Title
Regulation of Biomineralization in Magnetotactic Bacteria

Author
Browne, Patrick Jeffrey

Publication Date
2017

Peer reviewed|Thesis/dissertation

Regulation of Biomineralization in Magnetotactic Bacteria

By

Patrick J Browne

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Microbiology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Arash Komeili, Chair
Professor Kathleen R. Ryan
Professor David Savage
Professor Britt Glaunsinger

Fall 2017

Abstract

Regulation of Biomineralization by an HtrA Protease in Magnetotactic Bacteria

by

Patrick J Browne

Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Arash Komeili, Chair

Magnetotactic bacteria are a unique group of bacteria characterized by their ability to manufacture magnetic crystals for the purpose of aligning with geo-magnetic fields. They offer a unique opportunity for merging the study of bacterial cellular biology, classic molecular genetics, biochemistry, and evolution, both on the small scale of individual proteins, and a larger system of proteins with a defined complex function. This dissertation will first look at the genetics and evolution of magnetotactic bacteria as a whole before focusing on MamE, an HtrA protease conserved throughout magnetotactic bacteria. HtrA proteases are ubiquitous throughout all domains of life and serves several different functions. This dissertation will show that MamE is a multifunctional protein and will also highlight other magnetosome related genes that either regulate MamE's activity or are targets of proteolysis.

The first chapter of this dissertation is a to-be-submitted review article and introduces the genetics of magnetotactic bacteria (MTB), a paraphyletic group of gram-negative bacteria that use around a hundred genes, located in a defined gene island, to create magnetosomes, membrane-bound compartments containing a magnetic iron crystal. By organizing magnetosomes into a chain, cells are able to align with geo-magnetic fields. It will discuss the most well-conserved genes whose general functions are known. It also describes the interesting evolutionary history of the magnetosome gene island (MAI) which despite being an island seemingly capable of horizontal gene transfer, does not contain other hallmarks of recent HGT and closely matches the phylogenetic tree.

The second chapter, a published primary research article (Hershey et al., PLOS Biology 2016, and) is focused on MamO, a protein that also appears to also be an HtrA protease similar to MamE. However, closer examination shows that is contains no protease activity and is instead a metal-binding protein most likely involved in crystal nucleation. It is also shown to not only be a target of MamE proteolysis but also regulate MamE's activity. Finally, it contains an analysis of MamOs throughout MTB and shows that they are not actually a single protein family. Instead, four separate lineages of MTB have all acquired inactive proteases in a convergent manner.

The third chapter, a published primary research article (Hershey et al., Journal of Biological Chemistry 2016) is the first to focus primarily on MamE. Unlike MamO, this chapter shows that MamE is a bonafide protease and identifies several targets of MamE, including MamE itself, MamO, and MamP. It also interrogates the kinetics of MamE's protease activity and shows that MamE processes its targets in a regulated manner.

The fourth chapter, to be submitted as a primary research article, also investigates MamE and its targets, this time beginning with a proteomic approach. Cells carrying an inactive form of MamE have deformed smaller crystals. Using a novel method of cell lysis and magnetosome separation, wild-type and mutant magnetosome are compared. This method uncovered a few potential additional targets of MamE proteolysis, one of which was confirmed *in vitro*. This protein, MamD was also further investigated and determined to possibly function as a biomineralization inhibitor. This chapter also discusses the proteome of the magnetosome as a whole, defining proteins enriched in the magnetosome

Finally, the fifth chapter contains unpublished material related to MamE and ends with concluding thoughts and perspectives. It includes a preliminary investigation of how magnetosome proteins reach the magnetosome, a poorly understood topic, but one that MamE also plays a role in. It will also identify one further potential target of MamE proteolysis, MamT, and investigates how the magnetochrome domain, shared by MamE, MamP, and MamT is crucial to the function of all three proteins but not in identical ways. Finally, it will include thoughts on ways to move forward in the field of magnetosome genetics and molecular biology.

Chapter 1

The evolution of magnetotactic bacteria and genetics of magnetosome production

Patrick Browne[1] and Arash Komeili[1,2,3]

1 Department of Plant & Microbial Biology, 2 Department of Molecular & Cellular Biology, 3 California Institute for Quantitative Biosciences

Abstract
Magnetotactic bacteria are a paraphyletic group of bacteria that respond to magnetic fields using magnetosomes, organelles unique to bacteria that contains magnetic iron particles with precise sizes and shapes. All identified magnetotactic bacteria derive their magnetosomes from gene islands, easily identifiable due to the presence of some well conserved genes. Magnetotactic bacteria and their gene islands provide a unique opportunity for melding the study of gene mechanism and gene evolution. Magnetosomes are complex structures, requiring the correct function of over two dozen genes, but they are also an ancient bacterial feature, and existing extant magnetotactic bacteria and their respective magnetosome islands have diverged for millennia. Here, we will review the most well conserved genes from magnetosome gene islands based on general function as well as explore the evolutionary relationships between magnetotactic bacteria.

Introduction

Magnetotactic bacteria (MTB) are a ubiquitous aquatic group of bacteria defined by their response to external magnetic fields. They are relatively easy to enrich for due to their movement towards strong magnets, and were first identified in 1963(91). While several possible functions have been proposed, the most well accepted is that magnetosomes allow magneto-aerotaxis (92). Aquatic oxygen concentrations vary by depth but are uniform laterally. By aligning to geo-magnetic fields bacteria are able to tax in one dimension, allowing them to locate desired oxygen concentrations in a more efficient manner (93).

Extensive genetic analysis of MTB was first performed in the *α-Proteobacterium Magnetosprillum magneticum* AMB-1 and its close relative *Magnetosprillum gryphiswaldense* MSR-1. Both of these bacteria are relatively genetically tractable and remain the most common models for studying magnetosome formation today (16, 94). Recently, the eighteen-gene operon (MamAB) was shown to harbor numerous genes crucial to magnetosome biogenesis (12, 21). While genes outside of the AB operon play a role in insuring robust magnetosomes, the AB operon is sufficient to make magnetosome membranes with some biomineralized crystal, even when the rest of the MAI in AMB-1 is deleted (50).

More recently, extensive sequencing of MTB outside the *α-Proteobacteria* has revealed that at least five of the Amb-1 and MSR-1 mamAB genes (MamA, MamB, MamE, MamI, and MamK) are clearly found in all sequenced MTB within *Proteobacteria* and *Nitrospirae* (95, 96). Some of the remaining mamAB genes are found in some but not all MTB while others are only conserved in closely related *α-Proteobacteria*. Other accessory operons found in AMB-1 and MSR-1 are also conserved through-out the *α-Proteobacteria* MTB (MamGFDC, MamXY, and Mms6F) but not found outside of *α-Proteobacteria* (97). Roughly half of AMB-1's MAI is made up of genes that are poorly conserved, even in its close relative MSR-1. None of these genes, other than ones that are clearly recent duplications of other conserved genes, have been shown to play a crucial role in magnetosome biogenesis and what their functions are, if they have any, remain unclear. Recently, the mamAB operon and other conserved alpha MAI genes from MSR-1 were shown to be sufficient for regenerating magnetosomes in *Rhodospirillum rubrum,* a closely related but non-magnetotactic bacteria (60).

Evolution of MTB

While the majority of genetic experimentation has been confined to alpha-proteobacteria MTB, recent work in the delta-proteobacteria RS-1 hints that the most well conserved mamAB genes likely perform the same central role in all MTB. However, other MAI genes not conserved in AMB-1 and MSR-1 were also shown to be crucial (61). Furthermore, the finding of mutants with only one magnetosome per cell (a phenotype never observed in AMB-1 or MSR-1) implies that how magnetosomes are generated in *δ-Proteobacteria* MTB might be fundamentally different. Extensive sequence analysis has also shown that there are several genes conserved through-out *δ-Proteobacteria* and *Nitrospirae* MTB (Mad genes). While the function of most of these genes remains unknown it can be presumed that at least some of them play a role in the different crystal morphologies seen in these clades. Finally, a third class of genes appear only in *Nitrospirae* (Man genes). Partial sequencing of the magnetotactic *Omnitrophica* SKK-01 shows it also at least contains *mamE, mamM, mamB,* and *mamK*, as well as *Mad11*.

Because the genes for magnetosome biogenesis are located within a gene island and the simple fact that the clear majority of bacteria are not magnetotactic, it is often assumed that the MAIs are acquired through HGT. However, sequence analysis of the core MAI genes largely matches the phylogenetic trees, built on housekeeping such as the 16S ribosomal RNA. Furthermore, common markers of HGT such as different GC content and codon bias that deviate largely from the whole genome, are not observed. This suggests that magnetosome biogenesis largely has a monophyletic origin, and while some amount of ancient HGT cannot be ruled out, it is likely that the presence of magnetosome genes was far more prevalent among ancient bacteria (15).

One clear and informative exception can be found in *Desulfamplus magnetovallimortis* BW-1, a *δ-Proteobacterium*. *δ-Proteobacteria* MTB fall in two general lineages. One contains the magnetite producing strains found in the Desulfovibrionales order, and the other contains the greigite producing strains found in the Desulfobacterales order (phylogenetically very distinct lineages despite the similar name). While BW-1 clearly falls in the latter group by 16S, it is able to produce magnetosomes containing greigite or magnetite (105). Sequencing shows that the strain actually contains two core mamAB operons, one matching its close relatives within Desulfobacterales and the other matching the magnetite producers of the Desulfovibrionales. While both operons are largely left intact, the strain does only contain one copy of *mamK*, a gene coding for a protein that, unlike most MAI proteins, operates in the general cytoplasm and not in the magnetosomes themselves. This operon organization hints that, with the exception of *mamK*, BW-1 MAI genes have evolved to fit their specific mineral and are not interchangeable. Greigite producers are also the only MTB to not contain any heme-binding magnetochrome domain containing proteins or TauE like proteins. While the lack of a workable genetic model makes further analysis of the greigite producers difficult, they offer a unique look into what is absolutely necessary for magnetosome biosynthesis.

Delta-proteobacteria MAIs demonstrate an interesting mix of nitrospirae and alpha/gamma-proteobacteria MAIs. On the one hand, they share several conserved genes with the Nitrospirae not found in the alpha/gamma-proteobacteria. Recent re-analysis of the bacterial tree of life shows that delta-proteobacteria most likely diverged far earlier from the other clades of proteobacteria and are in fact closer related to nitrospirae (101). These finding are reflected by

the recent analysis of core nitrospirae MAI genes, which are shown to be more closely related to delta-proteobacteria MTB than the core genes from alpha- and gamma-MTB are. On the other hand, delta-proteobacteria MAIs contain certain similarities to alpha- and gamma- MAI such as the presence of repurposed serine proteases (discussed below).

While most conserved MamAB proteins appear to have been inherited vertically, other conserved features were acquired independently. Specifically, genes annotated as *MamO* was recently shown to be a non-monophyletic gene. While alpha-proteobacteria MamOs are true distinct protein family, genes annotated as MamO within gamma-proteobacteria, magnetite producing delta-proteobacteria, and greigite producing delta-proteobacteria arose from duplications from those lineages' respective MamEs. In other words, in four different occasions during the evolution of magnetotactic bacteria, serine proteases have been repurposed into a non-proteolytic function. While there is no experimental evidence that the degraded proteases of the gamma- and delta-proteobacteria fulfill the same function of AMB-1's MamO (metal-binding) they all appear by sequence analysis to not be functional proteases. Finally, nitrospirae MAIs contain only one (presumably functional) protease, homologous to mamE (73).

Not only is the independent repurposing of serine proteases a potential example of a protein family convergently evolving to fulfill a new role but also begs the investigation into whether other magnetosome proteins may have arisen in a similar fashion. For example, while all MTB contain FeoAB proteins, the genes found in alpha and gamma proteobacteria appear distinct from the genes found in delta-proteobacteria and nitrospirae. Furthermore, while most MTB contain two related cation transporters (MamB and MamM) the gene annotated as *mamM* in the greigite producing MTB appears to actually be a recently duplicated copy of its own *mamB*. Finally, the TauE domain that is fused to the inactive protease in alpha-proteobacteria seems evolutionarily distinct from the TauE protein found in delta-proteobacteria and nitrospirae.

Membrane Formation and Chain Organization

Magnetosome biogenesis begins with the invagination of parts of the inner membrane. Invagination occurs at disparate points throughout the cells and magnetosomes are not organized into chains initially. In AMB-1 the budding vesicle does not fully pinch off from the general membrane. In the very closely related MSR-1, however, the magnetosomes eventually fully separate from the membrane to form distinct vesicles (20, 16).

The mechanism of how MTB trigger their membranes to remodel remains relatively mysterious. There are no clear homologs to eukaryotic proteins that are involved in membrane remodeling within the MAI although a subset of genes is specifically required. In AMB-1, four genes are necessary for membrane formation (mamB, I, L, and Q), while in MSR-1 only mamB seems absolutely necessary for at least some amount of membrane formation (102, 103).

In both species, these four genes are not sufficient for the formation of membranes and require additional MAI genes. Extensive work in MSR-1 has shown that is no clear small set of proteins that can reconstitute magnetosome membranes. For example, the set of mamBILQEMO is not sufficient to generate membranes alone. However, when combined with either the remaining mamAB genes or the mamGFDC and mms6F accessory operons cells will

produce magnetosome membranes. These accessory operons however can normally be deleted with slight crystal defects but with no apparent effect to membranes. In other words, even though very few proteins are absolutely required for membrane remodeling, many MAI proteins, including some outside of the core operon are capable of providing some overlapping function necessary to produce a magnetosome membrane (103).

MamB, the one gene shown to be absolutely crucial to membrane generation in both AMB-1 and MSR-1 is also one of the most well-conserved genes. However, it is a predicted cation transporter suspected to play a role in iron uptake and perhaps a surprising candidate to be instrumental in membrane remodeling (104). On the other hand, it does interact with other MAI proteins *in vitro.* Taken together with the fact that just having a majority of conserved MAI proteins is necessary, it may be that magnetosome membrane generation is triggered by protein crowding and lateral protein-protein interactions that mamB provides just happen to be non-redundant.

Protein Localization

The majority of MAI genes required for mature magnetosome generation localize to the magnetosome itself, either within the soluble compartment itself, the surrounding membrane, or within the cytoplasm immediately outside the membrane (mamK). While there are no identified magnetosome specific sorting motifs, several conserved MAI proteins have domains predicted to be involved in protein-protein interaction.

MamE and MamP both contain PDZ domains and deleting MamE does cause several other MAI proteins to mislocalize (25). However, magnetosome membranes still form in a *mamE* deletion, hinting that proteins involved in the early steps of membrane formation still localize correctly. MamA contains TPR domains and forms *in vitro* reactions with itself and other MAI proteins. Predicted to be a cytoplasmic protein, it coats the outside of the magnetosome (19). While its deletion does affect the localization of MamC, the most notable phenotype is a severe drop in the number of overall magnetosomes. However, the magnetosomes that are present appear to be WT in size. As indicated above, it is likely that MAI proteins form several non-specific and overlapping lateral interactions with each other, and perhaps MamA fulfills a coordinating action that can be sporadically compensated for by other MAI protein-protein interactions. Recent work shows that magnetosomes membranes are very enriched for magnetosome proteins and largely exclude non-MAI proteins, further supporting that MAI proteins find each other through lateral interactions within the membrane and initialization of formation might proceed through a lipid raft like model (106).

Chain Organization

One of the most well conserved features of MTB is that their magnetosomes are precisely arranged into chains. While the number of chains per cell, organization of those chains, and number of magnetosomes per chain vary widely throughout MTB, all MTB have a consistent arrangement. Because displacement of magnetosomes results in a less efficient magnetic potential MTB need a mechanism to maintain chain integrity (20, 107-109).

Early cryo-electron tomograms of MTB revealed bundled protein filaments flanking magnetosomes and these filaments were soon shown to be made of MamK, one of the universally conserved magnetosome proteins. MamK is related to MreB, the well-studied actin-like cytoskeletal bacterial protein involved in maintaining cell shape. Similar to MreB, MamK is an ATPase and forms chains *in vitro* and *in vivo.* Its deletion in AMB-1 and MSR-1 results in a loss of chain organization, leaving magnetosomes scattered throughout the cell. Other proteins, such as MamJ as well as MamK-like, a homolog of MamK located outside of the main MAI of AMB-1 have also been shown to interact with MamK as well as play a role in chain organization (108). However, the fact that MamK-like is not even present in MSR-1 and the fact that deletions of *mamK* or *mamJ* produce different chain phenotypes in AMB-1 and MSR-1 hint that the specifics of chain maintenance might be a relatively less well conserved aspect of MTB.

Nucleation and Biomineralization

In AMB-1 and MSR-1, magnetosomes contain precise sized cubo-octahedral crystals. One reason MTB are found interesting is because of their ability to manufacture these nano-crystals in a more consistent manner than common material science practices. Similar to the above functions, there are no MAI genes that bear any similarity to known eukaryotic biomineralization genes. However, many MAI proteins have biomineralization specific phenotypes when deleted. Broadly speaking, these genes can be divided into two categories, those involved in crystal nucleation and those involved in complete and functional biomineralization (12,13).

Four genes within the *mamAB* operon result in fully developed magnetosome membranes with no crystals when deleted (MamM, N, O and E). MamM, MamN, and the tauE-like domain of MamO are all predicted cation transporters. While MamM might be responsible for transporting the large amount of iron necessary for the crystals themselves, MamN and MamO likely transport other ions. While the exact chemical environment within a magnetosome as opposed to the cytoplasm or periplasm, remains mostly undefined, it is likely that a specific pH and other ion concentrations must be maintained for nucleation. Finally, MamE, which as previously mentioned, is an HtrA protease with multiple functions, plays a role in correct protein localization and its absence likely results in other proteins directly necessary for nucleation to not successfully reach magnetosome (25).

Several MAI genes can be categorized as biomineralization genes. Within the core MamAB operon, the deletion of mamP, mamT, mamS, and mamR all cause severe defects to magnetosome crystals and severely negatively impact a cell's ability to align with a magnetic field. Furthermore, while MamE serves several purposes its actual protease activity is directly implicated in biomineralization (25). MamE has been shown to process several MAI proteins, including itself, MamP, and MamO (112). Finally, the proteins from the mamGFDC and mms6F operons also play an accessory but less crucial role biomineralization. These proteins are not conserved through-out all MTB and may be partially key to understanding why crystals occur with such widely different morphologies throughout MTB (63).

Conclusion

The significant progress in understanding of the genetics of magnetosomes in the last fifteen years, has also revealed the complexity and diversity of these organelle. While conserved genes

can be broadly sorted into basic functions based on their deletion, the actual mechanism of action of all but a few of these genes remains largely not understood. Even the most well-studied genes such mamK, mamE, and mamO have revealed previously unpredicted functions and a large number of interacting partners. Furthermore, the fact that even the first step of magnetosome biogenesis (membrane formation) requires a large but not precisely defined amount of conserved MAI proteins shows that these proteins are often multi-functional and in some cases partially or fully redundant. While it is always tempting to categorize a gene as having one precise function, this supposition is increasingly being proven untrue in MTB.

Finally, the process of magnetosome formation in MTB outside of alpha-proteobacteria remains largely mysterious. Work in RS-1 has shown that while certain gene functions may be conserved, the timing and nature of how magnetosomes are made may be wildly different. The collection of MAI sequences shows that relatively few genes are completely conserved throughout all MTB and that certain types of genes have been absorbed by MAIs in a convergent manner. These diverse MAIs provide a fascinating opportunity for studying how replaceable certain proteins and/or specific functions are, while also preserving the overall purpose of a collection of genes.

CHAPTER 2

MamO is a repurposed serine protease that promotes magnetite biomineralization through direct transition metal binding in magnetotactic bacteria

David M. Hershey[1], Xuefeng Ren[2,3], Ryan A. Melnyk[5], Patrick J. Browne[1], Ertan Ozyamak[1], Stephanie R. Jones[4], Michelle C. Y. Chang[2,4,5], James H. Hurley[2,3,6] and Arash Komeili[1,2,3]

1 Department of Plant & Microbial Biology, 2 Department of Molecular & Cellular Biology, 3 California Institute for Quantitative Biosciences, 4 Department of Chemistry, University of California, Berkeley, CA 94720, USA; 5 Physical Biosciences Division, 6 Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

ABSTRACT
Many living organisms transform inorganic atoms into highly ordered crystalline materials. An elegant example of such biomineralization processes is the production of nano-scale magnetic crystals in magnetotactic bacteria. Previous studies implicated the involvement of two putative serine proteases, MamE and MamO, during the early stages of magnetite formation in *Magnetospirillum magneticum* AMB-1. Here, using genetic analysis and X-ray crystallography we show that MamO has a degenerate active site, rendering it incapable of protease activity. Instead, MamO promotes magnetosome formation through two genetically distinct, non-catalytic activities: activation of MamE-dependent proteolysis of biomineralization factors and direct binding to transition metal ions. By solving the structure of the protease domain bound to a metal ion, we identify a surface exposed di-histidine motif in MamO that contributes to metal binding and show that it is required to initiate biomineralization *in vivo*. Finally, we find that pseudoproteases are widespread in magnetotactic bacteria and that they have evolved independently in three separate taxa. Our results highlight the versatility of protein scaffolds in accommodating new biochemical activities and provide unprecedented insight into the earliest stages of biomineralization.

INTRODUCTION
Biomineralization is the widespread phenomenon by which living organisms transform inorganic atoms into highly ordered, crystalline structures. Controlling the size and shape of such materials requires specialized protein machinery that can define the nano-scale trajectory of crystal growth (3). Incorporating biochemical principles uncovered from studying biomineralization has the potential to improve the design and synthesis of nanomaterials *in vitro* (4). In addition to the well-known examples of tooth, bone, and shell production by multicellular eukaryotes, a number of bacteria have the ability to biomineralize small magnetic crystals within subcellular compartments called magnetosomes (1, 2). These particles allow the cells to passively align in the earth's magnetic field, facilitating the search for their preferred oxygen environments (5). Although these magnetotactic bacteria have drawn longstanding interest due to their ability to manipulate transition metals, the biochemical details of how they transform iron into magnetite (Fe_3O_4) remain poorly understood.

Magnetotactic organisms are phylogenetically diverse. Nearly all isolates come from the α-, δ- or γ- classes of *Proteobacteria,* but representatives from the *Nitrospirae* and *Omnitrophica* phyla have recently been identified (6). The genes responsible for making magnetosomes are often contained in a genomic region called the magnetosome island (MAI) (7, 8, 9, 12, 13). Comparative genomic and phylogenetic studies have identified a set of core genes that appears to have been assembled a single time and inherited vertically, indicating that magnetosome formation likely predates the divergence of the *Proteobacteria* (14, 15). The MAI seems to have formed by incorporating elements from other cellular processes, as the majority of the core factors have homology to ancient and widespread protein domains (16, 17). Uncovering the biochemical functions encoded in the MAI in relation to its evolutionary history provides a unique opportunity to understand how new cellular processes evolve.
Due to the availability of genetic systems, α-*Proteobacteria* like *Magnetospirillum magneticum* AMB-1 are used as models for studying the molecular biology of magnetosome formation (18). AMB-1 contains 15-20 magnetite crystals, each formed within a cytoplasmic membrane invagination and organized in a chain spanning the length of the cell (19, 20). By making

deletions within the MAI and characterizing the ultrastructure of the mutant cells, specific genes have been assigned roles in various stages of magnetosome formation (12, 13, 21, 22). Genes whose deletions produced empty magnetosome compartments or compartments with abnormally small magnetite crystals were termed biomineralization factors. It is the proteins encoded by these genes that hold the secrets of how magnetotactic organisms interface with solid magnetite. Two genes in the MAI, *mamE* and *mamO*, are homologous to the HtrA proteases, a ubiquitous family of trypsin-like enzymes that functions using His-Asp-Ser catalytic triads (23). An additional pair of genes, called *limE* and *limO,* with homology to the protease domains of *mamE* and *mamO* exists in a secondary genomic region termed R9 (12). Disrupting *mamE* or *mamO* causes cells to produce empty magnetosome membranes, but removing R9 has no effect, showing that *mamE* and *mamO* are required for biomineralization and that *limE* and *limO* are not (Figure 2-1) (12, 24). Adding variants of *mamE* or *mamO* with all three predicted active site residues mutated to alanine could not restore normal biomineralization in the Δ*mamO*Δ*R9* or Δ*mamE*Δ*R9* strains but complemented single Δ*mamO* or Δ*mamE* deletions (25). These genetic analyses show that *mamE* and *mamO* are required for the initiation of magnetite biomineralization. Furthermore, *limE* and *limO* are partially redundant in that they can cross-complement the active site dependent crystal maturation defects of their respective orthologs.

Here we use a combination of *in vivo* and *in vitro* approaches to reveal an unexpected dual role for MamO. It promotes MamE-dependent proteolysis of three biomineralization factors through the use of its C-terminal transporter domain. Separately, the protease domain has lost the ability to carry out catalysis and has instead been repurposed to bind transition metal ions. Two surface-exposed histidine residues that contribute to this metal-binding function are required for initiating magnetite biosynthesis *in vivo*. Bioinformatic analysis shows that similar pseudoproteases evolved independently in the three major taxa of magnetotactic organisms, highlighting a unique evolutionary mechanism behind microbial nanoparticle synthesis.

MATERIALS & METHODS

Strains, plasmids and culture conditions

The strains and plasmids used in this study are listed in Tables 2-4 and 2-5, respectively. For general maintenance and genetic manipulation, *M. magneticum* AMB-1 was grown in MG medium supplemented with ferric malate (30μM). 0.7% agar was used in plates and kanamycin was used for antibiotic selection at a concentration of 7 μg/mL (solid) or 10μg/mL (liquid). For sucrose counter-selection, MG plates contained 2% sucrose. Cultures for magnetic response measurements, western blotting or transmission electron microscopy (TEM) were grown in 10 mL MG medium containing 25mM HEPES buffer (pH 7.2) and ferric malate under a 10% oxygen atmosphere. For comparing the temperature dependence of magnetic response, the strains were treated as above except that they were grown under anaerobic atmosphere. The magnetic response of each culture was assessed using the Coefficient of Magnetism (Cmag), which was measured as described (56).

Genetic manipulation

For complementation of deletion mutants, we used a modified form of pAK253(47). This plasmid contains a neutral region of the AMB-1 genome and integrates as a single copy at this site. Each allele is inserted under the control of the *mamAB* promoter, allowing the constitutive expression of each protein. To create the Δ*mamE*Δ*mamO*Δ*R9* (Δ*E*Δ*O*) strain, plasmid pAK243

(*sacB* based counterselection system for deleting *mamE*) was transformed into strain AK94 (Δ*mamO*Δ*R9*), and initial integrations were selected using kanamycin. The resulting strains were grown to stationary phase in MG medium without antibiotic selection and plated on sucrose for counter selection. Deletion of *mamE* was confirmed by antibiotic sensitivity, PCR analysis and complementation of the magnetic phenotype upon reintroduction of both *mamE* and *mamO*.

Protein analysis
Cultures of AMB-1 were grown to late-log phase and harvested by centrifugation at 6k x *g*. The resulting pellets were resuspended in PBS. The cell suspensions were mixed with an equal volume of 4x SDS Loading buffer and heated for 10 min at 75°C. The lysates were separated on SDS-PAGE and transferred to PVDF. The membranes were blotted and visualized using standard western blotting techniques. Polyclonal antibodies to MamE and MamP were raised in rabbits against recombinant forms of the soluble portion of each protein (26). The monoclonal anti-FLAG antibody was purchased from Sigma.

Transmission electron microscopy (TEM)
For TEM, cultures were grown to late-log phase. 1mL from each culture was pelleted at 16k x *g*, and the pellet was resuspended in the residual medium. Cell suspensions were spotted on formvar coated copper grids, rinsed, dabbed dry and stored at room temperature until imaging. Imaging was performed with a FEI Tecnai 12 TEM at an accelerating voltage of 120kV. For each strain, 15-20 cells totaling >200 crystals were analyzed.

Protein expression
The protease domain of *Mm*MamO (resdiues 45-261) was cloned into mcsII of pETDuet to create pAK876 for expression without a tag. BL21 Codon Plus cells transformed with pAK876 were grown at 37°C in 2xYT with carbenicillin (100 μg/mL) and chloramphenicol (25 μg/mL) until the OD_{600} reached 1.0. The cultures were then equilibrated at 20°C for 30 min, induced with 0.125mM IPTG and grown overnight. Cells were harvested by centrifugation, resuspended in Buffer A (25mM Tris-HCl pH7.4, 400mM NaCl, 5mM Imidazole, 10% glycerol) supplemented with 1μg/mL pepstatin A, 1μg/mL leupeptin and 0.5mM DTT, frozen in liquid nitrogen and stored at -80°C.

Protein purification
For crystallography, the frozen cell suspension was thawed on ice, lysed by sonication and clarified at 13,000 x *g* for 30 min. The supernatant was loaded on a 3 mL Ni-NTA column, which was then washed with 5 column volumes of Buffer B (25mM Tris-HCl pH7.4, 500mM NaCl, 10mM Imidazole, 10% glycerol). The protein binds to Ni-NTA with its native metal binding site. After elution with 1 column volume Buffer C (25mM Tris-HCl pH7.4, 250mM NaCl, 250mM Imidazole, 10% glycerol), the sample was dialyzed overnight against Buffer D (25mM Tris-HCl pH8.0, 50mM NaCl, 10% glycerol, 0.5mM EDTA) and loaded onto a 1mL HiTrap CaptoQImpRes column. The column was developed to Buffer E (25mM Tris-HCl pH8.0, 1M NaCl, 10% glycerol), and the peak fractions containing the MamO protease domain were pooled, exchanged into Storage Buffer (25mM Tris-HCl pH7.4, 300mM NaCl, 10% glycerol), concentrated to ~20mg/mL and frozen in liquid nitrogen for storage at -80°C.

For tmFRET, the expression was performed as above except that the various mutants of the

*Mm*MamO protease domains were expressed as fusions to a C-terminal strepII tag. The lysate was prepared as above and loaded onto a 1mL StrepTrap HP column, which was washed with 5 column volumes of Buffer F (25mM Tris-HCl pH7.4, 250mM NaCl, 10% glycerol) and eluted with 3 column volumes of Buffer F supplemented with 2.5mM desthiobiotin. The eluate was concentrated and loaded onto a 16/60 Superdex 200 column and developed in Storage Buffer. The peak fractions were pooled, concentrated and used immediately for fluorescent labeling. The purity of all proteins was confirmed by SDS-PAGE.

Crystallization and structure determination

Frozen aliquots of untagged MamO protease domain were thawed on ice and exchanged to Buffer G (20mM Tris-HCl pH7.4, 150mM NaCl, 5% glycerol, 0.5mM DTT) while adjusting the protein concentration to 5mg/mL in a 10kDa cutoff Amicon ultrafilter. Crystals grew in the hanging drop vapor diffusion format after mixing the protein with an equal volume of well solution (50mM Na-Acetate pH4.6, 3.6M NH4Cl, 5% glycerol) and equilibrating against 1mL well solution at 18°C. Cubic crystals of MamO appeared in 1-2 days and grew to their full size after 4-6 weeks. Each crystal was cryo-protected with a solution of 50mM Na-Acetate pH 4.6, 3.6M NH4Cl and 22% glycerol before plunge-freezing in liquid N2.

For nickel soaking, NiCl2 was added directly to drops containing fully-grown crystals for a 10mM final concentration. The wells were re-sealed and the crystals soaked overnight. At harvest, each crystal was passed through three 1 min soaks in a buffer containing 50mM Tris-HCl pH 8.0, 3.6M NH4Cl, 22% glycerol, 10mM NiCl2 before plunge-freezing.
Diffraction data was collected on beamline 8.3.1 at the Advance Light Source (Lawrence Berkeley National Laboratory, Berkeley CA). Indexing and scaling was performed using HKL2000 (27). The apo- structure was solved by molecular replacement in Phaser with the protease domain of *Ec*DegP (1KY9) as a search model (28). The Ni-bound structure was solved using the apo-MamO structure as a molecular replacement search model. Model building and refinement were carried out with alternating cycles of COOT (29) and phenix.refine (30). Placement of the Ni ion was confirmed with an anomalous map from SAD data collected at the Ni absorption edge.

Fluorescein-5-maleimide (F-5-M) labeling

For fluorescent labeling, variants of the MamO protease domain were diluted to 50μM in 1mL of Buffer H (50mM NaPhosphate pH7.2, 150mM NaCl, 10% glycerol). Fluorescein-5-maleimide (dissolved at 50mM in DMSO) was added to a final concentration of 1mM, and the reaction was incubated overnight at 4°C. The reaction was quenched with DTT, exchanged into Buffer D using a PD-10 column, and loaded on a 1mL HiTrap Q FF. The column was developed to Buffer E, and each protein eluted as a single fluorescent peak. The peak fractions were pooled, treated with 1mM EDTA to remove any trace metal, exchanged extensively into Chelex-treated Storage Buffer and frozen in small aliquots. Labeling efficiency was measured based on A492/A280 and was between 50-55% for all preparations.

tmFRET
For metal binding experiments all buffers were prepared in acid-washed glassware and treated

with chelex resin. Each protein was diluted to 60-80nM in fluorescence buffer (10mM Tris-HCl pH 8.0, 150mM NaCl). Metal solutions were prepared at 10X concentration in chelex treated H2O. Protein was dispensed into 96-well plates, and the plates were scanned for fluorescence emission in a Tecan Infinite 200 plate reader in top-read mode (Ex: 492nm; Em: 505-570nm). Metal solutions were then diluted into each well at the appropriate concentration and the plate was re-scanned. Due to the spontaneous oxidation of ferrous iron in ambient atmosphere, all iron binding experiments were performed in the absence of oxygen. For iron binding, all solutions were prepared using anoxic liquids. The samples were prepared under anoxic atmosphere in a clear-bottom 96-well plate. To prevent introduction of oxygen, the wells were sealed by covering the plate with a black adhesive cover. The plate was then removed from the anaerobic chamber and scanned using bottom-read mode.

Each measurement was performed on independently prepared solutions in quadruplicate. The metal quenched fluorescence spectrum from each well was normalized to the fluorescence before metal addition. Fmetal/F represents the normalized fluorescence averaged from an 11nm window around the peak. After plotting Fmetal/F as a function of metal concentration, each curve was fit to a two-site model (below), where Kd2 and E represent the dissociation constant and FRET efficiency for metal binding by MamO, respectively. Kd1 represents a nonspecific, solution-based quenching component (31).

Ni-NTA binding assays
StrepII tagged forms of the MamO protease domain were purified as described for the tmFRET experiments with the omission of the F-5-M labeling and ion exchange steps. Proteins were diluted to ~15μM in Column Buffer (25mM Tris-HCl pH7.4, 300mM NaCl, 5mM Imidazole, 10% glycerol) and loaded onto a 0.5mL Ni-NTA column equilibrated in column buffer. The column was washed with 10 column volumes of Column Buffer and eluted with 2 column volumes Buffer C. Binding was assessed by separating the fractions on 12% SDS-PAGE and staining with Coomassie Blue.

Phylogenetic analysis
To understand the phylogeny of Mam proteases, we took a broad approach, characterizing their location within the trypsin-like protease family. The *Trypsin_2* Pfam (PF13365) was used to generate a hidden Markov model (HMM) that was used to search a database composed of protein sequences from ~2100 bacterial and archaeal genomes along with the Mam proteases (32). This search was performed using hmmscan (HMMER3.1b1, hmmer.janelia.org), retaining all hits with an e-value for the entire sequence less than 10-5, identifying 6104 proteins. The hits were clustered using CD-HIT with a sequence similarity cut-off of 0.8, yielding 3431 sequences (33). These were aligned using MUSCLE (v3.8 .31) with the maxiters parameter set to 2 (34). The resulting alignment was trimmed using Gblocks, using parameters appropriate for divergent datasets as described by Sassera *et al*(35, 36). This alignment was then used to generate a phylogeny using FastTree 2 (v2.1.7) with the default settings (37). However, this alignment contained only 6 informative positions. To improve the quality of the alignment, we iteratively removed long branches from the tree and regenerated the alignment. After removing 90 taxa over 4 iterations, we settled on an alignment with 18 conserved positions.

Two clear branches emerged from this analysis. One branch ('the Deg branch') contained DegP, DegQ, and DegS sequences from the γ –*Proteobacteria*, in addition to most of the MamE

sequences. The bottom branch contained the YdgD sequence from *E. coli*. We extracted the sequences from the Deg branch (843 sequences) and generated an alignment (46 positions) and tree using the methods described above (Figure 2-14A). In this tree, the MamE sequences appeared to be closely related, but their phylogenetic relationships were ambiguous. Additionally, we realized that a subset of δ-proteobacterial trypsin-like sequences were falsely excluded from the MamE branch due to substitutions in the catalytic triad that obscured the phylogenetic signal in the short alignment. Additional phylogenetic testing confirmed the relationship of these 4 sequences to the canonical MamE sequences, so they were merged into the MamE branch.

We extracted the MamE sequences and aligned them using 5 DegS sequences from the γ – *Proteobacteria*, which appeared to be closely related based on the phylogeny in Figure 2-14B. This alignment was much larger (172 positions) and the resulting tree from FastTree 2 gave well-supported interior nodes for the MamE branch. We used two other phylogenetic methods to test this phylogeny. First, we used ProtTest 3.0 to select the best substitution matrix (in this case, the WAG model) and performed 100 independent inferences with 300 bootstraps in RAxML (38, 39). Secondly, we used the non-parametric Monte Carlo Markov Chain algorithm PhyloBayes 3 to generate a tree not based on prior assumptions about the site-specific evolution of the MamE sequences (40). A summary tree integrating the results from PhyloBayes and RAxML is depicted in Figure 2-14B. Additionally, we rooted the MamE branch to 5 closely-related sequences from the Clostridiales and this too strongly supported the phylogeny in Figure 2-14B according to both PhyloBayes and RAxML.

RESULTS

The catalytic triad in MamO is not required for magnetite nucleation

Trypsin-like proteases utilize a histidine-aspartate pair to deprotonate the hydroxyl group of a serine residue, providing a nucleophile for catalysis. MamO is unusual in that it contains a threonine instead of serine as the predicted nucleophile. To further clarify the role of MamO's unusual active site we focused our initial efforts on assessing the contribution of each putative catalytic residue to biomineralization. We performed these initial studies using the Δ*mamO*Δ*R9* strain (referred to as Δ*O*Δ*R9*) to avoid cross-complementation from *limO*. Consistent with our previous findings, individual H116A and D149A mutations in MamO severely reduced the cells' ability to turn in a magnetic field (Figure 2-2A). Surprisingly, a T225A mutation had no defects in magnetic response (Figure 2-2A). Using transmission electron microscopy (TEM) we confirmed that $mamO^{H116A}$ and mam^{OD149A} cells have small magnetite crystals while $mamO^{T225A}$ crystals are indistinguishable from wild-type, mirroring the bulk magnetic response measurements (Figure 2-2B-D).

Given that the predicted nucleophile is dispensable for magnetosome formation, we found it curious that the other two catalytic triad mutations disrupted crystal maturation. Upon further examination, we found that the phenotypes associated with the $mamO^{H116A}$ and $mamO^{D149A}$ alleles were actually temperature dependent. Growing wild-type AMB-1 at room temperature instead of the standard 30°C did not dramatically alter the magnetic response. However, both the $mamO^{H116A}$ and $mamO^{D149A}$ alleles displayed improved complementation at the lower temperature. In particular, the $mamO^{D149A}$ mutant restored a nearly wild-type magnetic response to the Δ*O*Δ*R9* strain under these conditions (Figure 2-3). Our results suggest that protease

activity from MamO is not required for biomineralization. None of the three predicted catalytic residues are required for magnetite nucleation, and, although two of the three contribute to crystal
maturation, these effects are conditional suggesting that they are not central to the biomineralization process.

Proteolytic processing of biomineralization factors
During the course of our experiments, we examined the MamO variants by western blotting. Although no changes in overall protein abundance were present, we noticed that each MamO variant was proteolytically processed, having both a full-length and shorter form (Figure 2-5). This finding led us to examine whether other magnetosome proteins are similarly proteolyzed. We found that MamE and another biomineralization factor, MamP, are also proteolytic targets in AMB-1 cells by using antibodies targeted to each protein (Figure 2-4A). MamP is a *c*-type cytochrome, and its iron oxidase activity is required for the proper maturation of magnetite crystals (41, 26). Since these three biomineralization factors exist in both full-length and shorter forms, it is likely that proteolytic maturation plays a role in their function.

To examine the potential involvement of *mamE* and *mamO* in promoting these proteolytic events, we analyzed processing of MamO, MamE, and MamP in various genetic backgrounds. In both the Δ*O*Δ*R9* and Δ*E*Δ*limE* strains we observed only the full-length form of MamP. Similar analyses showed that MamO is required for the proteolysis of MamE and that MamE is required for the proteolysis of MamO. Thus, processing of each target requires both putative proteases (Figures 2-4A and 2-5A). Addition of the $mamE^{WT}$ allele restored processing of each target in the Δ*E*Δ*limE* strain, but the catalytically inactive form ($mamE^{PD}$) did not, suggesting that MamE participates in proteolysis directly. Surprisingly, both the $mamO^{WT}$ and $mamO^{H116A}$ alleles restored processing in the Δ*O*Δ*R9* strain (Figures 2-4A and 2-5). Therefore, the presence of MamO, but not its catalytic triad, appears to be required to promote the activity of MamE. While these results strongly suggest that MamE directly cleaves all three biomineralization factors, we cannot rule out the possibility that its activity contributes to a more complex targeting process.

Given that HtrA proteases are often regulated by the formation higher order oligomers, an attractive model could be that MamO activates MamE through an interaction involving both protease domains. To test this idea, we exploited the partial redundancy from genes in the R9 region. In addition to its N-terminal protease domain, MamO has a predicted TauE-like transporter domain on its C-terminus. *limO,* the partial duplication of *mamO* in R9, is 98% identical to the protease domain of *mamO*, but does not contain a C-terminal TauE domain (Figure 2-6). We confirmed that alleles with individual point mutations in the protease domain that reduce biomineralization in the Δ*O*Δ*R9* background do not have defects in the Δ*mamO* strain*,* reinforcing that LimO is a functionally redundant copy of the MamO protease domain (Figure 2-6). In contrast, processing of MamE and MamP is disrupted in both the Δ*mamO* and Δ*O*Δ*R9* strains, showing that a functional protease domain is insufficient to activate MamE and that activation requires the TauE domain (Figure 2-4B).

Structural analysis of the MamO protease domain
In the trypsin family, loop L1 contains both the nucleophilic serine and oxyanion hole, required for creating the acyl-enzyme intermediate and stabilizing the oxyanion, respectively (Figure 2-7)

(42). In addition to the threonine substitution, the entire L1 loop in MamO differs significantly from the trypsin family consensus, once again suggesting that MamO might not be capable of protease activity (Figure 2-7). To explore this possibility further, we determined the crystal structure of the protease domain (Table 2-1). It crystallizes as a monomer with the chymotrypsin fold and the catalytic residues properly placed (Figure 2-7).

Loop L1 of MamO adopts the inactive conformation seen in many HtrA proteases in which the main chain carbonyl of residue 192 (chymotrypsin numbering; W222 in MamO) prevents access to the oxyanion hole (Figure 2-8A). In other HtrAs, the inactive form is thought to be in equilibrium with an active conformation where the main chain flips approximately 180° opening the oxyanion hole (Figure 2-8A) (43, 44). Switching to the active state forces residue 193 into a configuration that is strongly disfavored for non-glycine residues. Although glycine is highly conserved at position 193 in the trypsin family and critical for catalysis, MamO contains a glutamate (E223) at this position (Figure 2-7B) (45). Therefore, the active configuration of MamO would contain a strong steric clash between the E223 side chain and the main chain carbonyl of W222.

To illustrate this steric constraint, we examined a set of trypsin-like protease structures and analyzed the configuration of residue 193. We plotted ϕ and Ψ values for this position on a Ramachandran plot along with the favored and allowed geometries for glycine and non-glycine residues (Table 2-2) (46). As expected, the main-chain torsion angles at position 193 form two groups, distinguished by an approximately180° shift in ϕ angle. The groups correspond to active and inactive forms of loop L1. The active configuration is strongly disfavored unless glycine occupies position 193, indicating that the E223 side-chain in MamO likely prevents the formation of an oxyanion hole (Figure 2-8B).

While refining the MamO structure, we observed a peptide bound near the predicted peptide-binding groove (Figure 2-9A). We suspect its source is the flexible N-terminal region from a neighboring MamO in the crystal that is not built into the model. However, due to the modest resolution, we cannot confirm the sequence. Despite this uncertainty, the peptide has an interesting mode of binding. Similar to other trypsin-like proteases, the peptide enters the binding cleft parallel to loop L2, splitting the two β-barrels of the chymotrypsin fold. However, the bulky side-chain of W222 in MamO seems to block the exit path between loops LA and LD, forcing the peptide away from the catalytic center (Figure 2-9B-D). Overall, the structural features of the L1 loop appear incompatible with protease activity: E223 provides an energetic barrier to catalysis while W222 serves as a physical barrier to productive substrate binding.

Direct transition metal binding by MamO

Given the results of the genetic studies and structural analysis, we conclude that MamO does not act as a protease during biomineralization. This forced us to consider the possibility that the protease domain promotes magnetite formation using a function not predicted from its primary sequence. While purifying MamO, we serendipitously observed that it consistently bound to immobilized metal affinity columns. Knowing that mutations in MamO cause defects in AMB-1's ability to transform iron, a metal, into magnetite, we speculated that direct metal binding played a role in biomineralization.

Although the MamO crystals grow in acidic conditions that disfavor metal binding, they could be soaked at pH 8.0 without affecting diffraction. We solved another structure of the protease domain using crystals soaked in NiCl2 at pH 8.0 and identified a metal binding site. Overall, the conformation of the protease domain is highly similar to the original structure (root-mean-square deviation of 0.17Å over 184 residues), and it contains the unidentified peptide. Additionally, a single Ni2+ ion binds between loop LC and helix 2, with H148 and H263 directly coordinating the metal (Figure 2-10A-C). We confirmed the placement of the ligand at this site using single wavelength anomalous dispersion (SAD) data collected at the Ni absorption edge (Figure 2-10A and B). MamO's di-histidine motif is highly reminiscent of the Zn^{2+} binding site in another trypsin-like protease, kallikrein-3, where Zn^{2+} attenuates protease activity by altering the position of catalytic triad residues H57 and D102 (Figure 2-11) (47, 48).

To confirm the putative metal binding site from the MamO structure, we used transition metal ion Förster Resonance Energy Transfer (tmFRET) to assay binding in solution (49). This technique measures fluorescence quenching of a cysteine-conjugated fluorophore by a metal bound at a nearby site. Guided by the structure, we targeted Q258 because of its optimal geometry relative to the metal. Adding a number of transition metals, including iron, to a purified Q258C mutant protease domain that was labeled with fluorescein-5-maleimide caused strong fluorescence quenching (Figures 2-12 and 2-13). Although we expect Fe^{2+} to be the physiological ligand, its propensity to oxidize in ambient atmosphere added significant error to the measurements. Instead, the quenching properties and resistance to oxidation of Ni^{2+} were most suitable for detailed analysis. MamO bound to Ni^{2+} with 2.5μM affinity, compared to 1.1μM in an H148A/H263A mutant. Additionally, the FRET efficiency was significantly lower in the mutant protein, demonstrating that disrupting these residues changes the metal binding properties of MamO (Figures 2-10D and 2-13).

Although the H148A/H263A mutant displays altered behavior in the tmFRET assay, it retains the ability to bind metals (Figure 2-13). We could not identify any other metal ions in our Ni2+ soaked crystals, but the two histidines we identified did not appear to be the only metal binding residues in MamO. Consistent with this, both the wild-type and H148A/H1263A forms of MamO bound Ni-NTA resin, confirming that despite its altered binding geometry the mutant still binds to metal (Figure 2-13). Taken together, our biochemical and structural investigations show that MamO binds to transition metal ions using H148 and H263, but that metal binding is not restricted to this motif.

Because disrupting H148 and H263 altered the metal binding behavior in vitro, we predicted that these residues were important for MamO dependent biomineralization. Indeed, ΔOΔR9 strains with $mamO^{H148A}$ or $mamO^{H263A}$ alleles had no magnetic response and failed to produce electron dense particles, implying that proper metal binding is required to initiate biomineralization (Figure 2-10E and F). Both mutants displayed normal stability and proteolytic processing, and the biomineralization defects were not conditional, as lowering the growth temperature did not allow for a magnetic response (Figures 2-3 and 2-5). Additionally, both alleles restored a magnetic response in the single ΔmamO background, showing that limO can provide the required metal binding activity independent of co-translation with the TauE domain (Figure 2-6). These are the most disruptive point mutations we have observed in MamO, and they completely

recapitulate the phenotype of a mamO deletion. We conclude that H148 and H263 contribute to a metal binding function that is required for magnetite nucleation in vivo.

Convergent evolution of pseudo-proteases in magnetotactic bacteria While the finding that MamO has lost its protease activity to become a metal binding protein in AMB-1 was quite intriguing, we wanted to know whether this mechanism was conserved in other organisms. Due to the fact that magnetotactic *Nitrospirae* and *Omnitrophica* have not been isolated in culture, we focused our analysis on the Proteobacteria for which numerous representatives are available in pure-culture. Examining available whole-genome sequences indicated that all magnetotactic strains from the α-, δ- and γ-*Proteobacteria* contain two predicted trypsin-like proteases in their MAIs (Table 2-3). We attempted to understand the evolutionary history of these proteins by including them in a large phylogenetic tree of the bacterial trypsin-2 superfamily (Materials and Methods). Within this tree, the MamO sequences from each magnetotactic α-Proteobacterium form a distinct, monophyletic clade (Figure 2-14A). Each protein has a degenerate catalytic triad, a non-glycine residue at position 193 and a bulky tryptophan in the L1 loop. The metal binding positions in LC and helix 2 also appear conserved. Although one strain has an aspartate at position 263, aspartates are also common metal coordinating residues, and we predict that this H-D motif can also participate in metal binding. We conclude that the MamO family evolved specifically in α-Proteobacterial magnetotactic organisms to be metal binding pseudo-proteases (Figure 2-15).

In addition to the MamO proteins, we identified a second magnetosome-specific clade in the trypsin tree (Figure 2-14A). We named this group the MamE clade because it contains the MamE sequence from the MAI of each representative of the α-Proteobacteria. This group also features two predicted trypsin-like proteases from the MAI of each species of the δ- and γ-Proteobacteria, re-named MamE1 and MamE2 here (Table 2-3). The sequence phylogeny indicates that the δ- and γ -Proteobacteria both experienced recent duplications of their MamE's independently. In fact, there appears to have been two rounds of duplication in the δ- family. Strikingly, each duplication event has led to a degenerate catalytic triad in one of the sequences, showing that duplication and loss of function occurred three separate times in the MamE family (Figures 2-14B and 2-16). While these inactive MamEs do not have the di-histidine motif identified in MamO, we cannot rule out the possibility that they bind metals by another mechanism. Regardless, our results imply the existence of selective pressure for pairing active and inactive proteases as each major clade of magnetotactic bacteria has evolved this feature independently.

DISCUSSION

Magnetotactic bacteria control the growth of their associated magnetite crystals with a level of precision that cannot be replicated in vitro. The molecular details of how they perform this task can reveal novel bioinorganic interfaces and be exploited for improved synthesis of nanomaterials. Genetic analysis has shown that magnetite biomineralization is surprisingly complex. It requires over 15 factors in AMB-1, nearly all of which are predicted integral membrane proteins (4, 50, 22). A subset of these is required for the initial crystallization of iron within the magnetosome compartment (12). While the key players for this step are known, their biochemical functions have only been inferred from sequence homology, leaving the mechanism of how magnetite biosynthesis begins a mystery.

Here we examined two magnetite nucleation factors, the putative HtrA proteases MamE and MamO. We find that the presence of both MamE and MamO is required for the proteolysis of three biomineralization factors, MamE, MamO and MamP. These events depend on an intact catalytic triad from MamE but not MamO, indicating that MamO activates MamE in a noncatalytic manner. Thus far, we have not been able to detect a physical interaction between MamE and MamO, but we have found that the C-terminal TauE-like transporter domain of MamO is required for activation. The putative ion transport activity of this domain could be the feature that promotes proteolysis. This model is attractive because it does not require a direct interaction between the two proteins. Indirect evidence suggests that TauE family proteins transport sulfite or sulfur containing organic ions, leading us to speculate that the concentrations of specific solutes in the magnetosome might control MamE's activity (51, 52).

Separately, we discovered that a metal-binding function in the protease domain of MamO is required for the initiation of magnetite biomineralization. In our structure, H148 and H263 directly coordinate a single metal ion. Disrupting these residues alters the binding behavior in a modified tmFRET assay, but the effect is unusual in that it lowers the FRET efficiency while slightly increasing the overall affinity for metal. Using the Förster equation and reported radius for Ni2+, we calculated that the fluorophore to metal distance changes from 12.8Å in wild-type to 15.2Å in the H148A/H263A mutant, suggesting that the binding geometry changes in a way that allows metal binding in the same vicinity (31). We favor the explanation that the di-histidine motif identified here is part of a more complex metal coordinating network. Our soaking strategy may have missed additional sites that are inaccessible in our crystal form, and separate attempts at characterizing a fully metal-bound state using co-crystallization have been unsuccessful. Nevertheless, we identified di-histidine motif using the structure that contributes to an unexpected transition metal-binding activity in MamO.

Despite the presence of other binding features, metal binding through H148 and H263 is absolutely required in vivo as disruption of either residue completely abolishes magnetite formation. Our structural analysis shows that MamO has lost its ability to perform proteolysis altogether, supporting the idea that metal binding is now the central function of the protease domain. Consistent with this, T225, the predicted catalytic nucleophile, is completely dispensable for biomineralization. Though disrupting H116 and D149 in the predicted catalytic triad causes conditional crystal maturation defects, magnetite nucleation is not affected. Interestingly, H116 and D149 participate in a hydrogen bond on the opposite face of loop LC from the H148/H263 metal binding motif, suggesting that the conditional phenotypes could be 34 due to temperature-dependent flexibility near the metal binding site (Figure 2-11). A potential link between the two motifs is consistent with the reported inhibition of protease activity through a highly analogous metal binding site in the kallikrein family that rearranges the H-D catalytic pair (Figure 2-11).

While templating of magnetite growth via an interaction between biomineralization factors and the mineral surface has been proposed, our findings with MamO emphasize that direct interactions with individual solute ions also play a role (53, 50, 54). One of the most fascinating aspects of MamO's metal ion interaction is that the H148A and H263A forms of MamO maintain the ability to bind metals but cannot support any magnetite biosynthesis in vivo. It

appears that binding is insufficient and that the precise coordination geometry must be maintained, leading us to speculate that MamO directly promotes nucleation by guiding individual iron atoms into the magnetite lattice. This model is consistent with the phenotypes observed in vivo, the modest binding affinity and the surface exposed nature of the simple dihistidine motif. Additionally, it agrees with topological predictions for MamO placing the protease domain in periplasm, which is continuous with the magnetosome lumen in AMB-1 (20). More broadly, our results define an unexpected mechanism for MamO in biomineralization. It appears to have lost the ability to perform serine protease activity and instead performs two noncatalytic functions: direct metal binding to promote magnetite nucleation and activation of MamE's proteolytic activity (Figure 2-17A).

In addition to the surprising mechanism for MamO in AMB-1, we uncovered a fascinating evolutionary expansion of the trypsin family within magnetotactic bacteria. Our analysis suggests that the ancestral MAI contained a single trypsin-like protease homologous to MamE. The δ- and γ –Proteobacteria experienced independent duplications of this ancestral enzyme, while the α-Proteobacteria appear to have acquired a second, distantly related trypsin-like protease. Despite these different origins, having two redundant proteases seems to have allowed one copy to lose its catalytic ability in all three clades (Figure 2-17B). In α-Proteobacteria, MamO specialized to promote biomineralization through the two non-catalytic activities identified here. While the pseudo-proteases in the other clades remain uninvestigated, the fact that inactive copies are retained strongly suggests that they also play important non-catalytic roles. The pathway that led to convergent evolution of pseudo-proteases in magnetotactic organisms highlights the critical role duplication and redundancy play in facilitating diversification of protein function (55).

Perhaps more intriguing is the fact that MamO's metal binding motif is placed at the same site on the chymotrypsin fold as the highly analogous zinc-binding site seen in the distantly related kallikreins (47, 58). This hints toward the possibility that the ability to bind metals may be a latent biochemical function carried within the fold. Such activities are absent in specific evolutionary states of a protein but can quickly surface under selective pressure (56, 57). Consistent with this, trypsin-like proteases utilize catalytic residues on loops that are well separated from the core, a property termed fold polarity that correlates with the capacity for functional diversification (58, 59). Perhaps, neofunctionalization of the trypsin scaffold within magnetotactic organisms is due to an inherent stability and adaptability in the fold that makes it a useful building block for biochemical innovation.

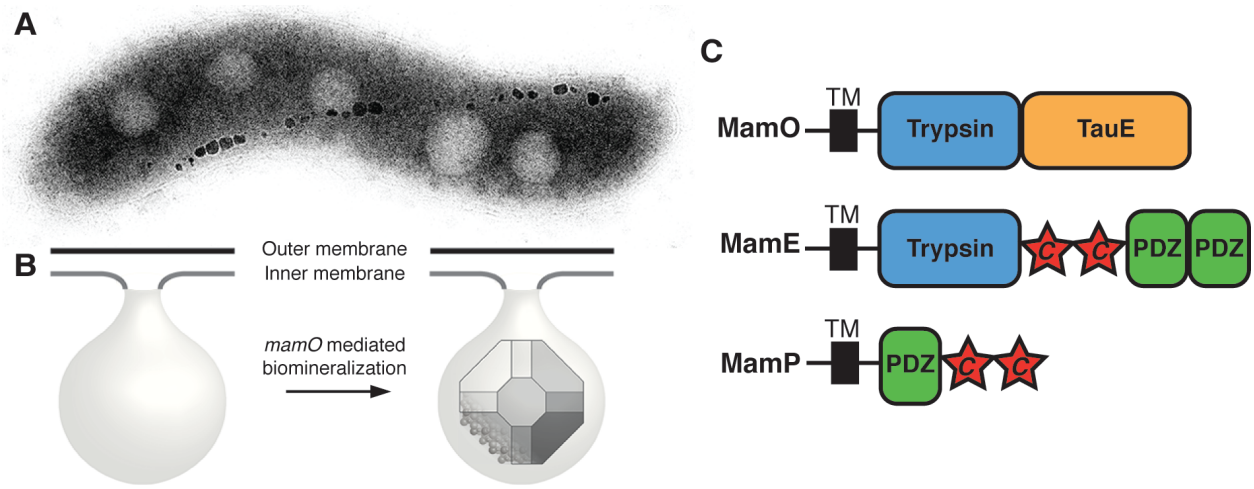


Figure 2-1. MamO promotes the nucleation of magnetosome crystals. (a) TEM micrograph of a wild-type AMB-1 cell. The electron dense particles make up a magnetosome chain. (b) Cellular organization of the magnetosome compartment. MamO promotes the nucleation of magnetite within inner membrane invaginations. (c) Domain structure of the three biomineralization factors discussed in the text. "*c*" represents a CXXCH *c*-type cytochrome motif.

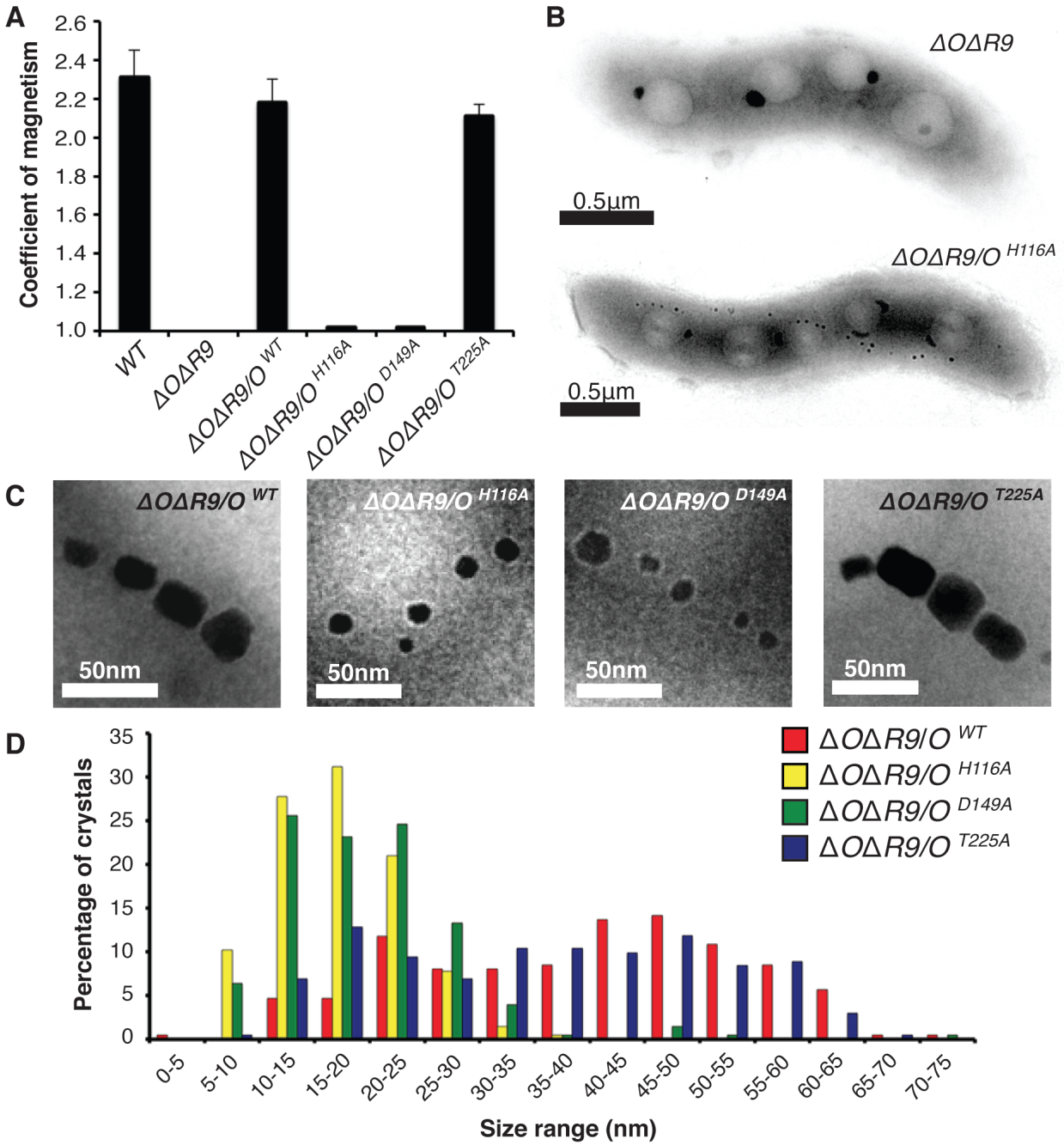


Figure 2-2. MamO promotes the nucleation of magnetosome crystals. (a) TEM micrograph of a wild-type AMB-1 cell. The electron dense particles make up a magnetosome chain. (b) Cellular organization of the magnetosome compartment. MamO promotes the nucleation of magnetite within inner membrane invaginations. (c) Domain structure of the three biomineralization factors discussed in the text. "c" represents a CXXCH c-type cytochrome motif.

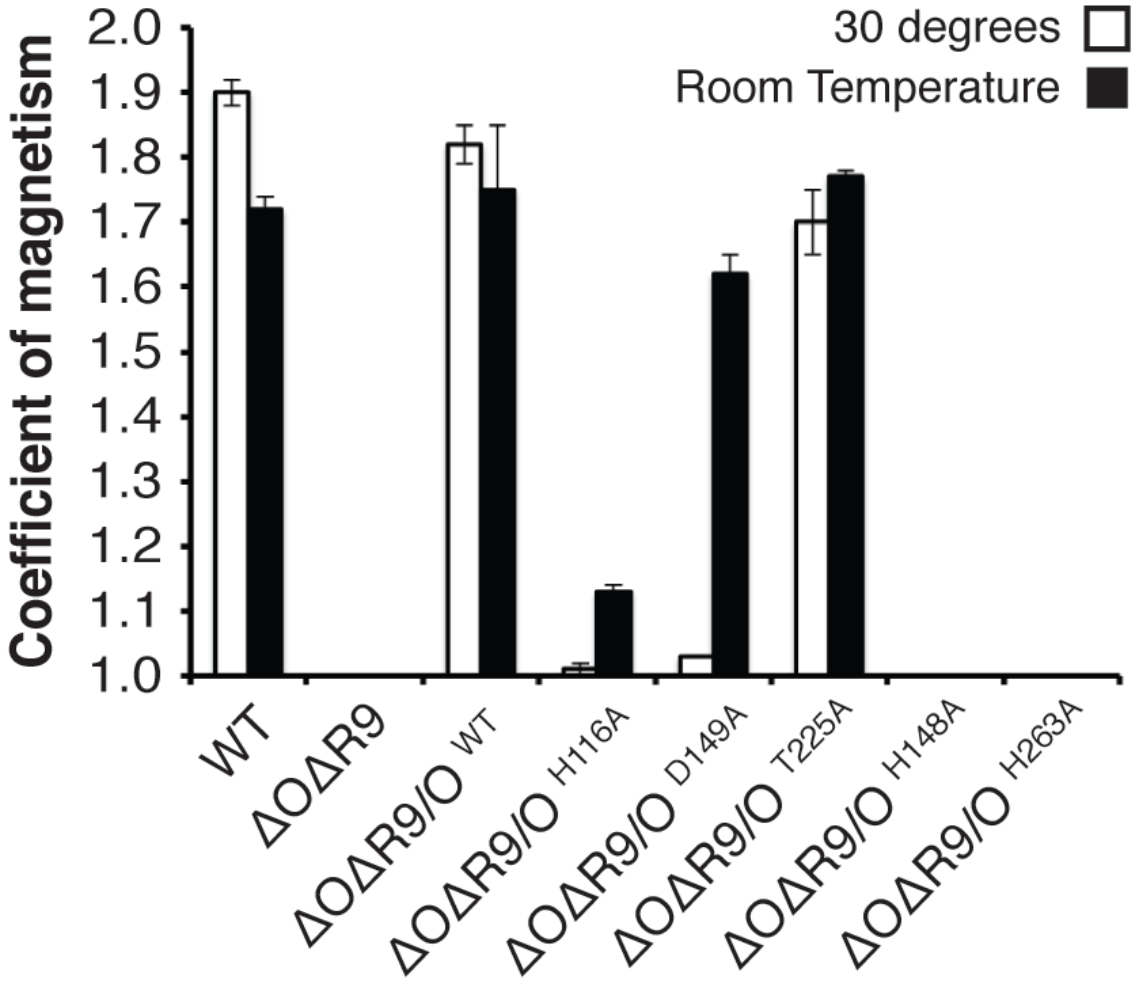


Figure 2-3. Temperature dependence of magnetic response for mamO alleles in this study.

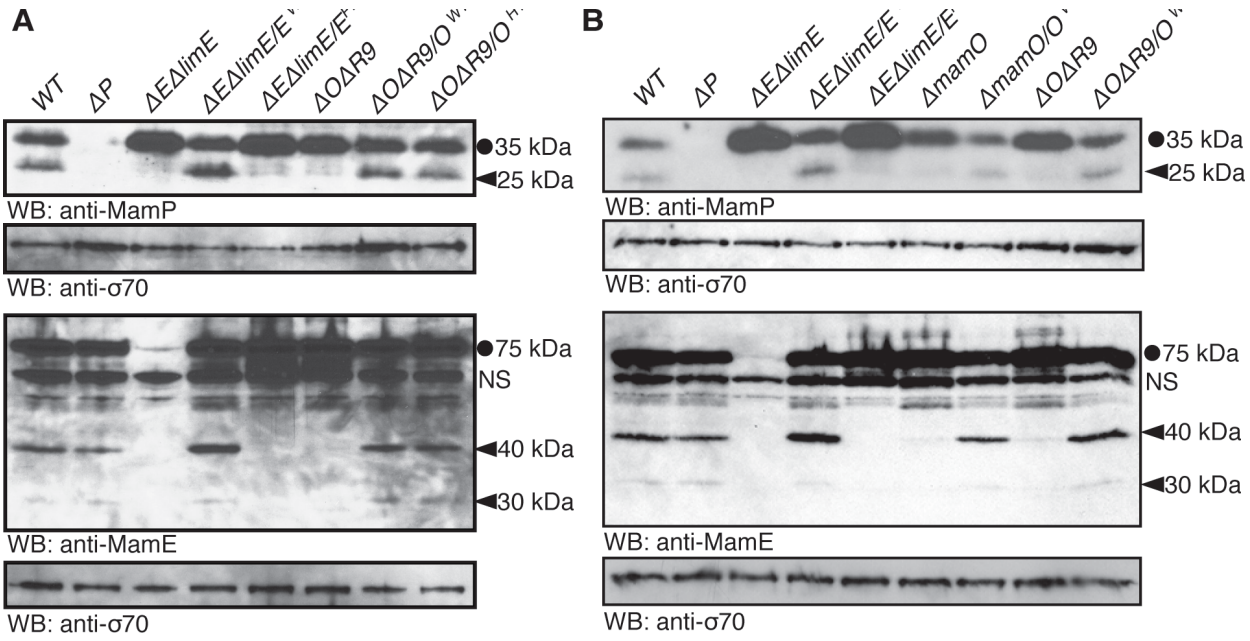


Figure 2-4. Genetic requirements for proteolytic processing of biomineralization factors. Full-length proteins are marked with a circle and proteolytic fragments with a carat. mamE^{PD} refers to the previously described allele with all three catalytic triad residues mutated to alanine. A nonspecific interaction with the anti-MamE antibody is marked with "NS". (**A**) Proteolysis of MamE and MamP depends on both MamE and MamO, The MamE active site is required, but the MamO active site is dispensable. (**B**) Efficient proteolysis of MamE and MamP requires the TauE domain of MamO. In both the ΔOΔR9 strain and the ΔmamO strain (which contains limO) proteolysis of the two targets is minimal.

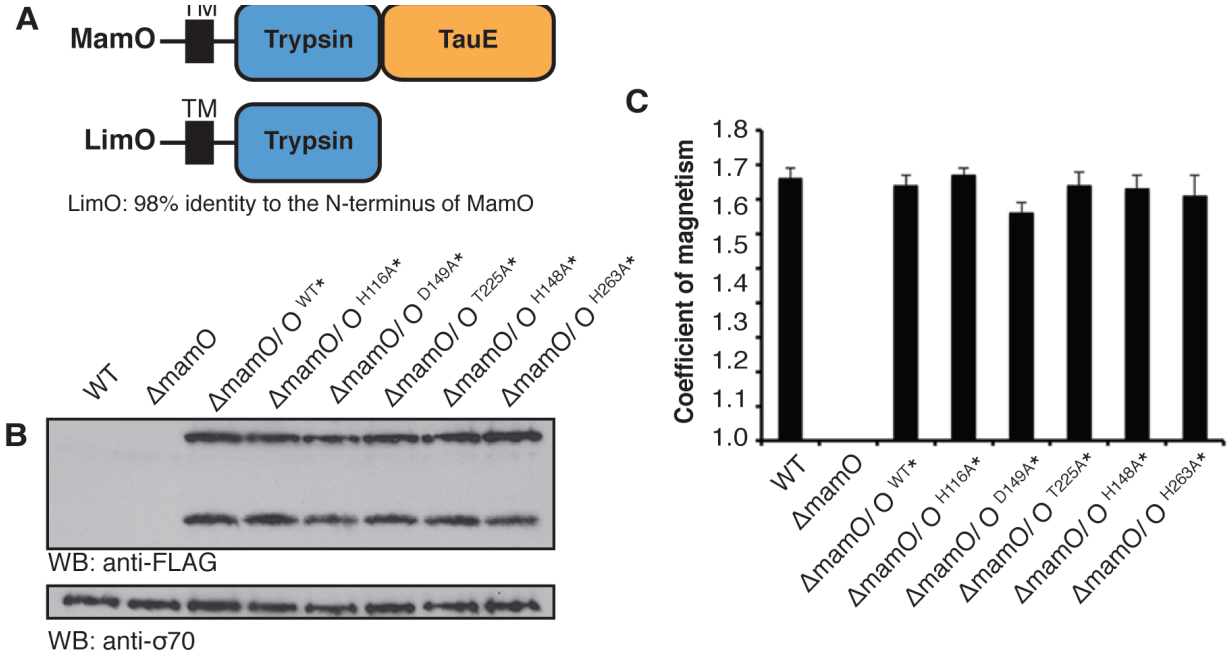


Figure 2-6. Cross-complementation of MamO protease domain mutants by *LimO*. (a) *limO* is contained within a partially duplicated region of the *mamAB* cluster termed R9. While *mamO* is predicted to have a trypsin like-protease domain and a TauE-like transporter domain, *limO* has only a predicted trypsin-like domain with 98% identity to the N-terminus of *mamO*. *LimO* contains all of the critical residues identified in MamO in this study. (b) MamO alleles are proteolytically processed identically in the single Δ*mamO* strain as they are in the Δ*O*Δ*R9* background. (c) All of the *mamO* alleles examined in this work restore wild-type biomineralization in the single Δ*mamO* background, showing that *limO* encodes a fully functional copy of the *mamO* protease domain.

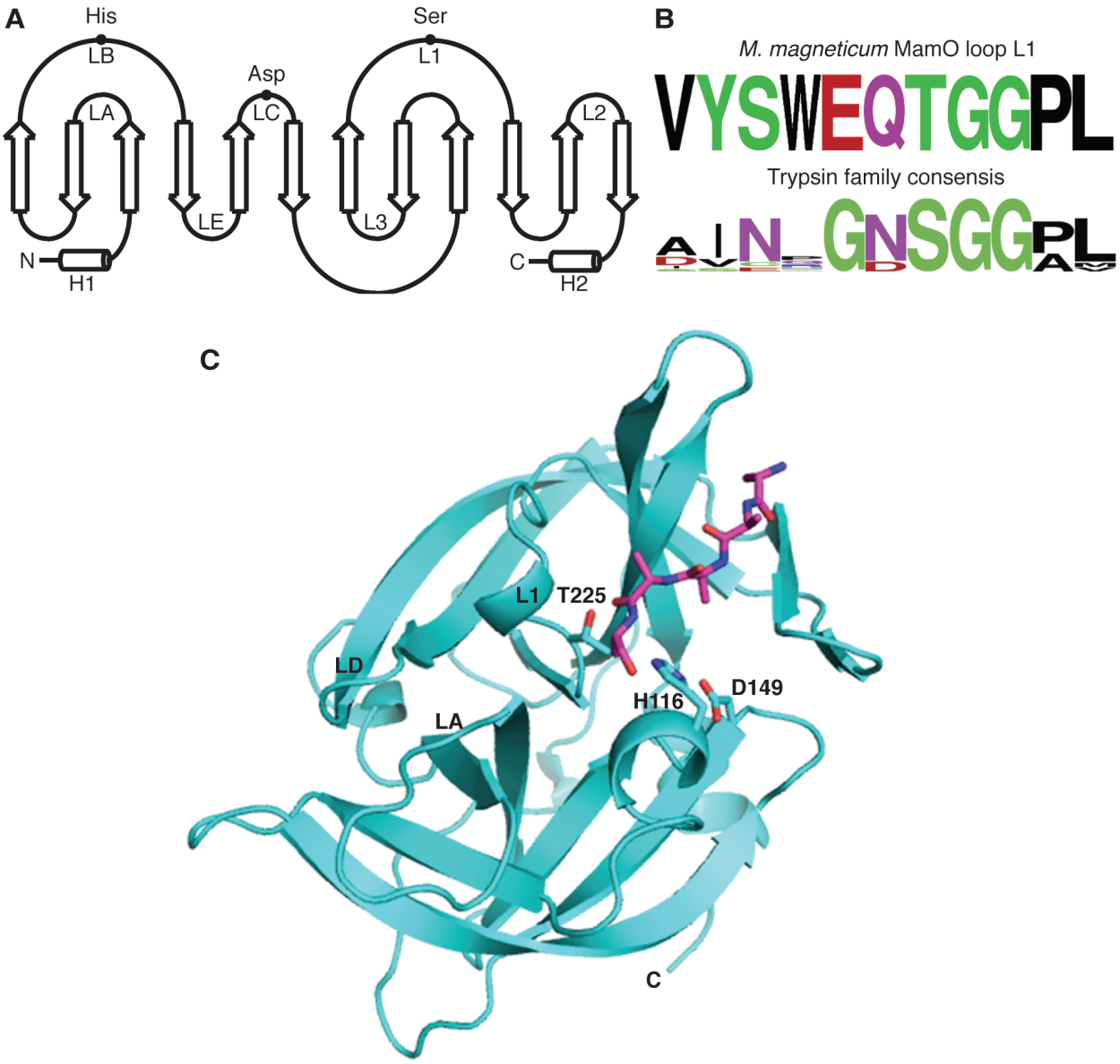


Figure 2-7. Structure of the MamO protease domain. (a) Schematic of the chymotrypsin fold with the loops and catalytic residues indicated. (b) Comparison of the L1 loop in MamO to the trypsin family consensus. (c) Overall structure of the MamO protease domain solved to 2.6Å. The catalytic residues and bound peptides are show in stick representation.

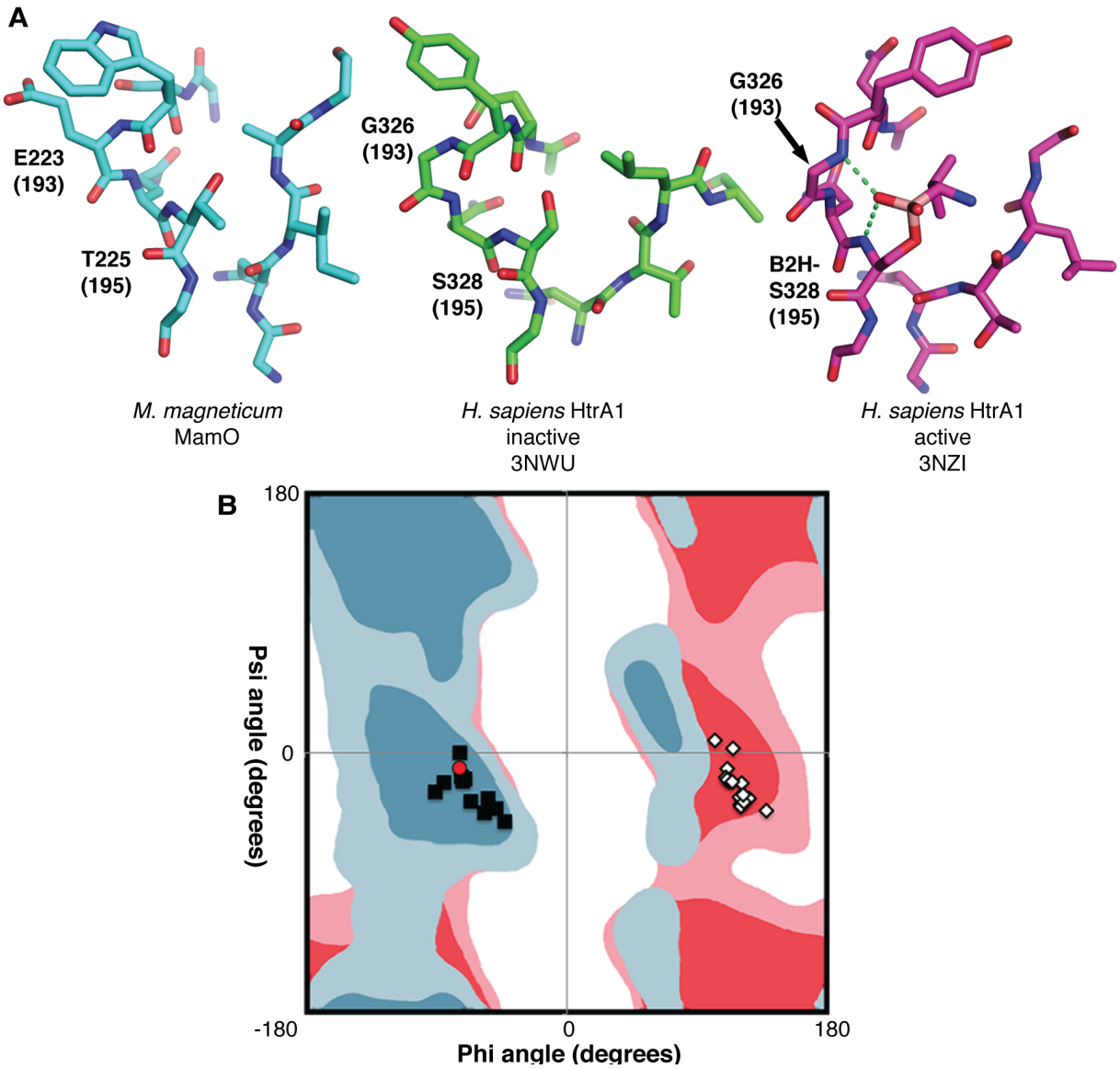


Figure 2-8. Steric constraints on the MamO active site. (a) Comparison of active site structures, showing MamO in the inactive conformation. Each residue's chymotrypsin numbering position is in parentheses. Dashes represent hydrogen bonds contributed by the oxyanion hole. (b) Ramachandran plot showing favored (dark shades) and allowed (light shades) geometries for non-glycine residues (blue) and glycine (red). The configuration at residue 193 for a set of trypsin-like structures are plotted. Black squares: inactive conformation; white diamonds: active conformation; red circle: MamO. The active conformation is disallowed for MamO.

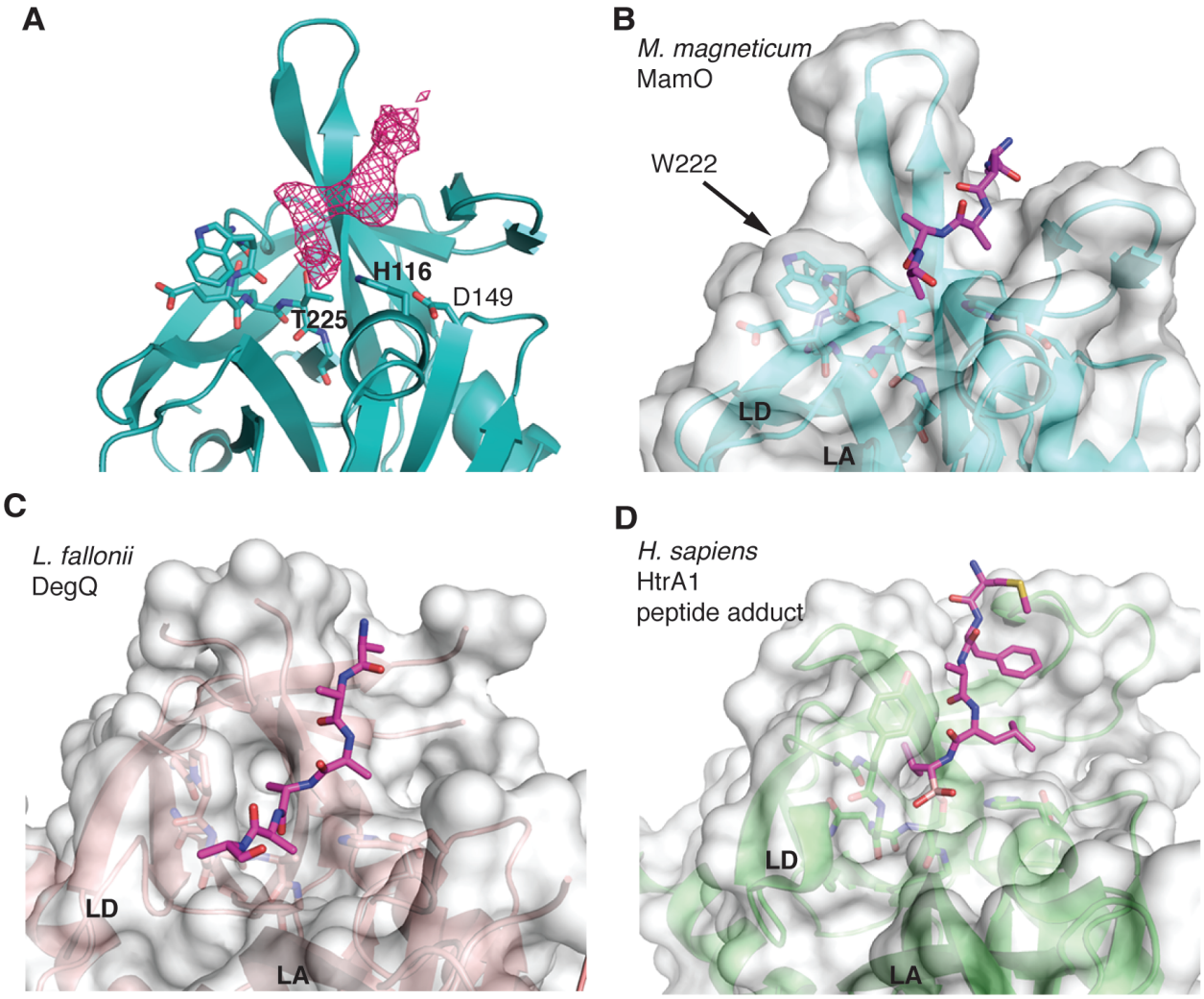


Figure 2-9. W222 in MamO forces the bound peptide away from the active site. (a) *Fo – Fc* omit map contoured at 3σ showing missing density for the peptide in the MamO crystal structure. (bd) Comparison of peptide binding pockets in HtrA proteases. W222 in MamO blocks the normal exit path between loops LA and LD and pushes the peptide away from the active site. PDB codes for each panel are as follows b: 5HM9; c: 3PV3; d: 3NZI. Loops LA and LD are marked.

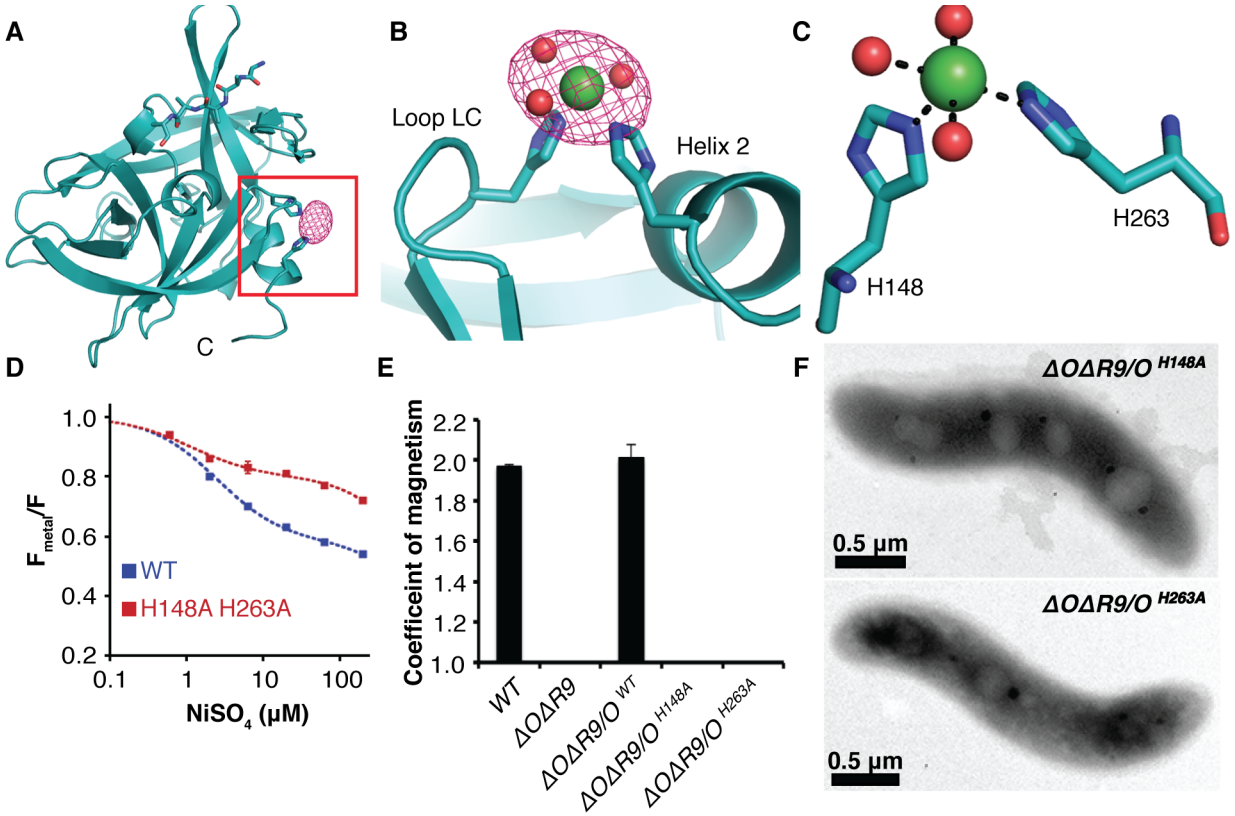


Figure 2-10. Characterization of the MamO metal binding site. (a) Anomalous map contoured at 5σ showing the placement of Ni2+ in MamO. (b) Ni2+ ion bound between loop LD and helix 2. (c) H148, H263, and three ordered water molecules participate in metal binding. (d) tmFRET analysis of Ni2+ binding by MamO labeled at Q258C. Each measurement represents the average from four replicates. Error bars represent the standard deviation of the replicates. (e) Magnetic response of strains with disrupted metal coordination sites. Error bars represent the standard deviation from three cultures. Each measurement represents the average of three biological replicates. Error bars represent the standard deviation of the replicates. (f) TEM analysis showing that the *mamOH*^{148}*A* and *mamOH*^{263}*A* strains lack detectable minerals.

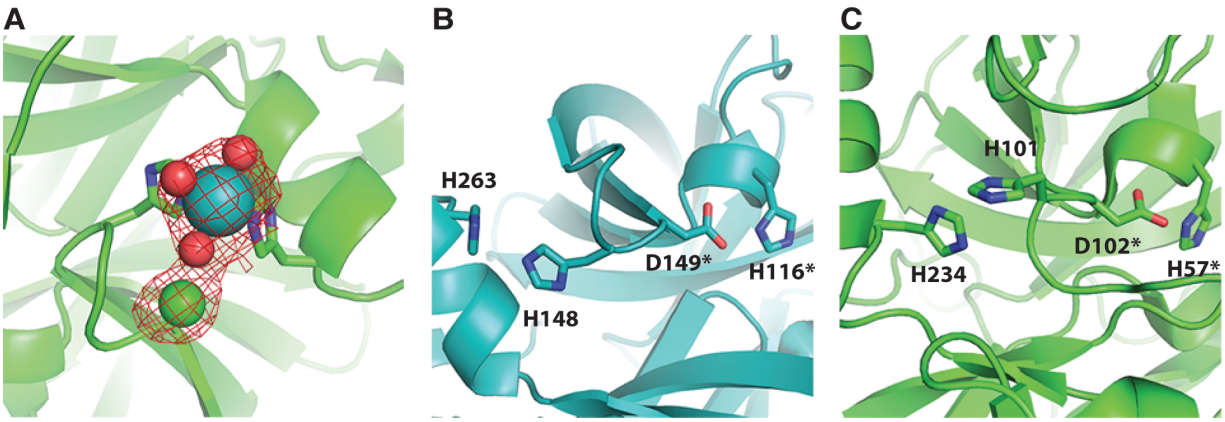


Figure 2-11. Similarity of the MamO metal binding site to equine kallikrein-3. (a) F_O-F_C omit map showing the bound Ni^{2+} ion in MamO. Blue: Ni^{2+}; red: H_2O; green: Cl^-. (b, c) Comparison of metal binding sites in MamO and equine kallikrein-3. Coordinates of the zinc-bound structure reported in Carvahlo et al.(47) were not deposited in the PDB.

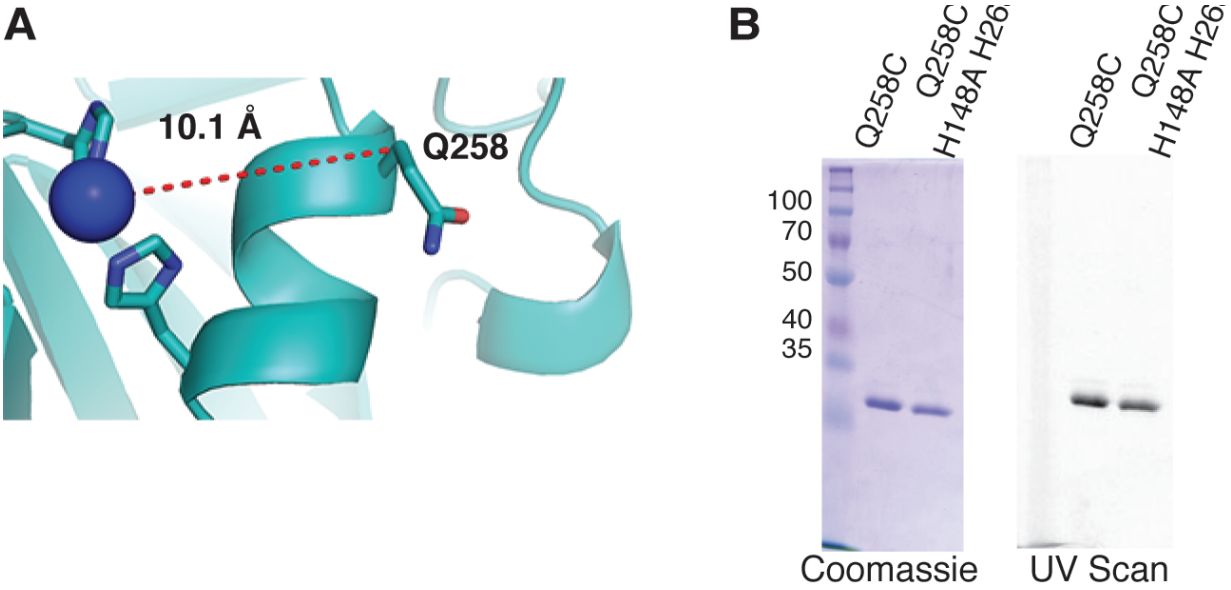
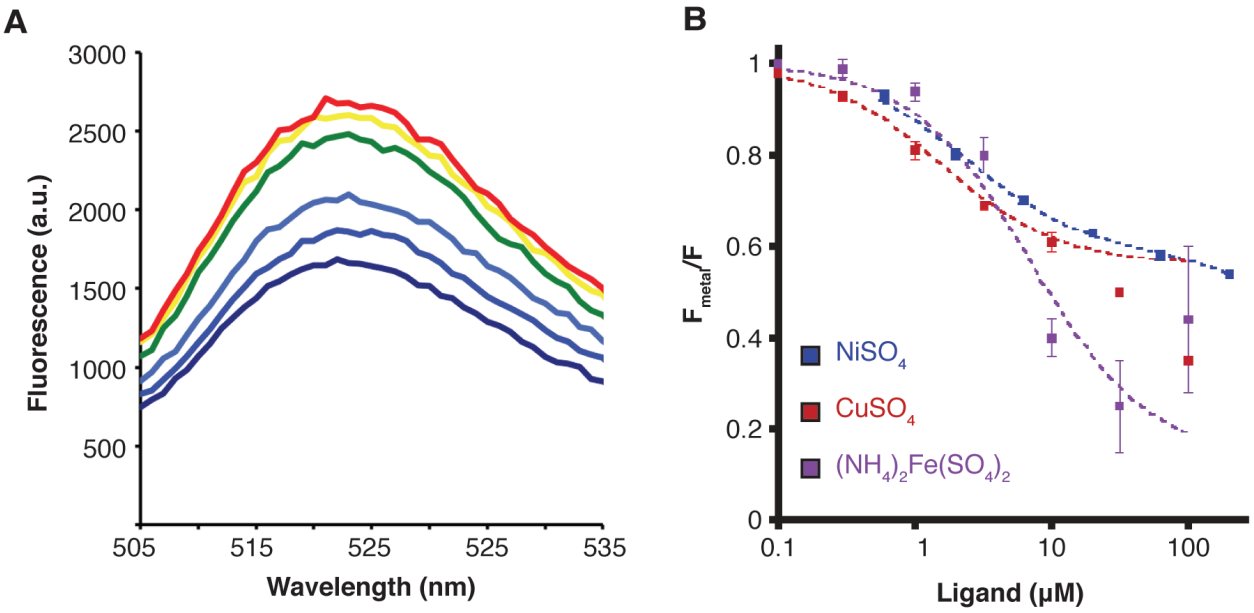
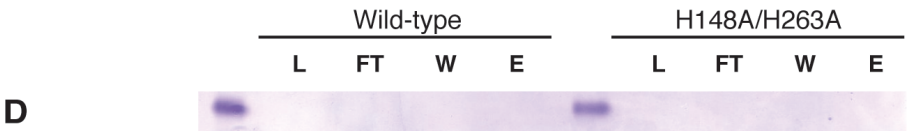


Figure 2-12 Fluorescein-5-maleimide labeling of MamO^{Q258C}. (a) The fluorescent labeling site in MamO was chosen based on the optimum FRET distance from Taraska et al. (49) (b) Purification and fluorescent labeling of MamO^{Q258C} and MamO$^{Q258C\ H148A\ H263A}$.



C

Ligand	Protein	Kd_2(μM)	E_{FRET}(%)	Kd_1(mM)
$NiSO_4$	Q258C	2.5 ± 0.3	41.6 ± 1.7	2.2 ± 1.1
$(NH_4)_2Fe(II)(SO_4)_2$	Q258C	7.1 ± 0.0	86.6 ± 21.0	2.53×10^{14}
$CuSO_4$	Q258C	1.6 ± 0.0	44.0 ± 21.0	1.63×10^{21}
$NiSO_4$	Q258C H148A H263A	1.1 ± 0.2	20.0 ± 0.1	1.7 ± 0.3



D

Figure 2-13. tmFRET analysis of metal binding. (a) Fluorescence quenching of MamO Q258C labeled with fluorescein-5-maleimide in the presence of increasing concentrations of NiSO4. (b) Binding of various transition metals to labeled MamO. Error bars represent the standard deviation from four independent measurements. The dotted lines are fits to the binding equation described in Materials and Methods. (c) Binding constants from tmFRET experiments. (d) NiNTA affinity assays with purified protease domains.

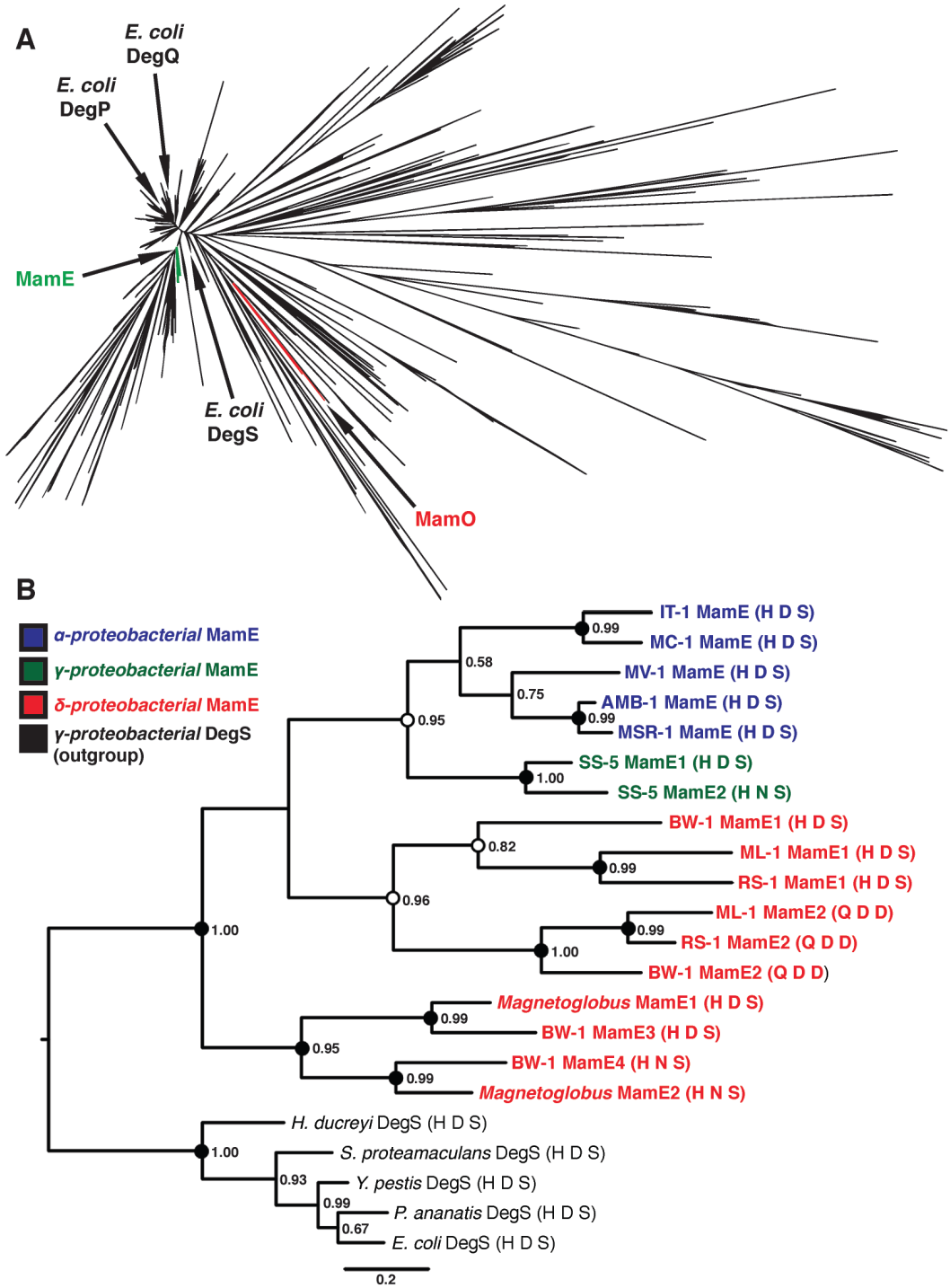


Figure 2-14. Phylogenetic analysis of magnetotactic trypsin-like proteins. (a) Phylogeny of the "Deg" branch created from the trimmed Trypsin-2 alignment. MamE and MamO clades and the three E. coli HtrA proteases are marked. (b) Phylogeny of the MamE clade of HtrA proteases. Numbers represent the posterior probability determined by PhyloBayes. Circles represent the degree of support from 300 bootstrap replicates in RAxML. Black: >90% support; white: >80% support. MamE sequences are colored based on the class of the associated organism, and the catalytic triad residues are shown in parentheses after each name.

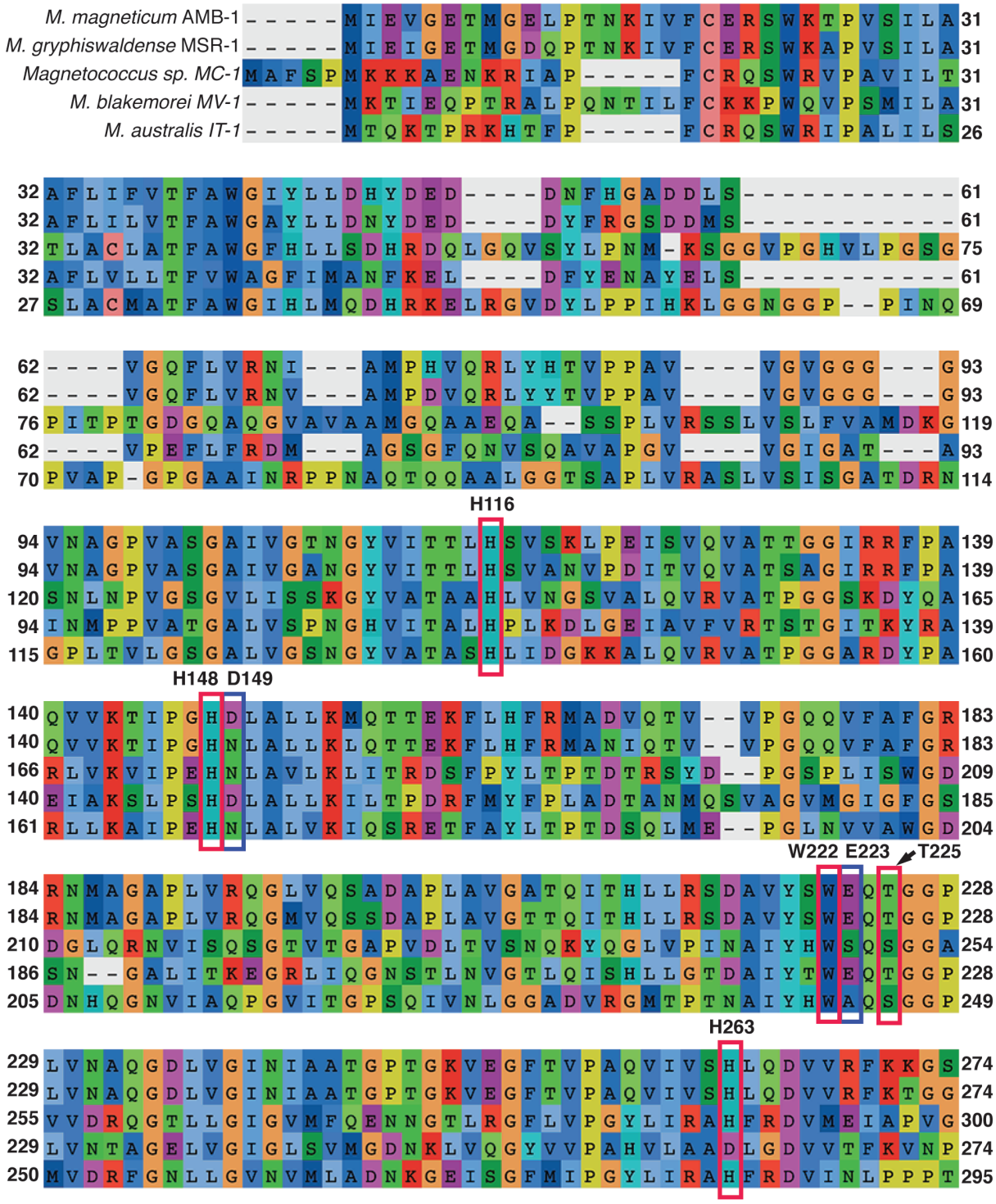


Figure 2-15. Alignment of the MamO family. The conservation of critical residues discussed in the text is indicated with colored boxes.

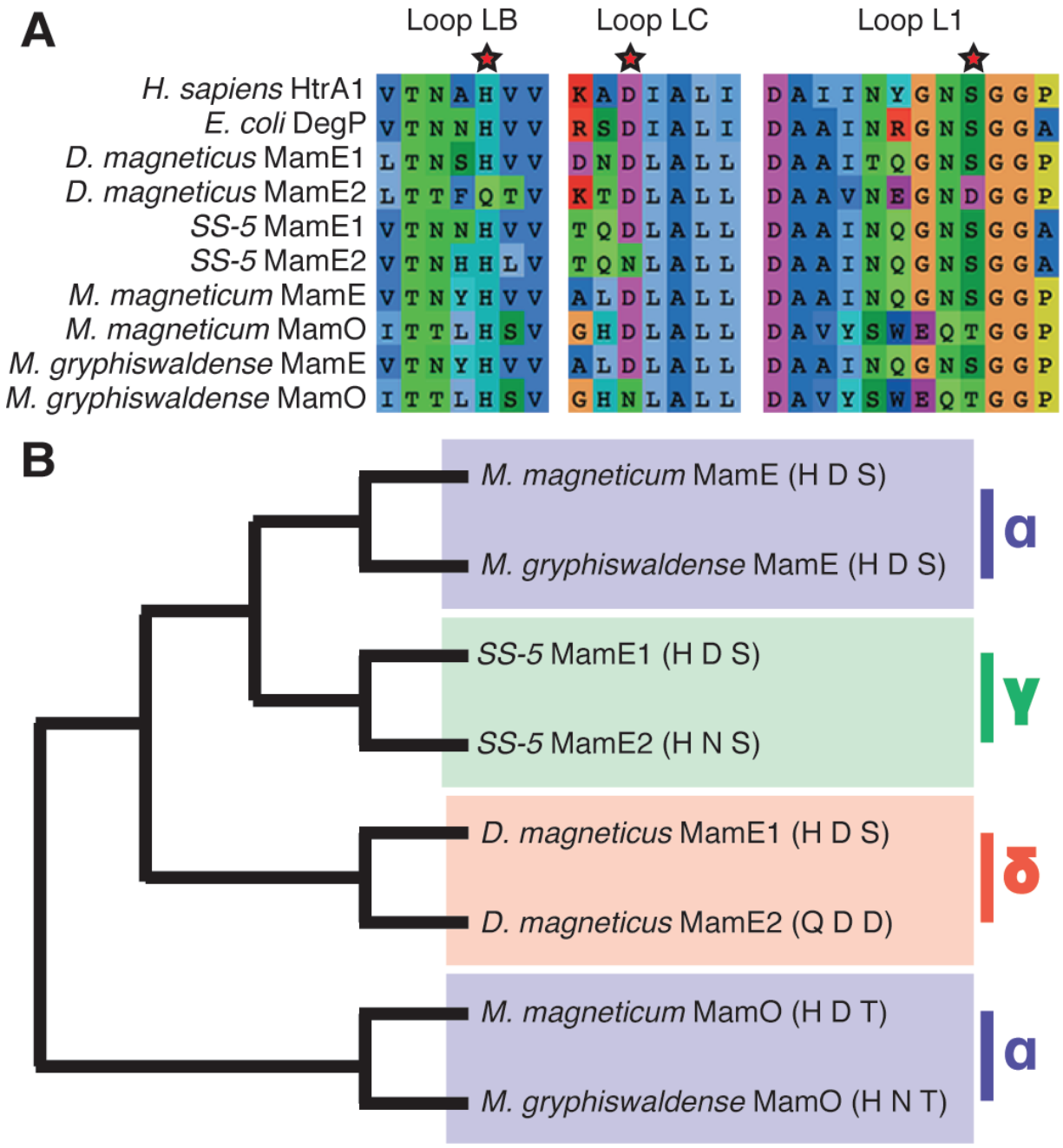


Figure 2-16. Sample analysis of representative trypsin-like sequences from magnetotactic bacteria. (a) Alignment of the catalytic loops from a set of trypsin like-sequences. The trypsin sequences from four magnetotactic organisms were aligned with two canonical HtrAs, *H. sapiens* HtrA1 and *E. coli* DegP. Positions of catalytic triad residues are marked with a star. (b) Phylogeny of the sequences inferred from the detailed analysis shown if Figure 2-14. The identities of the catalytic triad residues are shown in parentheses after each protein name. Boxes represent the class level taxonomy of the organism within the *Proteobacteria.*

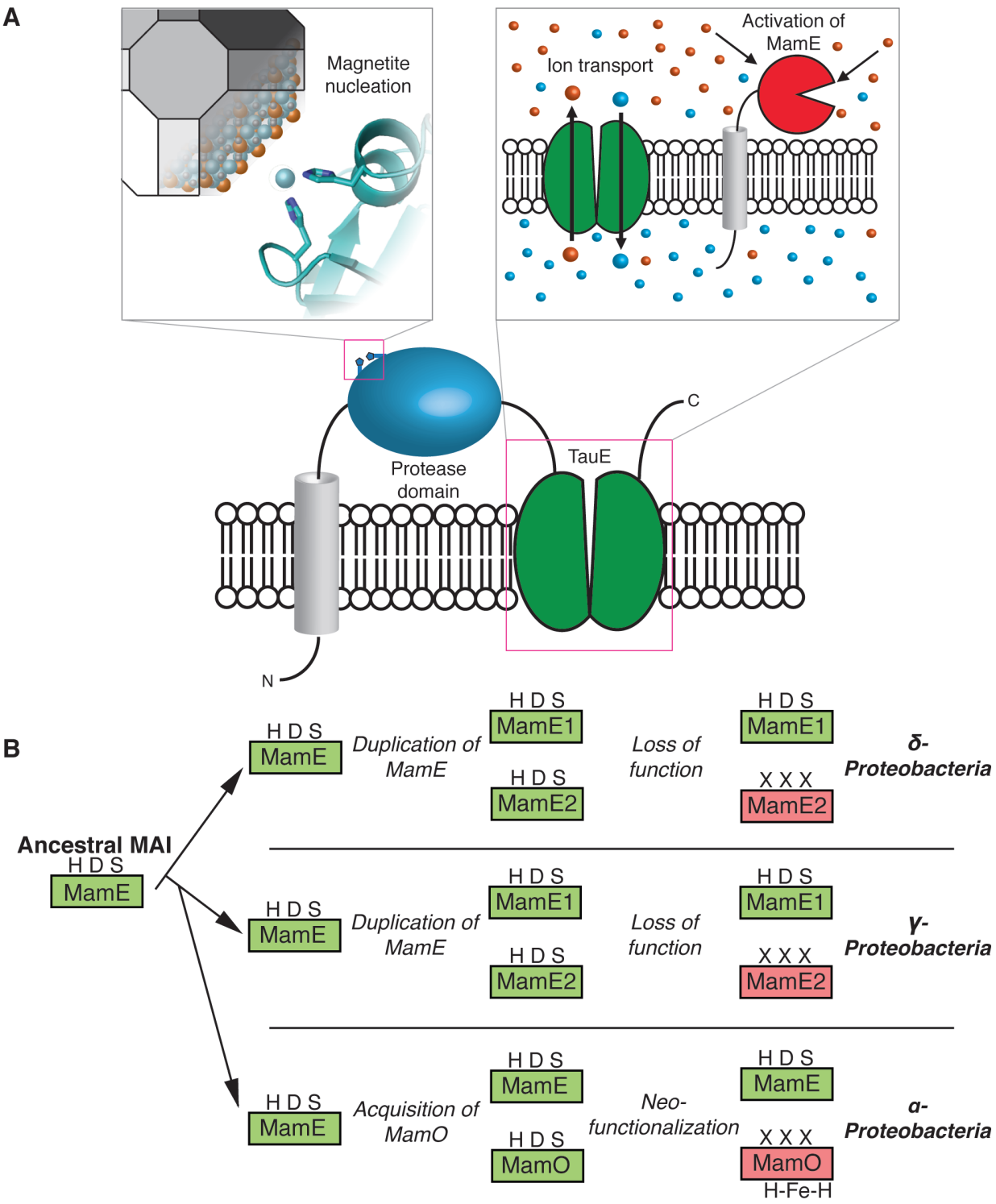


Figure 2-17. MamO in magnetosome formation and evolution. (a) A dual role in biomineralization. Distinct regions of the protein contribute to each activity separately. The protease domain promotes nucleation by binding iron and the TauE domain manipulates solute conditions that regulate MamE's activity. (b) Specialization of the trypsin-like protease family in magnetotactic bacteria through gene duplication and subsequent neofunctionalization.

Chapter 3

Magnetite biomineralization in Magnetospirillum magneticum is regulated by a switch-like behavior in the HtrA protease MamE

David M. Hershey[1] , Patrick J. Browne[1] , Anthony T. Iavarone[2,3], Joan Teyra[5] , Eun H. Lee[4] , Sachdev S. Sidhu[5] , and Arash Komeili[1,2,4]

1 Department of Plant and Microbial Biology, 2 California Institute for Quantitative Biosciences, 3 QB3/Chemistry Mass Spectrometry Facility, 4 Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA 5 Department of Molecular Genetics, Terrance Donnelly Centre for Cellular and Biomedical Research, University of Toronto, Toronto, Canada ON M5S 3E1

ABSTRACT
Magnetotactic bacteria are aquatic organisms that produce subcellular magnetic crystals in order to orient in the earth's geomagnetic field. The genetic basis for this process has been dissected in the model magnetotactic bacterium *Magnetospirillum magneticum* AMB-1, leading to the identification of biomineralization genes that control the formation and growth of magnetite crystals. One such factor, MamE, is a predicted member of the HtrA family of serine proteases, a widespread enzyme family that plays important roles in protein turnover and quality control. MamE was recently shown to promote the proteolytic processing of itself and two other biomineralization factors *in vivo*. Here, we have studied MamE-dependent proteolysis in detail. We analyzed the *in vivo* processing patterns of three proteolytic targets and used this information to reconstitute proteolysis with a purified form of MamE. MamE cleaves a custom peptide substrate with positive cooperativity, and its auto-proteolysis can be stimulated with exogenous substrates or peptides that bind to either of its PDZ domains. A constitutively active form of the protease causes biomineralization defects, showing that proper regulation of its activity is required during biomineralization *in vivo*. Our results demonstrate for the first time that MamE is a bonafide HtrA protease and that its activity is consistent with the previously proposed checkpoint model for biomineralization.

INTRODUCTION
Interactions between living organisms and insoluble inorganic compounds are an underexplored aspect of biology. Magnetotactic bacteria assemble magnetic crystals called magnetosomes into chains within their cells allowing them to passively align and navigate along magnetic fields(1, 2). Genetic analyses have shown that magnetosome formation is a complex developmental process controlled by a conserved set of genes(12, 13, 21, 60, 61). A subset of these, called biomineralization factors, whose deletions disrupt or eliminate magnetite crystal formation, have drawn increasing interest for their potential utility in biomedical applications(64). Understanding the biochemical mechanisms employed by these gene products can provide novel strategies for manipulating transition metal-based nanomaterials *in vitro*(62, 63, 65).

Two predicted trypsin-like proteases, MamE and MamO, are required to produce magnetite in the model magnetotactic organism, *Magnetospirillum magneticum* AMB-1(12). Disrupting either gene abolishes the formation of magnetite crystals without affecting the development of their surrounding membrane compartment, indicating that both proteins are required for magnetite nucleation within the magnetosome lumen(12, 20, 24). Furthermore, deleting *mamE* causes a number of proteins that normally localize at the magnetosomes to become dispersed throughout the cytoplasmic membrane. Adding either wild-type or catalytically inactive (E^{PD}) alleles of *mamE* complements the protein localization defect, but only the wild-type allele restores normal magnetite biosynthesis(25).

Cells with the E^{PD} allele show an intermediate biomineralization phenotype in which they produce small magnetite particles. While wild-type AMB-1 has a distribution of crystal sizes centered at 50-60nm in diameter, the size distribution in the E^{PD} cells is centered at ~20nm. Interestingly, 97% of the crystals in the E^{PD} strain are smaller than 35nm, the point above which magnetite particles become paramagnetic and can hold a stable magnetic dipole. The correlation between mineral sizes in the E^{PD} strain and the superparamagnetic to paramagnetic transition point lead to the speculation that MamE's putative protease activity regulates the transition to a

magnetotactic lifestyle. This so-called checkpoint model predicts that cells produce small superparamagnetic crystals until an unknown signal activates MamE, promoting maturation to paramagnetic particles(25).

MamE is a member of the HtrA/DegP family of trypsin-like proteases, a ubiquitous family of enzymes that controls various aspects of protein quality control(23). The family is characterized by domain structures consisting of an N-terminal trypsin-like domain and one or two C-terminal postsynaptic density 95/discs large/zonula occludens-1 (PDZ) domains(28, 66, 67). Like many proteases, intricate regulatory mechanisms are a hallmark of HtrA catalysis(43, 68). Based on structural and mechanistic investigations, the PDZ domains play several roles in regulating proteolysis including promoting assembly and activating the protease domain by binding to extended peptide motifs(44, 69, 70, 71). MamE has an unusual domain structure in which a tandem *c*-type cytochrome motif has been inserted between the protease and two PDZ domains (Fig. 1A)(25). Unfortunately, studies aimed at understanding the catalytic activity of MamE and its regulation has been hindered by an inability to obtain recombinant protein.

Recently, MamE was found to promote the *in vivo* proteolytic processing of itself, MamO, and another biomineralization factor named MamP in a manner that required the predicted MamE active site(72, 126, 73). Although MamO was also required for these proteolytic events, this effect did not require the predicted MamO active site. Subsequent structural analysis showed that MamO's protease domain was locked in an inactive state and incapable of catalysis, suggesting that it played a non-catalytic role in activating MamE(73). Despite the genetic evidence for MamE's role in proteolysis, its activity has not been confirmed directly using purified components. Here, we have characterized MamE-dependent proteolysis in detail and identified a number of regulatory mechanisms. Developing a method to purify MamE and analyzing *in vivo* proteolytic patterns of each target facilitated reconstitution of MamE-dependent proteolysis *in vitro*. Detailed analysis of its catalytic activity suggests a switch-like model in which the basal state of the protein is an inactive form that can be turned on through a number of routes. A constitutively active allele of MamE disrupts biomineralization, confirming that properly regulated proteolysis is critical to magnetosome formation.

MATERIALS & METHODS

Strains, plasmids and growth conditions

The strains and plasmids used in this study are described in Tables 3-1 and 3-2, respectively. AMB-1 was maintained in MG medium supplemented with kanamycin when necessary as previously described (12). Magnetic response was measured using the coefficient of magnetism as previously described (12). Standard molecular biology techniques were used for plasmid manipulation. *E. coli* strains were grown in LB medium supplemented with appropriate antibiotics. Plasmids were maintained in *E. coli* strain DH5α *λpir*. *E. coli* strain WM3064 was used for plasmid conjugations as described previously(12).

Immunoblotting

Whole-cell lysates of AMB-1 strains were prepared from 10mL cultures and analyzed as described previously(73). For analysis of the auto-cleavage reaction products, a 1 in 10 dilution of each time-point was separated on a 12% acrylamide gel for immunoblotting. The MamE and MamP antibodies have been described previously(73). The anti-6xHis (Sigma), anti-FLAG

(Sigma), anti-σ70 (Thermo Fisher) and anti-strep (Qiagen) antibodies were purchased from commercial sources.

Fractionation of MamO fragments
A strain with the genetic background Δ*O*Δ*R9*/*FLAG-O* was cultured without shaking at 30 °C in 2 L screw-capped flasks that were filled to the top with MG medium. The cells were harvested by spinning at 5k x g for 15 min, re-suspending in cold 20mL 25mM Tris-HCl pH7.4, respinning at 8k x g for 10 min and freezing the resulting pellet at -80°C until use. Cell pellets were thawed on ice and re-suspended in 5mL lysis buffer A (10mM Tris-HCl pH 8.0, 50mM NaCl, 1mM EDTA). Pepstatin A and leupeptin were added to a final concentration of 2µg/mL and PMSF was added to 2mM. Lysozyme was added from 50mg/mL stock to a final concentration of 0.5mg/mL and the cells were incubated at room temperature for 15 min. 15mL of lysis buffer B (20mM HEPES-KOH pH7.5, 50mM NaCl, 1.25mM CaCl2) was added along with DTT to 2mM and DNAse to 5µg/mL and the suspension was incubated for 15min at 4°C with agitation. The cells were sonicated twice for 10 seconds and the suspension was spun at 8k x g for 10min to isolate the magnetite-associated material.

The resulting pellet was re-suspended in 5.5 mL solubilization buffer (20mM BisTris-HCl pH7.0, 75mM NaCl, 10% glycerol) and CHAPS was added to 1% from a 10% stock solution. The suspension was incubated at room temperature for 15min with agitation followed by an incubation for 15min at 4°C with agitation. The suspension was spun at 16k x g for 15min. The resulting pellet was re-suspended in 5.5 mL solubilization buffer and the detergent extraction was repeated with 1% FosCholine. The FosCholine-soluble material was loaded on a 1mL HiTrap Q FF column (GE Healthcare) that had been equilibrated with solubilization buffer containing 0.03% *n*-dodecyl β-D-maltoside (DDM). The column was washed with 10mL of solubilization buffer with 0.03% DDM and eluted with 4mL of buffer Q1 (20mM BisTris-HCl pH7.0, 275mM NaCl, 10% glycerol) with 0.03% DDM followed by 4mL of buffer Q2 (20mM BisTris-HCl pH7.0, 400mM NaCl, 10% glycerol) with 0.03% DDM. The Q1 fraction was added to 50µL anti-FLAG M2 resin (Sigma) and incubated at 4°C with agitation for 3 hrs. The resin was isolated by spinning at 4k x g and washed with sequential 1mL washes of buffer Q2, buffer Q1 and solubilization buffer each containing 0.03% DDM. Bound proteins were eluted by 3 washes with 50µL of 0.2M glycine pH 2.8, which were pooled with 50µL of 1M Tris-HCl pH8.0.

Preparation of trypsin digests for liquid chromatography-mass spectrometry (LC-MS)
The concentrated FLAG elution fraction was separated on a 12% acrylamide gel and stained with colloidal Coomassie Blue. A ~3 x 10mm section of the gel corresponding to the processed MamO band was excised from the gel and chopped into small pieces. These were washed with 100mM NH4HCO3 followed by reduction and alkylation of cysteines with DTT and iodoacetamide. The gel pieces were then dehydrated by washing with increasing concentrations of acetonitrile in 100mM NH4HCO3 and dried under vacuum. A 0.1mg/mL solution of trypsin was used to re-swell the gel pieces, and they were incubated overnight at 37°C. The resulting peptides were extracted from the gel slices with successive washes of 0.1% formic acid solutions containing increasing concentrations of acetonitrile. The extracts were pooled in a fresh tube, concentrated under vacuum to remove the organic phase and stored at 4°C until analysis.

LC-MS

Trypsin-digested protein samples were analyzed using a Thermo Dionex UltiMate3000 RSLCnano liquid chromatograph that was connected in-line with an LTQ-Orbitrap-XL mass spectrometer equipped with a nanoelectrospray ionization (nanoESI) source (Thermo Fisher Scientific, Waltham, MA). The LC was equipped with a C18 analytical column (Acclaim® PepMap RSLC, 150 mm length Å~ 0.075 mm inner diameter, 2 μm particles, 100 Å pores, Thermo) and a 1-μL sample loop. Acetonitrile (Fisher Optima grade, 99.9%), formic acid (1-mL ampules, 99+%, Thermo Pierce), and water purified to a resistivity of 18.2 MΩÅEcm (at 25 °C) using a Milli-Q Gradient ultrapure water purification system (Millipore, Billerica, MA) were used to prepare mobile phase solvents. Solvent A was 99.9% water/0.1% formic acid and solvent B was 99.9% acetonitrile/0.1% formic acid (v/v). The elution program consisted of isocratic flow at 2% B for 4 min, a linear gradient to 30% B over 38 min, isocratic flow at 95% B for 6 min, and isocratic flow at 2% B for 12 min, at a flow rate of 300 nL/min.

Full-scan mass spectra were acquired in the positive ion mode over the range m/z = 350 to 1800 using the Orbitrap mass analyzer, in profile format, with a mass resolution setting of 60,000 (at m/z = 400, measured at full width at half-maximum peak height, FWHM), which provided isotopic resolution for singly and multiply charged peptide ions. Thus, an ion's mass and charge could be determined independently, i.e., the charge state was determined from the reciprocal of the spacing between adjacent isotope peaks in the m/z spectrum. In the data-dependent mode, the eight most intense ions exceeding an intensity threshold of 50,000 counts were selected from each full-scan mass spectrum for tandem mass spectrometry (MS/MS) analysis using collision-induced dissociation (CID). MS/MS spectra were acquired using the linear ion trap, in centroid format, with the following parameters: isolation width 3 m/z units, normalized collision energy 30%, default charge state 3+, activation Q 0.25, and activation time 30 ms. Real-time charge state screening was enabled to exclude unassigned and 1+ charge states from MS/MS analysis. Real-time dynamic exclusion was enabled to preclude re-selection of previously analyzed precursor ions, with the following parameters: repeat count 2, repeat duration 10 s, exclusion list size 500, exclusion duration 90 s, and exclusion mass width 20 ppm. Data acquisition was controlled using Xcalibur software (version 2.0.7, Thermo). Raw data were searched against the *Magnetospirillum magneticum* AMB-1 FASTA protein database using Proteome Discoverer software (version 1.3, SEQUEST algorithm, Thermo). Peptide identifications were validated by manual inspection of MS/MS spectra, i.e., to check for the presence of y-type and b-type fragment ions1 that identify the peptide sequences(74).

Expression and purification of MamE

pAK825 or pAK964 was transferred to C43 cells (Lucigen) that had been previously transformed with the pEC86 heme-loading plasmid(75). The transformed cells were maintained at 30°C due to toxicity of the construct at 37°C. An overnight liquid culture was inoculated into 600mL 2xYT medium supplemented with the appropriate antibiotics. The cultures were grown at 30°C until the OD600 reached ~0.5, at which point the culture was transferred to 20°C. After a 30 min equilibration, the culture was induced with 35μM IPTG. Expression was performed for 12.5-13 hrs at 20°C with shaking at 200 rpm.

Cells were harvested by immediately chilling the cultures on ice and spinning at 6k x g for 10 min. The resulting pellet was re-suspended in 50mL of cold osmotic shock buffer (50mM

NaPhosphate pH8.0, 1mM EDTA, 20% sucrose). Leupeptin (1.5µg/mL), pepstatin A (1.5µg/mL) and lysozyme (0.5mg/mL) were added, and the suspension was rocked at room temperature for 15 min. An equal volume of ice-cold H2O was added and the suspension was rocked on ice for 15 min before spinning at 8k x g for 10 min to remove debris.

The resulting supernatant was added to 3 mL Ni-NTA resin (Qiagen) and supplemented with NaCl (150mM), DNAseI (5µg/mL), NP-40 (0.1%) and MgCl2 (2.5mM). The slurry was rocked at 4°C for 30 min and the beads were allowed to settle. After decanting the upper phase, the slurry was poured into a column, washed with 10 column volumes of Ni wash buffer (25mM Tris-HCl pH7.4, 250mM NaCl, 10mM imidazole, 10% glycerol) and the bound proteins eluted with Ni elution buffer (25mM Tris-HCl pH7.4, 250mM NaCl, 250mM imidazole, 10% glycerol). The Ni-NTA eluent was loaded onto a 1mL StrepTrap column (GE Healthcare), which was then washed with 5mL strep wash buffer (25mM Tris-HCl pH7.4, 250mM NaCl, 10% glycerol). Bound proteins were eluted in strep wash buffer containing 2.5mM desthiobiotin. The purified protein was concentrated in 50kDa cutoff ultrafilter while simultaneously removing the desthiobiotin by repeated dilution and concentration with strep wash buffer. The concentration was determined by the Bradford method using bovine serum albumin to prepare a standard curve. Aliquots were frozen in liquid N2 and stored at -80°C until use.

Analysis of MamO1 peptide cleavage
A custom peptide with the sequence 5-carboxymethylfluorescein-Thr-Gln-Thr-Val-Ala-Ala-Gly-Ser-Lys(CPQ2)-D-Arg-D-Arg was obtained commercially (CPC Scientific). The peptide was dissolved in DMSO and stored at -20°C. 5X substrate solutions with various concentrations of the MamO1 peptide were prepared in assay buffer (50mM Tris-HCl pH8.0) containing 0.05% NP-40 and 1.6% DMSO. To initiate the reaction, 10µL samples of the substrate mix were added to 40µL of MamE protein solution that had been diluted to 125nM in assay buffer in a 96-well plate. The fluorescence was scanned (excitation: 485nm; emission 538nm) every 5 min for 2 hrs in a Tecan plate reader.

The slope was determined from the linear portion of each reaction. Cleavage rates were calculated by making a standard curve from a MamO1 cleavage reaction that had been incubated for 24hrs to allow for complete hydrolysis. Specific activities were determined by normalizing these cleavage rates to the enzyme concentration. Rates were plotted as a function of peptide concentration and fit to the Hill form of the Michaelis-Menten equation using the Kaleidagraph software package:

Analysis of MamE auto-proteolysis
25µL reactions were prepared by adding 1µL of activating peptide dissolved in DMSO at the appropriate concentration to 24µL of MamE diluted to 2µM in assay buffer. The reactions were incubated at 30°C, and 8µL aliquots were removed at the appropriate times. Each aliquot was quenched immediately by mixing with SDS sample buffer. Equal volumes of each aliquot were separated on a 12% acrylamide SDS-PAGE gel and stained with Coomassie Blue to visualize the processing pattern.

Expression and purification of PDZ domains
For all three PDZ domain constructs (EP1, EP2 and EP12), the appropriate plasmids for

expression as N-terminal 6xHis-MBP-TEV fusions were transformed into BL21 Codon Plus cells. Cultures were grown in 2xYT at 37°C until the OD600 reached ~0.8 at which point they were transferred to 20°C for 30min followed by induction with 0.1mM IPTG and expression overnight. The cells were harvested by centrifugation, resuspended in resuspension buffer (25mM Tris-HCl pH7.4, 800mM NaCl, 10mM imidazole, 10% glycerol) and frozen at -80°C until use.

For protein purification, the cells were thawed on ice and sonicated for three 30-second cycles. The lysate was clarified by spinning at 13k x g for 30 min. The resulting supernatant was loaded on a 3mL Ni-NTA column that had been equilibrated in resuspension buffer. After washing with 10 column volumes of resuspension buffer and 10 column volumes of wash buffer 2 (25mM Tris-HCl pH7.4, 400mM NaCl, 25mM imidazole, 10% glycerol), bound proteins were eluted with Ni elution buffer.

For the purification of the EP1 and EP2 proteins, the elution fractions were dialyzed overnight against AEX buffer A (25mM BisTris-HCl pH7.0, 75mM NaCl and 10% glycerol). The desalted protein was passed through a 1mL HiTrap QFF column (GE Healthcare) and the flow-through was concentrated in a 50kDa cutoff ultrafilter, injected onto a16/60 Superdex 200 column and developed in Storage Buffer. Each protein eluted as a single symmetrical peak. The peak fractions were concentrated in a 50kDa ultrafilter and small aliquots were frozen in liquid N2 and stored at -80°C for use in the phage display experiments.

For the purification of the EP12 protein, the elution fraction was dialyzed overnight against digest buffer (50mM NaPhosphate pH8.0, 75mM NaCl, 5mM imidazole, 10% glycerol) in the presence of 6xHis tagged TEV protease to remove the 6xHis-MBP tag. The resulting sample was passed through a 3mL Ni-NTA column that had been equilibrated in digest buffer. The flow-through fraction was concentrated in a 10kDa ultrafilter, passed through a 1mL HiTrap SP FF column and concentrated again before injection on a16/60 Superdex 200 column that was developed in storage buffer. The protein eluted as a single symmetrical peak. The peak fractions were concentrated in a 10kDa ultrafilter and small aliquots were frozen in liquid N2 and stored at -80°C for use in the fluorescence anisotropy experiments.

Phage display
C-terminally and N-terminally displayed peptide libraries were used to assess the peptide binding preferences of MamE PDZ1 and PDZ2. The C-terminal peptide library consisted of random decapeptides constructed using 10 consecutive NNK degenerate codons encoding for all 20 natural amino acids and fused to the C terminus of a mutant M13 bacteriophage major coat protein (2 Å~ 1010 unique members)(64, 75). The N-terminal peptide library consisted of random hexadecapeptides constructed using 16 consecutive mixes of 19 codon trimers (cysteine and STOP codons were excluded) and fused to the N terminus phage coat protein (2.4 Å~ 1011 unique members)(76).

Fluorescence anisotropy
Peptides with the following sequences were synthesized commercially with a fluorescein-aminocaproic acid group fused to the N-terminus: WSQEMEDWFWQMPLSG (PDZ1*) and MEDYGIFMTSPEGPWA (PDZ2*). Each peptide was diluted to a concentration of 40nM in

25mM Tris-HCl pH7.4 containing 0.25mg/mL bovine serine albumin. A dilution series of EP12 protein was prepared in storage buffer. 6μL of the ligand solution was added to 18μL of the appropriate protein solution in a 384-well plate, and the mixture was allowed to equilibrate at room temperature for 15 min. Polarization measurements were made at 535nm using a Perkin Elmer Victor 3V 1420 plate reader. The resulting anisotropy values were plotted as a function of protein concentration and fit to a single site binding model using the Kaleidagraph software package:

RESULTS

Mapping MamE-dependent cleavage patterns in vivo

To learn more about the context of MamE-dependent proteolysis, we attempted to map the cleavage patterns of MamE, MamP and MamO *in vivo* using an epitope tagging approach. Strains with N- or C-terminally 3xFLAG-tagged alleles of each gene were added to their respective deletion strains. Each of the tagged alleles complemented the biomineralization defects of the deletions with the exception of the C-terminally tagged *mamO* allele (*O-FLAG*). Analysis of cell extracts from these strains by Western blotting was used to assess the processing patterns for each proteolytic target.

For MamE, a number of N-terminal proteolytic fragments but no C-terminal fragments are observed (Figure 3-1B), indicating that short segments are sequentially removed from the C-terminus. Interestingly, this seems to culminate in a stable ~45kDa fragment that corresponds to the MamE protease domain and the N-terminal membrane helix separated from the rest of protein. For MamP, one N-terminal band and two C-terminal bands are observed that correspond to a full-length protein and a protein truncated by approximately 10kDa at the N-terminus, indicating the removal of the membrane anchor from the predicted soluble domains (Figure 3-1B). For MamO, a full-length and a shorter band are observed for both the N- and C-terminally tagged proteins (Figure 3-1B). The pattern predicts that MamE-dependent proteolysis separates the N-terminal protease and C-terminal TauE domains of MamO (Figure 3-1A).

In many MamE and MamO blots, there are small (i.e. ~20 kDa) bands near the bottom of the gel when analyzing the N-terminally tagged proteins. Although these bands appear in many experiments, the levels and even the presence of these two signals are inconsistent. This potentially suggests that short segments are removed from the N-terminus of each protein to produce unstable fragments that are quickly degraded. Due to the inconsistency in these bands we have not focused on them in our analysis of the processing pattern.

Identification of a putative cleavage motif in MamO

Analysis of the *in vivo* processing patterns strongly indicated that MamE processes itself, MamO and MamP(73). We reasoned that one strategy for reconstituting its proteolytic activity *in vitro* could be to design substrates from motifs that are cleaved in a MamE-dependent fashion. Based on epitope tagging, MamO is cleaved between the protease and TauE domains, which suggests that there is a MamE-dependent cleavage site at the mature C-terminus of the protease domain (Figure 3-2B). Cell pellets from 2L cultures of the Δ*O*Δ*R9*/*FLAG-O* strain were used for biochemical fractionation. Enzymatic lysis and sonication followed by a low speed (8000 x g) spin were used to isolate material associated with the dense magnetite particles. A number of detergents were tested for their ability to dissolve the MamO fragments. Most classes of

detergents are ineffective or only partially effective in the initial solubilization step. Lipid-like zwitterionic detergents including lauryldimethylamine oxide (LDAO) and FosCholine-12 extract the fragments from the membrane, but they disrupt binding to the α-FLAG affinity resin. However, once the initial extraction step is complete, the detergent requirements to maintain solubility become less stringent.

Based on the solubility information, the low speed pellet was pre-washed with CHAPS before extracting the MamO fragments from the membranes with FosCholine-12. In order to facilitate binding to the affinity resin, the FosCholine soluble material was loaded on an anion exchange column and exchanged to the detergent DDM by extensive washing before eluting with salt. This fraction was then used as the input for an α-FLAG affinity isolation to yield a final fraction enriched in N-terminal fragments of MamO (Figure 3-2A and B).

The concentrated α-FLAG elution was separated on an SDS-PAGE gel, stained with colloidal Coomassie Blue and the region around 35kDa was excised. After performing in-gel trypsin digestion, peptides were extracted and concentrated for LC-MS/MS analysis. A number of peptides from the MamO sequence are consistently detected, and, as expected, they map almost exclusively to the protease domain of MamO (Figure 3-2C). In all of the samples, the protein sequence coverage drops off sharply in the linker between the protease and TauE domains (Figure 3-2C). A peptide with the sequence GSATAPGQPQTQTV is routinely detected at the C-terminal edge of the peptide coverage (Figure 3-2D). This peptide results from a predicted tryptic cleavage on the N-terminus but has a non-tryptic C-terminus, which suggests that it contains the C-terminal sequence of the mature MamO protease domain.

Purification of MamE

In order to reconstitute its proteolytic activity *in vitro*, we developed a method to express and purify appreciable amounts of MamE. In the expression construct, the N-terminal membrane anchor is replaced with the OmpA signal peptide to produce a soluble protein that can still undergo heme loading in the periplasm(78). Initial fractionations with an N-terminally 6xHis-tagged form of the protein had significant contamination due to what appeared to be truncated fragments caused by auto-cleavage during expression. To eliminate this problem, a strep tag was added on the C-terminus to allow for a sequential affinity isolation of full-length protein. Finally, MamE has a predicted region of 60-70 disordered residues downstream of the N-terminal membrane anchor and upstream of the trypsin-like domain. Removing this region dramatically improved the solubility.

The expression conditions also had to be carefully optimized in order for MamE to accumulate. Maximum expression occurs with low levels of IPTG induction in C43 cells at 20 °C. The precise details of the expression and purification procedure are described in the Materials and Methods. It allows for the soluble region of MamE as well as a catalytically inactive form (MamES297A) to be purified with yields approaching ~1mg/L. Importantly, both preparations appear red and display absorbance spectra characteristic of *c*-type cytochromes (Figure 3-3).

Direct proteolysis of the MamO cleavage motif

We designed a fluorogenic peptide (peptide MamO1) to test as a substrate for purified MamE. The substrate contains 8 residues that make up the putative cleavage motif identified in MamO

(Figure 3-2) flanked by an upstream fluorophore and a downstream fluorescence quencher. Normally, the peptide has low fluorescence due to interaction between the fluorophore and quencher. If the peptide is cleaved, the two fragments will separate, resulting in an increase in fluorescence. Upon addition of the O1 peptide to purified MamE, there is a linear increase in fluorescence. Importantly, the $MamE^{S297A}$ protein does not alter the fluorescence indicating the signal is due to serine protease activity from MamE (Figure 3-4A). MamE hydrolyzes the O1 peptide with a k_{cat} of 0.64 ± 0.03 min^{-1} and a K_M of 6.1 ± 0.5 μM. As for other HtrA proteases, the reaction is positively cooperative, displaying a Hill coefficient of 1.5 ± 0.1 (Figure 3-4B) (43, 79). These values are similar to those reported for cleavage of peptide substrate by other trypsin-like proteases and confirm that MamE can efficiently cleave the motif identified in MamO (80). Combined with the *in vivo* analysis these results confirm that MamO is a direct proteolytic target of MamE.

Reconstitution of MamE auto-proteolysis

Analysis of MamE processing in AMB-1 along with the extensive auto-proteolysis during its expression in *E. coli* suggested that MamE is capable of auto-proteolysis. However, purified MamE is relatively stable such that after an hour of incubation at 30°C, nearly all of the protein remains intact (Figure 3-5A). The positive cooperativity observed for the steady-state kinetics of O1 peptide cleavage indicated that MamE's catalytic activity could be stimulated by substrates. This mode of regulation might also lead to peptide-induced activation of auto-cleavage. Indeed, the MamO1 peptide stimulates degradation of full-length MamE in a dose-dependent manner, confirming that MamE's activity can be stimulated by the presence of substrate (Figure 3-5A).

Taking advantage of the distinct tags used to purify MamE, we examined the auto-cleavage fragmentation pattern by Western blotting. Numerous truncated proteins are detected by blotting for the N-terminal 6xHis tag, the smallest of which is a ~27 kDa fragment presumed to be the protease domain (Figure 3-5B). In contrast, only the full-length protein can be seen when blotting for the C-terminal strep tag (Figure 3-5C). The pattern indicates that the reaction proceeds via sequential removal of small fragments from the C-terminus. Furthermore, it matches the pattern seen by examining epitope-tagged alleles of MamE expressed *in vivo* and confirms the successful reconstitution MamE-dependent proteolysis *in vitro.*

Activation through the PDZ domains

PDZ domains of other HtrA proteases regulate proteolytic activity by binding to extended peptide motifs(68). Phage display has been a productive approach for identifying peptide ligands that bind to PDZ domains(81-84). Each of the MamE PDZ domains was purified and used as bait in phage display selections. Both bait proteins showed phage enrichment for specific particles with a library displaying peptides on the N-terminus of the coat protein, but no enrichment was observed in libraries displaying C-terminal fusions. This suggests that, unlike those associated with other HtrA proteases, the MamE PDZ domains do not display a preference for C-terminal peptides(85). Interestingly, phage selections for both domains showed a strong preference for internal regions (Tables 3-3 and 3-4). However, a single clone dominated both pools making the identification of consensus motifs hard to interpret.

Peptides corresponding to the sequence that dominated each selection were synthesized and labeled with a fluorescent dye. In addition, the C-terminal region of MamE containing only the PDZ domains (EP12) was purified and used to test for direct binding to the phage-derived

ligands. Fluorescence anisotropy experiments demonstrate that both phage-derived peptides bind to the C-terminus of MamE (Figure 3-6A). The PDZ1 peptide shows ~10-fold tighter binding than the PDZ2 peptide, but both affinities are comparable to those seen for other PDZ domains(84). Addition of either peptide to full-length MamE, results in a dose-dependent activation of auto-processing (Figure 3-6B). Importantly, the activation threshold for the PDZ2 peptide is higher than the PDZ1 peptide, mirroring the equilibrium binding data. These results show that, like other HtrA proteases, MamE's protease activity can be regulated through peptide binding to its PDZ domains.

Misregulation of MamE disrupts biomineralization in vivo

MamE's low basal activity and its stimulation by various peptides define a switch-like behavior that toggles a high activity and a low activity state of the protein. This *in vitro* behavior suggests that modulation of the two states is important during biomineralization *in vivo.* To date, we have not been able to identify growth conditions that prevent *in vivo* processing of MamE, MamO or MamP. Thus, we utilized a genetic approach to examine whether proper regulation of MamE-dependent proteolysis was required for biomineralization.

Work with the model HtrA protease DegS has shown that mutating residue 192 (chymotrypsin numbering) in the oxyanion hole to proline increases basal cleavage rates by shifting the allosteric equilibrium toward the active state (42, 43). We introduced an allele with the corresponding mutation of MamE (Q294P) into the *mamE* null strain (Figure 3-7A). This strain displayed significantly less MamE and MamP as assessed by western blotting, indicating increased proteolytic activity from MamE. Additionally, the $mamE^{Q294}P$ allele partially circumvents the requirement for MamO in promoting MamE-dependent proteolysis. MamE and MamP appeared to be processed when the $mamE^{Q294}P$ allele was introduced into a strain lacking both *mamE* and *mamO*, though it did not restore processing of MamE to wild-type levels (Figure 3-7B and C). These results indicate that $mamE^{Q294}P$ produces a constitutively active, unregulated protease *in vivo*.

We next examined this allele for its ability to complement the biomineralization defects seen in the *mamE* null strain. While wild-type $mamE^{WT}$ allele completely restores the magnetic response, the $mamE^{Q294P}$ strain has a significantly lower response (Figure 3-7D). Thus, both the inactive ($mamE^{PD}$) and constitutively active forms ($mamE^{Q294P}$) of the protease disrupt biomineralization *in vivo*. The magnetic response of the $mamE^{Q294P}$ strain is higher than the negligible signal measured in the $mamE^{PD}$ strain, indicating that the biomineralization process is stalled at a different stage when MamE's activity is unregulated. These results demonstrate that complete biomineralization of magnetosome crystals relies not only on the occurrence of MamE-dependentproteolysis but also careful regulation of the activity.

DISCUSSION

Biochemical principles underlying the biomineralization of magnetite by magnetotactic bacteria represent a model for understanding how biological molecules manipulate inorganic compounds (4). Recent advances describing the genetic basis for this process have paved the way for mechanistic studies of the factors that promote mineral formation (12). The HtrA protease MamE has emerged as a central biomineralization factor in the model organism *Magnetospirillum magneticum* AMB-1. In addition to promoting crystal nucleation and protein

localization to the magnetosome, MamE regulates a transition from small superparamagnetic crystals to full-sized paramagnetic particles. This crystal maturation phenotype was linked to MamE's putative protease activity suggesting a model where its catalytic activity controls crystal maturation (25).

Here, we have studied MamE's serine protease activity in detail. We mapped the proteolytic patterns of three *in vivo* targets both at the domain level and, in one case, at the individual residue level. Using this information, we reconstituted a number of aspects of MamE-dependent proteolysis with purified components. We show that MamE directly cleaves a motif from the linker between MamO's protease and TauE domains in a positively cooperative fashion. Furthermore, we show that purified MamE has low basal activity, but that it can be activated in a number of ways including the presence of substrate and peptide binding to either of its PDZ domains. This behavior is consistent with a switch-like mode of regulation in which the protease requires activation by environmental cues. These results also show for the first time that MamE is a serine protease that is capable of degrading itself and other biomineralization factors.

Nearly all studied members of the HtrA family behave as trimers or multiples thereof(28, 66, 44 67, 86, 87). In other systems with two PDZ domains, the first PDZ seems to regulate protease activity directly while the second is thought to mediate rearrangements of core trimers into higher order oligomers(44, 69, 71). Peptide binding to either the first or the second PDZ domain of MamE can activate proteolysis, although the activation through PDZ2 is much weaker. Furthermore, the protein behaves as a monomer as indicated by gel filtration. Transitions between a monomer and higher order assemblies are rare in the HtrA family, but the positive cooperativity observed for MamE suggests that the active form is indeed a larger assembly(88). The protein production method described here should enable future structural studies aimed at understanding the potential for novel assembly behavior as well as an unusual regulatory role for the second PDZ domain.

The switch like activation of MamE's activity suggested that allosteric regulation of its protease activity was required for proper crystal maturation. Using a mutation reported to stabilize active forms of other HtrA proteases, we showed that, like the catalytically inactive form, a constitutively active form of MamE had defects in crystal maturation(43). These results confirm that both the active and inactive states are important during the process of magnetosome formation. Similar experiments with DegP in *E. coli* indicated that the proper balance between active and inactive forms is required for fitness during heat stress(89). Our results show that, in addition to maintaining fitness during stress, this mode of regulation can be used to control a developmental process.

Taken together, our results are consistent with the checkpoint model for MamE-dependent proteolysis in regulating the maturation of magnetite crystals. However, the detailed mechanism by which that activity promotes crystal growth remains unclear. One possible mechanism could be by controlling the size of the surrounding membrane. A recent study demonstrated a link between the growth of magnetosome membrane compartments and growth of the magnetite crystals within. This finding led to the proposal that there is a checkpoint regulating a second stage of membrane growth after the onset of biomineralization(90). Although *mamE* deletions have intact magnetosome membranes, the sizes of the membranes have not been quantified. It is

tempting to speculate that proteolysis controls this switch by linking membrane growth to crystal growth. In this scenario, crystal nucleation in the *EPD* cells would not lead to membrane growth, while the *EQP* cells would initiate membrane growth before crystals had grown sufficiently, leading to stunted particles in both cases (Figure 3-8).

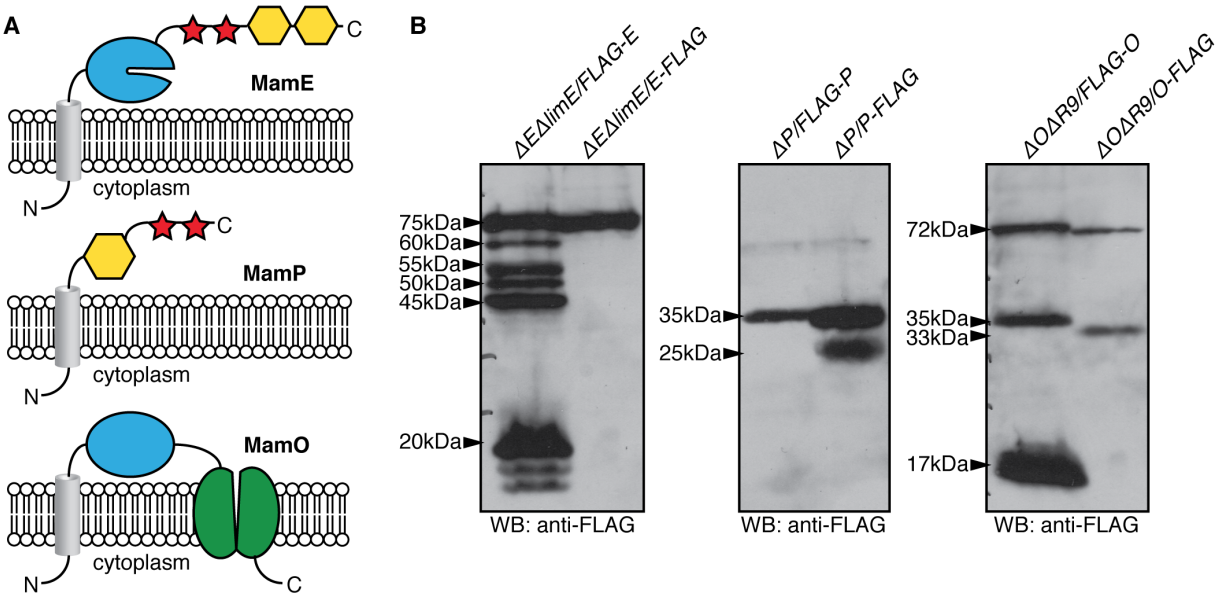


Figure 3-1. *In vivo* proteolytic processing of MamE, MamP and MamO. (a) Predicted domain structures of the three proteolytic targets. Grey cylinder: transmembrane helix; blue: trypsin-like domain; red: *c*-type cytochrome; yellow: PDZ domain; green: TauE domain. (b) Proteolytic processing patterns observed through epitope tagging. The inferred sizes of each fragment are indicated. The fragments were observed in each of at least four independent experiments with the exception of the 20kDa N-terminal fragment of MamE and the 17kDa N-terminal fragment of MamO, which varied dramatically between experiments.

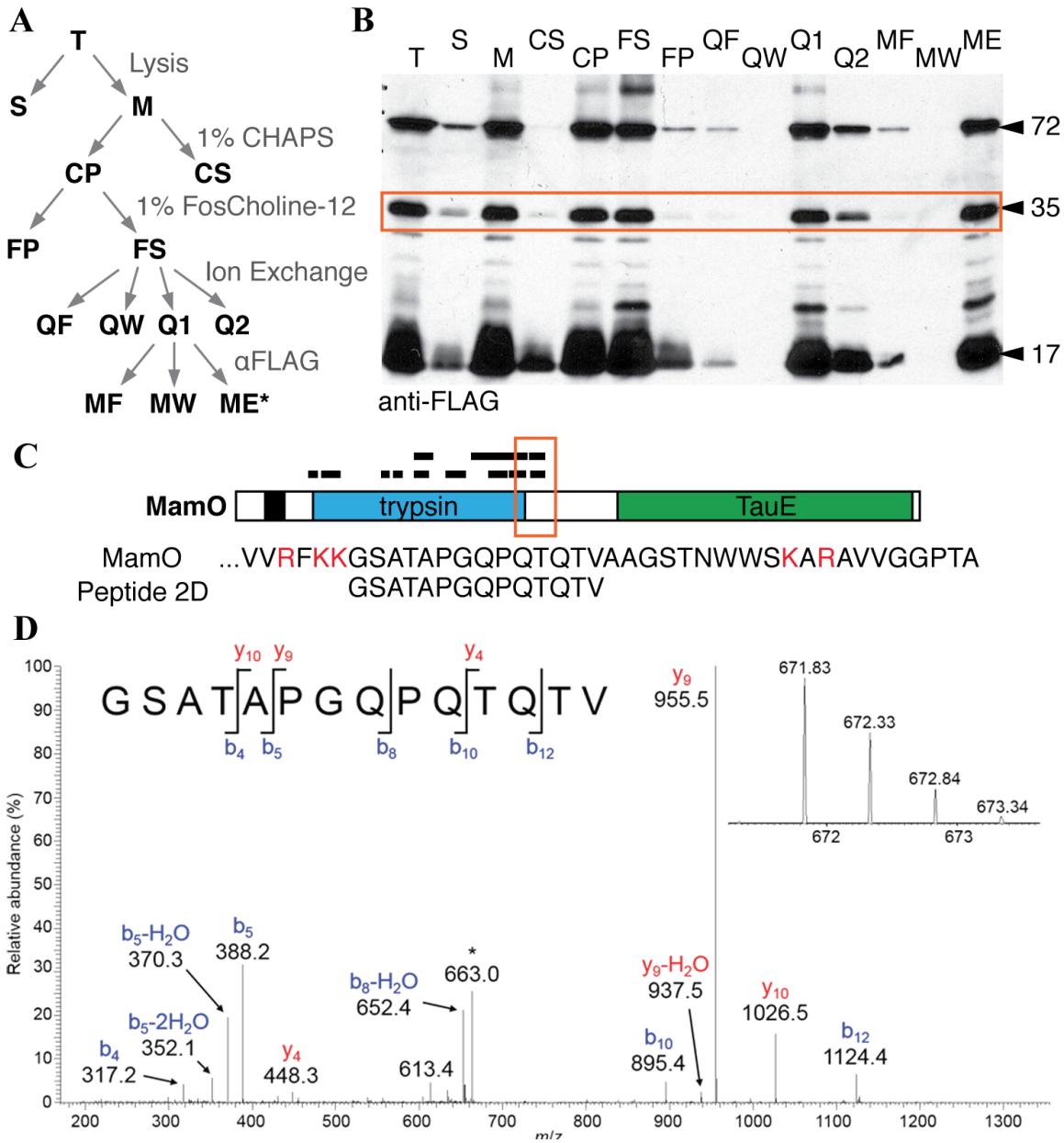


Figure 3-2. Biochemical fractionation to enrich N-terminal MamO fragments. (a) Schematic of the enrichment procedure. (b) Western blot of each fraction from Panel a. The predicted protease domain fragment in marked with a red box. (c) Peptides identified in a representative LCMS/MS analysis. The red letters in the MamO sequence represent predicted tryptic cleavage sites. The coverage pattern is characteristic of analyses for three separate preparations. (d) Tandem mass spectrum from collision-induced dissociation of the [M+2H]2+ ion of the peptide, GSATAPGQPQTQTV, corresponding to amino acid residues 273-286 of MamO. The inset shows detail for the isotopically resolved, unfragmented peptide precursor ion. The fragment ion at m/z = 663 (denoted by the asterisk) is due to precursor ion that has undergone neutral loss of a molecule of water. This peptide was detected in each of three biological replicate experiments.

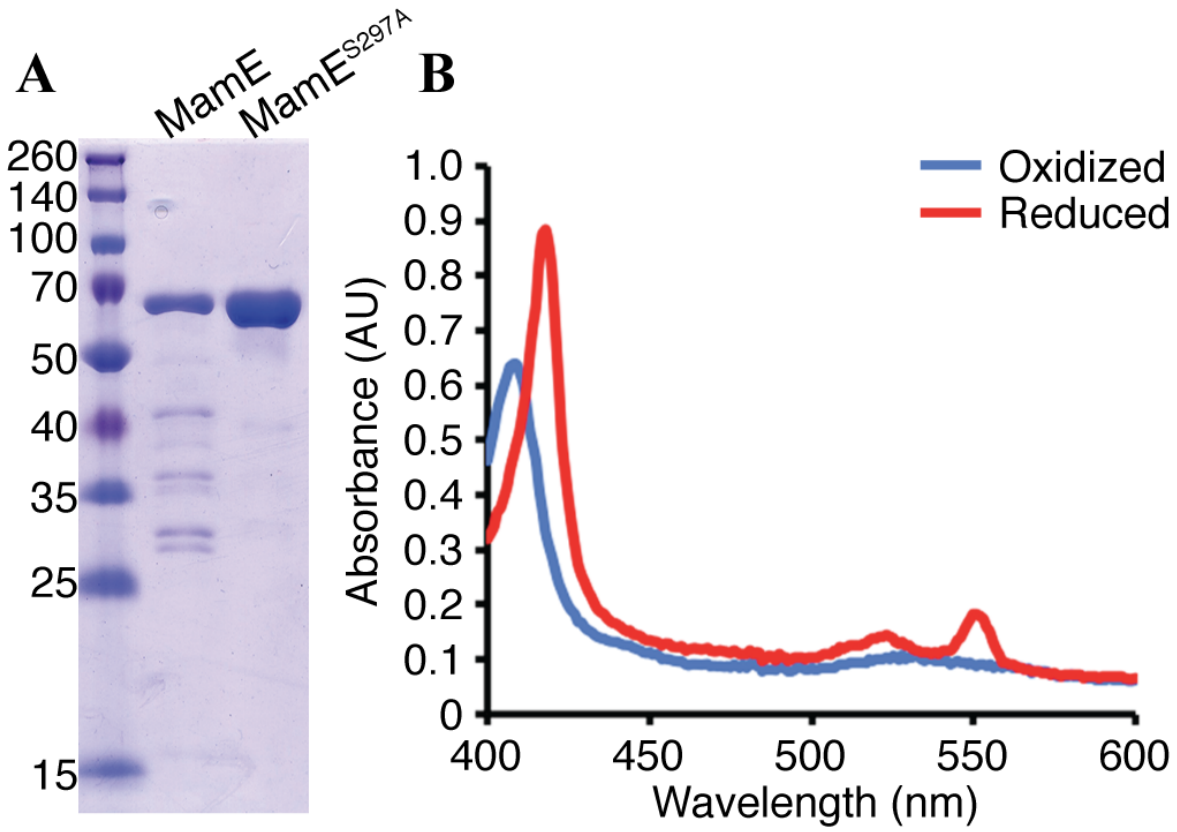


Figure 3-3. Purification of MamE. (a) MamE and MamE^{S297A} (residues 108-728) were purified as a fusion to an N-terminal 6xHis and C-terminal strep tag. (b) Absorbance spectrum of MamES297A in the oxidized (blue) and dithionite reduced (red) forms.

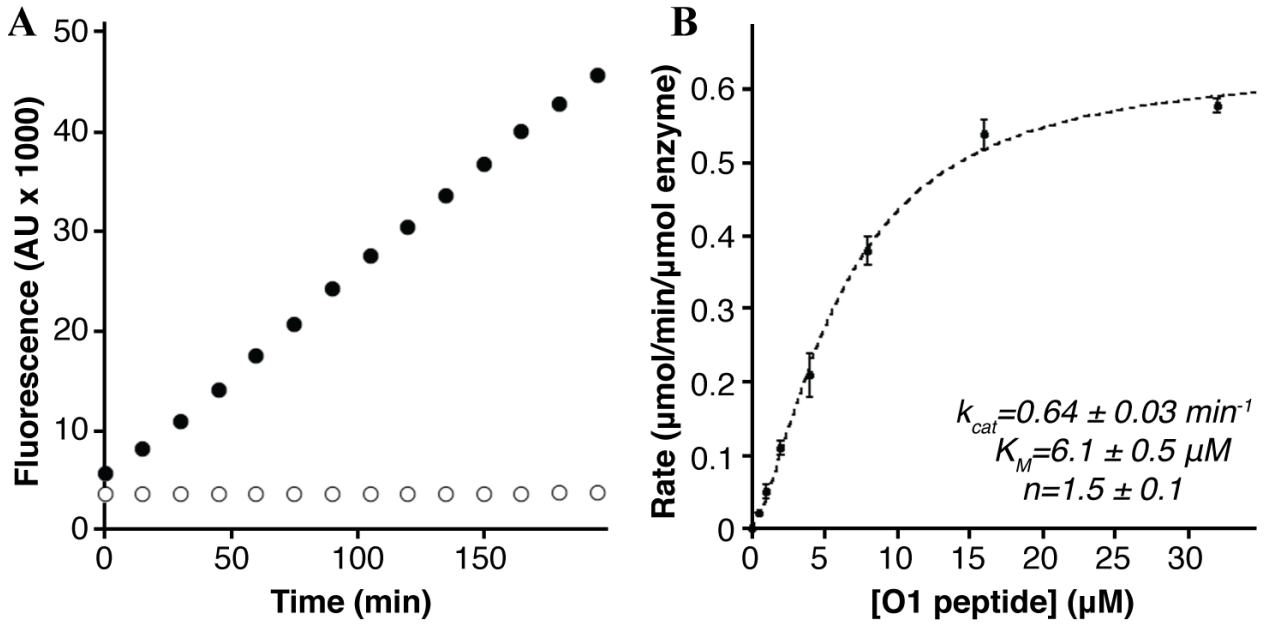


Figure 3-4. Cleavage of the MamO1 fluorogenic substrate. (a) Linear increase in fluorescence upon addition of 20µM MamO1 peptide to 200nM MamE (black circles). No increase is seen with peptide addition to MamES297A (white circles). (b) Steady-state kinetics of O1 cleavage by MamE. The dotted line represents a fit to the Hill form of the Michelis-Menten equation. Error bars represent the standard deviation from three technical replicates. The plot is characteristic of the data seen in five biological replicates.

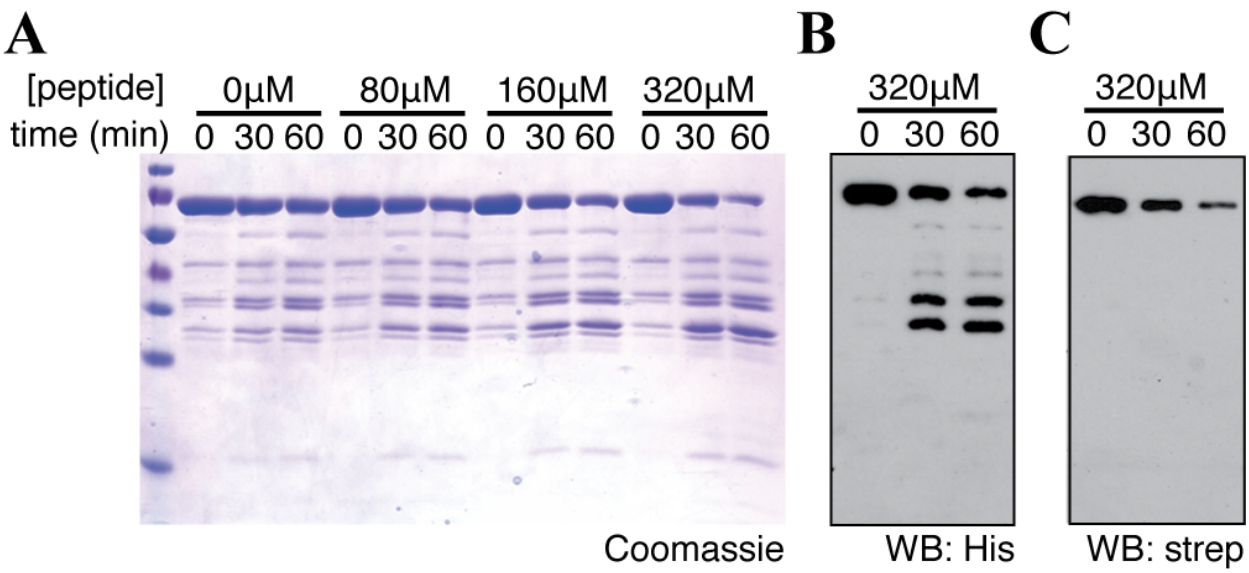


Figure 3-5. Reconstitution of MamE auto-processing. (a) The MamO1 substrate induces auto-cleavage of MamE. The processing pattern was assessed by Western blots (b, c) of the reaction containing 320μM O1 for the indicated tag on MamE. The experiment is representative of three biological replicates.

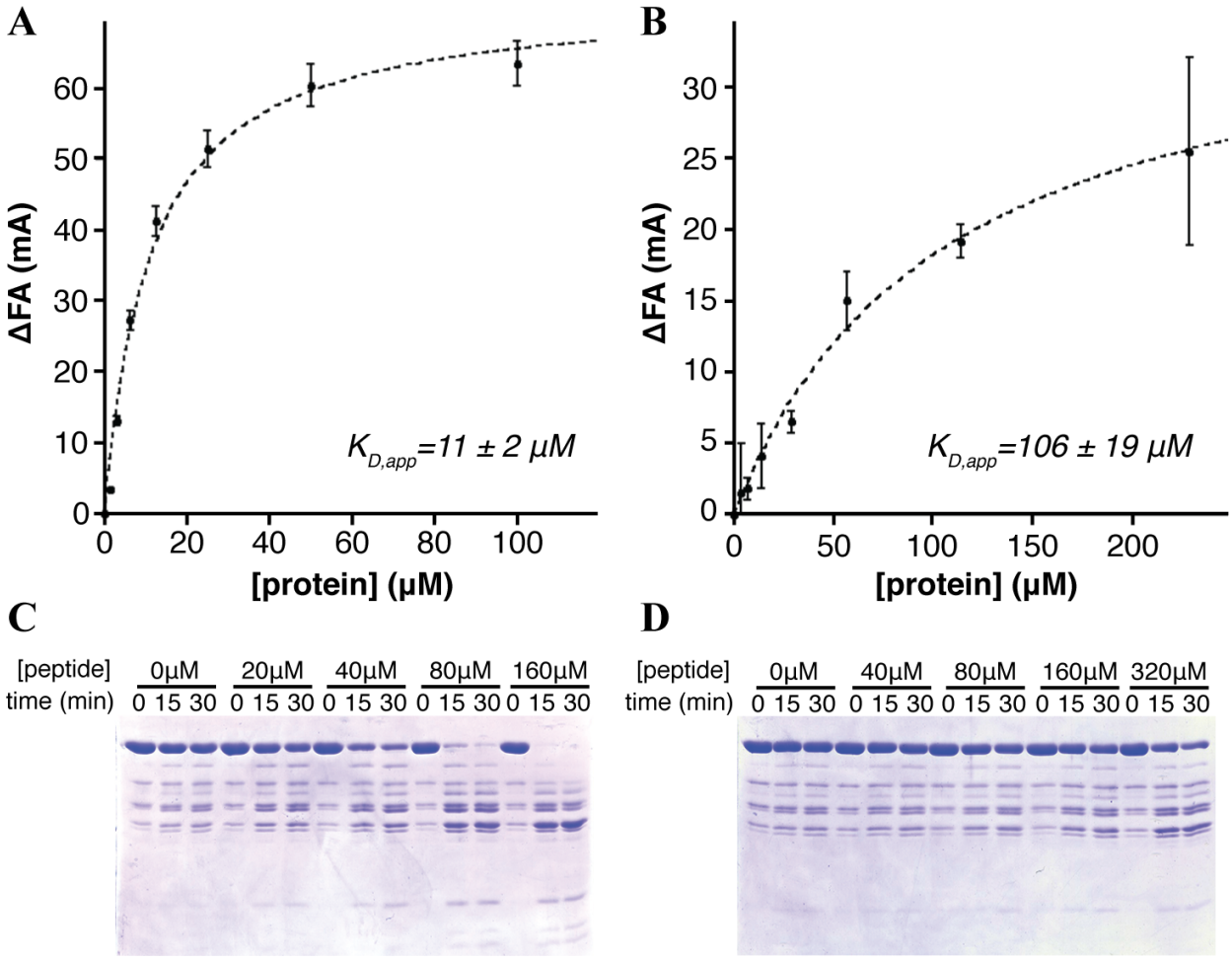


Figure 3-6. Peptide binding to the PDZ domains activates MamE. (a) Fluorescence anisotropy showing binding of the EPDZ1* peptide to the EP12 protein. The error bars represent the standard deviation from three technical replicates. The dotted line represents a fit of the data to a single-site binding model. (b) Binding of the EPDZ2* peptide to the EP12 protein. (c) Activation of MamE auto-cleavage by the EPDZ1* peptide. (d) Activation of MamE auto-cleavage by the EPDZ2* peptide. Each experiment was repeated a minimum of three times.

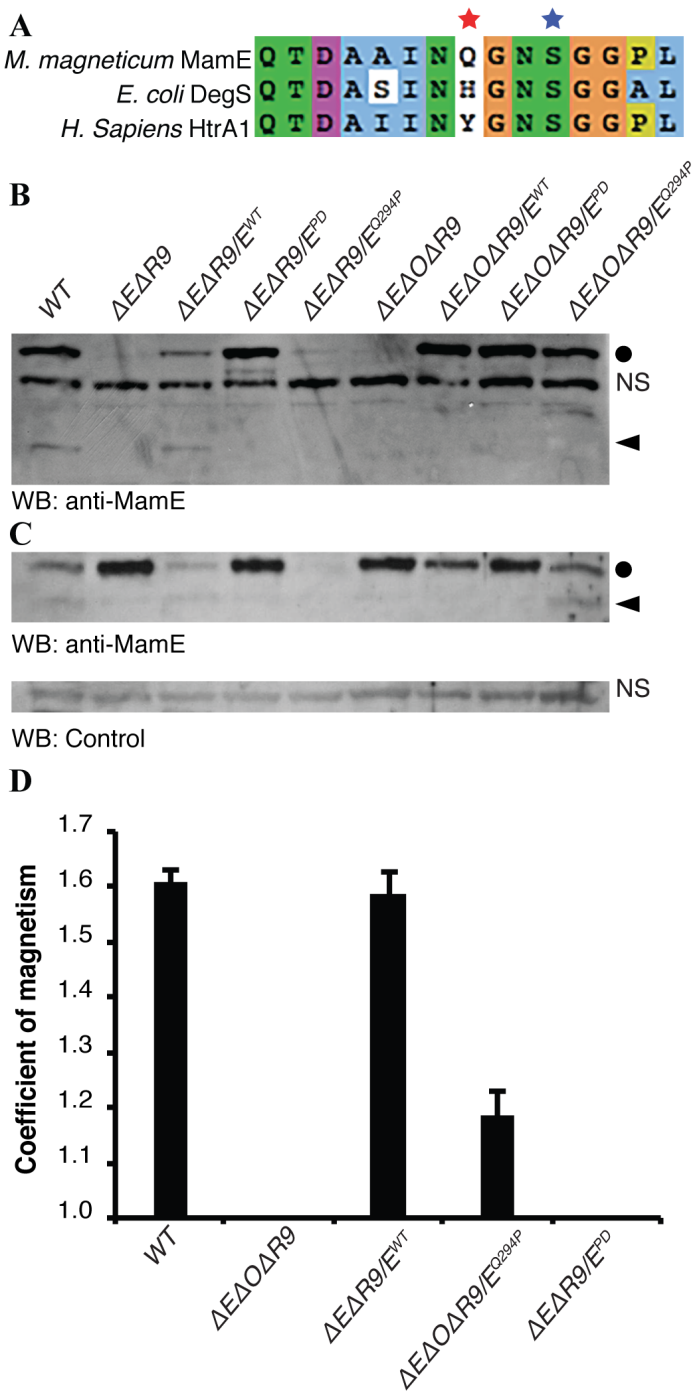


Figure 3-7. A constitutively active form of MamE disrupts biomineralization. (a) Alignment of HtrA proteases. The red star marks position 192 in the oxyanion hole (residue 198 in DegS), and the blue star marks the catalytic serine nucleophile. (b, c) Immunoblot analysis of AMB-1 lysates probed for MamE (b) and MamP (c). Circles mark full-length proteins and carats mark proteolytic fragments. NS marks nonspecific bands reacting with each antibody preparation. (d) Magnetic response of AMB-1 cultures with the indicated genetic background. Biological replicates represent independent cultures of each strain and each measurement represents the average and standard deviation from three independent experiments.

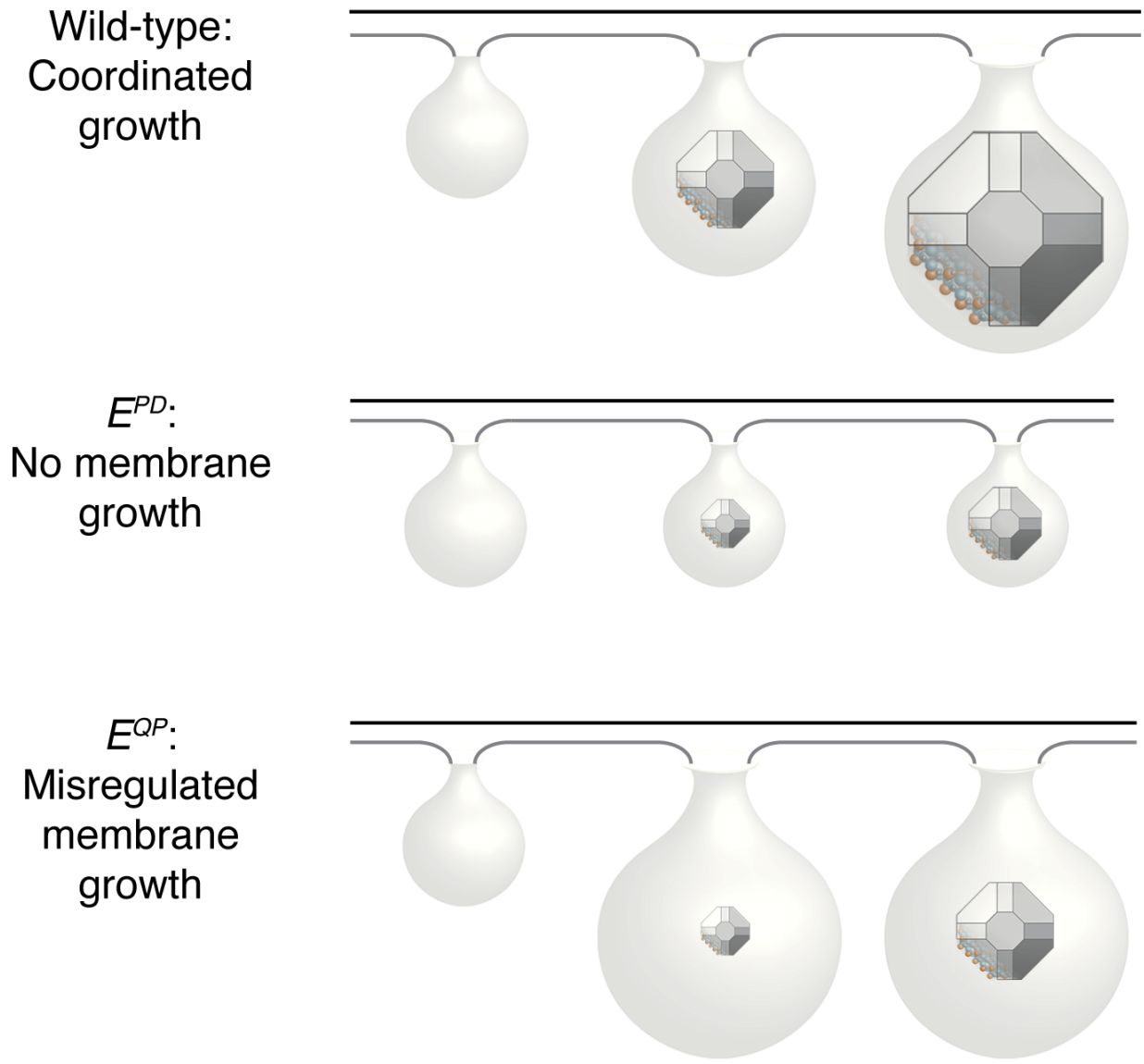


Figure 3-8. Model for MamE dependent membrane remodeling to promote crystal maturation. The model illustrates the predicted membrane remodeling defects in the inactive and misregulated *mamE* alleles that lead to the crystal maturation defects in Figure 3-7.

Chapter 4

A proteomic analysis of wildtype and mutant magnetosomes reveal new molecular functions of conserved magnetosome proteins

Patrick J. Browne[1], Anthony T. Iavarone[2,3], David M Hershey[1] and Arash Komeili[1,2,4]

1 Department of Plant and Microbial Biology, 2 California Institute for Quantitative Biosciences, 3 QB3/Chemistry Mass Spectrometry Facility, 4 Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA

Abstract
Magnetosomes are bacterial sub-cellular compartment that contain magnetic crystals that allow magnetotactic bacteria to align with geomagnetic fields. These compartments are coded by conserved genes and make an ideal model for studying both the cell biology and genetics of bacterial organelles. Here, a proteomic approach was used to study the general protein composition of magnetosomes as well as compare differences between magnetosomes from wild-type and mutant cells, specifically, a strain where MamE, a predicted HtrA protease conserved throughout all magnetotactic bacteria, and central to magnetosome biogenesis was inactivated. Previous studies have shown that MamE is involved in several key steps of magnetosome biosynthesis and targets several MAI proteins for proteolysis in a regulated manner. This study, in addition to further clarifying the protein composition of the magnetosome, uncovers additional potential targets for MamE proteolysis, both within and outside the magnetosome gene island (MAI). One of these targets, MamD, is confirmed as a MamE substrate *in vitro,* and genetic analysis reveals its potential function as an inhibitor of biomineralization.

Introduction
Despite being defined as less complicated than eukaryotes, bacteria and archaea are increasingly understood to have several organized aspects of their cell biology, including the presence of organelles. Bacterial organelles, like their eukaryotic counterparts, are subcellular compartments that serve to carry out a specific process separate from the general cytoplasm. A variety of bacterial protein and membrane-bound compartments exist in nature (113). Magnetosomes are one such bacterial organelle. An individual magnetosome contains a magnetic particle, made of magnetite (Fe_3O_4) or greigite (Fe_3S_4) and is surrounded by a membrane derived from the inner cell membrane (114). Magnetotactic bacteria (MTB) organize linear chains of magnetosomes in order to orient around magnetic fields (20). Previous studies have shown that magnetosomes are coded for by well characterized genetic islands (MAI). MAI's are typically made up of roughly 100 genes, more than half of which are poorly conserved even in closely related MTB (12, 13). A core set of 5-10 genes is well conserved throughout all MTB and are generally grouped together in one core (*mamAB*) operon. Additional genes are somewhat conserved throughout closely related MTB and are generally found in the main operon or in secondary operons within the MAI (115). Genetic manipulations of individual genes in some MTB of the α- and δ-*proteobacteria* have shown that the most well conserved genes are central to magnetosome formation. Additionally, the transfer of the conserved genes from an MTB to a non-magnetotactic bacterium is sufficient to create magnetosomes (12, 13, 61 60). One of the most well conserved MAI genes is MamE, an HtrA protease. While MamE serves multiple purposes within a magnetosome, its protease function has been found to be necessary for complete maturation of magnetosome crystals (25). Recently several targets of MamE proteolysis, all conserved MAI proteins, have been identified (112).

While proteomic analysis on magnetosomes was one of the earliest methods to identify some of the genes responsible for magnetosome synthesis (16, 116), advances in chemical separation and peptide identification have allowed for a greatly enhanced view of the magnetosome proteome. To better understand the protein profile of magnetosomes as whole, we developed a new process to enrich magnetosomes from *Magnetospirillum magneticum* AMB-1, with their membranes intact, separate from the general inner membrane and cytoplasmic fractions. We show that

almost all well conserved MAI genes are present and enriched for in the magnetosome. Furthermore, proteins coded outside the MAI are largely not enriched in this fraction. We then compare the proteomic profiles of wild-type magnetosomes from to those isolated from a strain where MamE protease activity is inactivated (*mamE*^{PD}) strain. There we show that one MAI protein (MamD) as well as three proteins encoded by a single operon outside the MAI are greatly enriched in magnetosomes in the absence of MamE protease activity. We show that MamE is capable of processing MamD *in vitro* and present data that processing of MamD is required for magnetosome maturation *in vivo*. Together, our results suggest that MamD is an inhibitor of biomineralization whose processing by MamE allows for progression of magnetite formation in AMB-1.

Materials and Methods

Strains, plasmids and culture conditions

The strains and plasmids used in this study are listed in Tables 4-2 and 4-3, respectively. For general maintenance and genetic manipulation, *M. magneticum* AMB-1 was grown in MG medium supplemented with ferric malate (30μM). 0.7% agar was used in plates and kanamycin was used for antibiotic selection at a concentration of 7 μg/mL (solid) or 10μg/mL (liquid). Cultures for magnetic response measurements and western blotting were grown in 10 mL MG medium (pH 6.9) and ferric malate under a 10% oxygen atmosphere at 30°C. For proteomic growth, MG was supplemented with 25mM HEPES buffer (pH 7.2) The magnetic response of each culture was assessed using the Coefficient of Magnetism (Cmag), which was measured as described (56).

Separating AMB-1 magnetosomes from the cell and general membrane

Initial 10mL cultures were grown in MG+kanamycin for 48 hours in 10% oxygen atmosphere. All 10mL were then transferred to 2L bottle containing MG with HEPES buffer. No kanamycin was added. The bottles were flushed with nitrogen for ten minutes before inoculation. After inoculation, cells were grown for an additional 48 hours at 30°C and 10% oxygen. Cells were harvested at 8k x *g* for ten minutes and in resuspended 8mL MG. These mixtures were then passed through a french press cell press three times at 1000psi. The resulting lysate was harvested at 2K x *g* for five minutes. The pellet was resuspended in 550μL MG. 50μL was retained as the general lysate sample. and carefully added to the top of percoll gradients. The gradients had previously been made by mixing percoll and MG (50% each) and spinning for 30 minutes at 8k x *g*. The gradient with deposited floating mixture was spun for four hours 0.5K x *g* in a swing bucket rotor. The top half of the gradient disposed of by pipetting. The band containing magnetosomes, settled near the bottom of the gradient, was then harvested by needle (to minimize contamination) to MG and washed 3X by centrifugation to remove the percoll. After the third wash, the pellet was resuspended in 1mL MG. This mixture was sonicated at low output for 10 seconds and then added to a magnetized MACS LS column that had previously been equilibrated to MG. After washing the column 3X with MG, the magnets were removed and the magnetosome fraction was eluted in 5mL of MG. All steps after harvesting of cultures were kept on ice or performed at 4°C when possible.

Preparation of trypsin digests for liquid chromatography-mass spectrometry (LC-MS)

After measuring protein concentration by the Bradford method using bovine serum albumin to prepare a standard curve, portions of the both the lysate and magnetosome samples were

harvested at 8k x *g* for 2 minutes and resuspended in 50µL MG for equal protein concentration. 10µL of 50m mM NH_4HCO_3, pH 7.5 was added. Then 25µL of 0.2% RapiGest SF and the mixture was heated at 80 ^oC for 15 minutes. 2.5 µL 100mM dithiothreitol was added and heated 60 ^oC for 30 minutes. 2.5µL was added, the mixture was vortexed and then left in the dark at room temperature for 30 minutes. 10µL of Promega trypsin gold was added and digested at 37 ^oC overnight. The next morning 10µL of 5% trufluoroacetic acid was added to hydrolyze the Rapigest. The samples were incubated at 37 ^oC for 90 minutes, and then centrifuged at 14,000 RPM, 4°C for 30 minutes. The resulting supernatant was transferred to a USA Scientific low adhesion vial and evaporated by speed vac. The samples were then resuspended in 15 µL of Optima LC/MS water.

Determination of enriched magnetosome protein
The resulting proteins were ranked by according to fold enrichment. Proteins that contained less than 3 tryptic peptides per 100 amino acids were removed. All proteins remaining with a 20 fold or higher over abundance are listed in the table below.

Genetic manipulation
For complementation of deletion mutants, we used modified form of pAK605. This plasmid contains a neutral region of the AMB-1 genome and integrates as a single copy at this site. Each allele is inserted under the control of the *mamAB* promoter, allowing the constitutive expression of each protein. FLAG-tagged versions of MamD were made by altering previously described pAK701 and pAK702 (73).

Immunoblotting
Cultures of AMB-1 were grown to late-log phase and harvested by centrifugation at 8k x *g*. The resulting pellets were resuspended in 2x SDS loading buffer and heated for 10 min at 95°C. The lysates were separated on SDS-PAGE and transferred to PVDF. The membranes were blotted and visualized using standard western blotting techniques. Polyclonal antibodies to MamP, and MamT, were raised in rabbits against recombinant forms of the soluble portion of each protein. The MamE, MamT and MamP antibodies have been previously described (164). The anti-6xHis (Sigma), anti-FLAG (Sigma), anti-σ70 (Thermo Fisher) and anti-strep (Qiagen) antibodies were purchased from commercial sources.

Preparation of trypsin digests of gel extracted fragments of mamD for LC/MS
After proteolysis, the mixture was separated on a 12% acrylamide gel and stained with colloidal Coomassie Blue. A ~3 x 10mm section of the gel corresponding to the two processed MamD bands were excised from the gel and chopped into small pieces. These were washed with 100mM NH4HCO3 followed by reduction and alkylation of cysteines with DTT and iodoacetamide. The gel pieces were then dehydrated by washing with increasing concentrations of acetonitrile in 100mM NH4HCO3 and dried under vacuum. A 0.1mg/mL solution of trypsin was used to re-swell the gel pieces, and they were incubated overnight at 37°C. The resulting peptides were extracted from the gel slices with successive washes of 0.1% formic acid solutions containing increasing concentrations of acetonitrile. The extracts were pooled in a fresh tube, concentrated under vacuum to remove the organic phase and stored at 4°C until analysis.

Results

Magnetosomes are enriched with conserved MAI genes

Analysis of general AMB-1 lysates were done using two different LC/MS instruments. The Orbitrap identified 1,168 unique peptides, twenty-one of which mapped to the MAI(1.80%). Seven total genes were identified. A more advanced Synapt was able to identify 33,176 unique peptides, 613 of which mapped to the MAI or islet (1.85%).

In AMB-1 magnetosomes remain attached to the general inner membrane (20). While isolating crystals free from all cellular material is relatively straightforward, isolating magnetosomes with enough membrane intact and free from general cellular content proved difficult. A method was developed where cells were lysed by relatively gentle means resulting in a mixture of magnetosomes with intact membrane and unlysed cells. This mixture was spun at low speeds, the supernatant was removed, and the pellet was resuspended in buffer. This mixture was passed through a percoll gradient in order to separate the very dense magnetosomes from the less dense whole cells. The magnetosome fraction was then further disturbed by sonication in order to free the magnetosomes from some of the general membrane and passed through a reversible magnetic matrix (Figure 4-1). The magnetic fraction was then analyzed by LCMS. Even after this separation, the magnetic fraction contained many non-MAI proteins. Using the Orbitrap, MAI and islet peptides count increased roughly 100X compared to the lysates, and made up 20-25% of the total sample over several.

Because of the Synapt's enhanced ability to detect peptides, MAI and islet peptides make up a smaller percentage of total peptides compared to the Orbitrap. However, the Synapt's ability to integrate peaks and compare relative abundance was used to distinguish proteins truly enriched in the magnetosome when compared to the lysate and corrected for total protein concentration. When ranked by enrichment over their abundance in the lysate fraction and filtered for having a sufficient number of peptides (see materials and methods), thirty proteins are found to have a 20X enrichment or higher in magnetosomes (Table 1).

Seventeen of the twenty-eight enriched proteins are encoded by conserved genes located within the MAI (*mamG, mamF, mamD, mamC, mmsF, mms6, mamE, mamJ, mamM, mamA, mamB, mamT, limJ, mamX*, *mamY*, *mmsF2 (amb1026*) and *mms5 (amb1027*). MamO also barely missed the 20X fold enrichment cutoff (Table 2). It has previously shown that MamO does act within the magnetosome although perhaps it does not localize as strongly as other proteins (73). Six other conserved proteins (MamI, MamR, MamS, MamV, MamN and MamZ) also appear to be enriched in the magnetosome fraction but did not make the list due to having too few representative peptides. MamI, MamS and MamR are very small with few even theoretical tryptic peptides, while MamV has no known function and may be slightly expressed. Four proteins, all relatively small (MamL, MamP, MamU, and FeoA*)* were completely unobserved. Finally, MamH, MamK, MamQ, FtsZ-like, and FeoB were observed in lysates but were not found to be enriched in the magnetosome fraction. MamK is a known cytoplasmic protein involved in forming a chain to organize the magnetosomes (although MamJ and MamA, two other predicted cytoplasmic proteins, remained enriched on the magnetosome, reflecting their tighter association with the magnetosome membrane). FtsZ-like, named because of its clear homology to the tubulin like protein FtsZ has previously been shown to influence biomineralization under nitrogen starvation (117). MamQ is crucial to the early stages of

magnetosome membrane formation (12, 13). A membrane protein, it is possible that it does not enter the magnetosome but instead is only involved in early stages. MamH and FeoB are transporters with no clearly required role in magnetosome formation although FeoB is well conserved among MTB (121).

Two of the remaining nine proteins (Amb0400 and Amb0412) are homologs to proteins mentioned above (MamD and MamF respectively) found outside of the MAI in a region called the islet, where eight homologs of MAI genes are found scattered around a 45KB region (118). Recently, the islet's MamK homolog was shown to also play a role in magnetosome chain maintenance (108). Appropriately, it was observed but not enriched in the magnetosome fraction. MamJ's islet homolog also was shown to be enriched in the magnetosome fraction but had only one representative peptide. None of the remaining five homologs were observed and it should be noted that all of the four islet proteins that were observed, appear to be underrepresented compared to their MAI homologs.

Of the nine most abundant proteins not localized to the MAI or island, five are annotated as Phasins, proteins involved in polyhydorxybutyrate (PHB) granule synthesis. It was previously reported that one Phasin, Mms16, is one of the most abundant magnetosome proteins. However, it was recently shown to play no actual role in magnetosome biogenesis (119). Whether these proteins truly do locate to the magnetosome or their appeared abundance is an artifact of the isolation technique remain unknown. Additionally, the final four proteins all are annotated as cytoplasmic. It is likely that their abundance is also an artifact, although one of them, Amb0399, is located directly next to the islet's MamD homolog, Amb0400.

Interestingly, none of the sixty-one remaining poorly conserved MAI proteins are found to be enriched in the magnetosome fraction. In fact, only ten of the proteins were observed at all and none are enriched in the magnetosome fraction. Compared to the genome as a whole, this is a significant underrepresentation as over 1/3 of AMB-1's coded proteins have at least one represented peptide when analyzed by the Synapt.

Protease Dead Magnetosomes are enriched for four proteins

Additional proteomics were carried out on several biomineralization mutants. Because the methods of magnetosome isolation described above depend on the presence of a dense particle, only mutants that still contain a crystal were analyzed. Most mutants displayed no clear proteomic differences (other that the loss of a single protein in cases where the corresponding gene was deleted). MamE has been previously shown to act on three other MAI proteins, MamP, MamO, and itself (112). While MamP was not abundant enough to be seen at all by proteomic analysis, MamE and MamO do not appear to be significantly altered between wildtype and $mamE^{PD}$ strains. Potentially, this is because those processed isoforms remain active and functional within the magnetosome. The $mamE^{PD}$, however did contain four proteins with clearly different abundancies (Fig 4-2).

Three of these proteins, Amb3286, Amb3287, and Amb3288 are encoded from a single operon and comprise the structural components of a predicted tetrathionate reductase. These differences are so apparent that in in Δ*mamE* and $mamE^{PD}$ strains, Amb3286 is one of the most abundant proteins in the general lysate by peptide count. Predicted to localize to the periplasm, at least

some amount of these proteins find themselves in the magnetosomes of *mamE*^{PD} cells. Curiously, these genes appear to be a result of horizontal gene transfer from *β-proteobacteria* and are not conserved in other *α-proteobacterial* MTB, including MSR-1. Also, all proteins are predicted to be moved to the periplasm via the twin-arginine secretion method (TAT). While most MAI proteins also are predicted to reach the periplasm, none are predicted to use the TAT pathway. The role these proteins play in magnetosome biogenesis, if any, and how they impact MamE remains an open question and possible topic of future study.

The fourth protein found to be enriched in *mamE*^{PD} cells is MamD*,* also known as Mms7. This protein is encoded by the *mamGFDC* operon which previously shown to play a minimal role in crystal formation as the deletion of the complete operon has a minimal effect on biomineralization and alignment of cells in magnetic fields (50). However, \ previous work on MamD shows that differing the amounts of its expression can alter the shape of the magnetite particles (120).

MamE cleaves MamD both in vitro and in vivo

The soluble portion of MamD was purified from *E. coli.* When incubated with the active soluble part of MamE it was degraded over time (Fig 4-3). Mapping of the two resulting bands showed that the protein was cut near the C-terminus of the soluble region. MamD has a single C-terminal transmembrane helix and the bulk of the protein is assumed to operate inside of the magnetosome. MamE belongs to a family of specific endopeptidases and was previously shown to cut its other targets at specific locations, including freeing the soluble portions of MamP and itself from their membrane anchors (112).

Furthermore, using antibodies raised to the soluble portion of MamD, abundance of full-length MamD was analyzed *in vivo* (Fig 4-4)). While MamD is clearly abundant in all strains lacking a functional MamE it is actually most abundant in *mamE*^{PD} strain, even when compared to Δ*mamE* strain. Additionally, MamD appears to be more abundant in the *mamE*^{Q294P} strain compared to the wildtype. The *mamE*^{Q294P} strain was previously shown to be an overactive protease, that processes its other full-length targets more completely (112). It is possible that MamD, unlike MamP and MamE itself, represents a later target of proteolysis and because the MamE^{Q294P} protein consumes itself too rapidly it is unable to act on MamD.

MamD can inhibit biomineralization

To better understand the context of MamE-dependent proteolysis of MamD, strains were created with N- or C-terminally 3xFLAG alleles in a wildtype, *ΔmamE,* and *ΔmamGFDC* backgrounds. Unlike previous MamE targets (MamP, MamO, and MamE itself) no bands corresponding to stable processed forms of MamD were observed. However, it is likely that processed version of MamD are further degraded due to their decreased presence as seen in the proteomics.

Surprisingly, the N-terminal fusion of MamD showed a severe decrease in Cmag (Fig. 4-5). MamD contains a single C-terminal trans-membrane helix and its larger N-terminal bulk of the protein is predicted to localize to the inside of the magnetosome. Proteomics and previous work showing that fragments of MamD are one of the tightest bound proteins on the magnetosome crystal itself support this model. The FLAG-MamD protein severely limits biomineralization in both a wildtype and *ΔmamGFDC* background. While pellets of the cells are pliable by strong

magnets, the resulting strains have almost no Cmag, similar to the *mamE*^{PD} strain and some other biomineralization mutants. TEM images of the strains showed that cells have less magnetosomes compared to wildtype and that the magnetosome crystals are much smaller that their wild-type counterparts.

As shown by western, FLAG-MamD is not over expressed relative to native MamD (Figure 4-4). The presence of FLAG-MamD does appear to increase the amount of native MamD in an otherwise wildtype background. However, FLAG-MamD has the same negative effect on biomineralization whether or not native MamD is also present which means the effect is an outcome of the recombinant protein.

Tagging proteins has never previously been shown to have an effect on biomineralization and several other proteins, including all previously reported targets of MamE proteolysis. Knowing that MamD is processed by MamE and that it may only be targeted in later stages of magnetosome development, we hypothesize that the FLAG-tagged version is not proteolyzed by MamE. Its new stability and continued presence do not allow biomineralization to fully occur. Unfortunately, we were unable to purify the FLAG-tagged version. Furthermore, attempts at mapping exactly where MamE cuts MamD were unsuccessful so we were unable to generate recombinant version of MamD unable to be processed in order to test the above hypothesis. Therefore, further work is needed to show that MamD, which has no clearly defined function, acts as an early inhibitor of biomineralization that is removed by MamE.

Discussion

Magnetotactic bacteria control the growth of their associated magnetosomes using a unique and somewhat conserved set of between 15 and 30 genes. While several of these genes have been given distinct roles based off of their respective deletions, how their corresponding proteins' molecular interactions with each other and the cell as a whole is an ongoing topic of study.

Here, we examined the magnetosome system as a whole. We show that MTB are able to create invaginations from their inner membrane with a wholly distinct protein profile from the membrane. This profile consists almost fully of MAI proteins. How these proteins find each other and how non-MAI proteins are largely excluded remains an open question. Additionally, the fact that none of the MAI genes from the more poorly conserved regions are enriched in the magnetosome, and they seem to be expressed at a much lower rate relative to the rest of the genome, raises the question as why these genes are present. It is possible that they may play an accessory role in conditions not represented in the laboratory or they may simply be an artifact of evolution.

A recent study also analyzed the proteomic content of magnetosome membranes from *Magnetospirillum gryphiswaldense* MSR-1, AMB-1's closest magnetotactic relative and another model of MTB. While magnetosomes in MSR-1 fully detach from the cells' inner membranes and the methods used to isolate the membranes differed greatly from our own, their findings closely align with ours. Magnetosomes were greatly enriched for conserved MAI proteins and less conserved proteins were not represented. They also report that MamQ, FtsZ-like and FeoB are not enriched in magnetosome membranes. Mms16, the PHB granule Phasin commonly observed to locate to the magnetosome was also enriched in their samples. However, while both

studies identify some other non-MAI proteins that are enriched in the magnetosomes, there is no additional overlap between the two groups, even though many of them have homologs in the other species (106).

By looking at magnetosomes from a mutant strain, specifically one that has been previously suggested to represent a potential checkpoint of magnetosome development, we sought to observe differences in protein profiles over the lifetime of a magnetosome. However, the profiles were largely identical with two exceptions. The first was the abundance of three non-MAI three proteins that compose the structural subunits of a tetrathionate reductase. These proteins are incredibly abundant in *ΔmamE* and $mamE^{PD}$ strains but largely absent in all other analyzed strains. How these proteins impact magnetosome generation and whether they represent true targets of MamE proteolysis remains unclear.

We also find a clear increase in MamD, a protein found in an accessory operon conserved in *α-proteobacterial* MTB but not found in other phyla. Because the loss of its whole operon causes only a minor disruption of biomineralization it was not believed to play an important role. Here we show that it is a direct target of MamE proteolysis. This protein is the first confirmed target of MamE found outside of the core *mamAB* operon. It also presents a novel relationship with MamE dependent proteolysis as it is not targeted when MamE is altered to have higher activity rates suggesting it is not targeted until later in the development of the magnetosome. MamE was already shown to be crucial in several stages of magnetosome development but here we see that proteolysis may also occur at distinct stages.

We also show the unexpected finding that a FLAG-tagged version of MamD has a severe impact on biomineralization. This finding further suggests that MamD does play a role in biomineralization inhibition. The fact that its deletion has a minor effect may be because it plays a negative regulatory role. Perhaps, this a useful role in AMB-1's natural environment where resources are more scarce compared to the laboratory and a cell may not want to constitutively produce large particles at a high energetic cost. However, at this point, MamD's role remains a hypothesis. It is possible that by FLAG tagging MamD we simply altered how it interacts with other proteins and began a chain reaction leading to permanently stunted magnetosomes.

As a whole, our findings support the portrait of the magnetosome as a complex bacterial organelle filled with its own associated proteins and largely depleted of other proteins, including ones found elsewhere in the inner membrane. It also details previously unknown interactions between MAI proteins and suggests a role for MamD, which despite is conservation, did not have a clearly defined purpose.

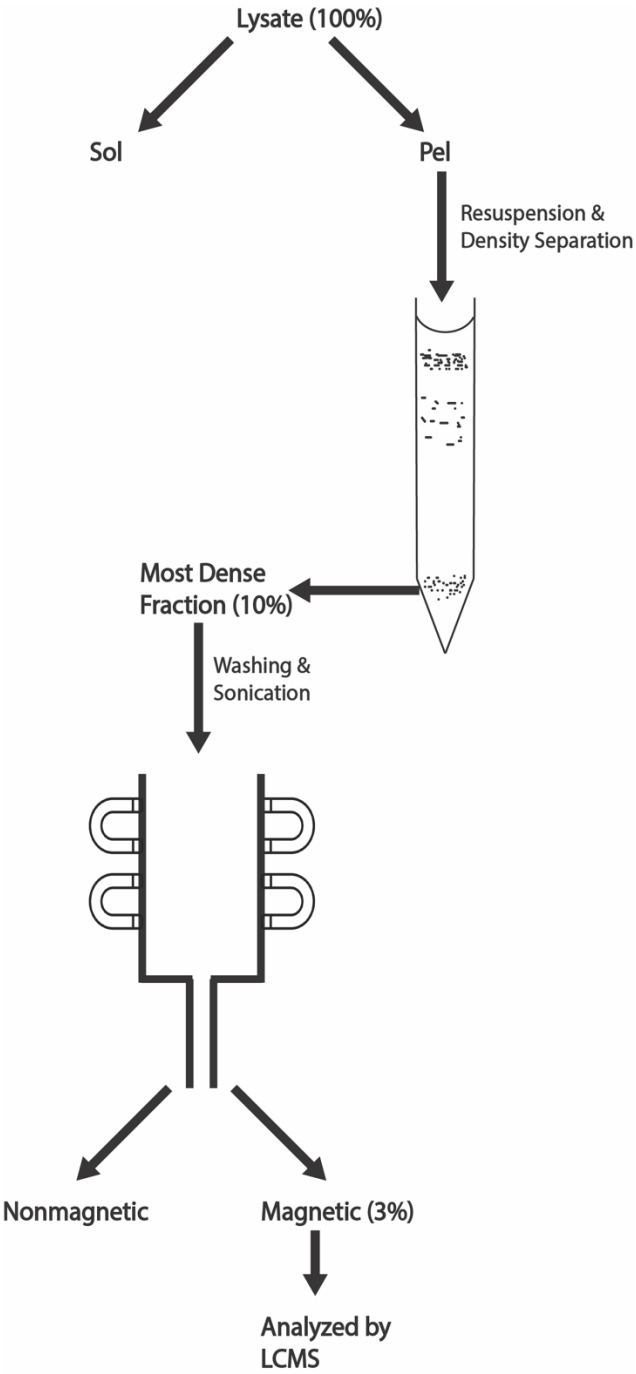


Figure 4-1 Fractionation of magnetosomes from unlysed cells and other material by density fractionation, before further disruption and increased separation from the general inner membrane by sonication and passing through a reversibly magnetic column.

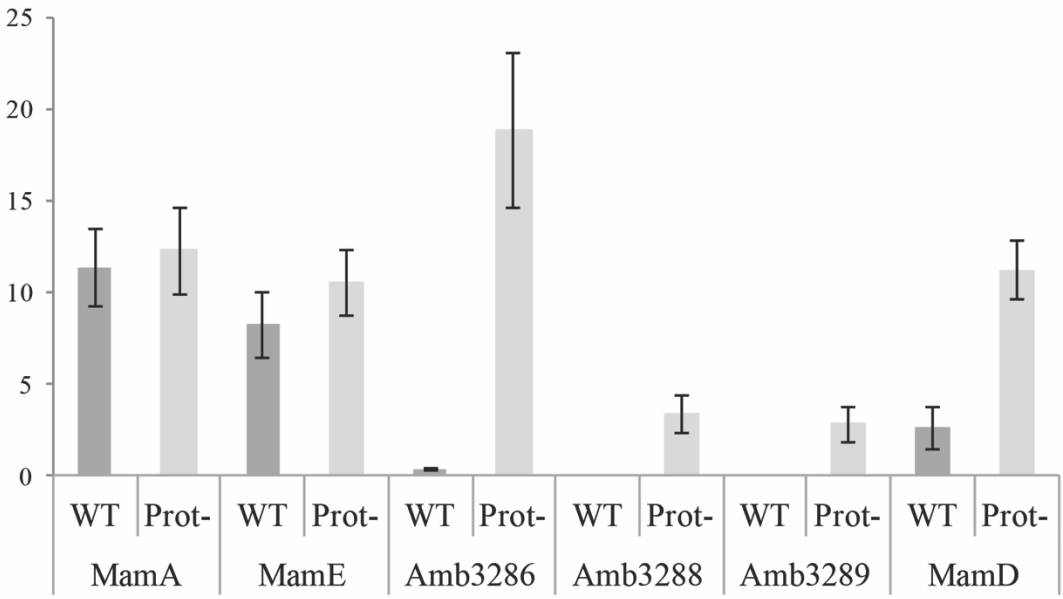
Rank Enrichment	Max fold change	Protein Name	Unique peptides	Genomic Location	Proposed Function
1	229	Mms16	19	General Genome	PHB Granule Associated Protein
2	141	MamF	4	MAI	Magnetite Biomineralization
3	112	Mms6	6	MAI	Magnetite Biomineralization
4	89	Amb0571	16	General Genome	PHB Granule Associated Protein
5	69	MmsF	5	MAI	Magnetite Biomineralization
6	68	Mms5	6	MAI	Magnetite Biomineralization
7	61	MamD	11	MAI	Magnetite Biomineralization
8	60	MamG	3	MAI	Magnetite Biomineralization
9	55	Amb1121	15	General Genome	PHB Granule Associated Protein
10	53	MamF-Like	4	Magnetosome Islet	Magnetite Biomineralization
11	52	MamZ	3	MAI	Magnetosome Redox Control
12	51	MamM	8	MAI	Magnetosome Iron Transport
13	51	MamD-Like	10	Magnetosome Islet	Magnetite Biomineralization
14	43	Amb3386	7	General Genome	PHB Granule Associated Protein
15	42	MamF2	5	MAI	Magnetite Biomineralization
16	38	MamC	7	MAI	Magnetite Biomineralization
17	35	Amb0839	10	General Genome	PHB Granule Associated Protein
18	33	LimJ	3	MAI	Magnetosome Chain Localization
19	31	Amb0399	5	Magnetosome Islet	Hypothetical
20	31	MamB	4	MAI	Magnetosome Iron Transport
21	31	Amb3775	21	General Genome	Hypothetical
22	30	MamY	21	MAI	Magnetosome Membrane Control
23	30	MamT	6	MAI	Magnetite Biomineralization
24	28	MamE	11	MAI	Multifunctional Magnetosome Protein
25	27	Amb1616	16	General Genome	Flavodoxin
26	24	MamJ	7	MAI	Magnetosome Chain Localization
27	24	Amb1957	12	General Genome	Phospholipase
28	20	MamA	19	MAI	Magnetosome Activation

Table 1 All twenty-eight proteins that are 20X enriched in the magnetosome fraction compared to the lysate

Rank Enrichment	Max fold change	Protein Name	Unique peptides	Genomic Location	Proposed Function
37	18	MamO	15	MAI	Magnetite Nucleation, Activation of MamE
NA	78	MamI	1	MAI	Magnetosome Membrane Control
NA	77	MamR	2	MAI	Magnetite Biomineralization
NA	21	MamN	2	MAI	Magnetite Nucleation
NA	20	MamX	1	MAI	Magnetosome Redox Control
NA	18	MamV	1	MAI	Unknown, Putative Transporter
NA	13	MamJ-Like	1	Magnetosome Islet	Magnetosome Chain Localization
No Data	NA	FeoA	0	MAI	Unknown, Putative Transporter
No Data	NA	MamL	0	MAI	Magnetosome Membrane Control
No Data	NA	MamP	0	MAI	Magnetite Biomineralization
No Data	NA	MamU	0	MAI	Unknown, Putative Kinase
Not Enriched	NA	MamH	5	MAI	Unknown, Putative Transporter
Not Enriched	NA	MamQ	6	MAI	Magnetosome Membrane Control
Not Enriched	NA	MamK	20	MAI	Magnetosome Chain Localization
Not Enriched	NA	FeoB	15	MAI	Unknown, Putative Transporter
Not Enriched	NA	FtsZ-Like	3	MAI	Magnetosome Redox Control
Not Enriched	NA	MamK-Like	6	Magnetosome Islet	Magnetosome Chain Localization

Table 2 All remaining conserved magnetosome proteins and the two other islet proteins also identified through proteomics.

A



B

Protein	Putative Function	Bacterium with closest homolog	E Value	Identity
MamD	magnetite biomineralization	*Magnetospirillum magnetotacticum* MS-1	0.00	99%
Amb3286	tretrathionate reductase subunit A	*Rhodocyclaceae bacterium* Paddy-1	0.00	65%
Amb3288	tretrathionate reductase subunit C	*Thiobacillus denitrificans*	1E-131	62%
AMB3289	tretrathionate reductase subunit B	*Janthinobacterium lividum*	4E-116	68%

Figure 4-2 A. A comparison of peptide counts mapped of various genes from magnetosome preps from either $\Delta E/mamE^{WT}$ or $\Delta E/mamE^{PD}$ strains. MamA like the vast majority of MAI proteins had no noticeable differences. MamE which processes itself in wildtype strains may be enriched in the inactive protease strain, but only slightly. Amb3286, Amb3288, Amb3289, and MamD are clearly enriched in inactive protease magnetosomes. Numbers are averages from five distinct preps of magnetosomes. B. Descriptions of the four genes enriched in inactive protease magnetosomes. *MamD*, like the vast majority of genes from AMB-1 have closely related homologs in both MSR-1 and MS-1. The operon containing the tetrathionate reductase, however, seems to have arisen in betaproteobacteria.

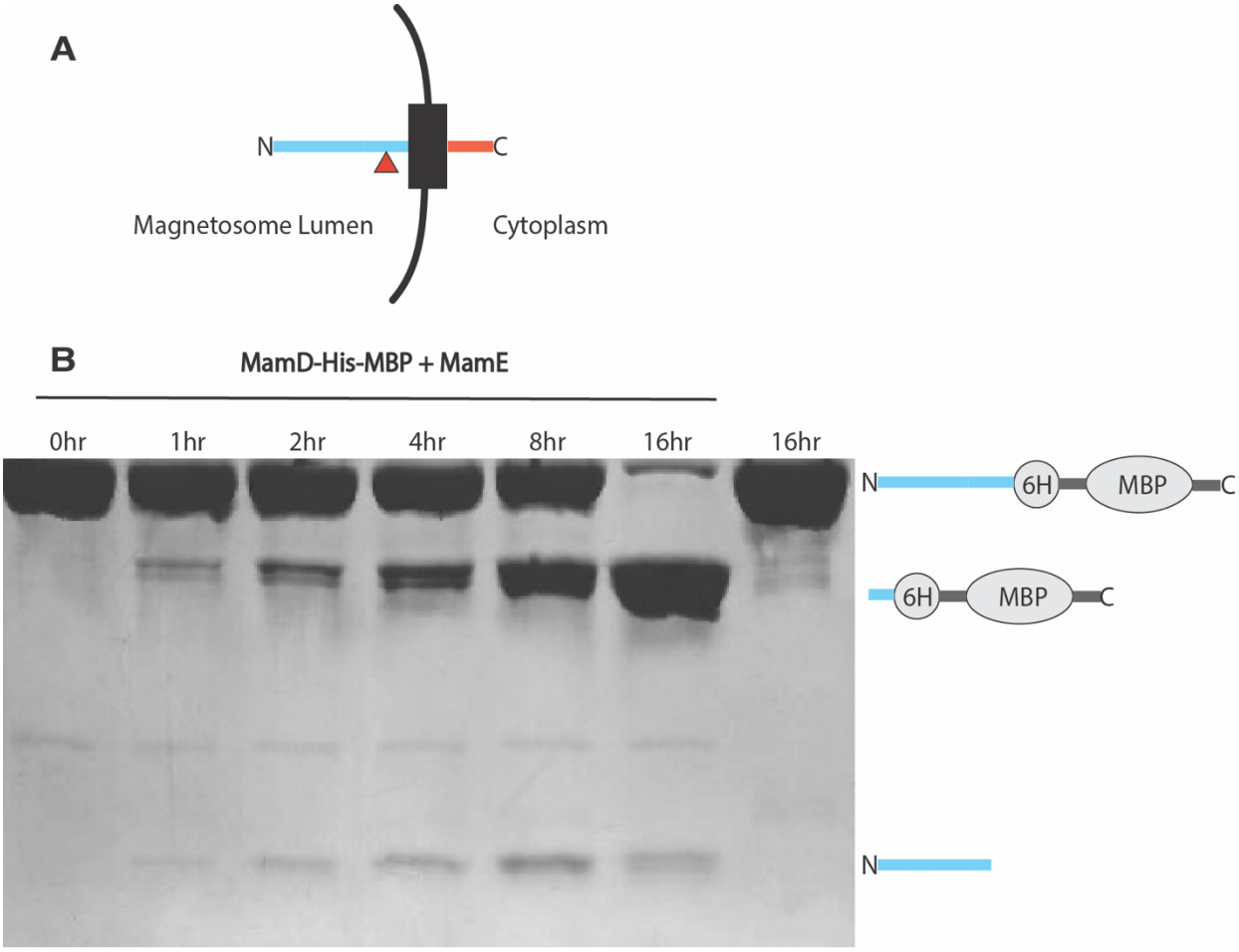


Figure 4-3 MamD is processed by MamE to free the soluble n-terminal region from its transmembrane anchor. A. Topology of MamD, the red triangle indicates the rough processing point. B. MamD incubated with the soluble trypsin domain of MamE incubated over time for 16 hours. Note that processed bands appear in couplets, indicating more than one possible processing point of MamD.

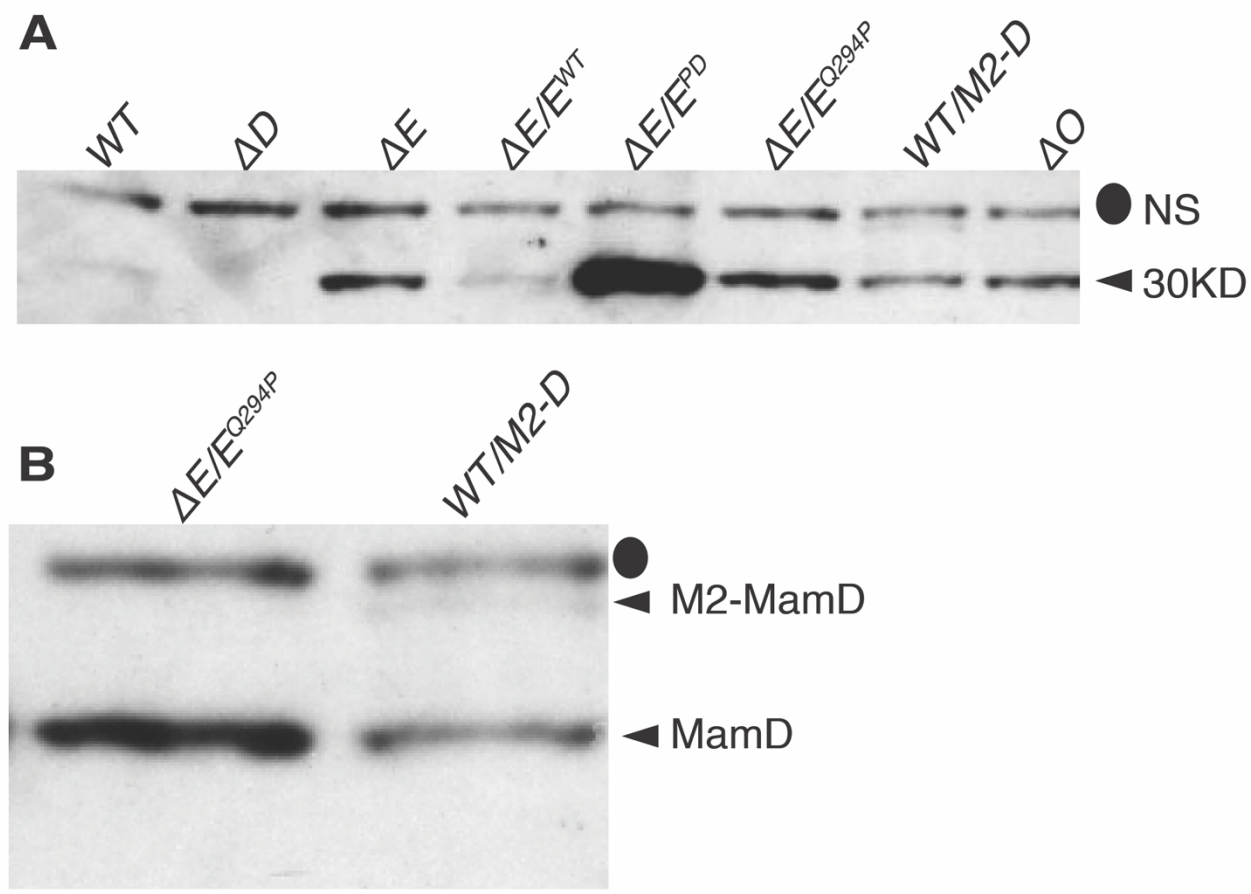


Figure 4-4 A. MamD is enriched in all strains, not containing a fully functional MamE. In wildtype and the complemented deletion of *mamE,* MamD is barely visible. Even though the E^{Q2949P} has been shown to further degrade itself and MamP, MamD is also enriched in this strain, potentially indicating that it is a later target of MamE proteolysis.

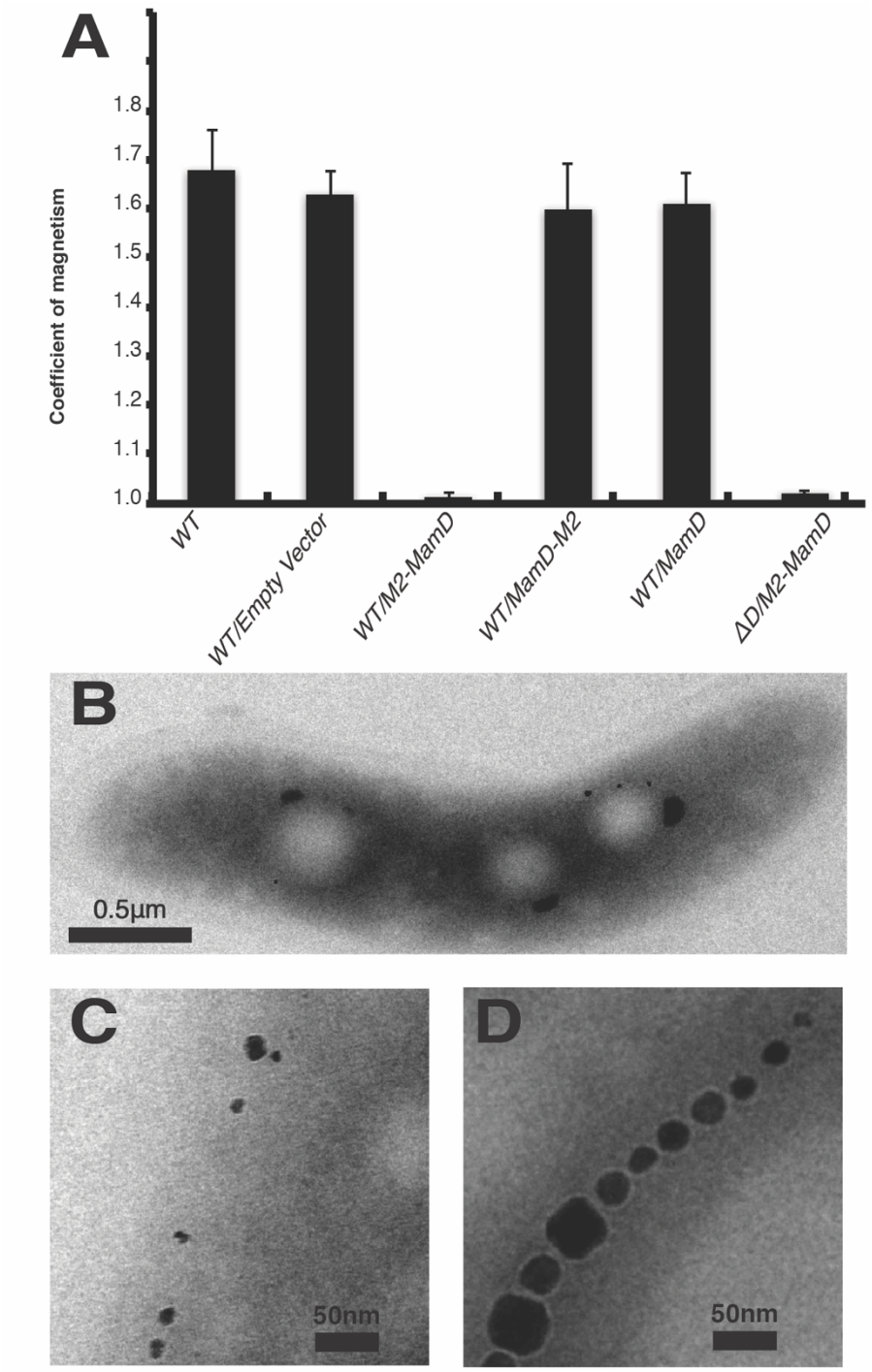


Figure 4-5 N-terminally tagged FLAG-MamD hinders biomineralization. A. Cmags of various strains, both strains containing FLAG-M2 have minimal Cmag. B,C. EMs of FLAG-MamD in an otherwise wildtype cell. D. WT magnetosomes.

.

Chapter 5.1

The investigation of the N-terminal domains of MAI proteins and how they relate to function and sorting of MAI proteins (unpublished data)

Introduction

One unanswered question of magnetosome biogenesis is how magnetosome associated proteins are sorted to magnetosomes. While several interactions between various MAI proteins have been characterized and some proteins, including MamE, have been implicated in localization. There is no identified "magnetosome signal" and the majority of magnetosome proteins are membrane proteins or soluble proteins predicted to reach the periplasm through the classic SecYEG system. No MAI proteins have predicted TAT signals.

Many of the proteins mentioned in the chapters above have a single N-terminal transmembrane helix with the larger C-terminal region predicted to be outside of the cytoplasm and assumed to be in the magnetosome. Specifically, MamE, MamT, and biomineralization protein MamS all have signals strongly predicted to be stable membrane anchors. MamP however, has a signal that is ambiguously predicted to be either a stable membrane anchor or one where once translocated across the membrane, would be cleaved.

In order to determine whether the n-terminal region containing the predicted signal sequences were required for correct function and further understand the ambiguous nature of MamP's signal a series of plasmids were constructed where a BamHI sequence was inserted three or four bases downstream of where the predicted transmembrane helix ends. BamHI was selected due to its coding for a glycine-serine pair of amino acids, which is unlikely to disrupt the function of a protein. All proteins with the inserted BamHI signal (MamE, MamT, and MamP) complemented their deletions when integrated into AMB-1's genome.

A series of swaps were then made with the n-terminal signal of one protein would be replaced with the N-terminal of different MAI proteins. MamS's N-terminal TM was also tested. Additionally, the N-terminal regions of additional non-MAI proteins were also tested. Specifically, the AMB-1 homolog of PbpC was selected as an example of stable membrane anchors, TolB*,* was selected as a model for a soluble periplasmic protein that uses the SecYEG channel, and NapA was selected due to its TAT signal. Finally, additional plasmids were constructed where the N-terminal region was completely deleted in order to test what happens when the proteins remained in the cytoplasm. While not every potential hybrid was constructed, several clear trends were uncovered and MamE's potential central importance to magnetosome biogenesis was underscored.

Materials and Methods

Strains, plasmids and culture conditions

The strains and plasmids used in this study are listed in Tables 5-1 and 5-2, respectively. For general maintenance and genetic manipulation, *M. magneticum* AMB-1 was grown in MG medium supplemented with ferric malate (30μM). 0.7% agar was used in plates and kanamycin was used for antibiotic selection at a concentration of 7 μg/mL (solid) or 10μg/mL (liquid). Cultures for magnetic response measurements and western blotting were grown in 10 mL MG medium (pH 6.9) and ferric malate under a 10% oxygen atmosphere. The magnetic response of each culture was assessed using the Coefficient of Magnetism (Cmag),

which was measured as described (56).

Genetic manipulation
For complementation of deletion mutants, we used a modified form of pAK669, a modified version of pAK605 where the BamH1 in the multi-cloning site was removed by quickchange. This plasmid contains a neutral region of the AMB-1 genome and integrates as a single copy at this site. Each allele is inserted under the control of the *mamAB* promoter, allowing the constitutive expression of each protein.

Protein analysis
Cultures of AMB-1 were grown to late-log phase and harvested by centrifugation at 8k x *g*. The resulting pellets were resuspended in 2x SDS loading buffer and heated for 10 min at 95°C. The lysates were separated onSDS-PAGE and transferred to PVDF. The membranes were blotted and visualized using standard western blotting techniques. Polyclonal antibodies to MamE, MamP, and MamT, and were raised in rabbits against recombinant forms of the soluble portion of each protein.

Results
MAI protein need to reach and fold inside the periplasm to function correctly
None of the hybrid proteins where the N-terminal regions were completely removed showed any signs of complementation. Additionally, all hybrids containing the NapA N-terminal TAT signal demonstrated no complementation through Cmag (Figure 5-1 and Table 5-2). Because TAT proteins fold completely within the cytoplasm before being translocated, it is not surprising that these proteins did not function. Furthermore, the cytoplasmic and NapA hybrid versions of MamE and MamP were not observed when viewed by western (Figure 5-3).

The specific N-terminal regions of MamP and MamT are not required for function
While every potential pair of hybrid gene was not created (Table 5-1) all hybrids of MamP and MamT (other that than the cytoplasmic and NapA hybrids mentioned above) demonstrated at least some level of complementation of their respective deletions as tested through Cmag. This includes signals taken from either within the MAI or from PbpC and TolB, from outside of the MAI.

MamE's specific N-terminal signal is required for its function
Unlike the MamP and MamT, MamE did not demonstrate a flexible N-terminal region. While the positive control fully complemented *mamE*'s deletion, none of the hybrids did. The MamE hybrids containing the N-termini of MamP, MamS, MamT, PbpB, and NapA all had no magnetic response, identical to the *ΔmamE* strain. Interestingly the TolB/MamE hybrid did have a small magnetic response, far lower than wild-type but easily distinguishable from *ΔmamE,* due to the cell pellets being pliable when a magnet is applied outside of an Eppendorf tube, unlike the *mamE* deletion and all other hybrids.

Protein amounts and processing patterns can vary while still maintaining protein function
MamP, as discussed in chapters 3 and 4 exhibits a specific processing pattern depending on the genetic status of the cell. Wildtype AMB-1 cells have bands corresponding to full-length protein

and a lower C-terminal band. When MamE processing is disrupted, the lower band largely disappears while the full-length band is clearly enriched.

All of the MamP hybrids demonstrated clearly different patterns from both wildtype and the control strain, keeping MamP's N-terminus but containing the glycine-serine linker. The hybrids with N-termini of MamS, MamE, and PbpC all have larger amounts of full-length protein. The TolB hybrid, however, has no visible full-length protein. Strangely, the MamT hybrid showed no observable MamP at all, appearing to be a deletion. However, all of these strains had a clearly increased magnetic response from the *ΔmamP* strain (Figure 5-5).

Discussion

The localization and targeting of MAI proteins remains largely a mystery but this work demonstrates the centrality of MamE, one of the truly most well conserved MAI genes. In addition to its protease activity, the presence of its n-terminal region is absolutely required for magnetosome biogenesis. It is possible that this region is required for the MamE's proper protease function. Alternatively, it is also possible that the n-terminal region interacts laterally with other MAI proteins and its absence disrupts how many MAI proteins find each other.

Finally, the disparate appearances of full-length and processed MamP does somewhat elucidate whether native MamP is a membrane bound or soluble periplasmic protein. We know from previous work that some amount of MamP is cleaved by MamE in native cells. This work shows that even a strain (*tolB/P*) where the protein appears to be constitutively cleaved upon translocation into the periplasm is functional. Furthermore, the amount of full-length MamP can vary widely without any noticeable difference in function.

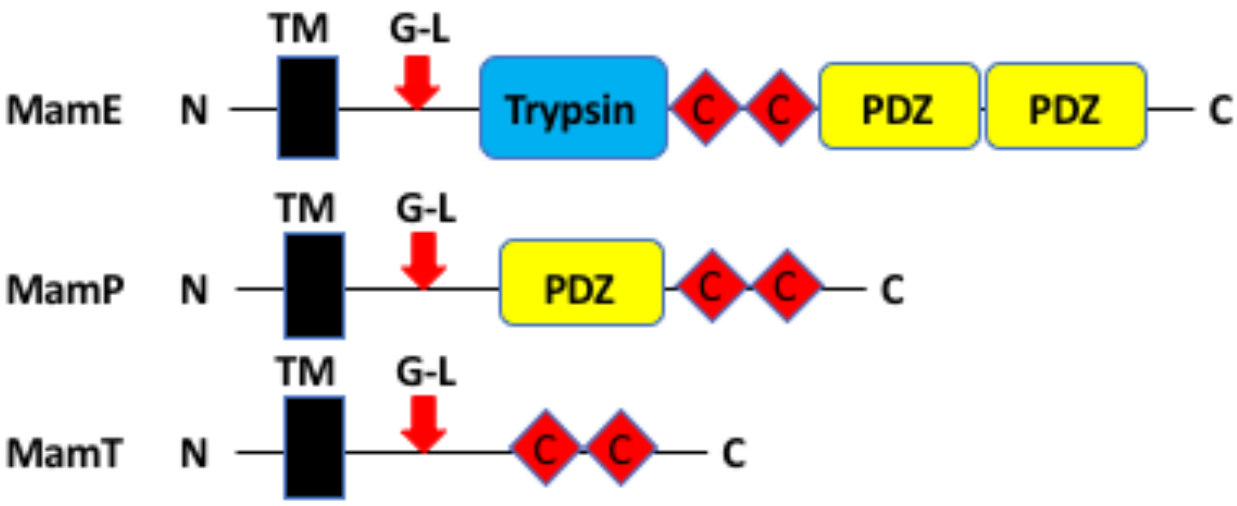


Figure 5-1 The Domain structures of MamE, MamP, and MamT. The red arrows labeled G-L refer to the insert points of a BamHI site.

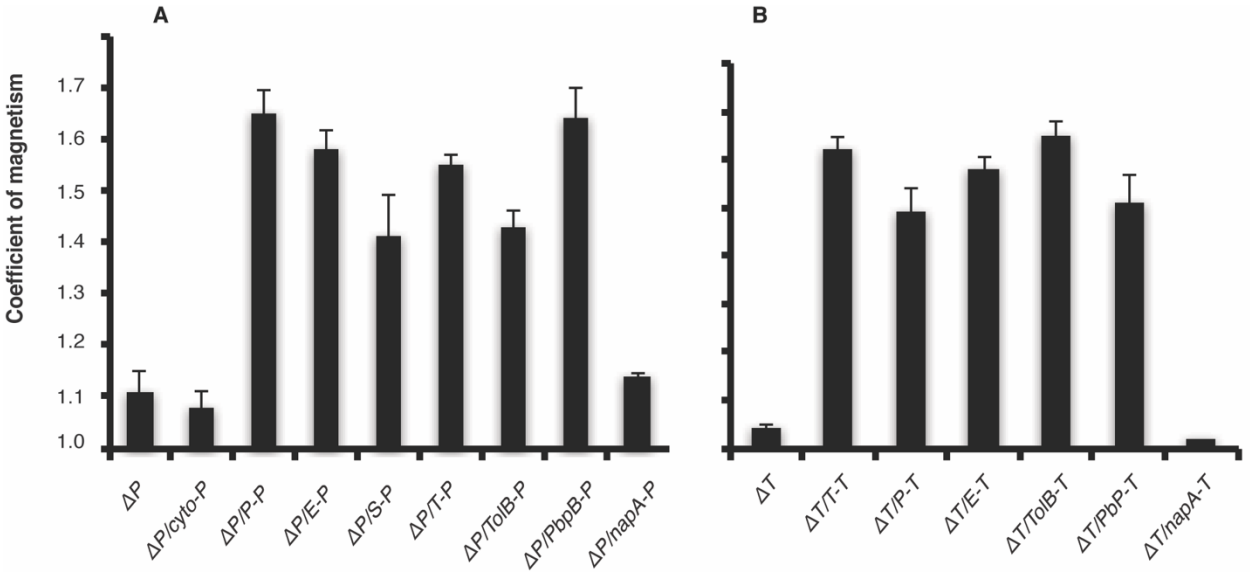


Figure 5-2. The N-terminal regions of MamP and MamT are replaceable A. Cmags of various hybrids of MamP, all in a *ΔmamP* background. B. Cmags of various hybrids of MamT, all in a *ΔmamT* background.

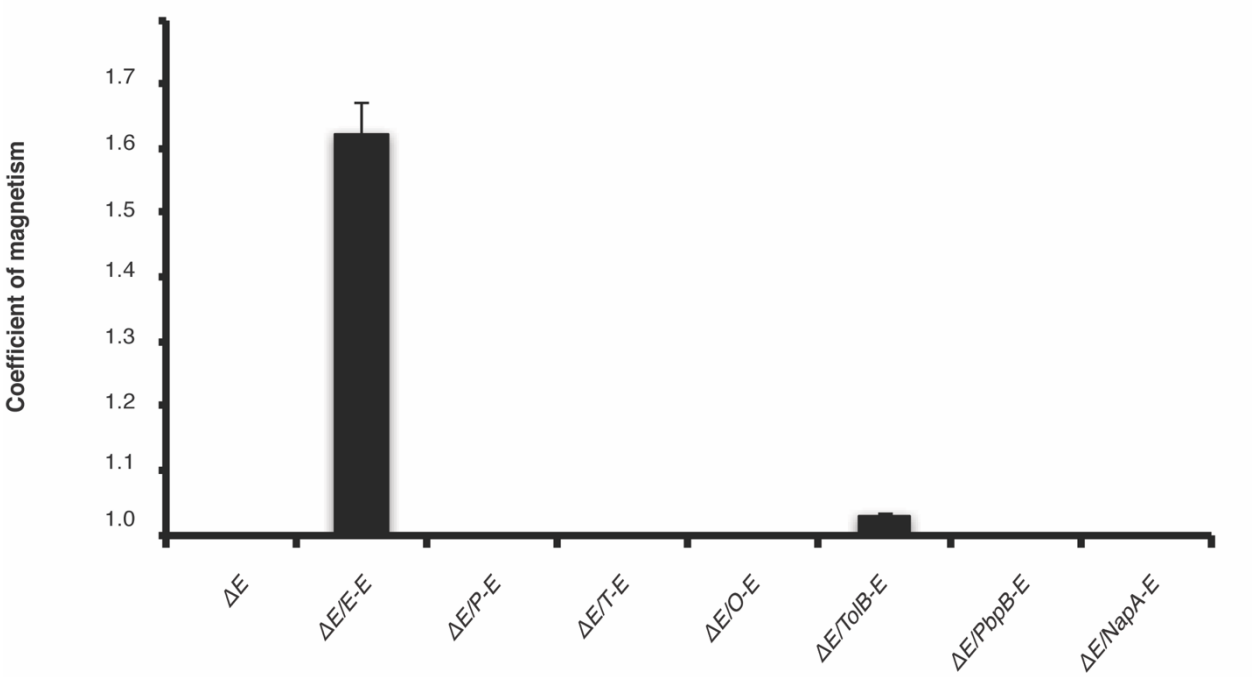


Figure 5-3 MamE's N-terminus is required for complete function. Only the hybrid protein with TolB demonstrates any complementation. Cmags for various hybrids are shown.

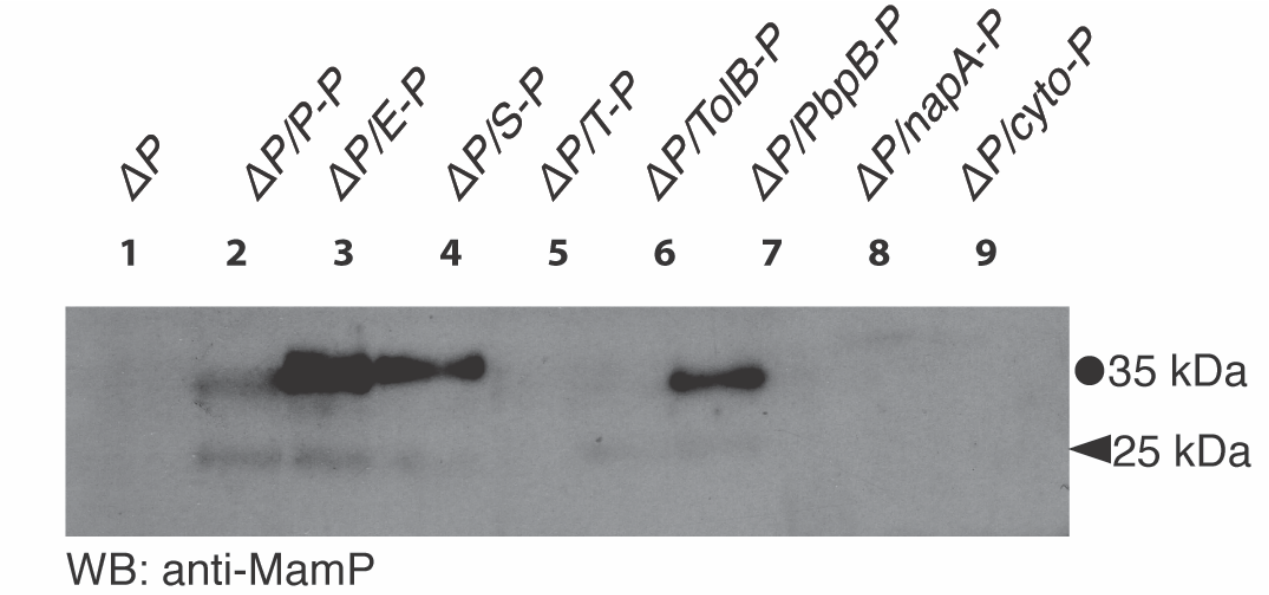


Figure 5-4 Modified versions of MamP display a variety of amounts of full-length and processed protein. Despite the fact that lanes 2-7 all have wildtype magnetic responses the amounts of processed and full-length protein vary wildly**.**

Chapter 5.2

Analysis of magnetochrome domains and how they relate to protein function (unpublished data)

Introduction

The magnetochrome domain is defined by two nearby CXXCH sequences and is found in all magnetite biomineralizing MTB, although how many magnetochrome domains and what specific genes they are found in vary from species to species. In AMB-1, *mamE, mamP, mamT,* and *mamX* all contain the magnetochrome domain. Previous work shows that the magnetochrome domain of the *mamE* is required for completely correct function as strains carrying a mutation in either of the CXXCH motifs display a lowered Cmag. However, these mutants display a mixture of mature wild-type looking magnetosomes as well as stunted ones. Furthermore, less full-length MamE was observed through western analysis. Therefore, it was speculated that *mamE's* magnetochrome domain aided either the correct folding of the protein or helped stabilize the protein. When disrupted MamE proteins could still carry out their functions but were simply less stable.

Here, we interrogate the magnetochrome domains of *mamT* and *mamP*. We show that the presence of complete and unaltered magnetochrome domains are absolutely required for the function of both proteins. However, the resulting mutant proteins seem altered in different ways with MamT appearing to not fold in a stable way, while MamP is observable but accumulates the full-length protein does not appear to be able to be processed, a necessary part of its function.

We also show that MamT is a potential target of MamE proteolysis. While stable processed versions of MamT are not observed, amounts of full-length protein greatly vary depending on the presence of an active MamE.

Materials and Methods

Strains, plasmids and culture conditions

The strains and plasmids used in this study are listed in Tables 5-3 and 5-4, respectively. For general maintenance and genetic manipulation, *M. magneticum* AMB-1 was grown in MG medium supplemented with ferric malate (30μM). 0.7% agar was used in plates and kanamycin was used for antibiotic selection at a concentration of 7 μg/mL (solid) or 10μg/mL (liquid). Cultures for magnetic response measurements and western blotting were grown in 10 mL MG medium (pH 6.9) and ferric malate under a 10% oxygen atmosphere. The magnetic response of each culture was assessed using the Coefficient of Magnetism (Cmag), which was measured as described (56).

Genetic manipulation

For complementation of deletion mutants, paK605. This plasmid contains a neutral region of the AMB-1 genome and integrates as a single copy at this site. Each allele is inserted under the control of the *mamAB* promoter, allowing the constitutive expression of each protein. Quickchange was used to alter the magnetochrome domains

Immunoblotting

Cultures of AMB-1 were grown to late-log phase and harvested by centrifugation at 8k x *g*. The resulting pellets were resuspended in 2x SDS loading buffer and heated for 10 min at 95°C. The lysates were separated on SDS-PAGE and transferred to PVDF. The membranes were blotted and visualized using standard western blotting techniques. Polyclonal antibodies to

MamP, and MamT, were raised in rabbits against recombinant forms of the soluble portion of each protein.

Results

Any mutation in the magnetochrome domains of mamT or mamP cause complete loss of function

Both *mamT* and *mamP* are considered biomineralization proteins. Their respective deletions result in strains with severe loss of magnetic response but do contain magnetic particles. While wild-type copies of mamT and mamP are fully able to complement their respective mutations, all strains containing any mutation in either of the CXXCH motifs cause complete loss of function. The Cmags of these strains are considerable lower than wildtype *AMB-1* and when given versions of the proteins containing mutated versions of either of their CXCCH motifs, no increase in Cmag is observed (Figure 5-5).

Heme mutations can cause stability issues and effect interactions with other MAI proteins

While the functional results of *mamT* and *mamP* magnetochrome proteins are identical, analysis by western blot shows differences in the final result mutated proteins. Magnetochrome modified MamT is not seen at all by western. Most likely, loss of the heme groups results in incorrect folding in the periplasm and quick removal by housekeeping proteases. This is a similar result to what has been previously observed in in *mamE* heme mutants (figure 5-6).

Heme mutations in *mamP* meanwhile however cause an opposite issue. Not only is full-length protein observed, far more of it seen than in wildtype strains. However, unlike wildtype strains, no processing is observed. These westerns are remarkably similar to the *ΔmamE* and $mamE^{PD}$ strains where MamE is not available to process MamP. (Figure 2-4, Figure 5-6).

MamT varies depending on the presence of MamE

Because MamP and MamE are both targets of MamE proteolysis it was investigated whether MamT could potentially be another target. Whole cell westerns revealed that full-length MamT did in fact vary depending on the presence of MamE. Furthermore, in a $mamE^{PD}$ strain MamT was also enriched. However, unlike MamP, MamO, and MamE, no stable processed forms of MamT were observed. It is possible that once processed MamT further degrades, similar to MamD, but it is also possible its varying amounts is simply a downstream result of not having a fully functioning magnetosome (Figure 5-7).

Discussion

This work underlies that not only is the magnetochrome domain conserved throughout MTB, its serves a crucial function in proper function of the genes containing them. Unlike the previous work done in *mamE,* it cannot be said that the heme binding motifs only serve an auxiliary role. Furthermore, it underlies the importance of MamE related processing and links the functionality of MamP with its processing.

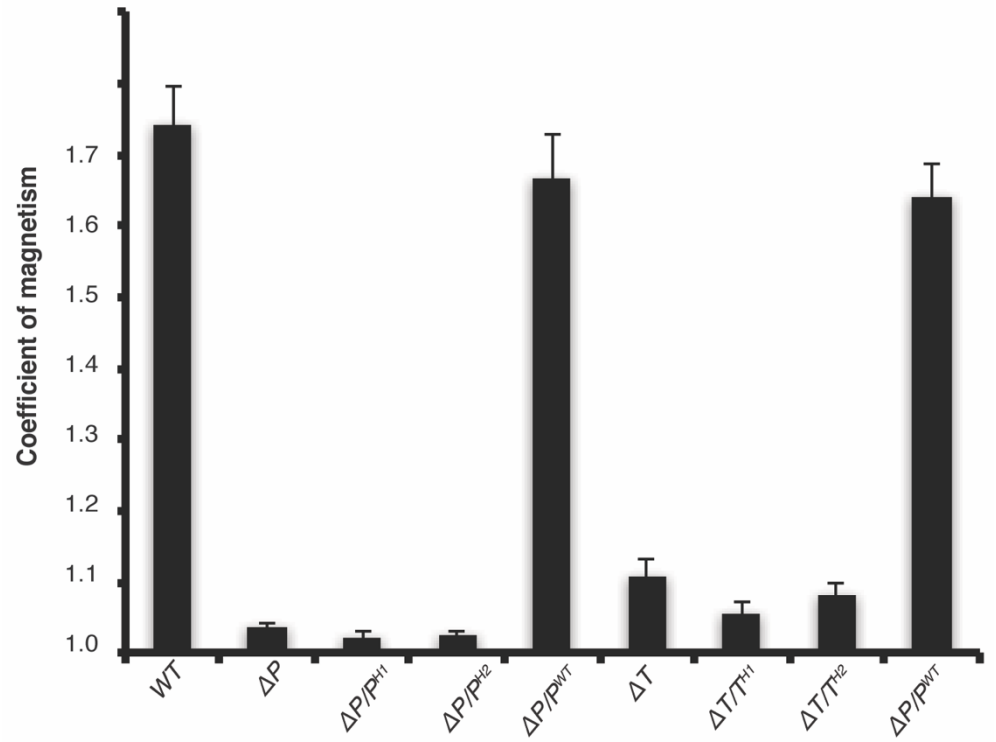


Figure 5-5 Mutation of either heme group causes complete loss of function in *mamT* and *mamP*

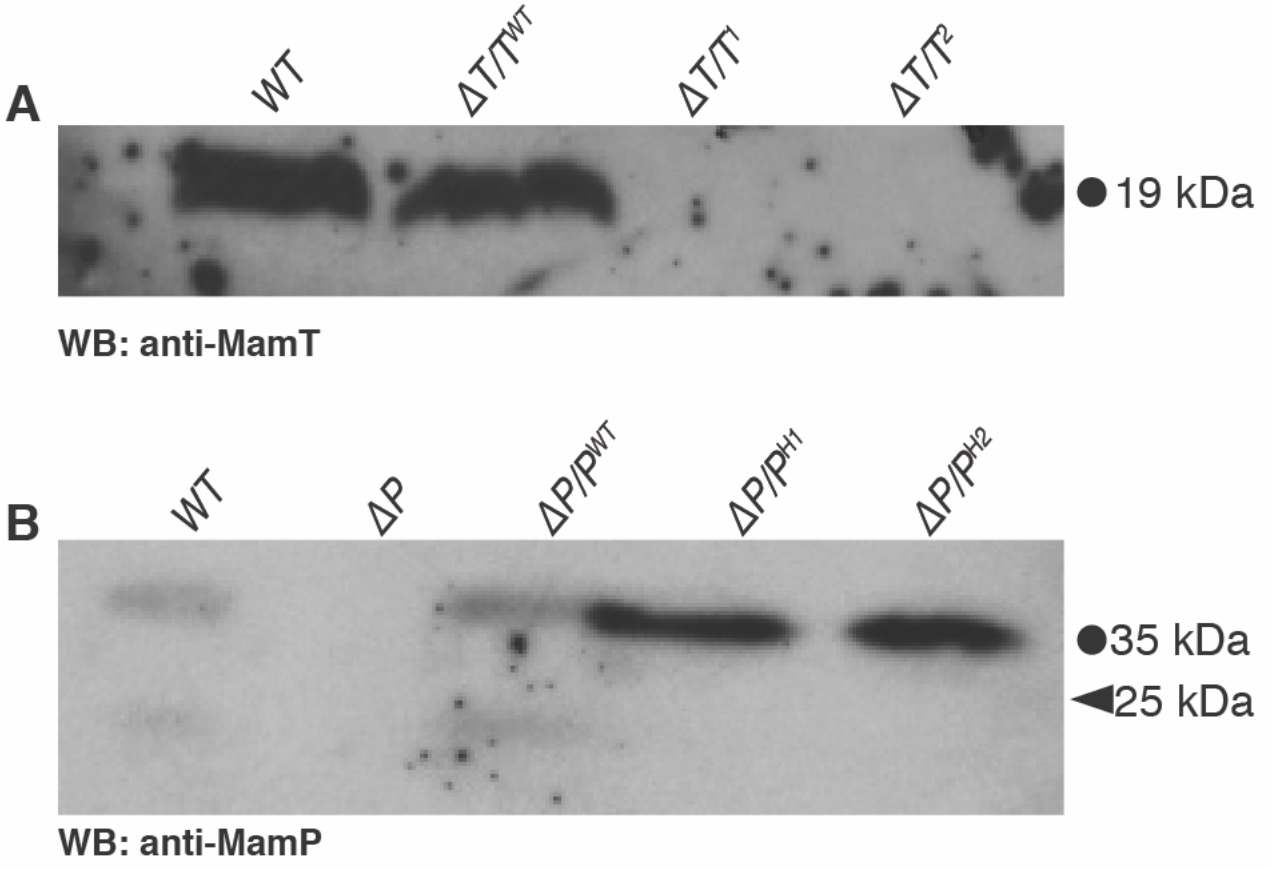


Figure 5-6 Mutation of either CXXCH motif in *mamT* and *mamP* disrupts the protein

A. Western of MamT showing that losing either CXXCH motif causes a complete loss of protein
B. Western of MamP showing that losing either CXXCH motif causes full-length protein to accumulate and not be properly processed as is necessary for function.

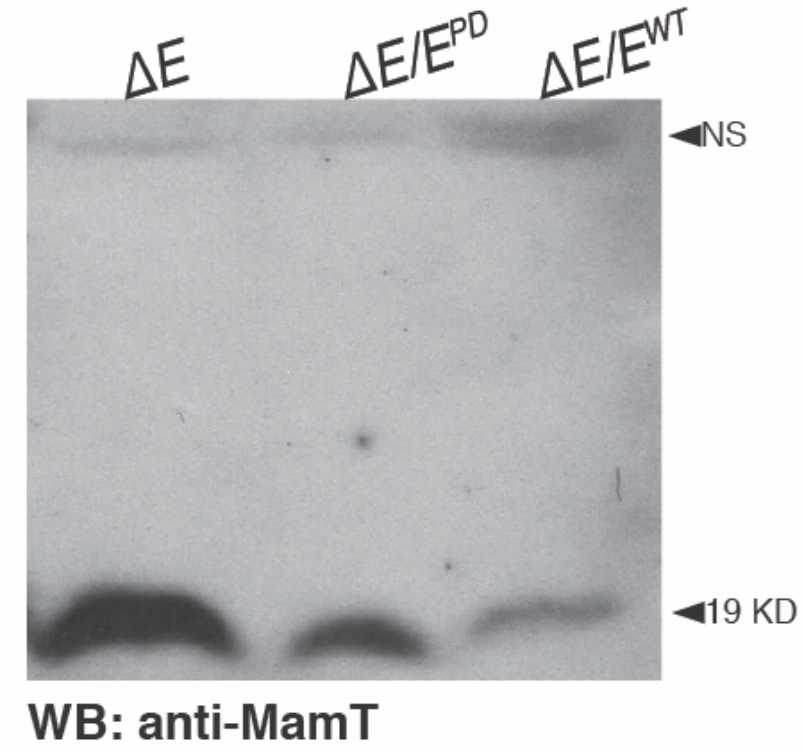


Figure 5-7 Full-length MamT (19KD) appears to be enriched in *mamE*^{PD} *ΔE* strains. However,

Chapter 5.3

Conclusions and perspectives

Patrick Browne

Introduction

This works began as an investigation of what MamE, an HtrA protease, conserved throughout all known magnetotactic bacteria, acted on and how this activity contributes to the biomineralization of the magnetosome. However, parallel work being done on a related magnetosome gene, MamO, also an apparent HtrA protease, demonstrated early on that the two proteins interacted. Further work showed that the idea that we would find one target of MamE and be able to show how processing that target impacted biomineralization was perhaps naïve.

The increasing complexity of magnetosome generation

Bacteria, long thought to be simpler cousins of eukaryotic life, are often found most useful because they contain a more manageable amount of information. However, recent work in microbial biology has increasingly underscored that bacteria can be as complex as eukaryotic life in many regards. They also represent a much larger diversity of life, in both metabolic and environmental exploitation contexts. Furthermore, many bacteria have been shown to form complex cell structures. While some bacterial structures are in fact relatively simple from a genetic standpoint, able to be traced to single operon with only a handful of genes, magnetosomes, as well as other structures have been shown to be remarkably complex.

The ability to create deletions in magnetotactic bacteria was a landmark achievement. It allowed for the identification of genes crucial to magnetosome development. It immediately demonstrated that uncovering how magnetosome are generated would be a difficult one, as well over a dozen genes played crucial roles. Furthermore, while all important genes were located in a single gene island, this island contained roughly a hundred genes. Further experiments have shown, however, that the complexity of magnetosome generation goes far beyond the number of necessary genes.

For example, MamK, a bacterial actin and MreB homolog, was one of the first MAI proteins to be better studied. While it can be conclusively shown that MamK is involved in maintenance of the magnetosome chain by forming filaments, even the most closely studied and very closely related AMB-1 and MSR-1 show marked differences in how those filaments are formed and maintained. AMB-1 has two MamKs that interact with each other while MSR-1 has only one. How they interact with MamJ, another protein involved in chain eminence also seems to vary.

MamO

MamO also is a fascinating example of how complex a role a single protein can play in forming a magnetosome. A multi-domain protein it was one of the few MAI proteins that could be phylogenetically identified as belonging to a protein family. One domain was clearly a trypsin like protease and the other was some sort of transporter. Also, its deletion resulted in a complete loss of the magnetosome crystal so it must be involved in magnetite nucleation. However, as shown in chapter two, it is far from simple.

First, it was conclusively shown that MamO cannot function as a protease due to mutations that physically confine it to one of the two states a trypsin must fluctuate between in order to function correctly. Instead, this domain seems to function as a metal binding domain. Both *in vitro* and *in vivo,* disruption of key sites, not found in other proteases, severely limits the protein's ability to bind metal. Secondly, despite being a pseudo protease, and the fact that it was shown early on to

play some role MamE dependent proteolysis, it appears that that it is MamO's second, proposed transporter, domain that plays a role in activating MamE. At this time, it is unknown what it is that is being transported into or outside of the magnetosome. Better understanding this transporter could provide great insight into how MamE functions as well as provide a better idea of the interior chemical environment of the magnetosome.

However, studying the evolutionary history of MamO uncovered a new layer of complexity. Previously believed to be among the better conserved magnetosome proteins it was shown to actually be a polyphyletic group of proteins. In four distinct occasions throughout the history of magnetotactic bacteria, serine proteases have been repurposed. In three of those cases it was the MTB's preexisting functional serine protease that was duplicated, but in *α-Proteobacteria* a wholly different protein seems to have been absorbed into the MAI. This work revealed that a ubiquitous protein family, one that is among the better understood protein families, is able to adopt completely different functions and showcases the HtrA family as a potential case study for protein evolution.

MamE

Prior to the work shown in this thesis, it was already known that MamE played multiple roles in magnetosome generation. While its deletion resulted in empty magnetosome, disruption of the catalytic triad in the protease domain resulted in stunted crystals. Finally, it was also shown to play a role in properly localizing some proteins to the MAI.

Chapters three and four partially completed the original goal of this project as both chapters identified targets of MamE proteolysis. Beginning the work, I began with the hypothesis that MamE was either activating a biomineralization factor by proteolysis or deactivating a biomineralization inhibitor. Overall, as many as eight have been identified. MamP, MamD, and MamE itself have all been confirmed *in vitro* and a peptide containing a region of MamO was also cleaved. While no conclusive data has been gathered for MamT or any of the three components of the tetrathionate reductase has been acquired, all four proteins are enriched in the $mamE^{PD}$ strain. MamO is involved in proper MamE dependent proteolysis, MamP and MamT are both biomineralization proteins, and MamD is a potential inhibitor of biomineralization. Its activity on itself is most likely regulatory. Finally, if it does act on Amb3286, Amb3288, and Amb3289 it might do so completely randomly. Alternatively, something about these horizontally transferred proteins might cause them to mislocalize to the magnetosome and MamE, as several other HtrA proteases might act as a house cleaning mechanism.

Taken together, this work shows that MamE not only plays a role in protein localization, crystal nucleation, and regulation, it also plays a role in several distinct stages of biomineralization. Chapter 5.2 also illustrated the n-terminal domain of MamE is crucial to its function, unlike several other MAI proteins.

Biomineralization factors

The work on *mamD* here further increases the number of known biomineralization factors as well as redefining the role of potential biomineralization factors. Because its deletion has such a

minimal effect on magnetosome biomineralization, MamD was not previously considered to be a major biomineralization protein. However, this work, combined with the recent work showing how the misregulation of MamD can lead to oddly shaped magnetosomes, shows that a protein can dramatically impact biomineralization even if its respective deletion causes no clear effect.

This work also underscored the complicated relationship between previously known biomineralization factors such as *mamP* and *mamT* have with other MAI proteins. While early work showed that these genes' deletions cause magnetosome crystals to become significantly malformed. However, this works shows that it is more complicated than the hypothesis that these proteins bind the crystal and simply help model and shape the crystal. They are regulated by MamE and processed at specific times during the "lifespan" of the magnetosome. They also contain heme-binding domains crucial to their function but how exactly these factors impact the stability of the proteins varies gene to gene. Finally, the proteomic work, showed that both of the factors, along with other MamAB biomineralization factors, MamR and MamS, are present in incredibly small amounts even compared to other MAI proteins. How these factors actually impact biomineralization remains an open question but it is likely through a complex regulatory manner and not through direct action with the crystal.

References

1. Blakemore, R. (1975) Magnetotactic Bacteria. *Science*. **190**, 377–379
2. Balkwill, D. L., Maratea, D., and Blakemore, R. P. (1980) Ultrastructure of a magnetotactic spirillum. *J Bacteriol*. **141**, 1399–1408
3. Weiner, S. (2008) Biomineralization: A structural perspective. *Journal of Structural Biology*. **163**, 229–234
4. Komeili, A. (2011) Molecular mechanisms of compartmentalization and biomineralization in magnetotactic bacteria. *FEMS Microbiology Reviews*. **36**, 232–255
5. Addadi, L., Joester, D., Nudelman, F., and Weiner, S. (2006) Mollusk Shell Formation: A Source of New Concepts for Understanding Biomineralization Processes. *Chem. Eur. J.* **12**, 980–987
6. Lefevre, C. T., and Bazylinski, D. A. (2013) Ecology, Diversity, and Evolution of Magnetotactic Bacteria. *Microbiology and Molecular Biology Reviews*. **77**, 497–526
7. Schubbe, S., Kube, M., Scheffel, A., Wawer, C., Heyen, U., Meyerdierks, A., Madkour, M. H., Mayer, F., Reinhardt, R., and Schuler, D. (2003) Characterization of a Spontaneous Nonmagnetic Mutant of *Magnetospirillum gryphiswaldense* Reveals a Large Deletion Comprising a Putative Magnetosome Island. *J Bacteriol*. **185**, 5779–5790
8. Fukuda, Y., Okamura, Y., Takeyama, H., and Matsunaga, T. (2006) Dynamic analysis of a genomic island in *Magnetospirillum sp.* strain AMB-1 reveals how magnetosome synthesis developed. *FEBS Letters*. **580**, 801–812
9. Schubbe, S., Williams, T. J., Xie, G., Kiss, H. E., Brettin, T. S., Martinez, D., Ross, C. A., Schuler, D., Cox, B. L., Nealson, K. H., and Bazylinski, D. A. (2009) Complete Genome Sequence of the Chemolithoautotrophic Marine Magnetotactic Coccus Strain MC-1. *Appl Environ Microbiol*. **75**, 4835–4852
10. Ullrich, S., Kube, M., Schubbe, S., Reinhardt, R., and Schuler, D. (2005) A Hypervariable 130-Kilobase Genomic Region of *Magnetospirillum gryphiswaldense* Comprises a Magnetosome Island Which Undergoes Frequent Rearrangements during Stationary Growth. *J Bacteriol*. **187**, 7176–7184
11. Richter, M., Kube, M., Bazylinski, D. A., Lombardot, T., Glockner, F. O., Reinhardt, R., and Schuler, D. (2007) Comparative Genome Analysis of Four Magnetotactic Bacteria Reveals a Complex Set of Group-Specific Genes Implicated in Magnetosome Biomineralization and Function. *J Bacteriol*. **189**, 4899–4910
12. Murat, D., Quinlan, A., Vali, H., and Komeili, A. (2010) Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. *Proc Natl Acad Sci U S A*. **107**, 5593–5598
13. Lohße, A., Ullrich, S., Katzmann, E., Borg, S., Wanner, G., Richter, M., Voigt, B., Schweder, T., and Schüler, D. (2011) Functional Analysis of the Magnetosome Island in *Magnetospirillum gryphiswaldense*: The mamAB Operon Is Sufficient for Magnetite Biomineralization. *PLoS ONE*. **6**, e25561–10
14. Faivre, D., and Godec, T. U. (2015) From Bacteria to Mollusks: The Principles Underlying the Biomineralization of Iron Oxide Materials. Angew. Chem. 127, 4810–4829
15. Lefèvre, C. T., Trubitsyn, D., Abreu, F., Kolinko, S., de Almeida, L. G. P., de Vasconcelos, A. T. R., Lins, U., Schüler, D., Ginet, N., Pignol, D., and Bazylinski, D. A.

(2013) Monophyletic origin of magnetotaxis and the first magnetosomes. Environ Microbiol. 15, 2267–2274
16. Grunberg, K., Muller, E. C., Otto, A., Reszka, R., Linder, D., Kube, M., Reinhardt, R., and Schuler, D. (2004) Biochemical and Proteomic Analysis of the Magnetosome Membrane in Magnetospirillum gryphiswaldense. Appl Environ Microbiol. 70, 1040–1050
17. Zarivach, R. (2014) Structure prediction of magnetosome-associated proteins. Front Microbiol. 10.3389/fmicb.2014.00009/abstract
18. Rahn-Lee, L., and Komeili, A. (2013) The magnetosome model: insights into the mechanisms of bacterial biomineralization. Front Microbiol. 10.3389/fmicb.2013.00352/abstract
19. Komeili, A., Vali, H., Beveridge, T. J., and Newman, D. K. (2004) Magnetosome vesicles are present before magnetite formation, and MamA is required for their activation. Proc Natl Acad Sci U S A. 101, 3839–3844
20. Komeili, A., Li, Z., Newman, D. K., and Jensen, G. J. (2006) Magnetosomes Are Cell Membrane Invaginations Organized by the Actin-Like Protein MamK. Science. 311, 242–245
21 Lohße, A., Borg, S., Raschdorf, O., Kolinko, I., Tompa, E., Posfai, M., Faivre, D., Baumgartner, J., and Schuler, D. (2014) Genetic Dissection of the mamAB and mms6 Operons Reveals a Gene Set Essential for Magnetosome Biogenesis in Magnetospirillum gryphiswaldense. J Bacteriol. 196, 2658–2669
22. Arakaki, A., Yamagishi, A., Fukuyo, A., Tanaka, M., and Matsunaga, T. (2014) Coordinated functions of Mms proteins define the surface structure of cubo-octahedral magnetite crystals in magnetotactic bacteria. Mol Microbiol. 93, 554–567
23. Clausen, T., Kaiser, M., Huber, R., and Ehrmann, M. (2011) HtrA proteases: regulated proteolysis in protein quality control. Nat Rev Mol Cell Biol. 12, 152–162
24. Yang, W., Li, R., Peng, T., Zhang, Y., Jiang, W., Li, Y., and Li, J. (2010) mamO and mamE genes are essential for magnetosome crystal biomineralization in Magnetospirillum gryphiswaldense MSR-1. Res Microbiol. 161, 701–705
25. Quinlan, A., Murat, D., Vali, H., and Komeili, A. (2011) The HtrA/DegP family protease MamE is a bifunctional protein with roles in magnetosome protein localization and magnetite biomineralization. Mol Microbiol. 80, 1075–1087
26. Jones, S. R., Wilson, T. D., Brown, M. E., Rahn-Lee, L., Yu, Y., Fredriksen, L. L., Ozyamak, E., Komeili, A., and Chang, M. C. Y. (2015) Genetic and biochemical 97 investigations of the role of MamP in redox control of iron biomineralization in Magnetospirillum magneticum. Proc Natl Acad Sci U S A. 112, 3904–3909
27. Otwinoski, Z., and Minor, W. (1997) Processing of X-Ray Diffraction Data Collected in Oscillation Mode. Meth Enzymol. 276, 307–326
28. Krojer, T., Garrido-Franco, M., Huber, R., Ehrmann, M., and Clausen, T. (2002) Crystal structure of DegP (HtrA) reveals a new protease-chaperone machine. Nature. 416, 455–459
29. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 66, 486–501
30. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S.,

Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. 66, 213–221
31. Yu, X., Wu, X., Bermejo, G. A., Brooks, B. R., and Taraska, J. W. (2013) Accurate High-Throughput Structure Mapping and Prediction with Transition Metal Ion FRET. Structure. 21, 9–19
32. Eddy, S. R. (1998) Profile hidden Markov models. Bioinformatics. 14, 755–763
33. Li, W., and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 22, 1658–1659
34. Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research. 32, 1792–1797
35. Castresana, J. (2000) Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. Mol Biol Evol. 17, 540–552
36. Sassera, D., Lo, N., Epis, S., D'Auria, G., Montagna, M., Comandatore, F., Horner, D., Pereto, J., Luciano, A. M., Franciosi, F., Ferri, E., Crotti, E., Bazzocchi, C., Daffonchio, D., Sacchi, L., Moya, A., Latorre, A., and Bandi, C. (2011) Phylogenomic Evidence for the Presence of a Flagellum and cbb3 Oxidase in the Free-Living Mitochondrial Ancestor. Mol Biol Evol. 28, 3285–3296
37. Price, M. N., Dehal, P. S., and Arkin, A. P. (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLoS ONE. 5, e9940
38. Abascal, F., Zardoya, R., and Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics. 21, 2104–2105
39. Stamatakis, A. (2014) RAxML Version 8: a Tool for Phylogenetic Analysis and PostAnalysis of Large Phylogenies. Bioinformatics. 30, 1312–1313
40. Lartillot, N., Lepage, T., and Blanquart, S. (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinformatics. 25, 2286–2288
41. Siponen, M. I., Legrand, P., Widdrat, M., Jones, S. R., Zhang, W.-J., Chang, M. C. Y., Faivre, D., Arnoux, P., and Pignol, D. (2014) Structural insight into magnetochromemediated magnetite biomineralization. Nature. 502, 681–684
42. Perona, J. J., and Craik, C. S. (1997) Evolutionary Divergence of Substrate Specificity within the Chymotrypsin-like Serine Protease Fold. J Biol Chem. 272, 29987–29990
43. Sohn, J., Grant, R. A., and Sauer, R. T. (2010) Allostery Is an Intrinsic Property of the Protease Domain of DegS: IMPLICATIONS FOR ENZYME FUNCTION AND EVOLUTION. J Biol Chem. 285, 34039–34047
44. Wrase, R., Scott, H., Hilgenfeld, R., and Hansen, G. (2011) The Legionella HtrA homologue DegQ is aself-compartmentizing protease thatforms large 12-meric assemblies. Proc Natl Acad Sci U S A. 108, 10490–10495
45. Bobofchak, K. M., Pineda, A. O., Mathews, F. S., and Di Cera, E. (2005) Energetic and Structural Consequences of Perturbing Gly-193 in the Oxyanion Hole of Serine 98 Proteases. J Biol Chem. 280, 25644–25650
46. Lovell, S. C., Davis, I. W., Arendall, W. B., III, de Bakker, P. I. W., Prisant, M. G., and Richardson, J. S. (2003) Structure Validation by Cα Geometry: φ, ψ, and Cβ Deviation. Proteins. 50, 437–450

47. Carvalho, A. L., Sanz, L., Barettino, D., Romero, A., Calvete, J. J., and Romão, M. J. (2002) Crystal Structure of a Prostate Kallikrein Isolated from Stallion Seminal Plasma: A Homologue of Human PSA. J Mol Biol. 322, 325–337
48. Goettig, P., Magdolen, V., and Brandstetter, H. (2010) Natural and synthetic inhibitors of kallikrein-related peptidases (KLKs). Biochimie. 92, 1546–1567
49. Taraska, J. W., Puljung, M. C., Olivier, N. B., Flynn, G. E., and Zagotta, W. N. (2009) Mapping the structure and conformational movements of proteins with transition metal ion FRET. Nature Methods. 6, 532–537
50. Murat, D., Falahati, V., Bertinetti, L., Csencsits, R., Körnig, A., Downing, K., Faivre, D., and Komeili, A. (2012) The magnetosome membrane protein, MmsF, is a major regulator of magnetite biomineralization in Magnetospirillum magneticum AMB-1. Mol Microbiol. 85, 684–699
51. Weinitschke, S., Denger, K., Cook, A. M., and Smits, T. H. M. (2007) The DUF81 protein TauE in Cupriavidus necator H16, a sulfite exporter in the metabolism of C2 sulfonates. Microbiology. 153, 3055–3060
52. Mayer, J., Denger, K., Hollemeyer, K., Schleheck, D., and Cook, A. M. (2012) (R)-Cysteate-nitrogen assimilation by Cupriavidus necator H16 with excretion of 3-sulfolactate: a patchwork pathway. Arch Microbiol. 194, 949–957
53. Arakaki, A., Webb, J., and Matsunaga, T. (2003) A Novel Protein Tightly Bound to Bacterial Magnetic Particles in Magnetospirillum magneticum Strain AMB-1. J Biol Chem. 278, 8745–8750
54. Amemiya, Y., Arakaki, A., Staniland, S. S., Tanaka, T., and Matsunaga, T. (2007) Controlled formation of magnetite crystal by partial oxidation of ferrous hydroxide in 95 the presence of recombinant magnetotactic bacterial protein Mms6. Biomaterials. 28, 5381–5389
55. Espinosa-Cantú, A., Ascencio, D., Barona-Gómez, F., and DeLuna, A. (2015) Gene duplication and the evolution of moonlighting proteins. Front. Genet. 6, 1–7
56. Morrone, D., Hillwig, M. L., Mead, M. E., Lowry, L., Fulton, D. B., and Peters, R. J. (2011) Evident and latent plasticity across the rice diterpene synthase family with potential implications for the evolution of diterpenoid metabolism in the cereals. Biochem. J. 435, 589–595
57. Ben-David, M., Wieczorek, G., Elias, M., Silman, I., Sussman, J. L., and Tawfik, D. S. (2013) Catalytic Metal Ion Rearrangements Underline Promiscuity and Evolvability of a Metalloenzyme. J Mol Biol. 425, 1028–1038
58. Dellus-Gur, E., Toth-Petroczy, A., Elias, M., and Tawfik, D. S. (2013) What Makes a Protein Fold Amenable to Functional Innovation? Fold Polarity and Stability Trade-offs. J Mol Biol. 425, 2609–2621
59. Toth-Petroczy, A., and Tawfik, D. S. (2014) The robustness and innovability of protein folds. Current Opinion in Structural Biology. 26, 131–138
60. Kolinko, I., Lohße, A., Borg, S., Raschdorf, O., Jogler, C., Tu, Q., Pósfai, M., Tompa, É., Plitzko, J. M., Brachmann, A., Wanner, G., Müller, R., Zhang, Y., and Schüler, D. (2014) Biosynthesis of magnetic nanostructures in a foreign organism by transfer of bacterial magnetosome gene clusters. Nature Nanotechnology. 9, 193–197

61. Rahn-Lee, L., Byrne, M. E., Zhang, M., Le Sage, D., Glenn, D. R., Milbourne, T., Walsworth, R. L., Vali, H., and Komeili, A. (2015) A Genetic Strategy for Probing the Functional Diversity of Magnetosome Formation. PLoS Genet. 11, e1004811–18
62. Kashyap, S., Woehl, T. J., Liu, X., Mallapragada, S. K., and Prozorov, T. (2014) Nucleation of Iron Oxide Nanoparticles Mediated by Mms6 Protein in Situ. ACS Nano. 8, 9097–9106
63. Rawlings, A. E., Bramble, J. P., Walker, R., Bain, J., Galloway, J. M., and Staniland, S. S. (2014) Self-assembled MmsF proteinosomes control magnetite nanoparticle formation in vitro. Proc Natl Acad Sci U S A. 111, 16094–16099
64. Laura, R. P., Witt, A. S., Held, H. A., Gerstner, R., Deshayes, K., Koehler, M. F. T., Kosik, K. S., Sidhu, S. S., and Lasky, L. A. (2002) The Erbin PDZ Domain Binds with High Affinity and Specificity to the Carboxyl Termini of -Catenin and ARVCF. J Biol Chem. 277, 12906–12914
65. Wolff, A., Frese, K., Wißbrock, M., Eckstädt, K., Ennen, I., Hetaba, W., Löffler, S., Regtmeier, A., Thomas, P., Sewald, N., Schattschneider, P., and Hütten, A. (2012) Influence of the synthetic polypeptide c25-mms6 on cobalt ferrite nanoparticle formation. J Nanopart Res. 14, 1161–11
66. Wilken, C., Kitzing, K., Kurzbauer, R., Ehrmann, M., and Clausen, T. (2004) Crystal Structure of the DegS Stress Sensor: How a PDZ Domain Recognizes Misfolded Protein and Activates a Protease. Cell. 117, 483–494
67. Li, W., Srinivasula, S. M., Chai, J., Li, P., Wu, J.-W., Zhang, Z., Alnemri, E. S., and Shi, Y. (2002) Structural insights into the pro-apoptotic function of mitochondrial serine protease HtrA2/Omi. Nat. Struct Biol. 9, 436–441
68. Krojer, T., Sawa, J., Huber, R., and Clausen, T. (2010) HtrA proteases have a conserved activation mechanism that can be triggered by distinct molecular cues. Nat Struct Mol Biol. 17, 844–852
69. Kim, S., Grant, R. A., and Sauer, R. T. (2011) Covalent Linkage of Distinct Substrate Degrons Controls Assembly and Disassembly of DegP Proteolytic Cages. Cell. 145, 67–78
70. Sohn, J., Grant, R. A., and Sauer, R. T. (2007) Allosteric Activation of DegS, a Stress Sensor PDZ Protease. Cell. 131, 572–583
71. Krojer, T., Sawa, J., Schäfer, E., Saibil, H. R., Ehrmann, M., and Clausen, T. (2008) Structural basis for the regulated protease and chaperone function of DegP. Nature. 453, 885–890
72. Siponen, M. I., Legrand, P., Widdrat, M., Jones, S. R., Zhang, W.-J., Chang, M. C. Y., Faivre, D., Arnoux, P., and Pignol, D. (2014) Structural insight into magnetochromemediated magnetite biomineralization. Nature. 502, 681–684
73. Hershey, D. M., Ren, X., Melnyk, R. A., Browne, P. J., Ozyamak, E., Jones, S. R., Chang, M. C. Y., Hurley, J. H., and Komeili, A. (2016) MamO Is a Repurposed Serine Protease that Promotes Magnetite Biomineralization through Direct Transition Metal Binding in Magnetotactic Bacteria. PLoS Biol. 14, e1002402
74. Roepstroff, P. (1984) Proposal for a Common Nomenclature for Sequence Ions in Mass Spectra of Peptides. Biomed Mass Spectrom. 11, 601
75. Arslan, E., Schulz, H., Zufferey, R., Kunzler, P., and Thony-Meyer, L. (1998) Overproduction of the Bradyrhizobium japonicum c-Type Cytochrome Subunits of the

cbb3 Oxidase in Escherichia coli. Biochemical and Biophysical Research Communications
76. Held, H. A., and Sidhu, S. S. (2004) Comprehensive Mutational Analysis of the M13 Major Coat Protein: Improved Scaffolds for C-terminal Phage Display. J Mol Biol. 340, 587–597
77. Arita, Y., Allen, S., Chen, G., Zhang, W., Wang, Y., Owen, A. J., Dentinger, P., and Sidhu, S. S. (2016) Rapid isolation of peptidic inhibitors of the solute carrier family transporters OATP1B1 and OATP1B3 by cell-based phage display selections. Biochemical and Biophysical Research Communications. 10.1016/j.bbrc.2016.01.050
78. Londer, Y. Y., Pokkuluri, P. R., Orshonsky, V., Orshonsky, L., and Schiffer, M. (2006) Heterologous expression of dodecaheme "nanowire" cytochromes c from Geobacter sulfurreducens. Protein Expression and Purification. 47, 241–248
79. Mauldin, R. V., and Sauer, R. T. (2012) Allosteric regulation of DegS protease subunits through a shared energy landscape. *Nature Chemical Biology*. **9**, 90–96
80. Small, J. L., O'Donoghue, A. J., Boritsch, E. C., Tsodikov, O. V., Knudsen, G. M., Vandal, O., Craik, C. S., and Ehrt, S. (2013) Substrate Specificity of MarP, a Periplasmic Protease Required for Resistance to Acid and Oxidative Stress in *Mycobacterium tuberculosis*. *J Biol Chem*. **288**, 12489–12499
81. Fuh, G., Pisabarro, M. T., Li, Y., Quan, C., Lasky, L. A., and Sidhu, S. S. (2000) Analysis of PDZ Domain-Ligand Interactions Using Carboxyl-terminal Phage Display. *J Biol Chem*. **275**, 21486–21491
82. Zhang, Y., Appleton, B. A., Wu, P., Wiesmann, C., and Sidhu, S. S. (2007) Structural and functional analysis of the ligand specificity of the HtrA2/Omi PDZ domain. *Protein Sci.* **16**, 1738–1750
83. Runyon, S. T., Zhang, Y., Appleton, B. A., Sazinsky, S. L., Wu, P., Pan, B., Wiesmann, C., Skelton, N. J., and Sidhu, S. S. (2007) Structural and functional analysis of the PDZ domains of human HtrA1 and HtrA3. *Protein Sci.* **16**, 2454–2471
84. Ivarsson, Y., Arnold, R., McLaughlin, M., Nim, S., Joshi, R., Ray, D., Liu, B., Teyra, J., Pawson, T., Moffat, J., Li, S. S. C., Sidhu, S. S., and Kim, P. M. (2014) Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral phage peptidomes. *Proc Natl Acad Sci U S A*. **111**, 2542–2547
85. Lee, H.-J., and Zheng, J. J. (2010) PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun Signal*. **8**, 8–18
86. MohamedMohaideen, N. N., Palaninathan, S. K., Morin, P. M., Williams, B. J., Braunstein, M., Tichy, S. E., Locker, J., Russell, D. H., Jacobs, W. R., and Sacchettini, J. C. (2008) Structure and Function of the Virulence-Associated High-Temperature Requirement A of Mycobacterium tuberculosis. Biochemistry. 47, 6092–6102
87. Truebestein, L., Tennstaedt, A., Mönig, T., Krojer, T., Canellas, F., Kaiser, M., Clausen, T., and Ehrmann, M. (2011) Substrate-induced remodeling of the active site regulates human HtrA1 activity. Nat Rev Mol Cell Biol. 18, 386–388
88. Kley, J., Schmidt, B., Boyanov, B., Stolt-Bergner, P. C., Kirk, R., Ehrmann, M., Knopf, R. R., Naveh, L., Adam, Z., and Clausen, T. (2011) Structural adaptation of the plant protease Deg1 to repair photosystem II during light exposure. Nat Rev Mol Cell Biol. 18, 728–731

89. Kim, S., and Sauer, R. T. (2014) Distinct regulatory mechanisms balance DegP proteolysis to maintain cellular fitness during heat stress. Genes & Development. 28, 902–911
90. Cornejo, E., Subramanian, P., Li, Z., Jensen, G. J., and Komeili, A. (2016) Dynamic Remodeling of the Magnetosome Membrane Is Triggered by the Initiation of Biomineralization. mBio. 7, e01898–15–9
91. Bellini, S. *Ulteriori Studi Sui "Batteri Magnetosensibili."* (Pavia, 1963)
92. Frankel, R. B., Bazylinski, D. A., Johnson, M. S. & Taylor, B. L. Magneto-aerotaxis in marine coccoid bacteria. *Biophys. J.* **73**, 994–1000 (1997).
93. Lefévre, C. T. *et al.* Diversity of magneto-aerotactic behaviors and oxygen sensing mechanisms in cultured magnetotactic bacteria. *Biophys. J.* **107**, 527–538 (2014).
94. Grünberg, K., Wawer, C., Tebo, B. M. & Schüler, D. A large gene cluster encoding several magnetosome proteins is conserved in different species of magnetotactic bacteria. *Appl. Environ. Microbiol.* **67**, 4573–4582 (2001).
95. Lin, W. *et al.* Genomic insights into the uncultured genus 'Candidatus Magnetobacterium' in the phylum Nitrospirae. *ISME J.* **8**, 2463–2477 (2014).
96. Nakazawa, Hidekazu, Atsushi Arakaki, Sachiko Narita-Yamada, Isao Yashiro, Koji Jinno, Natsuko Aoki, Ai Tsuruyama et al. "Whole genome sequence of Desulfovibrio magneticus strain RS-1 revealed common gene clusters in magnetotactic bacteria." *Genome research* 19, no. 10 (2009): 1801-1808.
97. Schübbe, Sabrina, Timothy J. Williams, Gary Xie, Hajnalka E. Kiss, Thomas S. Brettin, Diego Martinez, Christian A. Ross et al. "Complete genome sequence of the chemolithoautotrophic marine magnetotactic coccus strain MC-1." *Applied and environmental microbiology* 75, no. 14 (2009): 4835-4852.
98. Kolinko, Sebastian, Michael Richter, Frank-Oliver Glöckner, Andreas Brachmann, and Dirk Schüler. "Single-cell genomics of uncultivated deep-branching magnetotactic bacteria reveals a conserved set of magnetosome genes." *Environmental microbiology* 18, no. 1 (2016): 21-37.
99. Descamps, Elodie CT, Caroline L. Monteil, Nicolas Menguy, Nicolas Ginet, David Pignol, Dennis A. Bazylinski, and Christopher T. Lefèvre. "Desulfamplus magnetovallimortis gen. nov., sp. nov., a magnetotactic bacterium from a brackish desert spring able to biomineralize greigite and magnetite, that represents a novel lineage in the Desulfobacteraceae." *Systematic and Applied Microbiology* (2017).
100. Komeili, Arash, Zhuo Li, Dianne K. Newman, and Grant J. Jensen. "Magnetosomes are cell membrane invaginations organized by the actin-like protein MamK." *Science* 311, no. 5758 (2006): 242-245.
101. Hug, Laura A., Brett J. Baker, Karthik Anantharaman, Christopher T. Brown, Alexander J. Probst, Cindy J. Castelle, Cristina N. Butterfield et al. "A new view of the tree of life." *Nature Microbiology* 1 (2016): 16048.
102. Cornejo, Elias, Poorna Subramanian, Zhuo Li, Grant J. Jensen, and Arash Komeili. "Dynamic remodeling of the magnetosome membrane is triggered by the initiation of biomineralization." *MBio* 7, no. 1 (2016): e01898-15.
103. Raschdorf, Oliver, Yvonne Forstner, Isabel Kolinko, René Uebe, Jürgen M. Plitzko, and Dirk Schüler. "Genetic and ultrastructural analysis reveals the key players and initial steps of bacterial magnetosome membrane biogenesis." *PLoS genetics* 12, no. 6 (2016): e1006101.

104. Uebe, René, Katja Junge, Verena Henn, Gabriele Poxleitner, Emanuel Katzmann, Jürgen M. Plitzko, Raz Zarivach et al. "The cation diffusion facilitator proteins MamB and MamM of Magnetospirillum gryphiswaldense have distinct and complex functions, and are involved in magnetite biomineralization and magnetosome membrane assembly." *Molecular microbiology*82, no. 4 (2011): 818-835.
105. Descamps, Elodie CT, Caroline L. Monteil, Nicolas Menguy, Nicolas Ginet, David Pignol, Dennis A. Bazylinski, and Christopher T. Lefèvre. "Desulfamplus magnetovallimortis gen. nov., sp. nov., a magnetotactic bacterium from a brackish desert spring able to biomineralize greigite and magnetite, that represents a novel lineage in the Desulfobacteraceae." *Systematic and Applied Microbiology* (2017).
106. Raschdorf, Oliver, Florian Bonn, Natalie Zeytuni, Raz Zarivach, Dörte Becher, and Dirk Schüler. "A quantitative assessment of the membrane-integral sub-proteome of a bacterial magnetic organelle." *Journal of Proteomics* (2017).
107. Ozyamak, Ertan, Justin Kollman, David A. Agard, and Arash Komeili. "The bacterial actin MamK in vitro assembly behavior and filament architecture." *Journal of Biological Chemistry*288, no. 6 (2013): 4265-4277.
108. Abreu, Nicole, Soumaya Mannoubi, Ertan Ozyamak, David Pignol, Nicolas Ginet, and Arash Komeili. "Interplay between two bacterial actin homologs, MamK and MamK-Like, is required for the alignment of magnetosome organelles in Magnetospirillum magneticum AMB-1." *Journal of bacteriology*196, no. 17 (2014): 3111-3121.
109. Toro-Nahuelpan, Mauricio, Frank D. Müller, Stefan Klumpp, Jürgen M. Plitzko, Marc Bramkamp, and Dirk Schüler. "Segregation of prokaryotic magnetosomes organelles is driven by treadmilling of a dynamic actin-like MamK filament." *BMC biology* 14, no. 1 (2016): 88.
110. Draper, Olga, Meghan E. Byrne, Zhuo Li, Sepehr Keyhani, Joyce Cueto Barrozo, Grant Jensen, and Arash Komeili. "MamK, a bacterial actin, forms dynamic filaments in vivo that are regulated by the acidic proteins MamJ and LimJ." *Molecular microbiology* 82, no. 2 (2011): 342-354.
111. Scheffel, André, Manuela Gruska, Damien Faivre, Alexandros Linaroudis, Jürgen M. Plitzko, and Dirk Schüler. "An acidic protein aligns magnetosomes along a filamentous structure in magnetotactic bacteria." *Nature* 440, no. 7080 (2006): 110-114.
112. Hershey, David M., Patrick J. Browne, Anthony T. Iavarone, Joan Teyra, Eun H. Lee, Sachdev S. Sidhu, and Arash Komeili. "Magnetite biomineralization in Magnetospirillum magneticum is regulated by a switch-like behavior in the HtrA protease MamE." *Journal of Biological Chemistry* 291, no. 34 (2016): 17941-17952.
113. Kerfeld, Cheryl A., Sabine Heinhorst, and Gordon C. Cannon. "Bacterial microcompartments." *Annual review of microbiology*64 (2010).
114. Uebe, René, and Dirk Schüler. "Magnetosome biogenesis in magnetotactic bacteria." *Nature Reviews Microbiology* 14, no. 10 (2016): 621-637.
115. Lefèvre, Christopher T., Denis Trubitsyn, Fernanda Abreu, Sebastian Kolinko, Christian Jogler, Luiz Gonzaga Paula Almeida, Ana Tereza R. Vasconcelos et al. "Comparative genomic analysis of magnetotactic bacteria from the Deltaproteobacteria provides new insights into magnetite and greigite magnetosome genes required for magnetotaxis." *Environmental microbiology* 15, no. 10 (2013): 2712-2735.

116. Fukuda, Yorikane, Yoshiko Okamura, Haruko Takeyama, and Tadashi Matsunaga. "Dynamic analysis of a genomic island in Magnetospirillum sp. strain AMB-1 reveals how magnetosome synthesis developed." *FEBS letters* 580, no. 3 (2006): 801-812.
117. Müller, Frank D., Oliver Raschdorf, Hila Nudelman, Maxim Messerer, Emanuel Katzmann, Jürgen M. Plitzko, Raz Zarivach, and Dirk Schüler. "The FtsZ-like protein FtsZm of Magnetospirillum gryphiswaldense likely interacts with its generic homolog and is required for biomineralization under nitrate deprivation." *Journal of bacteriology* 196, no. 3 (2014): 650-659.
118. Rioux, Jean-Baptiste, Nadege Philippe, Sandrine Pereira, David Pignol, Long-Fei Wu, and Nicolas Ginet. "A second actin-like MamK protein in Magnetospirillum magneticum AMB-1 encoded outside the genomic magnetosome island." *PLoS One* 5, no. 2 (2010): e9151.
119. Schultheiss, Daniel, René Handrick, Dieter Jendrossek, Marianne Hanzlik, and Dirk Schüler. "The presumptive magnetosome protein Mms16 is a poly (3-hydroxybutyrate) granule-bound protein (phasin) in Magnetospirillum gryphiswaldense." *Journal of bacteriology* 187, no. 7 (2005): 2416-2425.
120. Yamagishi, Ayana, Masayoshi Tanaka, Jos JM Lenders, Jarla Thiesbrummel, Nico AJM Sommerdijk, Tadashi Matsunaga, and Atsushi Arakaki. "Control of magnetite nanocrystal morphology in magnetotactic bacteria by regulation of mms7 gene expression." *Scientific reports* 6 (2016): 29785.
121. Taoka, Azuma, Chihiro Umeyama, and Yoshihiro Fukumori. "Identification of iron transporters expressed in the magnetotactic bacterium Magnetospirillum magnetotacticum." *Current microbiology* 58, no. 2 (2009): 177-181.