1991).
21. Xu, W. *et al.* Modeling and measuring the elastic properties of an archaeal surface, the sheath of *Methanospirillum hungatei*, and the implication of methane production. *J. Bacteriol.* **178,** 3106 3112 (1996).
22. Toso, D. B., Henstra, A. M., Gunsalus, R. P. & Zhou, Z. H. Structural, mass and elemental analyses of storage granules in methanogenic archaeal cells. *Environ. Microbiol.* **13,** 2587 2599 (2011).
23. Gunsalus, R. P. *et al.* Complete genome sequence of *Methanospirillum hungatei* type strain JF1. *Standard Genomic Sci.* **11,** 2 (2016).
24. Bardy, S. L. & Jarrell, K. F. Flak of the archaeon *Methanococcus maripaludis* possesses preflagellin peptidase activity. *FEMS Microbiol. Lett.* **208,** 53 59 (2002).
25. Bardy, S. L. & Jarrell, K. F. Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. *Mol. Microbiol.* **50,** 1339 1347 (2003).
26. Szabó, Z. *et al.* Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J. Bacteriol.* **189,** 772 778 (2007).
27. Albers, S.-V., Szabó, Z. & Driessen, A. J. M. Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J. Bacteriol.* **185,** 3918 3925 (2003).
28. Ng, S. Y. M., Chaban, B. & Jarrell, K. F. Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *J. Mol. Microbiol. Biotech.* **11,** 167 191 (2006).
29. Kelly, J., Logan, S. M., Jarrell, K. F., VanDyke, D. J. & Vinogradov, E. A novel N-linked flagellar glycan from *Methanococcus maripaludis*. *Carbohyd. Res.* **344,** 648 653 (2009).

30. Voisin, S. *et al.* Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of *Methanococcus voltae*. *J. Biol. Chem.* **280,** 16586 16593 (2005).
31. Sun, S. & Zhang, H. Identification and validation of atypical *N*-glycosylation sites. *Anal. Chem.* **87,** 11948 11951 (2015).
32. Banerjee, A., Neiner, T., Tripp, P. & Albers, S.-V. Insights into subunit interactions in the *Sulfolobus acidocaldarius* archaellum cytoplasmic complex. *FEBS J.* **280,** 6141 6149 (2013).
33. Jarrell, K. F., Bayley, D. P. & Kostyukova, A. S. The archaeal flagellum: a unique motility structure. *J. Bacteriol.* **178,** 5057 5064 (1996).
34. Tripepi, M. *et al. N*-glycosylation of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. *J. Bacteriol.* **194,** 4876 4887 (2012).
35. Ding, Y. *et al.* Effects of *N*-glycosylation site removal in archaellins on the assembly and function of archaella in *Methanococcus maripaludis*. *PLoS ONE* **10,** e0116402 (2015).
36. Meyer, B. H. *et al.* Agl16, a thermophilic glycosyltransferase mediating the last step of *N*-glycan biosynthesis in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. *J. Bacteriol.* **195,** 2177 2186 (2013).
37. Meyer, B. H., Birich, A. & Albers, S.-V. *N*-glycosylation of the archaellum filament is not important for archaella assembly and motility, although *N*-glycosylation is essential for motility in *Sulfolobus acidocaldarius*. *Biochimie* **118,** 294 301 (2015).
38. Turner, L., Ryu, W. S. & Berg, H. C. Real-time imaging of fluorescent flagellar filaments. *J. Bacteriol.* **182,** 2793 2801 (2000).
39. Yonekura, K., Maki-Yonekura, S. & Namba, K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424,** 643 650 (2003).
40. Craig, L. *et al.* Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol. Cell* **23,** 651 662 (2006).
41. Mortezaei, N. *et al.* Structure and function of enterotoxigenic *Escherichia coli* fimbriae from differing assembly pathways. *Mol. Microbiol.* **95,** 116 126 (2015).
42. Silverman, P. M. Towards a structural biology of bacterial conjugation. *Mol. Microbiol* **23,** 423 429 (1997).
43. McLaughlin, L. S., Haft, R. J. F. & Forest, K. T. Structural insights into the type II secretion nanomachine. *Curr. Opin. Struc. Biol.* **22,** 208 216 (2012).
44. Yu, X. *et al.* Filaments from *Ignicoccus hospitalis* show diversity of packing in proteins containing N-terminal type IV pilin helices. *J. Mol. Biol.* **422,** 274 281 (2012).
45. Henche, A.-L. *et al.* Structure and function of the adhesive type IV pilus of *Sulfolobus acidocaldarius*. *Environ. Microbiol.* **14,** 3188 3202 (2012).
46. Braun, T. *et al.* Archaeal flagellin combines a bacterial type IV pilin domain with an Ig-like domain. *Proc. Natl Acad. Sci. USA* **113,** 10352 10357 (2016).
47. Faguy, D. M., Koval, S. F. & Jarrell, K. F. Effect of changes in mineral composition and growth temperature on filament length and flagellation in the Archaeon *Methanospirillum hungatei*. *Arch. Microbiol.* **159,** 512 520 (1993).
48. Patel, G. B., Roth, L. A., Berg, L. v. d. & Clark, D. S. Characterization of a strain of *Methanospirillum hungatii*. *Can. J. Microbiol.* **22,** 1404 1410 (1976).
49. Edman, P. & Begg, G. A protein sequenator. *Eur. J. Biochem.* **1,** 80 91 (1967).
50. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7,** 539 (2011).
51. Altschul, S. F. *et al.* Protein database searches using compositionally adjusted substitution matrices. *FEBS J.* **272,** 5101 5109 (2005).
52. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292,** 195 202 (1999).
53. Suloway, C. *et al.* Fully automated, sequential tilt-series acquisition with Leginon. *J. Struct. Biol.* **167,** 11 18 (2009).
54. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10,** 584 590 (2013).
55. DeRosier, D. J. & Moore, P. B. Reconstruction of three-dimensional images from electron micrographs of structures with helical symmetry. *J. Mol. Biol.* **52,** 355 369 (1970).
56. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128,** 82 97 (1999).
57. Egelman, E. H. The iterative helical real space reconstruction method: surmounting the problems posed by real polymers. *J. Struct. Biol.* **157,** 83 94 (2007).
58. Ge, P. *et al.* Cryo-EM model of the bullet-shaped vesicular stomatitis virus. *Science* **327,** 689 693 (2010).
59. Scheres, S. H. W. A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415,** 406 418 (2012).
60. Clemens, D. L., Ge, P., Lee, B.-Y., Horwitz, M. A. & Zhou, Z. H. Atomic structure of T6SS reveals interlaced array essential to function. *Cell* **160,** 940 951 (2015).
61. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66,** 486 501 (2010).
62. Adams, P. D. *et al. PHENIX*: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213 221 (2010).
63. Pettersen, E. F. *et al.* UCSF chimera a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605 1612 (2004).

Acknowledgements

This project received support from National Institutes of Health grants GM071940 and AI094386, NIH/NCRR/NCATS UCLA CTSI grant UL1TR000124, from the UCLA-DOE Institute (DE-FC03-02ER6342) to R.P.G. and R.O.L., and NSF 1515843 to R.P.G. N.P. was supported in part by the NIH Biotechnology Training Grant Program (T32GM067555). P.G. was supported in part by an American Heart Association Western States Affiliates Postdoc Fellowship (13POST17340020). The authors acknowledge the use of instruments at the Electron Imaging Center for Nanomachines supported by UCLA and by instrumentation grants from NIH (1S10OD018111) and NSF (DBI-1338135). NIH support for mass spectrometry was provided by grant S10RR025600. The authors acknowledge computer time at the Extreme Science and Engineering Discovery Environment (XSEDE, grant MCB140140 to Z.H.Z.).

Author contributions

Z.H.Z. and R.P.G. designed the project. N.P., P.G., R.R.O.L. and H.H.N. performed the experiments and analysed the data. Z.H.Z., R.P.G. and N.P. wrote the paper. All authors contributed to editing the manuscript.

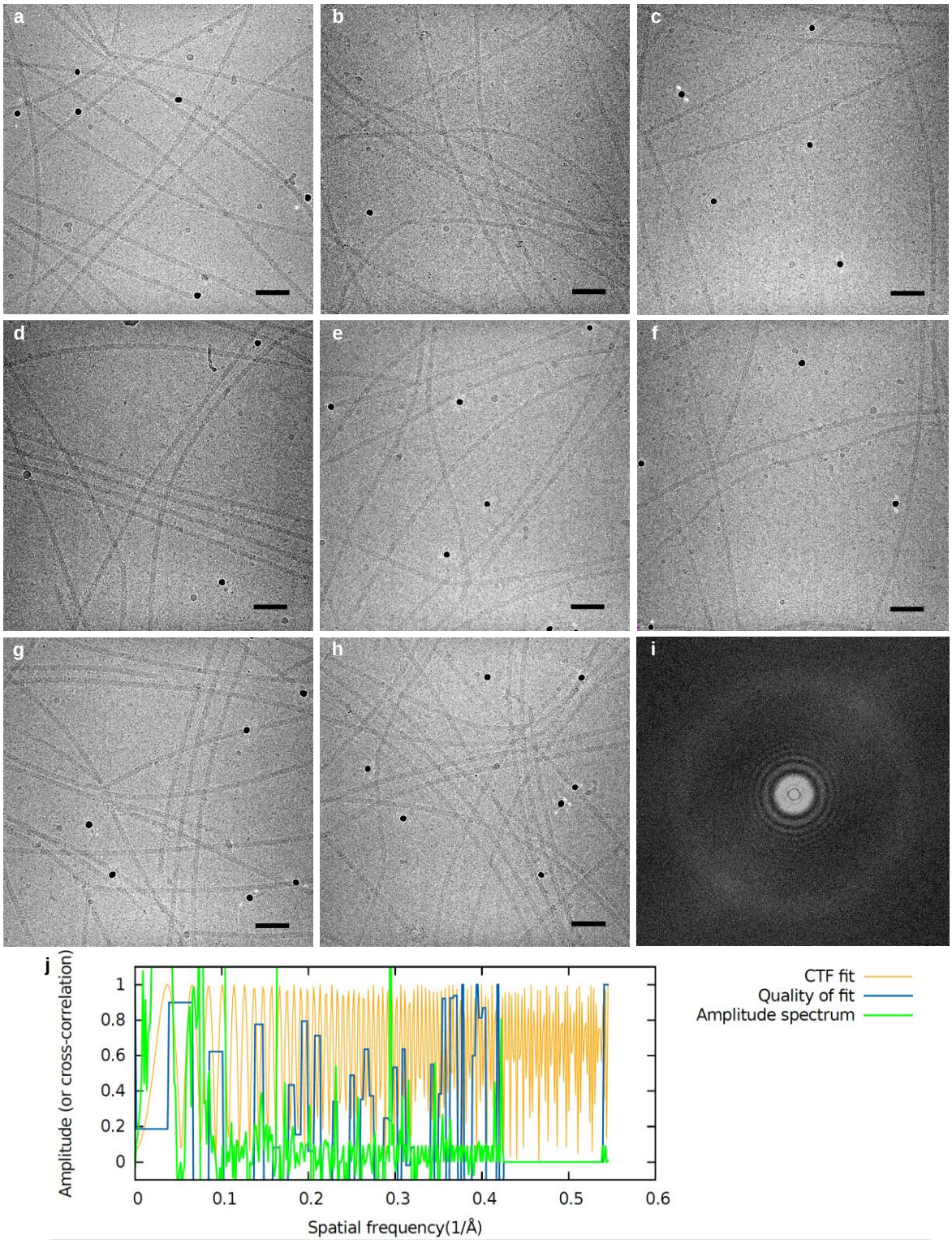
Additional information

Supplementary information is available for this paper. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Z.H.Z.

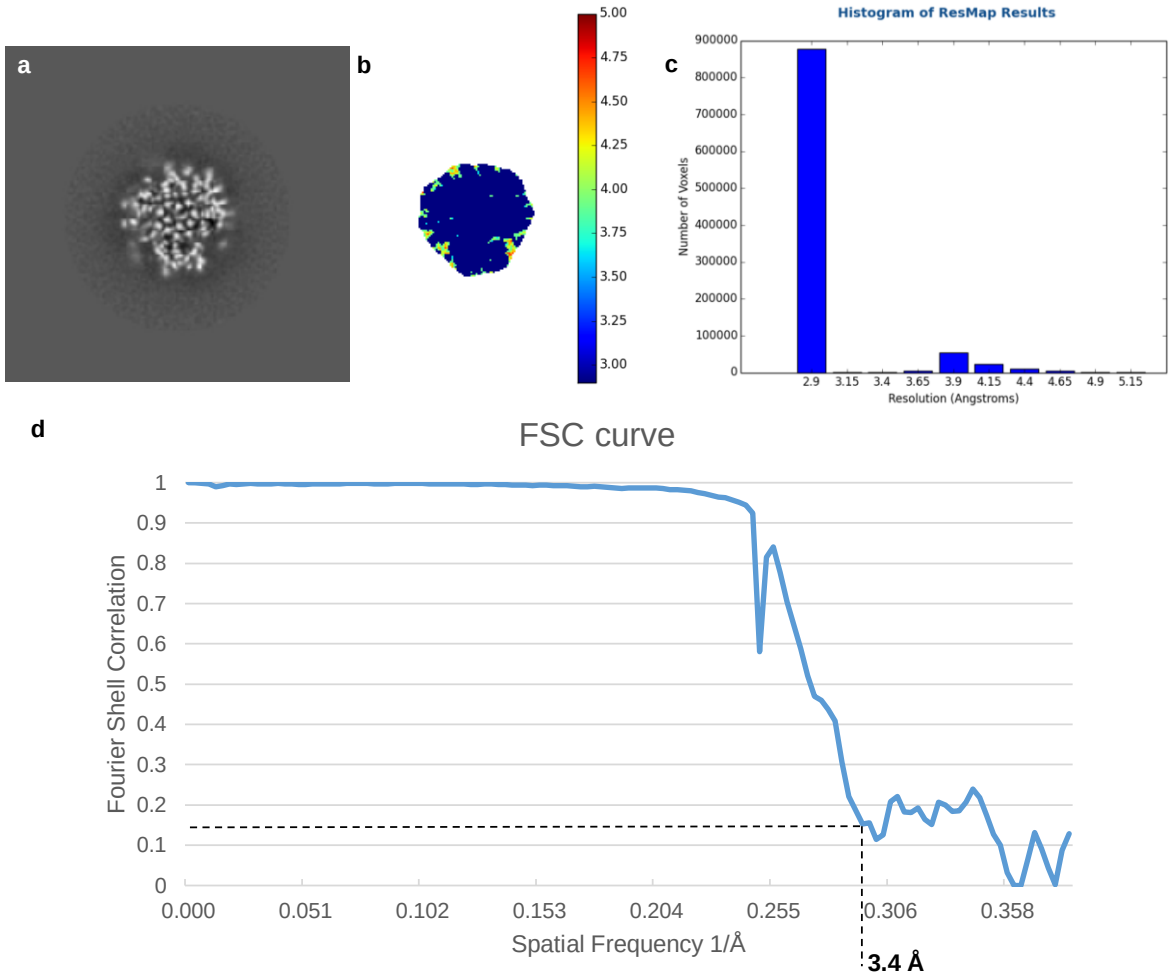
Competing interests

The authors declare no competing financial interests.

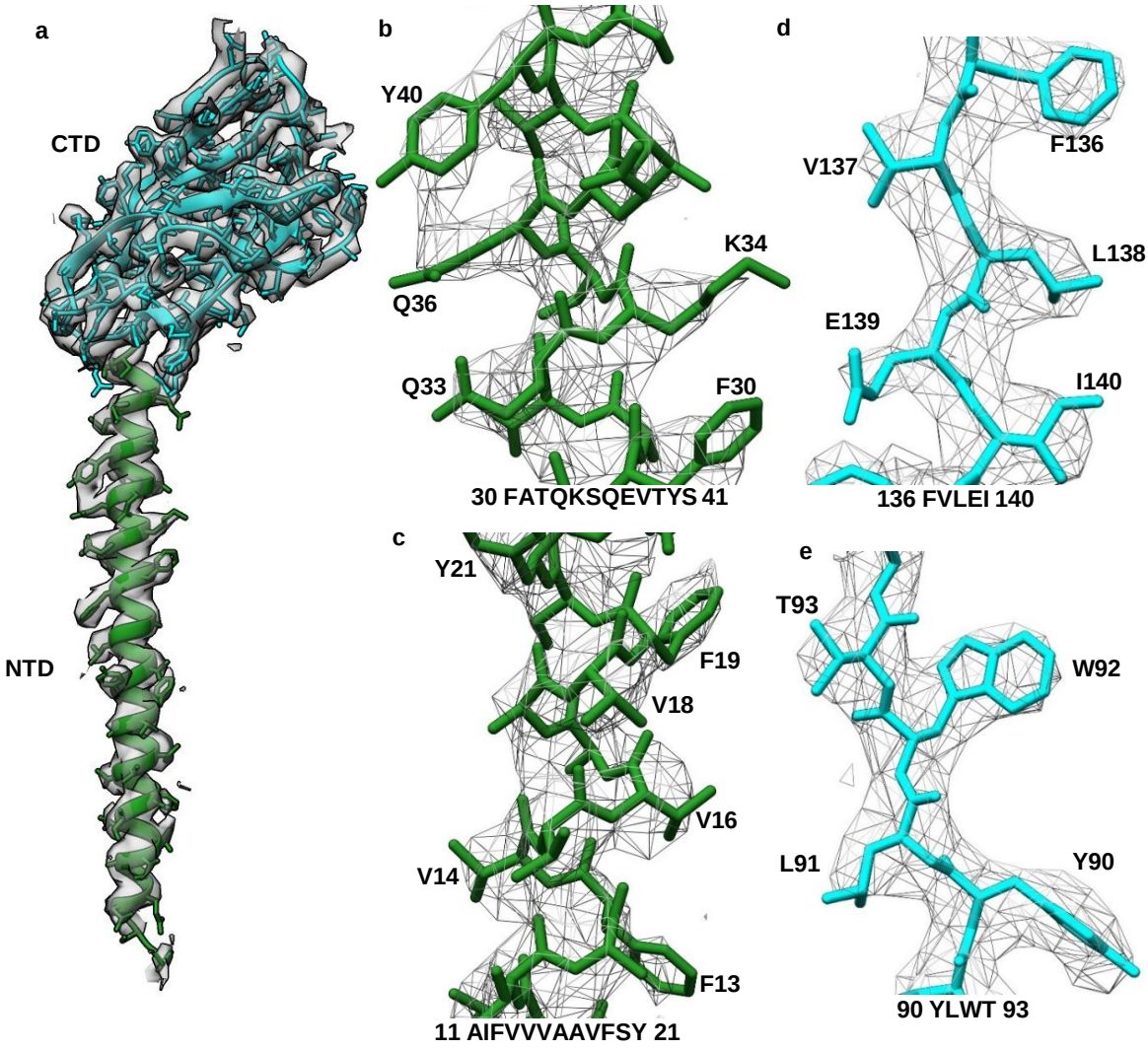
Supplemental Figures



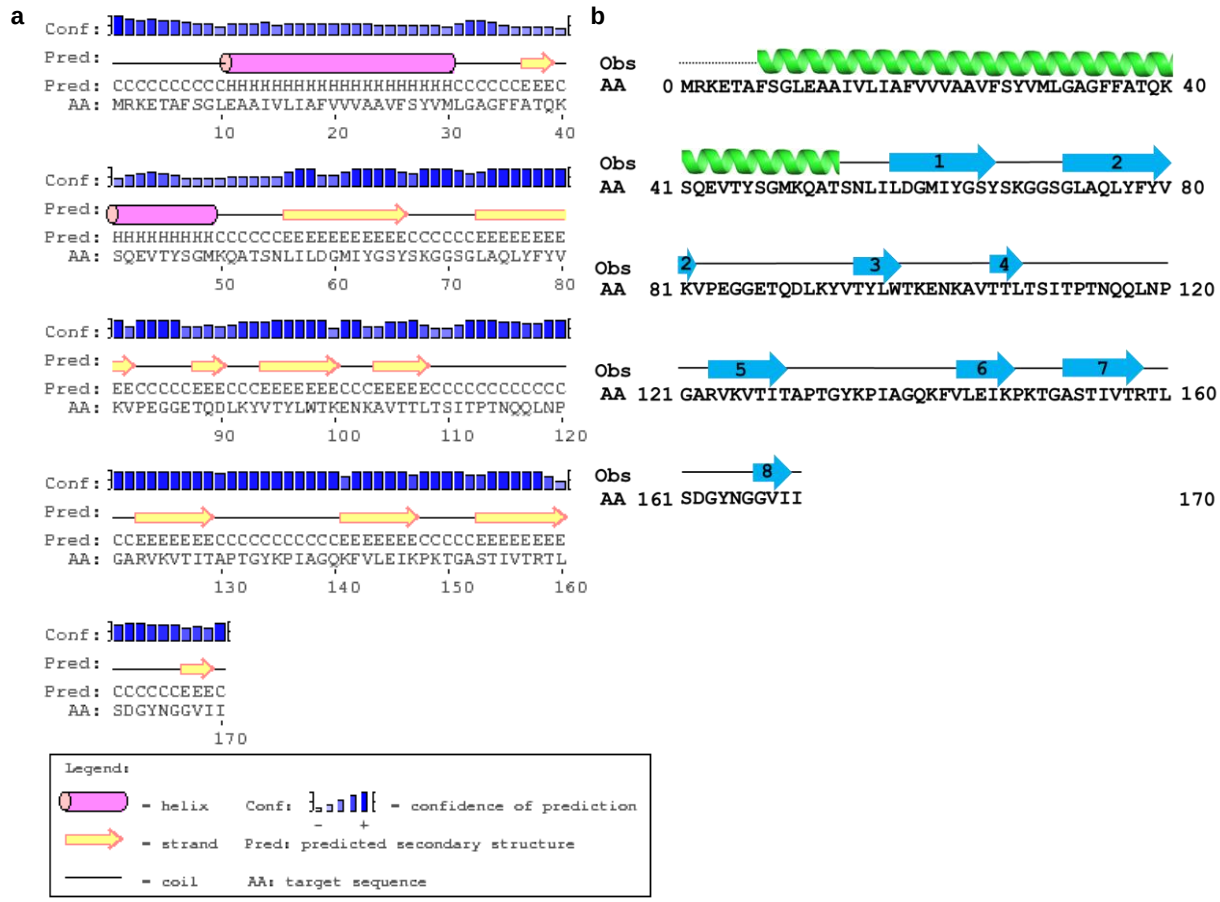
Supplementary Figure 1 | Additional representative micrographs from the K2 data set. a-h, Representative aligned movies, all scale bars 50 nm. **i,** Representative Fourier transforms of aligned movies in the dataset. **j,** A complementary rotational average of the Fourier transform shown in i.



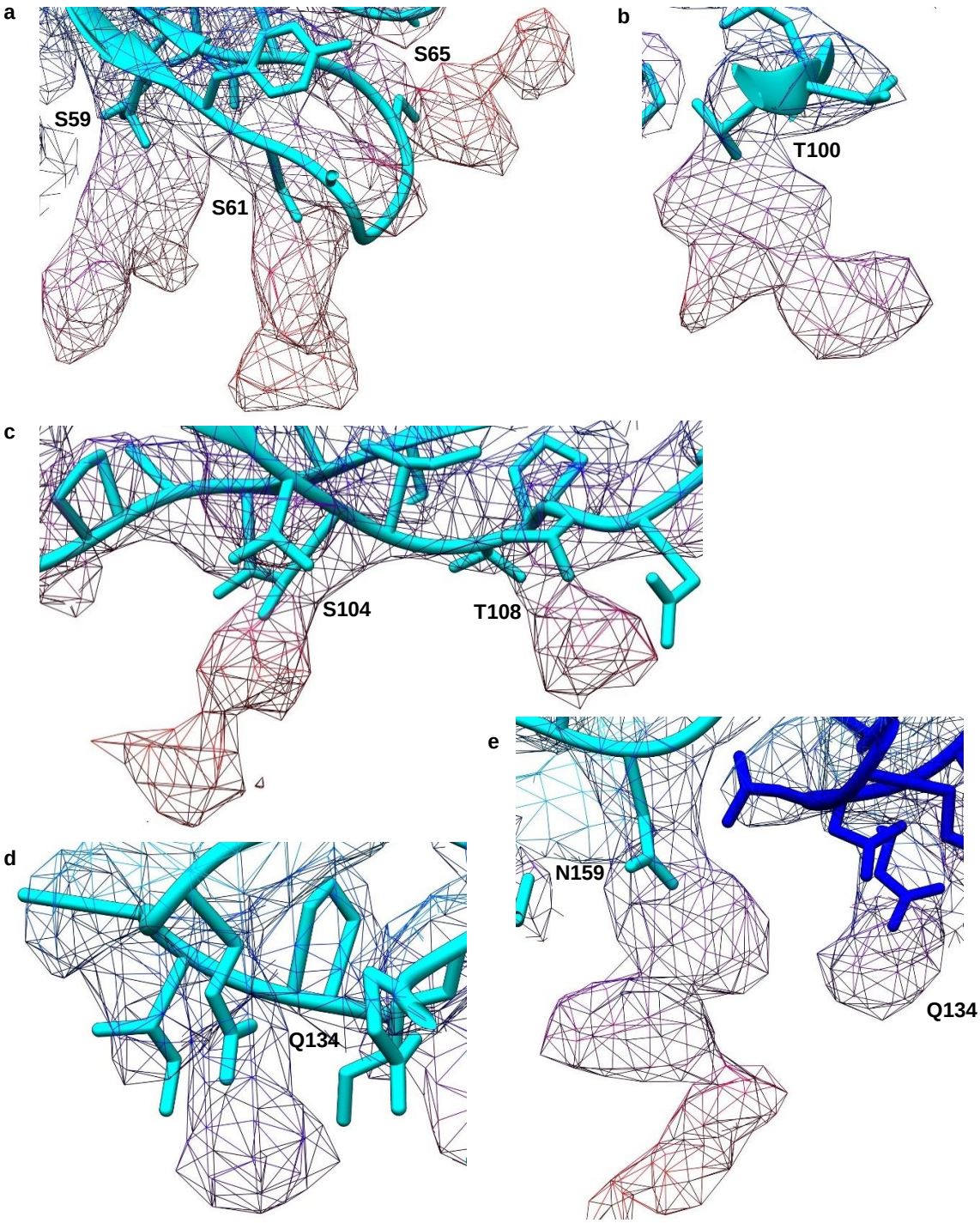
Supplementary Figure 2 | Resolution validation. a, Cross section of the *M. hungatei* flagella final refinement (before application of b-factors). **b,** Estimated resolution of the voxels from the slice depicted in a as determined by ResMap. **c,** Histogram of the resolution of each voxel in the reconstructed *M. hungatei* flagella as determined by ResMap. **d,** Fourier shell correlation curve from the refinement of the electron microscopy density map showing an average resolution for the structure of 3.4 Å.



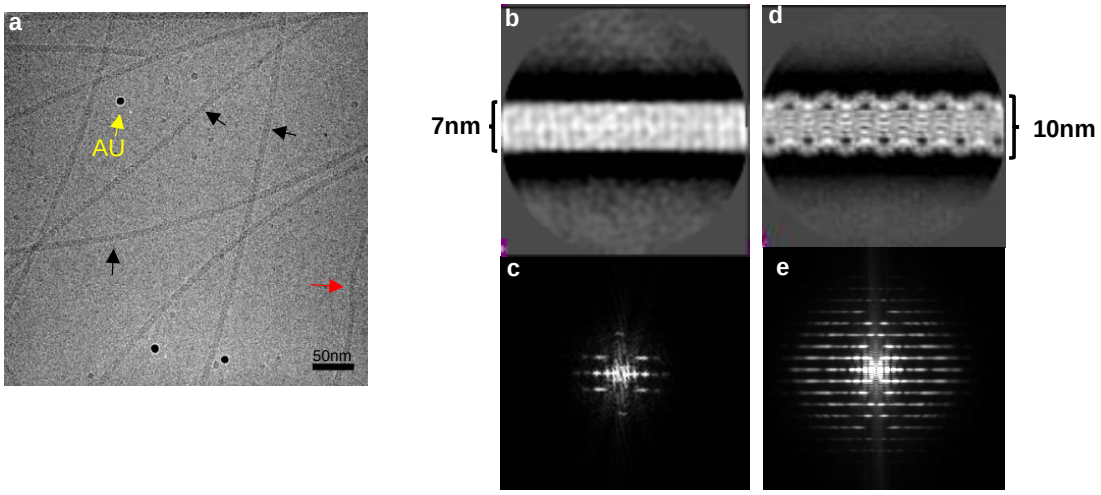
Supplementary Figure 3 | Additional regions of fit between the EM density and atomic model. **a,** One *M. hungatei* flagellin with the EM density map (grey) filtered to 2.9 Å. **b,c,** Atomic model fit to the EM density map in the alpha helix domain to demonstrate the fit of amino acid side chains at this resolution. **d,** Atomic model fit to the EM density map along beta strand #6 in the beta barrel region. **e,** Atomic model to the EM density map along beta strand #3 in the beta barrel domain.



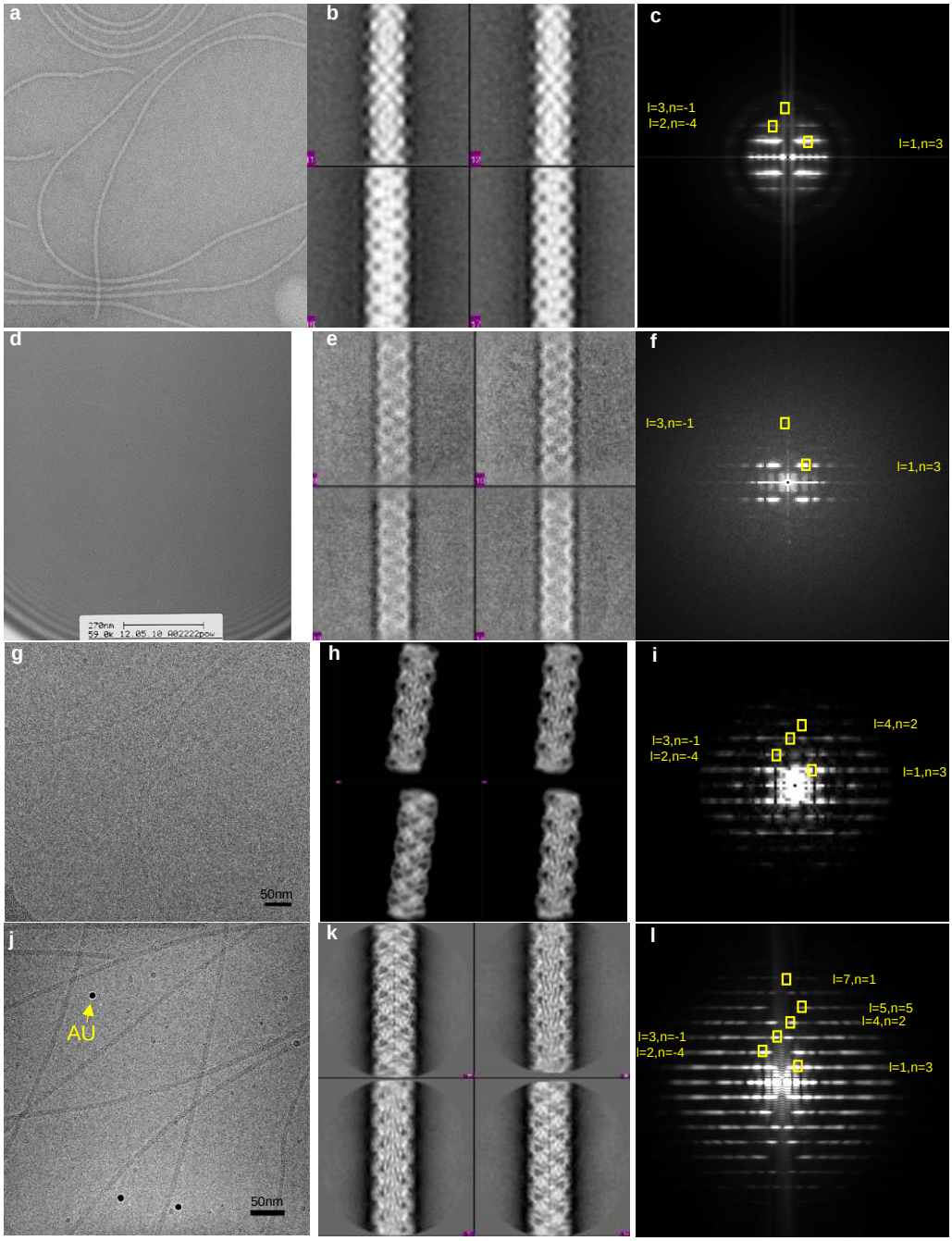
Supplementary Figure 4 | Comparison between predicted secondary structure and observed secondary structure of Mhun_3140. a, Predicted secondary structure of the Mhun_3140 gene product from psipred. **b,** Observed secondary structure of the *M. hungatei* flagella.



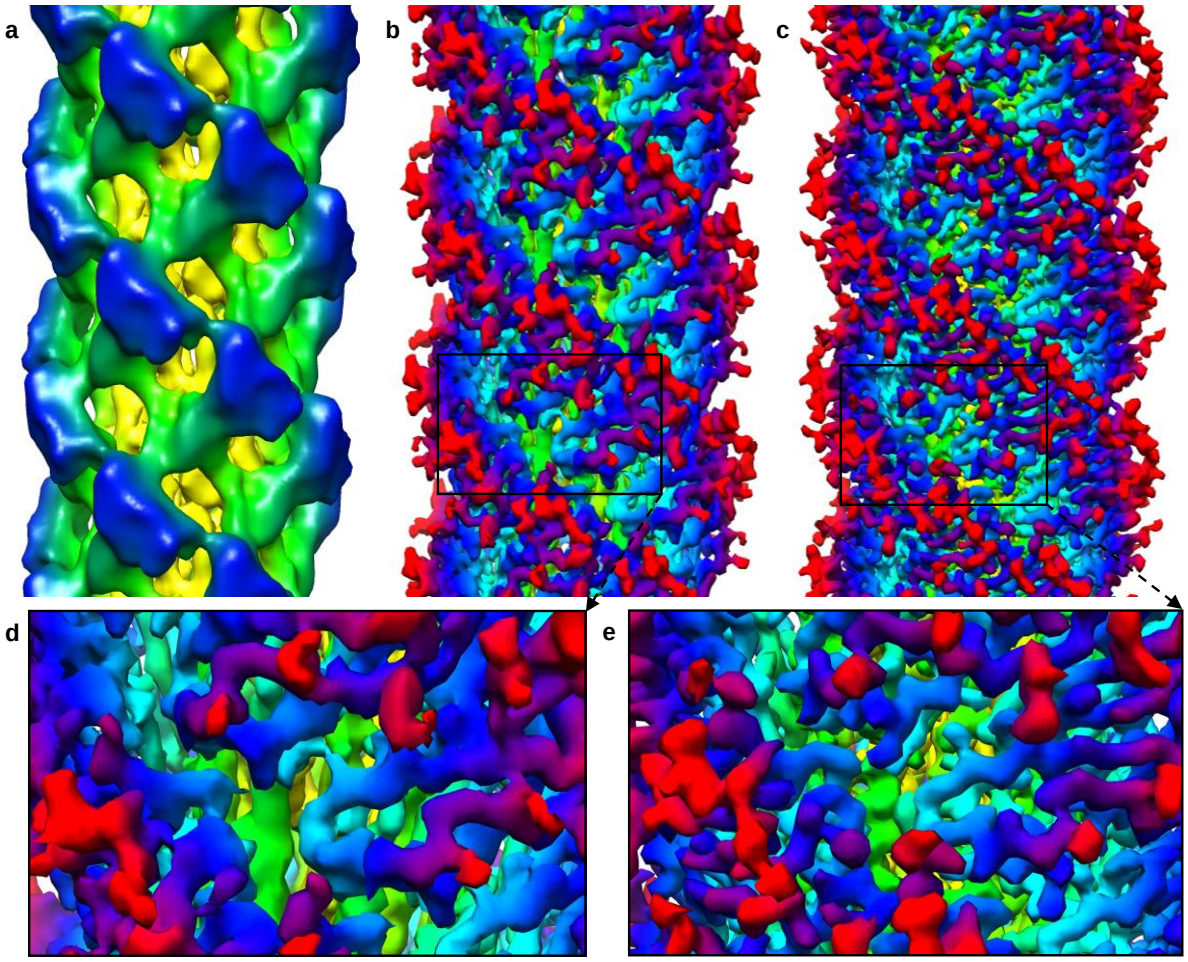
Supplementary Figure 5 | Local context of all post translational modifications. a, Archaeal flagellin monomer from glycine 58 to alanine 68 with extra cryoEM density extending from serine residues 59,61, and 65 clearly visible. The cryoEM mesh is colored red in surface exposed areas changing to blue and then cyan closer to the helical axis of the filament. **b,** Flagellin from threonine 100 to leucine 102 with extra cryoEM density over threonine 100 visible. **c,** Flagellin from serine 104 to aspargine 109 with extra cryoEM density around serine 104 and threonine 108. **d,** Flagellin from proline 130 to glutamine 134 with extra cryoEM density around glutamine 134. **e,** Two adjacent flagellins (cyan and blue) with extra cryoEM density around asparagine 159 (cyan) and glutamine 134 (blue).



Supplementary Figure 6 | Minority filament. a, Aligned movie from K2 cryoEM dataset with thinner filament indicated by a red arrow. **b,** Ctf corrected 2D classification of minority filament which is about 7 nm in diameter. **c,** Fourier transform of the class2D representation in b. **d,** Ctf corrected 2D classification of majority filament (flagellin) is 10 nm in diameter. **e,** Fourier transform of the class2D representation in d.



Supplementary Figure 7 | Representative data from four *M. hungatei* flagella datasets. a, Representative micrograph from the negatively stained flagella dataset. **b,** Two-dimensional class averages from the negatively stained dataset. **c,** Fourier transform from the negatively stained dataset with helical indexing. **d,** Representative micrograph from the cryoEM collected on photographic films dataset. **e,** Two-dimensional class averages from the film dataset. **f,** Fourier transform of the class averages from the film dataset with helical indexing. **g,** Representative aligned movie from the cryoEM dataset collected on the FEI Falcon II direct electron detector. **h,** Two-dimensional class averages from the Falcon II dataset. **i,** Fourier transform of class averages from the Falcon II dataset with helical indexing. **j,** Representative aligned movie from the cryoEM dataset collected on the Gatan K2 Summit direct electron detector. **k,** Two-dimensional class averages from the K2 dataset. **l,** Fourier transform of class averages from the K2 dataset with helical indexing.



Supplementary Figure 8 | Comparison between reconstructions of three cryo EM *M. hungatei* flagella datasets. a, CryoEM data collected on photographic films resolved to about 7Å resolution using the (incorrect) helicity determined from the negative stain dataset of 8 subunits per 3 turns. **b,** A small test cryoEM dataset collected on the FEI Falcon II direct electron detector and resolved to 4.8 Å resolution using the correct helicity of 10 subunits per 3 turns. **c,** CryoEM data collected with Gatan K2 Summit direct electron detector (data set used in this paper) at 3.4 Å resolution using the correct helicity of 10 subunits per 3 turns. **d,** Closer view of the Falcon II dataset reconstruction. **e,** Closer view of the K2 dataset reconstruction

K2 Summit Dataset	
Data collection	
Pixel size (Å)	1.325
Defocus range (μm)	1.5-3
Voltage (kV)	300
Particle overview	
Total particles boxed	175,801
Particles used in Class3D	154,278
Particles used in final refinement	90,118
Minority filament particles	6,441
Refinement	
Resolution	3.4
Map sharpening B factor (Å^2)	100
Model validation	
R-work	0.3575
R-free	0.3771
All-atom clashscore (percentile)	17.26
Poor rotamers (%)	2.7
Ramachandran favored (%)	95.1
Ramachandran allowed (%)	4.9
Ramachandran outliers (%)	0
r.m.s.d bonds (Å)	0.0068

Supplementary Table 1 | **Refinement and model statistics for the dataset collected on the K2 Summit direct electron detector.**

Log Number	gi|			Name	Protein Score	Unique Seqs	Single or Interesting Peptides
			Met. hungatei Flagella 35K				
F200608	88604366	Mhun_3140	YP_504544.1	flagellin	229	6	
F200608	88603503	Mhun_2257	YP_503681.1	methylenetetrahydromethanopterin reductase	163	5	
F200608	88602826	Mhun_1554	YP_503004.1	beta lactamase-like protein	78	2	
F200608	88602947	Mhun_1648	YP_503125.1	adenosine/AMP deaminase	40	1	R.YIYEALR.F
F200608	88601953	Mhun_0656	YP_502131.1	50S ribosomal protein L1	19	< 1+	K.IAVLGSGDIVTQAR.E (19) & K.TTMGPSVK.V ox (7)
Error Tolerant Below							
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.SQEVTYSGMK.Q ox
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.VPEGGETQDLK.Y
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.YVTYLWTK.E
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.YVTYLWTK.E +16
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.YVTYLWTK.E +32
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.YVTYLWTK.E +32
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		R.VKVTITAPTG.Y -0.984
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.VTITAPTGYK.P
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.VTITAPTGYKPIAG.Q
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.VTITAPTGYKPIAGQK.F
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.PKTGASTIVTR.T +14
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.TGASTIVTR.T
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.TGASTIVTR.T +28
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.TGASTIVTR.T +28
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		K.TGASTIVTR.T +96
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		KT.GASTIVTR.T + 141
F200998	88604366	Mhun_3140	YP_504544.1	flagellin	360=ET		KT.GASTIVTR.T + 184

Supplementary Table 2 | Raw Mass spec data form Mhun_3140 identification.

Cycle Number	Amino Acid	Amount Recovered (pmol)
1	X	
2	Ser	8.5
3	Gly	8.3
4	Leu	9.4
5	Glu	8.1
6	Ala	10.2

Supplementary Table 3 | **Edman sequencing data confirm N-terminal identification.** Six cycles of Edman degradation occurred; however, the first amino acid could not be conclusively determined based on mass. Amino acids 2-6 were recovered as SGLEA. The recovered N-terminus of the expressed flagellin is XSGLEA.

Supplementary Information Movie Legends

Movie 1 | Overview of the *M. hungatei* cryoEM density map. :00-:03, Full cryoEM density map with helical axis vertical. **:04-:09,** Cut away of cryoEM density map to show internal filament density with helical axis vertical. **:10-:15,** Rotate full cryoEM density map from vertical to horizontal view of helical axis. **:16-:20,** Cut away of cryoEM density map to show internal filament density with helical axis horizontal. **:21-:26,** Rotate full cryoEM density map from horizontal to vertical. **:27-:31,** Segmentation of a single subunit's density shown in color

Movie 2 | Overview of the *M. hungatei* FlaB monomer model. :00-:03, Single subunit segmentation in color with ribbon of atomic model fades to single subunit segmentation in transparent grey with amino acids of atomic model. **:04-:18,** Rotation of single subunit to show fit of amino acids in cryoEM density map. **:19-30,** Close up fly over alpha helix domain to show fit of amino acids in cryoEM density map. **:31-:50,** Close up fly over beta barrel domain to show fit of amino acids in cryoEM density map. **:51-:56,** Remove cryoEM density and show ribbon view to highlight domain structure. **:57-1:09,** Rotation of monomer in ribbon view to show domain structure.

Movie 3 | Posttranslational modifications – an overview of the extra densities in the cryoEM map. :00-:08, Fly from full filament reconstruction to focusing on regions of post-translational modification. CryoEM density map shown transparently in color with outermost regions in red/pink. Ribbon view of seven subunits with each subunit a different color and the primary subunit being focused on in cyan. Amino acid side chains of posttranslational modified residues are shown on all subunits. **:09-:31,** CryoEM

density of all sites of possible post-translational modification are shown in order of increasing amino acid sequence (i.e. Ser-59, Ser-61, Ser-65, Thr-100, Ser-104, Thr-108, Gln-134, and Asn-159). **:32-:41,** Return to original view.

Movie 4 | Overview of inter subunit interactions. :00-:09, Twenty-one subunits docked into cryoEM density map. **:10-:15,** Nine subunits which form a complete interaction unit for the central cyan residue. **:16-:35,** Showing nine subunits interacting with a focus on the seven subunits which interact along the alpha helical domains. **:36-1:03,** Showing the interaction of Phe-1 residues along the length of the helical axis in the central core of the flagellar filament.

Movie 5 | Focus on inter subunit hydrophobic interactions. This movie shows the hydrophobic interactions between seven adjacent subunits. **:00-:03,** Ribbon of the atomic model of the central cyan subunit which is present in all frames of this movie. **:04-:07,** Hydrophobic surface for the central subunit. **:08-:15,** The three interacting subunits labeled 2, central, and 5 in the cartoon in Fig. 4c in ribbon view and as hydrophobic surfaces. **:16-:22,** The three interacting subunits labeled 3, central, and 6 in the cartoon in Fig. 4c in ribbon view and as hydrophobic surfaces. **:23-:28,** The three interacting subunits labeled 1, central, and 4 in the cartoon in Fig. 4c in ribbon view and as hydrophobic surfaces. **:29-:40,** The entire seven subunit adjacent interaction unit as ribbons and hydrophobic surfaces. **:41-:45,** Hydrophobic surfaces of a twenty-one subunit proteofilament.

Movie 6 | Focus on inter subunit ionic interactions. The central cyan residue forms ionic interactions at four sites with six adjacent subunits as depicted in figure 5e-i (the seven subunits shown are the same interaction unit as shown in movie 5.)

Movie 7 | Comparison between a bacterial pilin, archaeal flagellin, and bacterial flagellin. :00-:14, Side by side comparison of hydrophobic surfaces of a single subunit from left to right of a bacterial pilin, an archaeal flagellin, and a bacterial flagellin. **:15-:26,** Side by side comparison of atomic models of the same bacterial pilin (orange), archaeal flagellin (cyan), and bacterial flagellin (pink). **:27-:39,** Comparison of most similar structural element between bacterial pilin, archaeal flagellin, and bacterial flagellin. **:40-:52,** Comparison of the similar alpha helix domain between bacterial pilin (orange) and archaeal flagellin (cyan).

Movie 8 | Comparison between protofilaments of a bacterial pili, archaeal flagella, and bacterial flagella. :00-:07, From left to right: bacterial pili (orange), archaeal flagella (cyan), and bacterial flagella (pink) protofilaments as seen with vertical helical axis. **:08-:12,** Three filaments viewed with horizontal helical axis. **:13-:20,** Filaments cut along the helical axis to show a slice one helical repeat tall.

CHAPTER 5: Conclusions

In the introduction of this thesis, we gave an overview of the electron microscopy techniques, recent advances in electron microscopy, and how those techniques and advances have been applied to studying microbial interactions. A huge variety of electron microscopy sample preparation and imaging techniques have been used to study microbial interactions. Microbiology as a field stands to benefit immensely from recent advances in electron microscopy as new hardware and software make it possible to study both smaller samples such as protein complexes and extracellular machinery and larger samples, such as whole bacterial and archaeal cells and how cellular components are organized within the cell itself.

For example, it is possible to use SEM and conventional TEM sample preparations to study large samples such as bacterial consortium and symbiosis. The second chapter of this thesis is a study which characterizes the epibiotic symbiosis between *Actinomyces odontolyticus* species XH001 and TM7x. In this study, a variety of techniques including light microscopy, genetic approaches, and SEM were used to describe the relationship between a well-known member of the microbiota, XH001, and a newly described, previously unculturable microbe, TM7x. This study showed how the well-established sample preparation and imaging techniques associated with SEM can still be useful for studying microbial interactions, especially, when used in conjunction with other microscopy and genetic tools. The work answered basic biological questions about the relationship between TM7x and XH001 by showing that TM7x is a parasite of XH001 and causes the host cell stress which adds to any environmental stress

experienced by the host. It also shows that TM7x divides by budding which was previously unknown.

Intermittently sized samples such as whole bacteria can be studied by conventional TEM and cryo electron tomography. The third chapter of this thesis described the cell envelope and identified the surface layer proteins of the bacterium *Syntrophomus wolfei*. Biochemical characterization by SDS-PAGE and mass spectroscopy complimented efforts to describe the cell envelope by negative stain TEM and whole cell cryoEM. This study shows how cryoEM can be used to better understand cell envelope architecture and answer other questions where studying whole bacterial cells rather than purified protein complexes is useful. Very few bacterial S-layer proteins have been characterized biochemically or genetically, and no syntrophic bacterial cell envelopes had previously been described. As syntrophs live in symbiosis or consortium with other prokaryotes, knowing how these cells interact with their environment is crucial.

In the final case study of this thesis, the atomic structure of the *Methanospirillum hungatei* flagellum was solved by cryoEM using helical reconstruction techniques. This is the first complete atomic structure of an archaeal flagellin, and solving it by cryoEM allowed us to elucidate the inter-subunit interactions within the flagellar filament. By using native protein for this study, we identified previously unknown post-translational modifications to the *M. hungatei* flagellin. Comparison between the archaeal flagellum, bacterial flagellum, and bacterial type IV pili showed that the archaeal flagellum is a structurally distinct filament (Table 5.1). Early attempts at solving the structure of the archaeal flagellum by cryoEM using photographic film for data collection were unable to

reach atomic resolution as the filaments were too thin and flexible. However, the creation direct electron detectors in the past five years has caused cryoEM to take a massive step forward and enabled the technique to solve the *de novo* structure of several previously unsolvable protein structures, including the protein structure of the archaeal flagella reported here. This structure is representative of the coming of age cryoEM has experienced during my doctoral career, and shows how cryoEM will become a major driving force in structural biology and the study of microbial interactions.

Electron microscopy has many uses for studying microbial interactions as illustrated by the three cases described in this thesis. Scanning electron microscopy is useful for characterizing the direct interactions between symbiotic partners such as XH001 and TM7x as described in chapter two. Transmission electron microscopy can be used to describe details of a microbial cell envelope as described in chapter three. In chapter four, I show how advances in TEM imaging technology and helical reconstruction techniques make it possible to solve protein structures at atomic detail with cryoEM. Electron microscopy has been an important tool in the microbiologist's toolbox for over half a century, and recent advances in hardware and software are making it indispensable.

It is a very exciting time to be working in the electron microscopy field. In addition to direct electron detectors, several other new advances in electron microscopy hardware, such as phase plates and correlative light and electron microscopy, are in development or newly available for commercial purchase. Advances in software, such as the Baysian statistical approaches discussed above and the development of GPU

image processing, will further advance the field by increasing the quality of final structural refinements and by decreasing the amount of time and computer resources it takes to produce the final refinements. New types of sample preparation techniques such as SEM based FIB milling are making it possible to examine protein complexes inside of cells. The next ten years promise to bring many amazing advances which can be applied to a variety of biological systems including the study of microbial interactions.

Property/feature	Archaeal flagellum	Bacterial flagellum	Bacterial pilus
Species used	*M. hungatei*	*S. enteracia*	*N. gonorrhoeae*
Gene	FlaB3	FlaF	PilE
Method of Structure Determination	CryoEM	CryoEM	CryoEM with docked crystal structure
Monomer size*	17.5 kDa	53.2 kDa	17.2 kDa
# of amino acids	164	494	158
Resolution	3.4 Å	4 Å	12.5 Å cryoEM map, 2.3 Å X-ray structure
Axial Rise	5.2 Å	4.7 Å	10.5 Å
Subunits per turn	3.3	5.5	3.6
Filament diameter	10 nm	24 nm	6 nm
Pore size	none	20 Å	6-10 Å
Protein domains	NTD, CTD	NDO, ND1a, ND1b, CDO, CD1, D1, D2, D3	NTD, CTD
% beta sheet	30.5	21.3	24.7
% alpha helix	28.6	52.0	32.3
alpha helix arrangement	single mild bend	coiled-coil	double kink
alpha helix angle to filament axis	20° offset	0°, parallel	variable offset
# of subunit helixes in an axial slice	12	20	7
subunits needed to span 26 nm	6	20	15
Post Translational Modifications	O-linked glycans N-linked glycans unknown	N-linked glycans	O-linked glycans

*molecular weight of protein without the N-terminal signal peptide and post-translational modifications

Table 5.1 | Comparison between archaeal flagellum, bacterial flagellum, and bacterial pilus structures

APPENDIX 1: Complete genome sequence of *Methanospirillum hungatei* type strain JF1

SHORT GENOME REPORT Open Access

Complete genome sequence of *Methanospirillum hungatei* type strain JF1

Robert P. Gunsalus[1,2*], Lauren E. Cook[1], Bryan Crable[3], Lars Rohlin[1], Erin McDonald[1], Housna Mouttaki[1,3], Jessica R. Sieber[3], Nicole Poweleit[1], Hong Zhou[1], Alla L. Lapidus[4,5], Hajnalka Erzsebet Daligault[4], Miriam Land[6], Paul Gilna[6], Natalia Ivanova[7], Nikos Kyrpides[7,8], David E. Culley[9] and Michael J. McInerney[3]

Abstract

Methanospirillum hungatei strain JF1 (DSM 864) is a methane-producing archaeon and is the type species of the genus *Methanospirillum*, which belongs to the family *Methanospirillaceae* within the order *Methanomicrobiales*. Its genome was selected for sequencing due to its ability to utilize hydrogen and carbon dioxide and/or formate as a sole source of energy. Ecologically, *M. hungatei* functions as the hydrogen- and/or formate-using partner with many species of syntrophic bacteria. Its morphology is distinct from other methanogens with the ability to form long chains of cells (up to 100 μm in length), which are enclosed within a sheath-like structure, and terminal cells with polar flagella. The genome of *M. hungatei* strain JF1 is the first completely sequenced genome of the family *Methanospirillaceae*, and it has a circular genome of 3,544,738 bp containing 3,239 protein coding and 68 RNA genes. The large genome of *M. hungatei* JF1 suggests the presence of unrecognized biochemical/physiological properties that likely extend to the other *Methanospirillaceae* and include the ability to form the unusual sheath-like structure and to successfully interact with syntrophic bacteria.

Keywords: *Methanomicrobiales*, Anaerobic, Motile, Methangenic archaea, Hydrogen, Formate, Syntrophic partnerships

Introduction

Strain JF1 (DSM 864 = ATCC 2790D-5) [1] is the type species for *M. hungatei* and represents the first isolated member of the *Methanospirillaceae* within the order *Methanomicrobiales* [2]. The species epithet derives from the Latin and honors Dr. R. E. Hungate, the inventor of methodologies for modern isolation and cultivation of strictly anaerobic bacteria and archaea [3, 4]. *M. hungatei* strain JF1 was isolated from a secondary anaerobic sewage treatment digestor in Urbana, Illinois, as part of a study of anaerobic aromatic hydrocarbon metabolism [5].

Here, we describe the genome sequence of *M. hungatei* strain JF1, a hydrogen- and formate-utilizing, methane-producing archaean. The genomic data provide insight towards defining the unique genes needed for anaerobic syntrophy [6], which occurs within a phylogenetically diverse range of bacteria, and for classifying genes identified by environmental DNA sequencing projects.

* Correspondence: robg@microbio.ucla.edu
[1]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA 90095-1489, USA
[2]UCLA DOE Institute for Genomics and Proteomics, University of California, Los Angeles, CA 90095-1489, USA
Full list of author information is available at the end of the article

Organism information

Morphology and physiology

Cells of *Methanospirillum hungatei* strain JF1 are narrow, curved rods (i.e., spirillum shaped) that measure ~0.5 μm by ~7 μm in size (Fig. 1, Table 1). The cells are contained within a sheath-like structure that contain one or more cells; the sheath may extend to over 100 μm in length depending on the nutritional conditions [1, 7]. Individual cells stain Gram-negative and are weakly motile by polar tufts of flagella. Cells also possess polyphosphate bodies or granules located at opposing cell ends [8]. Growth and metabolism is strictly anaerobic where hydrogen plus carbon dioxide and/or formate serve as the methanogenic substrate. Acetate is required as the major supply for cell carbon [1, 7]. Cells have no other organic nutritional requirements although addition of Casamino Acids or other plant/animal hydrolysis

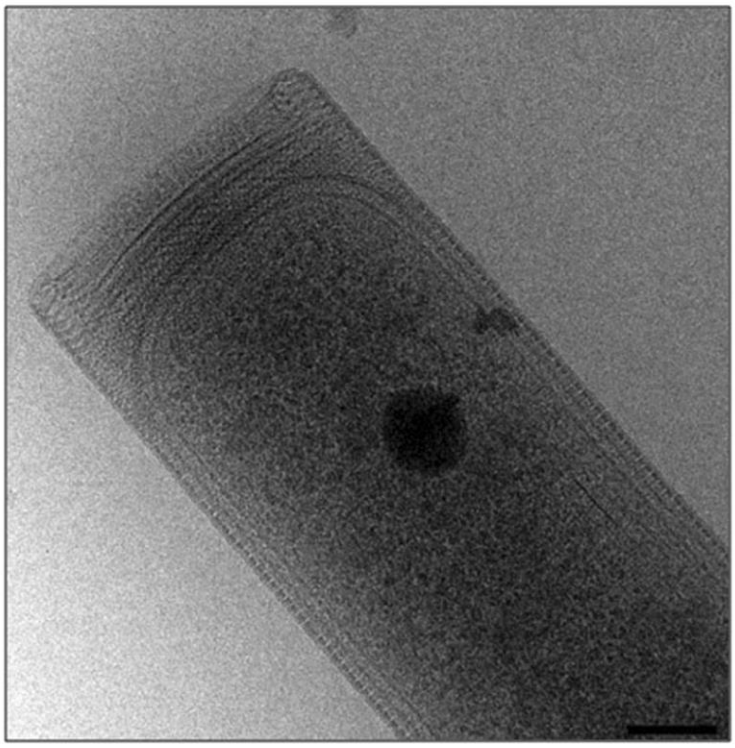


Fig. 1 Electron micrograph of *M. hungatei* strain JF1 cells and associated sheath structure. Scale bar corresponds to 100 nm

products speeds growth [1]. Temperature range for growth is 20–40 °C (optimum at 37 °C).

Biogenic methane production is important in the global carbon cycle and is used to treat sewage and other organic wastes and to produce biofuel from biomass [9, 10]. The degradation of fatty and aromatic acids is often the rate-limiting step in methanogenesis [6]. Fatty and aromatic acid degradation is thermodynamically favorable only when hydrogenotrophic methanogens such as *M. hungatei* strain JF1 maintain very low levels of hydrogen and/or formate in a process called syntrophy [10, 11]. Members of the genus *Methanospirillum* are often detected in ecosystems where syntrophy is essential [1, 12] and *M. hungatei* strain JF1 is the model partner in syntrophic cocultures of the propionate degrader *Syntrophobacter wolinii* [13], the butyrate degrader *Syntrophomonas wolfei* [14], and the benzoate degraders *Syntrophus buswellii* and *Syntrophus aciditrophicus* [15, 16].

Classification and features

The phylogenetic neighborhood of *M. hungatei* strain JF1 is shown in Fig. 2 for representative archaeal 16S rRNA sequences belonging to the order *Methanomicrobiales*. The four described *Methanospirillum* species form a well-defined cluster distinct from the other genera within the order where *Methanospirillum lacunae* and *Methanospirillum psychrodurum* form one subgroup and *M. hungatei* plus *Methanospirillum stamsii* form another. All strains of the genus *Methanospirillum* synthesize methane from hydrogen and carbon dioxide, though the ability to use formate is variable. None are able to ferment or respire by using other electron acceptors (i.e., with sulfate, nitrate, or iron). Certain species of other genera within the *Methanomicrobiales* also use formate, and some are reported to also metabolize short chain alcohols.

The analysis of the four 16S rRNA genes present in the *M. hungatei* JF1 genome reveled nearly identical nucleotide sequences but they differ from one another at two positions (nucleotide positions 937 and 1382) across the 1466 nucleotide length. The previously- published 16S rRNA gene sequences (AY196683 and AB517987) used in phylogenetic investigations were incomplete, i.e., 1271 and 1259 nucleotides, respectively [17, 18].

Chemotaxonomic data

The cell envelope of this Gram-negative cell wall type includes a surface layer coat, also known as a surface layer protein, which surrounds the cytoplasmic membrane, and an outermost sheath structure that encapsulates multiple cells, which are arranged in chains up to 0.1 mm in length [1, 8, 19]. Cytoplasmic membrane

Table 1 Classification and features of *Methanospirillum hungatei* strain JF1 according to MIGS recommendations [45] published by the genomic standards consortium [46] and the names for life database [47]

MIGS ID	Property	Term	Evidence code[a]
	Current classification	Domain *Archaea*	TAS [48]
		Phylum *Euryarchaeota*	TAS [49]
		Class *Methanomicrobia*	TAS [50]
		Order *Methanomicrobiales*	TAS [51]
		Family *Methanospirillaceae*	TAS [2]
		Genus *Methanospirillum*	TAS [1]
		Species *Methanospirillum hungatei*	TAS [1]
		Type strain JF-1	TAS [1]
	Gram stain	Negative	TAS [1]
	Cell shape	Curved rods 0.5 μM x 7.4 μM	TAS [1]
	Motility	Motile	TAS [1]
	Sporulation	Non-sporulating	TAS [1]
	Temperature range	30 °C-40 °C	TAS [1]
	Optimum temperature	37 °C	TAS [1]
	pH range; Optimum	6.5–10; 6.6–7.4	TAS [2]
	Carbon source	Carbon dioxide, formate, acetate	TAS [1]
	Energy source	Hydrogen, formate	TAS [1]
	Terminal electron receptor	Carbon dioxide	TAS [1]
MIGS-6	Habitat	Anaerobic sediments, sewage digesters	TAS [1]
MIGS-6.3	Salinity	Fresh to brackish water	TAS [1]
MIGS-22	Oxygen requirement	Strict anaerobe	TAS [1]
MIGS-15	Biotic relationship	Syntrophic	TAS [1]
MIGS-14	Pathogenicity	Non-pathogen	TAS [1]
MIGS-4	Geographic location	USA, Urbana, IL	TAS [1]
MIGS-5	Sample collection time	1972	TAS [1]
MIGS-4.1	Latitude	40.109°N	NAS
MIGS-4.2	Longitude	88.204°W	NAS
MIGS-4.4	Altitude	222 m	TAS [1]

These evidence codes are from the Gene Ontology project [52]
IDA Inferred from Direct Assay, *TAS* Traceable Author Statement (i.e., a direct report exists in the literature); *NAS* Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence)
[a]Evidence codes

lipids are composed primarily of biphytanyldiglycerol tetraether glycolipids [20]. *M. hungatei* strain JF1 lacks *b*-or *c*-type hemes, quinones, and methanophenazine (this study). The DNA G + C content was previously reported with 45 mol % [1].

Genome sequencing information

Genome project history

The *M. hungatei* strain JF1 genome was selected by DOE in 2004 as JGI sequencing project 364479 based on its phylogenetic position, its role in anaerobic decomposition of organic matter, and its ability to grow in co-culture with many syntrophic bacterial species [6]. The genome project is deposited in the Genomes OnLine Database [21] as project Id:Gc00350, and the complete genome sequence is deposited in GenBank. Sequencing, finishing, and annotation of the *M. hungatei* genome were performed by the DOE Joint Genome Institute [22]. A summary of the project information is shown in Table 2.

Growth conditions and genomic DNA preparation

M. hungatei strain JF1 was grown in basal medium under anaerobic conditions at 37 ° C as previously described [1]. High molecular weight genomic DNA was isolated from cell pellets (DSM 864 = ATCC 2790D-5) using the CTAB method described at the JGI's web site [22].

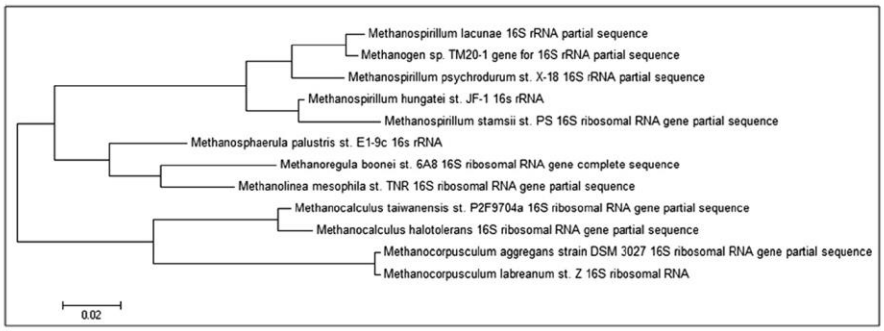


Fig. 2 Phylogenetic tree highlighting the position of *Methanosprillulm hungatei* strain JF1 relative to other type strains within the *Methanomicrobiales*. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model [43]. The tree with the highest log likelihood (−3033.8513) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 12 nucleotide sequences. Codon positions included were 1st + 2nd + 3rd + Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 789 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 [44]

Genome sequencing and assembly

The genome was sequenced at the Joint Genome Institute using a combination of 3 kb, 8 kb, and 40 kb DNA libraries. All general aspects of library construction and sequencing performed are described at the JGI's web site [22]. The Phred/Phrap/Consed software package [23] was used to assemble all three libraries and to assess quality [24, 25]. Possible miss-assemblies were corrected and gaps between contigs were closed by editing in Consed, custom primer walks, or PCR amplification (Roche Applied Science, Indianapolis, IN). The error rate of completed genome sequence of *M. hungatei* is less than 1 in 50,000. The sequence of *M. hungatei* can be accessed using the GenBank accession number CP000254.

Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Finished
MIGS-28	Libraries used	3, 8, 14 kb
MIGS 29	Sequencing platforms	Sanger
MIGS 31.2	Fold coverage	14.5X
MIGS 30	Assemblers	PGA
MIGS 32	Gene calling method	Prodical GenePRIMP
	Locus Tag	Mhun_0000
	Genbank ID	CP000254
	GenBank Date of Release	March 1, 2006
	GOLD ID	Gc00350
	BIOPROJECT	PRJNA13015
MIGS 13	Source Material Identifier	DSM 864 T
	Project relevance	Carbon cycle, energy production, bioreactors

Genome annotation

Genes were identified using Prodical [26] as part of the Oak Ridge National Laboratory genome annotations pipeline, followed by a round of manual curation using the JGI GenePRIMP pipeline [27, 28]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information nonredundant database, and the UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. Additional gene prediction analysis and functional annotation was preformed within the Integrated Microbial Genomes-Expert Review platform [29, 30]. Membrane transport protein analysis was done by IMG with additional analysis by TransportDB [31] TCDB [32] databases. Transcription factor analysis and prediction was by assisted by TBD database [33].

Genome properties

The genome statistics are provided in Table 3 and Fig. 3. The genome consists of one circular chromosome of 3,544,738 bp with 3,307 predicted genes of which 3,239 are protein-coding genes. Of these, approximately 61 % (2,018 genes) were assigned to a putative function while the remaining 37 % (1,221 genes) are without assigned functions. The genome is 45.15 G + C and 88.64 % coding. The distribution of genes into COGs functional categories is presented in Table 4. Of note, six CRISPER repeats were identified on the chromosome. The *M. hungatei* genome has 51 tRNA genes; 43 have identified functions, which cover all amino acids except His. The genes for histidine biosynthesis from pyruvate are

Table 3 Genome statistics

Attribute	Value	% of total
Genome size (bp)	3,544,738	100.00
DNA coding (bp)	3,142,074	88.94
DNA G + C (bp)	1,600,415	45.15
DNA scaffolds	1	100.00
Total genes	3,307	100.00
Protein coding genes	3,239	97.94
RNA genes	68	2.06
Pseudo genes	99	2.99
Genes in internal clusters	2172	65.68
Genes with function prediction	2,018	61.02
Genes assigned to COGs	1872	56.61
Genes with Pfam domains	2577	77.93
Genes with signal peptides	101	3.05
Genes with transmembrane helices	762	23.04
CRISPR repeats	6	

present with the exception that a gene for histidinol phosphate phosphatase (HisN) was not detected. Nutritional studies [1, 7] did not detect histidine auxotrophy, suggesting that *M. hungatei* has undescribed mechanisms for fulfilling the role of HisN and synthesizing His-tRNA.

Insights from the genome sequence

Methanogenesis pathway

The *M. hungatei* JF1 ORFs were organized into pathways where most pathways considered essential for viability of a typical archaeal cell were detected. The methanogenic pathway from hydrogen and carbon dioxide is highly conserved in methanogens and the genes for all the enzymes in the central methanogenic pathway were identified, including a soluble-type heterodisulfide reductase only (Fig. 4). The genome contains three gene sets for molybdenum (*fmd*) or tungsten (*fwd*) type formylmethanofuran (MFR) dehydrogenases (Mhun_1981-84, Mhun_1985-94 and Mhun_210612) that catalyze the ferredoxin-dependent first step of carbon dioxide reduction. There are three genes for methenyl–H_4MPT

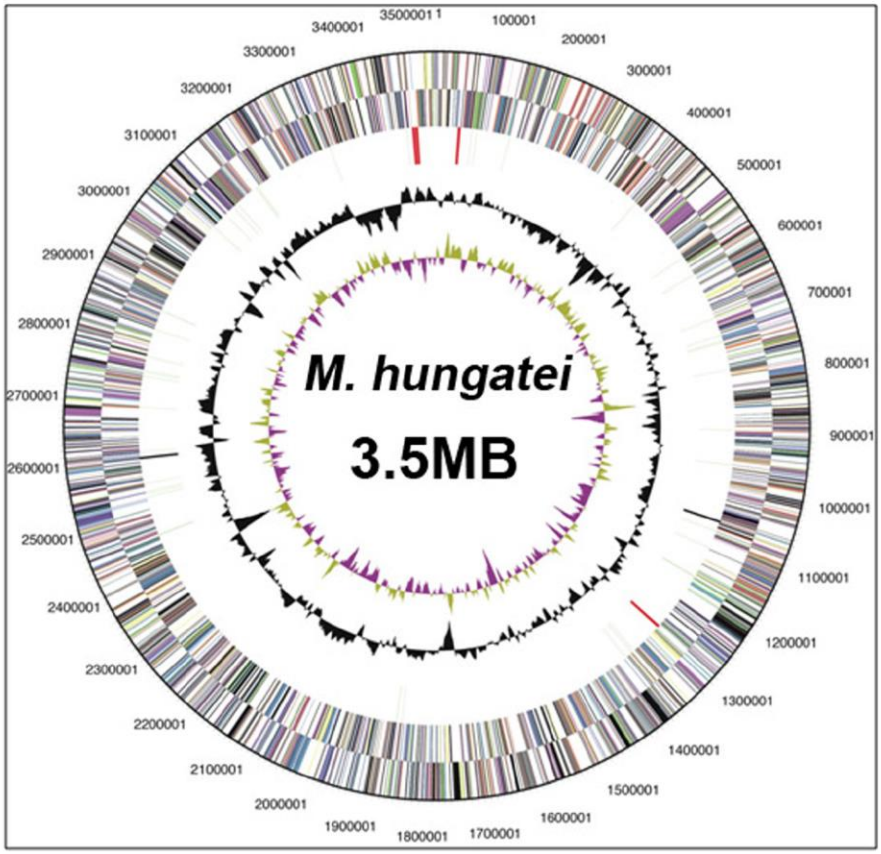


Fig. 3 Graphic circular map of the *M. hungatei* JF1 chromosome. The concentric circles from outside to inside indicate: genes on the forward strand, genes on the reverse strand, RNA genes (tRNA's green, .rRNA's red, other RNA's black), GC content, and GC skew

Table 4 Number of genes associated with general COG functional categories

Code	Value	% age	Description
J	180	8.65	Translation, ribosomal structure and biogenesis
A		0.00	RNA processing and modification
K	84	4.03	Transcription
L	82	3.94	Replication, recombination and repair
B	8	0.38	Chromatin structure and dynamics
D	16	0.77	Cell cycle control, Cell division, chromosome partitioning
V	53	2.55	Defense mechanisms
T	154	7.4	Signal transduction mechanisms
M	85	4.08	Cell wall/membrane biogenesis
N	54	2.59	Cell motility
U	17	0.82	Intracellular trafficking and secretion
O	89	4.27	Posttranslational modification, protein turnover, chaperones
C	186	8.93	Energy production and conversion
G	59	2.83	Carbohydrate transport and metabolism
E	165	7.93	Amino acid transport and metabolism
F	62	2.98	Nucleotide transport and metabolism
H	162	7.78	Coenzyme transport and metabolism
I	31	1.49	Lipid transport and metabolism
P	147	7.06	Inorganic ion transport and metabolism
Q	16	0.77	Secondary metabolites biosynthesis, transport and catabolism
R	217	10.42	General function prediction only
S	160	7.68	Function unknown
-	1435	43.39	Not in COGs

The total is based on the total number of protein coding genes in the genome

tetrahydromethanopterin (H_4MPT) cyclohydrolyase (Mch: Mhun_0022, Mhun_0444, Mhun _2384), which catalyze the third pathway step.

Single genes encode enzymes for the second, fourth, and fifth pathway steps, formylMFR:tetrahydromethanopterin formyl transferase (Ftr: Mhun_1808), methylene–H_4MPT dehydrogenase (Mtd: Mhun_2255) and methylene –H_4MPT reductase (Mer: Mhun_2257). The latter two enzymes employ reduced cofactor F_{420} as substrate. The remaining two enzymes in the pathway are multi-subunit complexes: H_4MPT Smethyltransferase (Mtr: Mhun_2168-75), and the type I methyl-CoM reductase (Mcr: Mhun_2144-2148). The CoM-S-S-CoB heterodisulfide reductase (Hdr: Mhun_1834-39) so named for the methanogenic co-enzymes M and B, reduces CoM-S-S-CoB hertodisulfide generated by Mcr. The reaction catalyzed by a soluble-type Hdr is likely an electron bifurcation, which couples the energetically favorable reduction of CoM-S-SCoB by formate and/or H_2 with the energetically unfavorable reduction of ferredoxin by formate and/or H_2 [34].

The oxidation of hydrogen or formate is needed to generate reduced ferredoxin and cofactor F_{420} used in several of the above reactions (Fig. 4). The oxidation of hydrogen or formate may be accomplished by one or more of the multiple hydrogenase and formate dehydrogenase enzymes. Five nearly identical gene clusters encode soluble formate dehydrogenase (Fdh) enzymes: Mhun_1813-1814, Mhun_1832-1833, Mhun_2020-2021, Mhun_2022-2023, and Mhun_3237-3238. There are two formate/nitrite-type transporters (Mhun_0075, Mhun_1811). The five hydrogenase gene clusters include *echABDDEF* (Mhun_1741-1747), *ehrABCDLS* (Mhun_1817-1822), *ehaABCDEFGHIJK* (Mhun_2094-2106), *frhADGB* (Mhun_2329-2332), and *mbhABCDEFGHIJKLMN* (Mhun_2579-2592). The *ech, eha, ehr*, and *mbh* gene clusters encode membrane-associated enzymes that likely reduce Fd. These are believed to employ ion gradients (Na^+ or H^+) to assist Fd reduction at low hydrogen levels. The remaining hydrogenase gene cluster (*frh*ADGB) encodes a soluble hydrogenase that reduces F_{420}.

Transporters, ion movement, and ATP synthesis

M. hungatei JF1 has 352 genes involved in membrane transport as determined by IMG/ER, which constitute 10.64 % of the genome. These include 34 multi-component ATP-binding cassette or ABC-type transporter genes plus related but unlinked genes (152 genes in total), sixty genes encoding secondary transporters, twelve genes for ion channels, seven genes for P-ATPases, one H^+ translocating pyrophosphatase (Mvp, H + PPase; Mhun_2414) gene, and four type II secretion systems. A highly unusual feature of the *M. hungatei* genome is the presence of three H^+ or Na^+ -translocating AoA_1-type ATP synthetase gene clusters encoded by 27 genes (Aha1, Mhun_1177-1185; Aha2, Mhun_1757-1765, and Aha3, Mhun_1768-1775). The gene order is conserved relative to the corresponding Aha complex in *Methanosarcina acetivorans* [35]. Although it is unknown whether these systems utilize protons or sodium ions, the *M. acetivorans* ortholog is believed to use sodium ions [35]. Likewise, the membrane-bound H_4MPT Smethyltransferase (Mtr) is predicted to be sodium dependent. Three genes encode $Na^{+/}H^+$ antiporters (Mhun_0680, Mhun_0841, Mhun_2803) that might maintain ion balance where the last differs by also possessing a Trk domain.

Cell biosynthesis

The genome of *M. hungatei* encodes an acetyl-CoA synthase/CO dehydrogenase complex (Cdh; Mhun_0686-0690). The role of Cdh is undefined at this time because *M. hungatei* must acquire acetate supplied in the medium for growth rather than synthesizing acetylCoA

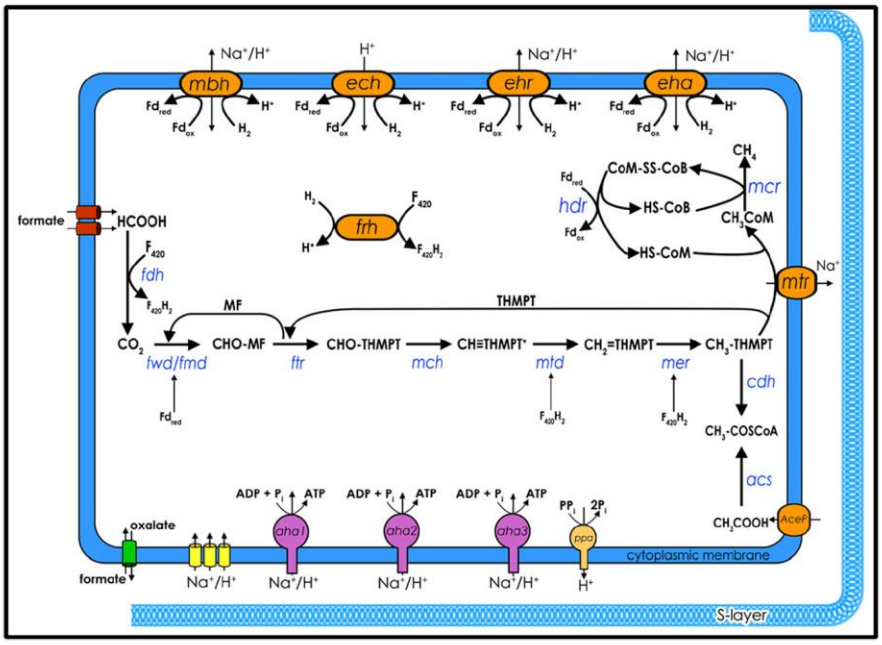


Fig. 4 Overview of central metabolism in *M. hungatei* strain JF1. The pathway for methane formation from hydrogen and formate is shown in black with key steps shown with gene/enzyme designations. Membrane proteins involved in energy transduction electron transport, and ion/solute translocation are arranged along the cytoplasmic membrane: archaeal ATP synthase, Aha; formate dehydrogenases, Fdh; hydrogenases (Mbh, Ech, Ehr, Eha Frh); formyl-methanofuran dehydrogenase, Fmd, Fwd; methenyl–H_4MPT tetrahydromethanopterin cyclohydrolyase, Mch; formylMFR:tetrahydromethanopterin formyl transferase, Ftr; methylene–H_4MPT dehydrogenase, Mtd; methylene–H_4MPT reductase, Mer; H_4MPT S-methyltransferase (Mtr; methyl-CoM reductasem Mcr; and CoM-S-S-CoB heterodisulfide reductase, Hdr

from CO_2, which is the usual role of Cdh in hydrogenotrophic methanogens. Uptake of acetate for incorporation into cell material is predicted to occur by the Mhun_0634 *aceP* gene product [35]. Five acetyl-CoA synthetase genes are present that could activate acetate to acetyl-CoA. Mhun_0352, Mhun_0567, and Mhun_1721 share > 62 % identity at the amino acid level with each other, but only share < 34.2 % amino acid identity with Mhun_0592 and Mhun_2392.

M. hungatei has two set of genes that could be used to carboxylate acetyl-CoA to pyruvate (Mhun_2393-2396 and Mhun_0450-0453). Oxaloacetate can be synthesized by carboxylation of pyruvate using pyruvate carboxylase (Mhun_3189-3190) or by conversion of pyruvate to phosphoenol pyruvate by pyruvate dikinase (Mhun_2610 or Mhun_1141) and carboxylation of phosphoenol pyruvate to oxaloacetate by phosphoenol pyruvate carboxylase (Mhun_0174). The genes necessary to convert oxaloacetate to 2-oxoglutarate by the reductive arm of the tricarboxylic acid cycle were detected (malate dehydrogenase, Mhun_1155; fumarate hydratase, Mhun_0089-0090; succinyl-CoA ligase, Mhun_0096-0095; 2-oxoglutarate synthase, Mhun_0091-0094 or Mhun_29922994; and fumarate reductase, Mhun_3052-3053). Complete biosynthetic pathways for the synthesis of all amino acids except histidine from pyruvate, oxaloacetate, and 2-oxoglutarate as the main starting materials were detected.

There are few genomic clues regarding the composition of the *M. hungatei* cell envelope. The genome contains a large number of PDK domain-containing genes (31 genes) as well as TRP domain-containing genes (41 genes). Many of these have transmembrane and/or SP signal elements that would suggest cell envelope associations but it is unknown if any of the proteins are significantly expressed. There are no clear protein candidates for the morphologically defined cell envelope structures containing a surface layer, sheath, and plugs [1, 8].

Stress

There appear to be few cellular adaptations in *M. hungatei* for stress response. Among those found are defense against oxygen damage: catalase (Mhun_2433), peroxidase (Mhun_2733), manganese/iron superoxide dismutase (Mhun_2974), heavy metal resistance (Mhun_1348, Mhun_3034), drug resistance (Mun_0598, Mhun_1195) and heat shock (Mhun_2436).

Regulation and signal transduction

The *M. hungatei* genome contains a typical set of archaeal RNA polymerase genes and one BRE recognition

factor analogous to eukaryotic transcription initiating factor B (Mhun_2481; Tfb) plus two TATA-box binding proteins or TBP's that confer promoter recruitment and specificity (Tbp1, Mhun_0568 and Tbp2, Mhun_0593). There are ~65 DNA-binding transcription factors identified that modulate gene expression. These belong to a variety of protein families common to bacteria but include few regulatory proteins typical of eukaryotes (e.g., homeodomain-like, zinc finger, SRF-like, or p53-like proteins). There are numerous bacterial-type two-component regulatory systems including 82 histidine kinase-type sensor transmitters, 41 response regulatory proteins, and 18 receiver-only domain proteins. Of the 82 histidine kinases, 55 are soluble and 27 are membrane-associated. They are generally unlinked genetically and thus do not suggest an interacting partner in sensory transduction.

Motility and taxis

M. hungatei JF1 possesses multiple archaeal-type flagella filaments at the cell ends [1, 8], now termed archaealla that resemble bacterial type IV pili [36, 37]. The genome contains one *flhGFHIJ* gene cluster (Mhun_0102-0105) encoding a basal body structure. Three FlaB-type pili genes make up the archaella filaments (Mhun_1238, Mhun_3139,

Mhun_3140). Although little is known about the chemotactic abilities of *M. hungatei*, other than movement towards an essential nutrient, acetate [38], there are multiple chemosensory genes present in the genome. These include 3 CheA, 4 CheB, 4 CheC, 1 CheD, 3 CheR, 1 CheY, and 14 CheW, genes plus 27 genes encoding MCP sensory proteins (methyl accepting chemotaxis proteins) that detect unknown attractants and/or repellants. Twelve MCPs are membrane-associated and 15 MCPs are soluble.

Multiple genes (~11 paralogs) are also present in the *M. hungatei* JF1 genome for archaeal-type pili like those seen in *Methanococcus maripaludis, Haloferax volcanii*, and *Sulfolobus acidocaldarius* [39]. These archaeal proteins, distinct from the bacterial pili-type proteins, were previously annotated as hypothetical genes (e.g., Mhun_0297). The *H. volcanii* pili proteins provide adhesion to surfaces and the orthologs in *M. hungatei* JF1 may function in cell-cell adhesion or in cell-cell communication, although such appendages have not been previously observed in EM micrographs. All but one of the eleven *M. hungatei* JF1 paralogs are in clusters of 2 to 3 genes each and often with ABCtype transport genes.

Comparison to other archaeal genomes

The 3.54 MB *M. hungatei* JF1 genome is the largest within the order *Methanomicrobiales* that have been sequenced thus far including *Methanosphaerula palustris* (2.92 MB) and *Methanocorpusculum labreanum* (1.80 MB). The *M. hungatei* JF1 genome is also among the largest within the *Archaea* domain: only three species sequenced thus far, belonging to the genus *Methanosarcina* (i.e., *Methanosarcina acetivorans*, 5.75 MB; *Methanosarcina barkeri*, 4.87 MB; and *Methanosarcina mazei*, 3.83 MB), plus one halophile, *Haloarcula marismortui* (4.27 MB), exceed it in size. The large genome of *M. hungatei* JF1 suggests the presence of unrecognized biochemical/physiological properties that likely extend to the other *Methanospirillaceae* and include the ability to form the unusual sheath-like structure and to successfully interact with syntrophic bacteria.

When *M. hungatei* ORFs were compared pair-wise to individual microbial genomes [40, 41], best reciprocal BLAST hits revealed closest associations to the taxonomically related archaea: *Methanoculleus marisnigri* (1395 reciprocal gene hits), *Methanosarcina acetivorans* (1203), and *Methanosarcina barkeri* (1150), and extending to *Haloquadratum walsbyi* (657) (Additional file 1: Figure S1). Thus, approximately 650 to 1,200 genes are similar and well-conserved across these 17 archaeal species whereby the remaining genes (ca. 1700 genes) represent a novel complement within the *M. hungatei* genome. Interestingly, seven of the next thirteen closest matches are bacterial species among which are many syntrophic microorganisms that likely grow in close association with *M. hungatei*. Strikingly, *Syntrophobacter fumaroxidans* strain MPOB exhibited 634 best reciprocal BLAST hits.

In another comparison, the best BLAST hit to any microbial gene product was determined (Additional file 2: Figure S2) and showed 1; 167; 277; and 142 ORFs closest hits in the genomes of *Methanoculleus marisnigri, Methanocorpusculum labreanum, and Methanosarcina barkeri*, respectively. Notably three bacterial genomes, *Syntrophus aciditrophicus, Syntrophobacter fumaroxidans*, and *Nostoc spp.* gave 21–19 best BLAST hits each, suggesting the possibility of lateral gene transfer events from these potential syntrophic partners. The occurrence of *Nostoc*-related genome sequences raises interesting questions concerning microbial interactions and lateral gene transfer with methanogens present in complex microbial communities [42].

Extended insights

The large genome of *M. hungatei* JF1 suggests the presence of unrecognized biochemical/physiological properties that likely extend to the other *Methanospirillaceae* and include the ability to form the unusual sheath-like structure and the ability to successfully interact with syntrophic bacteria. A number of genes may have been acquired by lateral gene transfer from its syntrophic partners or other microorganisms present in complex microbial communities. Also of particular note are multiple genes for archaeal type IV pili that may function in

cell-cell adhesion or cell-cell communication and genes for multiple hydrogenases and formate dehydrogenases to metabolize hydrogen and formate generated by its syntrophic partners. The core machinery of *M. hungatei* to produce methane from hydrogen and carbon dioxide and/or formate is typical of other hydrogenotrophic methanogens, except that *M. hungatei* has genes for three H^+ or Na^+-translocationg A_OA_1-type ATP synthases. *M. hungatei* has four 16S ribosomal RNA genes that each differ at two positions. Further understanding of the novel compliment of *M. hungatei* genes will likely provide a more thorough understanding of the multispecies interactions involved in syntrophy and the synthesis of complex structures such as the *M. hungatei* sheath, which is shared by multiple cells.

Conclusions

We report here an inventory of the genomic features of the methane-producing anaerobic archaeon, *Methanospirillum hungatei* strain JF1 (DSM 864), and describe its phylogenetic relationship to its neighbors. We further identify from the sizable genome of *M. hungatei* examples of genes involved in anaerobic syntrophy, and as the type strain of the *Methanospirillum*, suggest potential universal qualities of this genus. We hope this report aids and stimulates further study of this fascinating organism.

Additional files

Additional file 1: Figure S1. Best reciprocal protein hits for *M. hungatei* JF1 ORFs with other genomes. (DOCX 89 kb)

Additional file 2: Figure S2. Best Blast hit distribution of *M. hungatei* JF1 ORFs with other genomes. (DOCX 62 kb)

Competing interests

The authors declare they have no competing interests.

Authors' contributions

RPG and MJM contributed to the conception and design of this project. ALL, HED, ML, OG, NI, and NK were involved in the acquisition and initial analysis of the data; LEC, BC, LR, EM, HM, JRS, NP, HZ, RPG, and MJM were involved in the interpretation of the data. RPG prepared the first draft of the manuscript. All authors were involved in its critical revision and have given final approval of the version to be published and agree to be accountable for all aspects of the work.

Acknowledgements

This work was performed under the auspices of the US Department of Energy Office of Science Biological and Environmental Research Program. We also acknowledge support by the US Department of Energy Contract No. DE-FG02-96ER20214, from the Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences (MJM) for syntrophy functions, Department of Energy Grant DE-FG03-86ER13498 (RPG) the UCLA-DOE Institute of Genomics and Proteomics Grant DE-FC02-02ER63421 (RPG), and the National Science Foundation grant NSF 1244566 (MJM and RPG) for methanogen biochemistry.

Author details

[1]Department of Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, CA 90095-1489, USA. [2]UCLA DOE Institute for Genomics and Proteomics, University of California, Los Angeles, CA 90095-1489, USA. [3]Department of Botany and Microbiology, University of Oklahoma, Norman, OK 73019, USA. [4]Theodosius Dobzhansky Center for Genome Bionformatics, St. Petersburg State University, St. Petersburg, Russia. [5]Algorithmic Biology Lab, St. Petersburg Academic University, St. Petersburg, Russia. [6]Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA. [7]DOE Joint Genome Institute, Walnut Creek, CA, USA. [8]Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia. [9]Pacific Northwest National Laboratory, Richland, WA, USA.

Received: 2 October 2014 Accepted: 22 December 2015
Published online: 06 January 2016

References

1. Ferry JG, Smith PH, Wolfe RS. *Methanospirillum*, a new genus of methanogenic bacteria. Int J of Syst Bacteriol. 1974;24:465–9.
2. Boone DR, Whitman WB, Koga Y. Family III. *Methanospirillaceae* fam. nov. In: Boone DR, Castenholz RW, editors. Bergey's Manual of Systematic Bacteriology, vol. 1. 2nd ed. New York: Springer; 2001. p. 264.
3. Hungate RE. The anaerobic mesophilic cellulolytic bacteria. Bacteriol Rev. 1950;14:1–49.
4. Wolfe RS. Techniques for cultivating methanogens. Methods in Ezymology. 2011;494:1–22.
5. Ferry JG, Wolfe RS. Anaerobic degradation of benzoate to methane by a microbial consortium. Arch Microbiol. 1976;107:33–40.
6. McInerney MJ, Struchtemeyer CG, Sieber J, Mouttaki H, Stams AJM, Schink B, et al. Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism. In, Incredible Anaerobes: From Physiology to Genomics to Fuels. Edited by Wiegel J, Maier R, and Adams M. Anal NY Acad Sci. 2008;1125:58–72.
7. Ferry JG, Wolfe RS. Nutritional and biochemical characterization of *Methanospirillum hungatii*. Appl Environ Microbiol. 1977;34:371–6.
8. Toso DB, Henstra A-M, Gunsalus RP, Zhou ZH. Structural, mass, and elemental analyses of storage granules in methanogenic archaeal cells. Environ Microbiol. 2011;9:2587–99.
9. McInerney MJ, Sieber JR, Gunsalus RP. Syntrophy in anaerobic global carbon cycles. Curr Opin Biotechnol. 2009;20:623–32.
10. Schink B. Energetics of syntrophic cooperation in methanogenic degradation. Microbiol Mol Biol Rev. 1997;61:262–80.
11. Sieber JR, McInerney MJ, Gunsalus RP. Genomic insights into syntrophy: the paradigm for anaerobic metabolic cooperation. Annu Rev Microbiol. 2012;66:429452.
12. Qiu Y-L, Sekiguchi Y, Imachi H, Kamagata Y, Tseng I-C, Cheng S-S, et al. Identification and isolation of anaerobic, syntrophic phthalate isomer-degrading microbes from methanogenic sludges treating wastewater from terephthalate manufacturing. Appl Environ Microbiol. 2004;70:1617–26.
13. Boone D, Bryant M. Propionate-degrading bacterium, *Syntrophobacter wolinii* sp. nov., gen. nov., from methanogenic ecosystems. Appl Environ Microbiol. 1980;40:626–32.
14. McInerney MJ, Bryant MP, Hespell RB, Costerton JW. *Syntrophomonas wolfei* gen. nov. sp. nov., an anaerobic, syntrophic, fatty acid-oxidizing bacterium. Appl Environ Microbiol. 1981;41:1029–39.
15. Jackson BE, Bhupathiraju VK, Tanner RS, Woese CR, McInerney MJ. *Syntrophus aciditrophicus* sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen-using microorganisms. Arch Microbiol. 1999;171:107–14.
16. Mountfort D, Brulla W, Krumholz L, Bryant M. *Syntrophus buswellii* gen. nov., sp. nov.: a benzoate catabolizer from methanogenic ecosystems. Int J Syst Bacteriol. 1984;34:216–7.
17. Iino T, Mori K, Suzuki K. *Methanospirillum lacunae* sp. nov., a methaneproducing archaeon isolated from a puddly soil, and emended descriptions of the genus *Methanospirillum* and *Methanospirillum hungatei*. Int J Syst Evol Microbiol. 2010;60:2563–6.
18. Wright AD, Pimm C. Improved strategy for resumptive identification of methanogens using 16S riboprinting. J Microbiol Meth. 2003;55:337–49.
19. Zeikus JG, Bowen VG. Fine structure of *Methanosprillum hungatii*. J Bacteriol. 1975;121:373–80.
20. Kushwaha SC, Kates M, Sprott GD, Smith IC. Novel polar lipids from the methanogen *Methanospirillum hungatei* GP1. Biochim Biophys Acta. 1981;664:156–73.
21. Pagani I, Liolios K, Jansson J, Chen IMA, Smirnova T, Nosrat B, et al. The Genomes OnLine Database (GOLD) v.4: Status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res. 2012;40:D571–9.
22. DOE Joint Genome Institute [http://www.jgi.doe.gov]. Accessed 4 January 2016.

23. Phred/Phrap/Consed software package [http://www.phrap.com]. Accessed 4 January 2016.
24. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 1998;8:186–94.
25. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res. 1998;8:195–202.
26. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.
27. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, et al. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods. 2010;7:455–7.
28. GenePRIMP [http://geneprimp.jgi-psf.org]. Accessed 4 January 2016.
29. Markowitz VM, Ivanova NN, Chen IMA, Chu K, Kyrpides NC. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics. 2009;25:2271–8.
30. IMG-ER [http://img.jgi.doe.gov/er]. Accessed 4 January 2016.
31. TransportDB [http://www.membranetransport.org/]. Accessed 4 January 2016.
32. TCDB [http://www.tcdb.org/]. Accessed 4 January 2016.
33. TBD [http://www.transcriptionfactor.org]. Accessed 4 January 2016.
34. Welte C, Deppenmeier U. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. Biochim Biophys Acta. 1837;2014:1130–47.
35. Rohlin L, Gunsalus RP. Carbon-dependent control of electron transfer and central carbon pathway genes for methane biosynthesis in the Archaean, *Methanosarcina acetivorans* strain C2A. BMC Microbiol. 2010;10:62.
36. Jarrell KF, Albers SV. The archaellum: an old motility structure with a new name. Trends Microbiol. 2012;20:307–12.
37. Thomas NA, Bardy SL, Jarrell KF. The archaeal flagellum: a different kind of prokaryotic motility structure. FEMS Microbiol Rev. 2001;25:147–74.
38. Migas J, Anderson KL, Cruden DL, Markovetz AJ. Chemotaxis in *Methanospirillum hungatei*. Appl Environ Microbiol. 1989;55:264–5.
39. Esquivel RN, Xu R, Pohlschroder M. Novel archaeal adhesion pilins with a conserved N terminus. J Bacteriol. 2013;195:3808–18.
40. Plugge CM, Henstra AM, Worm P, Swarts DC, Paulitsch-Fuchs AH, Scholten JC, et al. Complete genome sequence of *Syntrophobacter fumaroxidans* strain (MPOB^T). Stand Genomic Sci. 2012;7:91–106.
41. Sieber JR, Sims DR, Han C, Kim E, Lykidis A, Lapidus AL, et al. The genome of *Syntrophomonas wolfei*: new insights into syntrophic metabolism and biohydrogen production. Environ Microbiol. 2010;12:2289–301.
42. Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, Martinez-Arias R, et al. The genome of *Methanosarcina mazei:* evidence for lateral gene transfer between Bacteria and *Archaea*. J Mol Microbiol Biotech. 2002;4:453–61.
43. Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molec Biol Evol. 1993;10:512–26.
44. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Molec Biol Evol. 2013;30:2725–9.
45. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. Towards a richer description of our complete collection of genomes and metagenomes "Minimum Information about a Genome Sequence" (MIGS) specification. Nat Biotechnol. 2008;26:541–7.
46. Field D, Amaral-Zettler L, Cochrane G, Cole J, Dawyndt P, Garrity GM, et al. The genomic standards consortium. PLoS Biol. 2011;9:e1001088.
47. Garrity GM. Names for Life Browser Tool takes expertise out of the database and puts it right in the browser. Microbiol Today 2010, 37:9.
48. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, Bacteria, and Eucarya. Proc Natl Acad Sci USA. 1990;87:4576–9.
49. Garrity GM, Holt JG. Phylum AII. *Euryarchaeota* phy. nov. In: Garrity GM, Boone DR, Castenholz RW, editors. Bergey's Manual of Systematic Bacteriology, vol. 1. 2nd ed. New York: Springer; 2001. p. 211–355.
50. Garrity GM, Bell JA, Lilburn T. The revised road map to the manual. In: Brenner DJ, Kreig NR, Staley JT, editors. Bergey's Manual of Systematic Bacteriology, vol. 2. 2nd ed. 2005. p. 159–220.
51. Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS. Methanogens: reevaluation of a unique biological group. Microbiol Rev. 1979;43:260–96.
52. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet. 2000;25:25–9.