This is the peer-reviewed version of the following article:
Raschdorf O., Bonn F., Zeytuni N., Zarivach R., Becher D., and Schüler D. (2018). A quantitative assessment of the membrane-integral sub-proteome of a bacterial magnetic
organelle. Journal of Proteomics 172:89-99,
which has been published in final form at https://doi.org/10.1016/j.jprot.2017.10.007

A quantitative assessment of the membrane-integral sub-proteome of a bacterial magnetic organelle

Oliver Raschdorf[1a], Florian Bonn[2b], Natalie Zeytuni[3c], Raz Zarivach[3], Dörte Becher[2] and Dirk Schüler*[1,4]

Affiliations:

1: Department of Microbiology, Ludwig Maximilian University of Munich, Germany

2: Department of Microbiology, Ernst Moritz Arndt University of Greifswald, Germany

3: Department of Life Sciences, The National Institute for Biotechnology in the Negev and Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer Sheva, Israel

4: Department of Microbiology, University of Bayreuth, Germany

***** Corresponding author

[a] Current address: Thermo Fisher Scientific (formerly FEI Company), Eindhoven, Netherlands

[b] Current address: Institute of Biochemistry II, Goethe University School of Medicine, Frankfurt am Main, Germany.

[c] Current address: Department of Biochemistry and Molecular Biology, University of British Columbia, Canada

Keywords

Bacterial organelles; Magnetosomes; Membrane integral sub-proteome; Protein quantification

Abstract

Magnetotactic bacteria produce chains of complex membrane-bound organelles that direct the biomineralization of magnetic nanoparticles and serve for magnetic field navigation. These magnetosome compartments have recently emerged as a model for studying the subcellular organization of prokaryotic organelles. Previous studies indicated the presence of specific proteins with various functions in magnetosome biosynthesis. However, the exact composition and stoichiometry of the magnetosome subproteome have remained unknown.

In order to quantify and unambiguously identify all proteins specifically targeted to the magnetosome membrane of the Alphaproteobacterium *Magnetospirillum gryphiswaldense*, we analyzed the protein composition of several cellular fractions by semi-quantitative mass spectrometry. We found that nearly all genuine magnetosome membrane-integral proteins belong to a well-defined set of previously identified proteins encoded by gene clusters within a genomic island, indicating a highly controlled protein composition. Magnetosome proteins were present in different quantities with up to 120 copies per particle as estimated by correlating our results with available quantitative Western blot data. This high abundance suggests an unusually crowded protein composition of the membrane and a tight packing with transmembrane domains of integral proteins. Our findings will help to further define the structure of the organelle and contribute to the elucidation of magnetosome biogenesis.

Significance

Magnetosomes are one of the most complex bacterial organelles and consist of membrane-bounded crystals of magnetic minerals. The exact composition and stoichiometry of the associated membrane integral proteins are of major interest for a deeper understanding of prokaryotic organelle assembly; however, previous proteomic studies failed to reveal meaningful estimations due to the lack of precise and quantitative data, and the inherently high degree of accumulated protein contaminants in purified magnetosomes. Using a highly sensitive mass spectrometer, we acquired proteomic data from several cellular fractions of a magnetosome producing magnetotactic bacterium and developed a comparative algorithm to identify all genuine magnetosome membrane-integral proteins and to discriminate them from contaminants. Furthermore, by combining our data with previously published quantitative Western blot data, we were able to model the protein copy number and density within the magnetosome membrane. Our results suggest that the magnetosome membrane is specifically associated with a small subset of integral proteins that are tightly packed within the lipid layer. Our study provides by far the most comprehensive estimation of magnetosomal protein composition and stoichiometry and will help to elucidate the complex process of magnetosome biogenesis.

Highlights

- First quantitative and unbiased assessment of the membrane-integral sub-proteome of bacterial organelles (magnetosomes).
- A comparative algorithm allows to distinguish genuine magnetosome integral protein from contaminants acquired during the isolation process, and to identify putative novel magnetosome membrane constituents.
- Surface modelling suggests that a very specific subset of integral proteins is highly enriched and tightly packed within the magnetosome membrane.

Introduction

The Alphaproteobacterium *Magnetospirillum gryphiswaldense* and related magnetotactic bacteria form intracellular, membrane-bounded crystals of a magnetic mineral, the magnetosomes, which serve as magnetic sensors to help to direct bacterial swimming towards growth-favoring suboxic zones in the sediments of natural waters [1]. Magnetosome biosynthesis comprises the invagination of magnetosome membrane vesicles from the cytoplasmic membrane [2,3], in which conditions are properly controlled for the biomineralization of nano-sized crystals of the iron oxide magnetite. Nascent magnetosomes are then aligned into linear chains along cytoskeletal filaments to achieve one of the highest structural levels found in prokaryotic cells [4].

Magnetosome biosynthesis is thought to involve the sorting of a complex set of proteins to the magnetosome membrane [3–5]. First comparative analyses suggested the presence of specific magnetosome membrane proteins in various quantities which co-purified with magnetosome particles isolated by magnetic collection and ultracentrifugation [6–8]. Attempts to assess the magnetosome membrane proteome of *M. gryphiswaldense* by denaturing one- and two-dimensional gel electrophoresis followed by Edmann degradation and mass spectroscopy in combination with comparative genomic analysis identified a set of about 25 *bona fide* magnetosome proteins termed Mam (magnetosome membrane) and Mms (magnetosome membrane specific) [7,9,10] which are thought to have key functions in magnetosome membrane biogenesis, iron transport, nucleation and crystallization of magnetite as well as the assembly of magnetosome chains [4]. The corresponding genes are clustered within four operons of a hypervariable genomic magnetosome island (MAI), namely *mamABop* (17 genes), *mamGFDCop* (4 genes), *mms6op* (4 genes)

and *mamXYop* (4 genes) [10,11]. Other genes within the MAI but located outside these clusters were later also implicated in magnetosome formation (*mamF2*, *mamD2*, *feoAB1*, *mamW*; [10], [R. Uebe, manuscript in preparation]. However, in addition to *bona fide* magnetosome proteins, these MAI gene clusters also predict a number of further proteins which remained undetected in previous approaches, and it is unknown whether those represent genuine integral magnetosome membrane constituents or are just loosely attached, whether they are targeted exclusively to this compartment or also present in the cytoplasmic membrane, or in some cases, whether they are expressed at all.

In addition, a multitude of other proteins encoded outside the well-established MAI gene clusters were found to co-purify with isolated magnetosomes [8,9,12], and so far it has yet not been resolved if some of these proteins represent either further functional integral constituents native to the magnetosome membrane, or contaminants that become bound upon cell disruption and isolation [8,13].

Moreover, previous proteomic approaches indicated that Mam and Mms proteins are present in the magnetosome membrane in vastly different quantities. However, attempts to estimate the abundance of several integral magnetosome membrane proteins by Coomassie-stained band intensities in protein gels [9] were inherently inaccurate. Thus, the complete protein complement of the magnetosome membrane still remains unknown, and it is not understood how and in which stoichiometry magnetosome proteins are assembled to form the structural framework required for biomineralization and organization of functional magnetic organelles.

In this study we performed a highly sensitive, semi-quantitative mass spectrometry analysis of purified magnetosomes and several other cellular fractions and used a comparative algorithm to reveal the genuine membrane-integral magnetosome sub-

proteome and to confidently estimate the relative abundances of the individual proteins.

With one exception we detected all previously predicted Mam and Mms proteins and demonstrate that several of them are highly and specifically enriched in the magnetosome membrane. Furthermore, we identified several novel putative genuine magnetosome-membrane proteins, of which one (MGR_4114) is encoded within the MAI. Our results also indicate that most of the proteins detected within the magnetosome-membrane fraction, but encoded outside the MAI are likely contaminants from other cellular compartments.

Correlation of our semi-quantitative proteomic data with available quantitative Western blot data allowed us to approximate the absolute copy numbers of magnetosome membrane proteins within the organelle. The magnetite-nucleating Mms6 was identified as the most abundant membrane-integral magnetosome protein, followed by MamC and MamD, together accounting for >40% of all genuine magnetosome membrane proteins. Using the predicted topology of magnetosome proteins, we could further estimate the membrane coverage of integral proteins, which hints towards an unusually crowded protein organization within the magnetosome membrane. In summary, our data allowed the most accurate estimation of protein composition of the complex magnetosome membrane up to date and will contribute to uncover the processes involved in biogenesis of this sophisticated bacterial organelle.

Experimental Procedures

Cultivation and cell harvesting of M. gryphiswaldense, C_{mag} determinations

Bacterial strains and plasmids used in this study are listed in suppl. Table S 7. *E. coli* strains were cultivated in lysogeny broth (LB) medium. When necessary, kanamycin (Km) was added to 25 µg mL^{-1}. *E. coli* BW29427 and WM3064 cultures were supplemented with 1 mM DL-α,ε-diaminopimelic acid. Media were solidified by addition of 1.5% (w/v) agar. *M. gryphiswaldense* cultures were grown at 30°C in modified flask standard medium (FSM) [14]. When appropriate, Km was added to 5 µg mL^{-1}. Optical density (OD) and magnetic response (C_{mag}) of exponentially growing cultures were measured photometrically at 565 nm as described previously [15]. Conjugations of plasmids were performed essentially as described earlier [16,17].

Cellular fractionation and purification

Cultivation and all further fractionation steps were performed in independent triplicates as described in the workflow chart of Figure 1. *M. gryphiswaldense* was cultivated and scaled up to 4,5 L culture in closed 5 L-Schott bottles (500 ml air in headspace) over-night at 30°C and harvested by centrifugation (10,000 g, 15 min, 4°C). The cell pellets were washed in buffer containing 20 mM HEPES (pH 7.4) and 5 mM EDTA, and frozen at -20°C. All further steps were carried out at 4°C. The cell pellets obtained from cell harvesting were resuspended in buffer containing 50 mM HEPES (pH 7.4), 1 mM EDTA and complete protease inhibitor (Roche, Germany) and lysed by a high-pressure cell disruption system. Cellular storage polyhydroxybutyrate (PHB) granules were removed by centrifugation (210 g, 10 min 4°C) of the lysate. The lysate was applied on MACS cell separation column type CS (Miltenyi Biotec, Germany), magnetized with two neodymium-iron-boron cube

magnets (gravity flow). The flow-through was applied a second time on the column and then collected as **total nonmagnetic lysate (non-mag)**. The fraction bound to the separation column was washed with 50 mL extraction buffer [10 mM HEPES (pH 7.4), 1 mM EDTA, 0.1 mM PMSF], two times 50 mL salt buffer [10 mM HEPES (pH 7.4), 1 mM EDTA, 200 mM NaCl, 0.1 mM PMSF] and again 50 mL extraction buffer by gravity flow. The magnets were removed; the magnetic fraction eluted from the column with approx. 10 mL H_2O and supplemented to final HEPES (pH 7.4), EDTA and PMSF concentrations of the extraction buffer. An ultracentrifugation tube was filled with 60% (w/w) sucrose (in extraction buffer) and overlaid with the magnetic fraction. After ultracentrifugation (100,000 g, 1.5 h), pellet was resuspended in 2 mL extraction buffer as **magnetically separated magnetosome fraction (mag)**.

The total nonmagnetic cellular lysate was centrifuged for 1 h at 100,000 g and the membrane pellet resuspended (central small white PHB pellet was omitted) and incubated in carbonate buffer (200 mM Na_2CO_3, 10 mM EDTA, 1 mM PMSF, pH 11.0) for 30 min under mild shaking. After centrifugation for 1 h at 100,000 g, the pellet was resuspended in high-salt buffer (20 mM Tris, 1 M NaCl, 10 mM EDTA, 1 mM PMSF, pH 7.5) and incubated under mild shaking for 30 min. After centrifugation for 1 h at 100,000 g, the pellet was resuspended in 50 mM TEAB (pH 7.8) and immediately pelleted for 1 h at 100,000 g. The pellet formed the total **nonmagnetic membrane fraction (mem)**.

The magnetically separated magnetosomes (mag) were centrifuged for 30 min at 100,000 g and the pellet resuspended in carbonate buffer. Subsequent purification was performed as described for the membrane fraction, with 30 min centrifugation runs between washes. The resulting pellet formed the **stringently washed magnetosome (mag.str)** fraction (Figure 1).

SDS-PAGE, tryptic digestion and mass spectroscopy analysis

All triplicate fractions were treated independently. For sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), all liquid samples were supplemented and all pelleted samples were dissolved in 2x SDS sample buffer [0.125 M Tris (pH 6.8), 4% SDS, 2% glycerol, 10% 2-mercaptoethanol, 0.004% (w/v) Bromophenol blue] and heated at 60°C for 5 min. Appropriate amounts of samples were determined empirically (Fig S 2). Electrophoresis of the protein samples was performed on 12% polyacrylamide gels. Staining, in gel tryptic digestion and LC-MS/MS was performed according to [18] with minor modifications. In brief, the in-gel digested peptides were separated with an easy nLC 2 (Thermo Fisher Scientific, MA, USA) column and analyzed with an LTQ Orbitrap Velos (Thermo Fisher Scientific). The 20 most intense precursor ions of each full scan were selected for collision induced dissociation fragmentation. A list of all open reading frames (ORFs) from the draft genome sequence of *M. gryphiswaldense* [10], was used as target database, supplemented with entries of recently assigned ORFs. The resulting output files were compiled with Scaffold 4 (Proteome Software, OR, USA). Proteins were only considered as identified if at least two unique peptides, matching solid quality criteria ($\Delta cN > 0.1$ and XCorr > 2.2; 3.3; 3.7 for doubly, triply, or higher charged peptides) have been assigned, resulting in a false positive rate below 0.2% on protein level. Only two reverse decoy peptides were assigned in the screen, indicating good filter criteria. Spectral counts for these two peptides were omitted from further analysis. Spectral counts from known contaminants (e.g. human source, trypsin) were also excluded from further analysis

Proteinase K membrane shaving

If not otherwise noted, all steps were carried out at 4°C. A schematic description of the process is presented in Figure 3. A 1 mL batch of magnetically separated magnetosomes (mag) was centrifuged for 30 min at 100,000 g. The pellet was resuspended in carbonate buffer and incubated for 1 h under mild shaking. Urea was added to final concentration of 8 M. For protein reduction, tris(2-carboxyethyl)phosphine hydrochloride was added to final concentration of 5 mM and sample incubated for 45 min at 60°C. Alkylation was performed by addition of 10 mM iodoacetamide and incubation for 15 min at room temperature in the dark. Proteinase K was added to a final concentration of 3 µg/ml and the sample was proteolytically digested for 15 h at 37°C under mild shaking. Samples were supplemented with 5% acetonitril, cooled down on ice and centrifuged for 1 h at 100,000 g. Supernatant was removed and pellet overlaid with 50 mM TEAB solution before centrifugation for 1 h at 100,000 g. The resulting pellet was frozen at -70°C. Pellet was resuspended in 180 µL of digestion buffer [50 mM TEAB, pH 7.8, 0.5% RapiGest (Waters, MA, USA)] and incubated for 30 min at 30°C under mild shaking. 6 µg of Chymotrypsin and 10 mM $CaCl_2$ were added and solubilized sample digested for 7 h at 30°C with mild shaking. HCl was added to final concentration of 250 mM and sample incubated at 37°C for 45 min to precipitate detergent. Sample was repetitively centrifuged for 12 min at 20,000 g, until no magnetosome and membrane pellet was visible. The supernatant formed the **shaved magnetosome transmembrane peptide** fraction and was analyzed by LC-MS/MS as described in [19].

Mass spectrometry proteomics Raw data deposition

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [20] partner repository with the dataset identifier PXD006166.

Data analysis

All calculations described in the main text were performed in Microsoft Excel and are included in the supplements (suppl. File **1**). To develop our working model, we made the following assumptions**:**

(1) The **relative abundance** of a single protein in a complex sample can be estimated by normalizing the assigned peptide spectra for this specific protein with the total number of peptide spectra measured in this sample and with the molecular weight (MW) of the protein. This normalization allows a size-independent comparison of protein abundance over several fractions. The estimated abundance (A) of protein N is defined by

$$\boldsymbol{A_N} = \boldsymbol{SpC_N} \Big/ [(\textstyle\sum_{\boldsymbol{i=1}}^{\boldsymbol{n}} \boldsymbol{SpC_i}) \times \boldsymbol{MW_N}]$$

whereas

N is the protein index

SpC is the number of unique peptide spectra matching the protein

MW is the molecular weight of the protein

n is the total number of proteins identified in the screen of a single sample

For further calculations, the A_N of all detected proteins was independently determined in all samples and averaged over the three replicate fractions of a sample type.

(2) A **genuine integral magnetosome membrane protein** has to fulfill the following conditions:

a. The protein is relatively more enriched in the membrane fraction than in the total non-magnetic lysate, *i.e.* it is a membrane protein. Determined by:

$A_{mem(N)}/A_{non\text{-}mag(N)} \geq 1$

b. The protein is comparatively more [or by the factor of 'f' much more] enriched in the magnetosome membrane than in the non-magnetic membrane fraction of the cell. Determined by:

$A_{mag(N)}/A_{mem(N)} > 1 > f$

c. The protein becomes more highly enriched in magnetosomes that were depleted from associated proteins and contaminations by stringent washing. Determined by:

$A_{mag.str(N)}/A_{mag(N)} \geq 1$

d. Optional condition: The protein has comparatively very high relative abundance in the magnetosome membrane fraction. Determined by:

$A_{mag.str(N)}/A_{mag.str(MamC)} > T$ (The estimated abundance of the protein has to at least meet threshold *T*, when compared to a known highly abundant magnetosome protein, here MamC)

Calculations for magnetosome protein abundance and surface coverage model

Based on quantitative Western blots with MamC-GFP labeled magnetosomes and the correlation with magnetite crystal size and density, a conservatively estimated number of approximately 100 molecules (estimated range: 80 - 250) of MamC were suggested for an average-sized magnetosome [21]. By calculating the $A_{mag.str(N)}/A_{mag.str(MamC)}$ value for every protein of interest, we estimated its copy

number in an average magnetosome, assuming a MamC reference copy number of 100. To calculate the transmembrane domain (TMD) coverage of the magnetosome membrane, we made the following assumptions: i) A simple transmembrane helix (TMH), perpendicularly inserted into the lipid bilayer, has a cross-section diameter of at least 1.1 nm [22]. This is a very conservative assumption, since contributing large amino acid residues and different insertion angles might increase the cross-section diameter ii) TMHs are surrounded by boundary lipids that interact with the hydrophobic protein domain. The major phospholipids in the cytoplasmic and magnetosome membrane of *M. gryphiswaldense* are phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) [9]. The lipid head cross-section area and the diameter of dilauroyl-PE and dilauroyl-PG are 0.39 nm^2 (0.70 nm) and 0.43 nm^2 (0.74 nm), respectively [23]. On average, the diameter of a lipid head group in the magnetosome membrane can therefore be estimated to be 0.72 nm. Hence, an annular boundary lipid shell would increase the diameter of an embedded TMH to 2.5 and 4.0 nm for one and two rings of boundary lipids, respectively. iii) TMHs of multi-membrane spanning proteins are more packed and without internal lipid boundary layers. According to Jacobson *et al.* [22], a diameter of 2.4 nm can be assumed for the whole TMD of a tetraspan-protein, and a diameter of 3.2 nm for a heptaspan-protein, excluding boundary lipids (Figure 3). We interpolated these values for magnetosome membrane proteins that exhibit 1 to 18 predicted TMH [5] by power regression. The average diameter of a magnetosome vesicle from *M. gryphiswaldense* is 45.5 nm [3]; the spherical surface area therefore can be calculated to be approx. 6450 nm^2. Taking into account the number of predicted TMHs and the here estimated absolute copy numbers of genuine magnetosome proteins, we calculated the integral protein occupancy of the membrane (surface) for both the

scenarios that all TMH are isolated and that TMHs of a single protein form a packed TMD (Figure 3).

To control our quantification results, we also used the protein abundance index (PAI) as an alternative method for quantifications [24]. The method does not take into account the molecular weight of the protein for normalization, but rather the number of peptides that are theoretically generated by the utilized protease (here: trypsin) and also excludes peptides that are too small or big to be measured by mass spectroscopy. We calculated three different PAI values for all assigned Mam and Mms proteins (PAI I: assuming theoretical tryptic peptides between 600 and 5000 Da, PAI II: assuming theoretical tryptic peptides with 7 to 25 amino acids and maximum molecular weight of 5000 Da and PAI III: assuming the number of tryptic peptides that were actually detected in our analysis). See also suppl. File **1**.

Molecular and genetic techniques

Oligonucleotides were purchased from Sigma-Aldrich (Germany) and are listed in suppl. Table S 8. Plasmids were constructed by standard recombinant techniques using enzymes from Thermo Fisher Scientific and Agilent Technologies (CA, USA) and confirmed using BigDye terminator v3.1 chemistry on an ABI 3700 capillary sequencer (Thermo Fisher Scientific). All plasmids used in this study are listed in suppl. Table S 7.

Plasmids pOR129 and pOR156 for markerless in-frame deletion of MGR_3691 and the MGR_4114 operon respectively, were created by PCR amplification and fusion of approximately 750 bp regions up- and downstream of the target sequences and cloning into pORFM. Genes were deleted as described in [25].

Magnetosome proteins were C- or N-terminally fused to enhanced green fluorescent protein (EGFP) and (over)expressed under control of the strong $P_{mamDC45}$ [21]

promotor either from replicative plasmids or by transposon mediated random integration in the genome. In all cases, the two fusion proteins were separated by a 25 amino acid alpha-helical linker region [LA(EAAAK)$_4$AAA] (HL) [3,26]. Alternatively, in-frame genomic fusions were constructed. Replicative plasmids pOR079, pOR089 and pOR099 were constructed by replacing *mamI* in pOR075 by PCR-amplified *mamW*, *mamR* or *mms6* sequences, respectively, using restriction digestion. Similarly, pOR085, pOR087, pOR088, were constructed by replacing *mamQ* in pOR086 by *mmsF*, *mamR* and *mamE* sequences, respectively. The P_{mamDC}-*mamE*-HL-*egfp* sequence form pOR088 was further cloned into the transposable plasmid pBAM-1 by restriction digestion, creating pOR148. A chromosomal *mamA*-*GFP* fusion was generated as described in [25], using plasmid pOR068. The plasmid was created by exchanging flanking regions of *mamC* in pFM236, by approximately 750 bp homologous flanking regions of *mamA* using restriction digestion.

Microscopy methods

For fluorescence microscopy, 3 µl samples of *M. gryphiswaldense* over-night cultures were immobilized on 1% (w/v) agarose pads with FSM medium salts. The samples were imaged with an BX81 microscope (Olympus, Japan) equipped with a 100×UPLSAPO100XO 1.4NA objective and an Orca-ER camera (Hamamatsu, Japan) and appropriate filer sets using Olympus Xcellence software. For transmission electron microscopy (TEM), unstained formaldehyde-fixed (0.075% w/v) *M. gryphiswaldense* cells were absorbed on carbon coated copper grids. Bright field TEM was performed on a Phillips (Netherlands) CM200 instrument using an accelerating voltage of 160 kV. Images were captured with an Eagle 4k CCD camera using EMMenu 4.0 (Tietz, Germany).

Results and Discussion

The four following different cellular fractions from *M. gryphiswaldense* were prepared for comparative mass spectroscopy analysis: (1) **Magnetically separated magnetosomes (mag**) following previously suggested protocols [27]; (2) **stringently washed magnetosomes (mag.str)**, additionally washed in high-salt and alkaline carbonate buffer to deplete contaminating proteins; (3) total **non-magnetic lysate (non-mag)** obtained as flow-through of magnetic column separation; (4) total enriched **non-magnetic membrane fraction (mem)**, washed in high-salt and alkaline carbonate buffer (Figure 1). Analysis of the mass spectroscopy data in total showed peptides of 2237 unique proteins in all fractions and replicates, which would account for approximately 53% off all predicted open reading frames (ORFs) of *M. gryphiswaldense* [28], and over 1000 proteins in the magnetosome fraction, indicating that the number of magnetosome-associated proteins would be unreasonably high without adequate filter algorithms to identify the most realistic subset of genuine magnetosome-associated proteins. Peptides of 1135 proteins were detected in magnetically separated magnetosomes, 1027 proteins in the stringently washed magnetosome fraction, 2031 proteins in the total non-magnetic fraction and 1305 proteins in the enriched non-magnetic membrane fraction.

Establishing and evaluating criteria to identify the genuine integral magnetosome membrane sub proteome

Genuine integral magnetosome proteins are bound to the magnetosome membrane *in vivo* and are specifically and exclusively enriched in this compartment. We developed a working model to identify these genuine integral magnetosome proteins and to discriminate them from contaminates by simply comparing and weighting the estimated relative abundance *A* of all proteins found within four cellular fractions. To fulfill our criteria, the protein has to be (I) a membrane-associated protein (determined by $A_{mem}/A_{non\text{-}mag}$-ratio of ≥ 1), (II) strongly enriched in the magnetosome

membrane compared to the non-magnetic membrane fraction (determined by A_{mag}/A_{mem}-ratio of > 1), and (III) would not become depleted from the magnetosome by stringent washing (determined by $A_{mag.str}/A_{mag}$-ratio of ≥ 1) (For more information, see data analysis section in experimental procedures).

In total, only 81 proteins (of which 23 were Mam and Mms annotated) were assigned genuine integral magnetosome membrane proteins if the parameters of our model were set to $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} > 1$ and $A_{mag.str}/A_{mag} \geq 1$; (Table 1). Application of each individual criteria contributed to the exclusion of proteins from the list of genuine magnetosome proteins, i.e. without taking into account every term, the number of assigned magnetosome proteins, especially those that are not encoded by the *mam* and *mms* operons, would have been unreasonably high (up to 651) (Table 1 and

suppl. Table S **1**). On the other hand, when $A_{mem}/A_{non\text{-}mag}$ and $A_{mag.str}/A_{mag}$ were set above the value of ≥ 1, e.g. to ≥ 3 and ≥ 2, respectively, the number of assigned proteins, and especially of those encoded outside the well-established *mam* and *mms* gene clusters shrank dramatically, indicating that the two terms should not exceed ≥ 1 (

suppl. Table S **1**).

Accordingly, the total number of assigned proteins further decreased when the A_{mag}/A_{mem} ratio was set more stringently, while the number of assigned Mam/Mms magnetosome membrane proteins only decreased by two when A_{mag}/A_{mem} was changed from > 1 to > 4 (Table 1 and suppl. Table S 2), indicating that the more stringent conditions are still sufficient to identify experimentally confirmed magnetosome membrane proteins. Therefore, parameter sets of $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} \geq 4$ and $A_{mag.str}/A_{mag} \geq 1$ seemed to be appropriate to predict the most likely genuine magnetosome proteins.

By comparing the estimated abundance of individual proteins within the stringently washed magnetosome fraction, the set of genuine magnetosome proteins might be

further refined. MamC was previously suggested to be the most abundant protein in the magnetosome membrane [7,29]. An $A_{mag.str(N)}/A_{mag.str(MamC)}$-threshold of 0.01 would indicate that for 100 estimated copies of MamC, at least one copy of the protein of interest N has to be present in the magnetosome membrane. The effect of supplementing the aforementioned conditions by the filter $A_{mag.str(N)}/A_{mag.str(MamC)} > T$ (T of 0.1, 0.01 or 0.001) to eliminate low abundant proteins is visualized in suppl. Table S 3. A threshold *T* of 0.1 further reduced the number of predicted non-magnetosome membrane proteins by 38 for $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} \geq 1$, $A_{mag.str}/A_{mag} \geq 1$ and by 4 for the more restrictive $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} \geq 4$, $A_{mag.str}/A_{mag} \geq 1$, while being alone not sufficient for adequate filtering (suppl. Table S 3). In combination with the parameters $A_{mem}/A_{non\text{-}mag} \geq 0$, $A_{mag}/A_{mem} \geq 4$, $A_{mag.str}/A_{mag} \geq 1$, a $A_{mag.str}/A_{mag.str(MamC)}$ threshold *T* of 0.01 predicted only 30 genuine magnetosome membrane proteins, of which 22 were previously identified Mam/Mms proteins. These values are comparable to the effects of $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} \geq 4$, $A_{mag.str}/A_{mag} \geq 1$ without MamC abundancy threshold and particular interesting as an alternative because the $A_{mem}/A_{non\text{-}mag}$ coefficient could potentially also exclude proteins that are so highly affine to the magnetosome membrane that they are completely undetectable in the nonmagnetic membrane fraction.

Mam and Mms proteins comprise the major fraction of the genuine magnetosome membrane proteome

With the exception of the small MamL, all other proteins encoded by the *mam* and *mms* gene clusters were identified in our proteomic data, including the recently identified MamF2 and MamD2 as well as MamX and MamI, which escaped detection in previous proteomic studies [13]. MamL is a small (approx. 9 kDa) protein that contains two predicted hydrophobic transmembrane domains. Although

two predicted tryptic peptides are in the theoretically detectable mass range, they were not detected in any of the samples of this study. However, as indicated by studies with MamL-GFP fusions and the strong magnetosome phenotype of gene deletion [3,30], the MamL protein is expressed and at least partially targeted to the magnetosome membrane, but was also never detected in previous proteomic analyses. In contrast, small proteins of comparable low mass such like MamR (approx. 8 kDa) and MamI (approx. 7 kDa) were confidently identified in this study. MamI also comprises two predicted trans-membrane domains, but more predicted tryptic peptides than MamL. The absence of MamL peptides from the data is an indication that for unknown reasons some proteins might be underrepresented or false-negatives in our screen.

With the stringent parameter setting of $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} \geq 4$, $A_{mag.str}/A_{mag} \geq 1$, the 21 assigned and also predicted genuine integral magnetosome-membrane proteins were: MamE, MamD, Mms6, MamO, MamM, MamC, MamB, MamY, MamP, MamF2, MmsF, MamF, MamS, MamH, MamT, MamZ, MamI, MamN, MamW, MamX and MamG. The two proteins MamQ and Mms48 (MGR_4070) met $A_{mem}/A_{non\text{-}mag} \geq 1$ and $A_{mag.str}/A_{mag} \geq 1$, but only showed an A_{mag}/A_{mem} value of 2.0 or 1.6 respectively, therefore barely escaping our classification as genuine magnetosome membrane proteins. It is thus possible that both proteins are indeed genuine magnetosome-membrane proteins that are similarly abundant in the cytoplasmic membrane; however, with an $A_{mag.str}/A_{mag.str(MamC)}$ value of 0.01, the abundance of Mms48 in the magnetosome membrane fraction is very low (i. e. 100 times less abundant than MamC). In the case of MamQ, this is consistent with the observation by microscopy that MamQ-GFP within the cells was mostly localized in the CM rather than the magnetosomes [3].

Several other Mam/Mms proteins also did not meet the criteria for genuine

magnetosome-membrane proteins. MamA, for example, failed by two criteria: Although its calculated A_{mag}/A_{mem} value of 51.2 was the second highest in the whole experiment, it did not fulfill the membrane protein threshold ($A_{mem}/A_{non\text{-}mag} = 0.6$) and was depleted from the magnetosome membrane in the purification process ($A_{mag.str}/A_{mag} = 0.3$). Consistent with its lack of predicted transmembrane helices (TMH), it is well-established that MamA only associates to the magnetosome surface and is not an integral part of the MM membrane, but becomes readily depleted with alkaline treatment [31–33]. Our results therefore are in good agreement with previous findings, and provide a further validation for the effectiveness of the selected parameter set. MamJ was excluded for the same reasons, however shows less clear parameter values ($A_{mem}/A_{non\text{-}mag}$ and $A_{mag.str}/A_{mag} = 0.9$). Since it also does not contain predictable transmembrane domains, MamJ *in vivo* is most likely strongly magnetosome-membrane associated, but not integral. The actin-like MamK protein was excluded since it was more abundant in the non-magnetic membrane (mem) than the magnetosome membrane (mag) and the total non-magnetic lysate (non-mag). The latter indicates that polymerized, high molecular weight MamK is either pelleted with the membrane fraction, or is bound to the cytoplasmic membrane.

Although MamD2 (like MamF2) was highly enriched in the magnetosome membrane, the protein did not meet the criteria since it became depleted from stringently washed magnetosomes ($A_{mag.str}/A_{mag} = 0.6$. The relevance of this finding remains unclear, but might indicate that the protein is not an integral magnetosome membrane protein, despite of its two predicted transmembrane domains. On the contrary, the 8 kDa MamR lacks any predicted transmembrane domain. However, it was virtually exclusively found in the magnetosome fraction, and could not be

depleted. With a $A_{mag.str(MamR)}/A_{mag.str(MamC)}$ value of 0.2, it was additionally highly abundant and therefore seems to be tightly magnetosome associated. Since it was undetectable in the non-magnetic membrane (mem) and almost undetectable in the soluble protein fraction (non-mag), MamR is an ambiguous case but formally had to be excluded. Although MamD and MamR are thus most likely no integral components, the proteins seem tightly bound to purified particles. Similar to Mms48, Mms36 was excluded because of its low A_{mag}/A_{mem} value of 1.0. Additionally, both proteins show a very low predicted abundance in the magnetosomes. Hence, Mms48 and Mms36, which were found to have a non-essential role in magnetite formation [34], rather are localized in the cytoplasmic membrane *in vivo*. The *mamXY* operon encoded protein FtsZm by far failed all criteria and was almost non-detectable in magnetosomes. It is therefore most likely also active in another cellular compartment *in vivo*, presumably the cytoplasm. Finally, the predicted soluble MamU was found to be expressed, but also failed to pass any of the set criteria and therefore is most likely not magnetosome associated *in vivo*, which is in line with the absence of a discernible magnetosome phenotype upon deletion of *mamU* [34]. The MAI-encoded iron transporter homologue FeoAB1 [R. Uebe, manuscript in preparation] was recently implicated in magnetite formation [35]. While FeoA1 was not found in our screen, FeoB1 was present with the same abundance in magnetic and non-magnetic membrane fraction (A_{mag}/A_{mem} = 1.0, $A_{mem}/A_{non\text{-}mag}$ = 2.7, $A_{mag.str}/A_{mag}$ = 1.2), and therefore not found to be specifically enriched in the magnetosome membrane, but might still be a potential constituent of it.

The predominant localizations of many Mam/Mms proteins in *M. gryphiswaldense* and other magnetotactic bacteria were already assessed by tagging and fluorescence microcopy analysis in previous studies [e.g 25,36–38]. In order to corroborate the findings of the proteome analysis, we investigated the intracellular localization of

some of those Mam/Mms proteins whose localization was not yet assessed previously in *M. gryphiswaldense*. GFP-labeled Mms6 and MmsF predominantly showed a strong linear fluorescent signal within the cells, resembling the organization of the magnetosome chain and indicating a strong enrichment in the magnetosome membrane (Figure 2A+B) as seen in our proteomic analysis. GFP-MamE also showed an accumulation of signal at mid-cell, however with higher cellular background (Figure 2C). Since MamW is only conserved in magnetospirilla, not part of one of the four major operons of the MAI, and its gene deletion did not shown any phenotype [13,30], its participation in magnetosome formation was yet not proved. The linear signal of the MamW-GFP fusion (Figure 2F) however is consistent with its proteomic detection and further suggests that MamW is specifically and genuinely magnetosome-integral. MamR and MamA GFP-fusions also showed a weak linear localization signal within the cell, along with a high cytoplasmic background (Figure 2D+E). These results corroborate the finding that both proteins were not assigned genuine integral magnetosome membrane proteins in this study, but were still found highly enriched in the magnetosome fraction, indicating strong magnetosome association.

Given the high sensitivity of detection, it is highly likely that most of the 1000 proteins that were detected in the magnetosome fraction represent contaminations from other cellular compartments, resulting from unspecific binding during cell disruption and purification. Comparable to previous proteomic studies, predicted outer membrane proteins and ATPase subunits were some of the most abundant proteins in the magnetosome membrane fraction (mag) [7–9] (Table 2). However, these proteins are known to be among the most abundant cytoplasmic membrane proteins in all bacteria [39] and did not meet our stringent filter criteria for genuine magnetosome-membrane proteins, showing that some highly abundant proteins in the magnetosome fraction are likely contaminants. Next, we assessed if novel candidates identified in our screen for genuine

magnetosome-membrane proteins could be linked to magnetosome biosynthesis: Besides the known Mam and Mms proteins discussed above, only few other proteins met our filter criteria or were highly enriched in the magnetosome membrane (Table 1 and **Fehler! Verweisquelle konnte nicht gefunden werden.**), but are not predicted to be magnetosome-associated (non Mam/Mms). Only one of those, the small (7 kDa) hypothetical transmembrane protein MGR_4114, is conspicuously encoded within the genomic magnetosome island, forming an operon together with two additional hypothetical proteins and one protein with similarities to ParA/MinD-like ATPases. However, our deletion of the entire operon failed to cause a discernible magnetosome phenotype (

Fig S **3**). Another candidate, MGR_3691 (also known as MM22 [9]) exhibited the highest magnetosome enrichment of all proteins identified in our screen (A_{mag}/A_{mem} = 64.8), but was not assigned genuine integral magnetosome protein, as it became depleted in stringently washed magnetosomes, and did not meet the integral membrane protein threshold ($A_{mag.str}/A_{mag}$ = 0.5, $A_{mem}/A_{non\text{-}mag}$ = 0.5). Indeed, deletion of the gene did also not cause any magnetosome phenotype (

Fig S **3**), indicating that it has no important function in magnetosome formation.

Among the most abundant proteins, the previously identified Mms16 (MGR_0659) was also identified, but not assigned genuine magnetosome protein (Table 2) due to an $A_{mem}/A_{non\text{-}mag}$ value of only 0.6 and an A_{mag}/A_{mem} value of 2.4. The protein was previously implicated in magnetosome formation [40], but later in fact shown to represent a phasin that rather functions in PHB metabolism of *M. gryphiswaldense* [41]. This indicates that contaminations can be successfully uncovered by our approach. While three of the other magnetosome assigned (according to our results), but non-magnetosome predicted proteins showed a $A_{mag.str(MamR)}/A_{mag.str(MamC)}$ below 0.01 (Table 2), five further proteins exhibited a value between 0.01 and 0.02. These are the hypothetical proteins MGR_2833, MGR_0916 and MGR_2730 as well as the histidine kinase MGR_0622 and the *ccb3*-type cytochrome oxidase maturation protein MGR_2552 (Table 2). Since their abundance in the magnetosome membrane

is very low in comparison with MamC and currently no functional connection to magnetosome formation is known for those proteins, it is highly likely that they represent false positives with respect to identified genuine magnetosome-membrane proteins.

Table 3 shows additional proteins that might be genuine magnetosome-membrane proteins, identified using less stringent filter parameters ($A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} \geq 1$, $A_{mag.str}/A_{mag} \geq 1$), but with a predicted abundance of at least two protein copies per magnetosome ($A_{mag.str(N)}/A_{mag.str(MamC)} \geq 0.02$). Besides MamQ, these proteins comprise putative translocases, peptidases, hypothetical proteins, acyl- or glycosyl-transferases. However, only MamQ shows a reasonably high A_{mag}/A_{mem} enrichment coefficient and also $A_{mag.str}/A_{mag.str(MamC)}$ abundance. Similarly, ignoring the A_{mag}/A_{mem} threshold, only MamR and a putative phasin could be additionally associated with high $A_{mag.str}/A_{mag.str(MamC)}$ abundance.

In summary, except few novel candidate proteins, of which only one showed higher confidence (MGR_4114), it therefore appears that the specific integral magnetosome-membrane sub-proteome is comprised of only previously predicted magnetosome-membrane proteins.

Proteinase K-shaved magnetosome membranes mostly contain MAI-encoded Mam/Mms protein fragments

To obtain additional information about the integral magnetosome-membrane proteome, we performed a Proteinase K membrane shaving assay on isolated magnetosomes, which should cleave off and digest all external protein domains, while buried transmembrane and membrane enclosed domains stay intact. Mass spectroscopy analysis of shaved magnetosomes membranes assigned peptide spectra to 61 proteins (suppl. Table S 4). Although we found previously undetected peptides

derived from predicted hydrophobic transmembrane helices of Mam and Mms proteins, most assigned spectra were from predicted soluble protein domains. However, among the 20 proteins with the highest exclusive un-normalized peptide spectra counts (SpC), 15 were Mam and Mms proteins (suppl. Table S 4). While 19% of all spectra were assigned to the magnetosome protein MamO, the top-10 proteins with highest SpC already covered 75% of all spectra (top-20 – 89%). Besides MamS and MamP, all identified genuine Mam and Mms magnetosome membrane proteins (according to Table 2) were also detected in the membrane shaving assay, however with highly variable SpC. Additionally, MamJ, MamA and MamD2 were also identified with comparably low SpC (suppl. Table S 4). Although with low SpC, only MGR_2730 and the MAI-encoded MGR_4114 were identified from the list of non-Mam/Mms proteins, but assigned genuine magnetosome proteins (Table 2). As in whole magnetosomes, MGR_3691 was again identified with a high SpC (suppl. Table S 4), emphasizing its role as a candidate magnetosome membrane protein. Six proteins previously not found in any of the analyzed fractions had assigned peptide spectra, of which MGR_1410, a predicted ammonia permease, had the highest SpC. Notably, also FeoB1 was detected with intermediate SpC. Most other identified proteins, most of them with a comparably low SpC, were components of transporters or of redox pathways and other proteins of the energy-metabolism. The results from the Proteinase K membrane shaving assay again suggest that magnetosome membranes are specifically enriched mainly with Mam and Mms proteins.

Predictions of protein stoichiometry suggest that the magnetosome membrane is densely packed with integral proteins.

Recently, based on quantitative Western blots with MamC-GFP labeled magnetosomes, a conservatively estimated number of approximately 100 molecules of MamC was suggested for an average sized magnetosome [21]. Based on this number and the calculated $A_{mag.str(N)}/A_{mag.str(MamC)}$-values, we estimated the putative copy numbers of the integral genuine magnetosome proteins for an average wild type magnetosome of 45.5 nm diameter [3] (Table 2). Within the membrane, TMHs of integral proteins are associated with one or two boundary lipid shells that interact with the hydrophobic protein domain. The diameter of the most prevalent lipid head groups in magnetosome membrane is approximately 0.72 nm. Hence, an annular boundary lipid shell would increase the diameter of an embedded TMH to 2.5 and 4.0 nm for one and two boundary lipid layers, respectively (Figure 3A and B). TMHs of multi-membrane spanning proteins might be packed without internal lipid boundary layers. According to Jacobson *et al*. [22], a diameter of 2.4 nm can be assumed for the whole transmembrane domain (TMD) of a packed tetraspan-protein, and a diameter of 3.2 nm for a packed heptaspan-protein (Figure 3C-E). We interpolated these values for magnetosome membrane proteins that exhibit 1 to 18 TMHs and calculated the average TMH-coverage of the magnetosome membrane (surface) based on the predicted copy numbers of the proteins and different boundary lipid assumptions (see experimental procedures for details) (Table 4). We chose to focus only on genuine Mam and Mms proteins, since the copy numbers of other potential integral magnetosome-membrane proteins are negligible, together accounting for only 2 % (Table 2).

According to this calculation, TMDs of magnetosome proteins already cover 18-20% of the magnetosome surface, without taking into account boundary lipids. If one shell

of boundary lipid is added, this coverage increases to 62-97% and further to impossible 131-238% if a second shell of boundary lipids is included. Assuming a mixed TMH-packing model in reality, this indicates that TMHs of all proteins seem to be in close contact to each other and in most cases can be only surrounded by one layer of boundary lipids. We used the alternative PAI quantification method to estimate protein abundance and membrane coverage (suppl. Table S 5 and suppl. Table S 6, full calculation can be found in suppl. File 1). Using PAI, we even estimated higher copy numbers for most proteins and up to 170% higher membrane coverages (suppl. Table S 6). Therefore, the magnetosome membrane seems to be very rigid and tightly packed with trans-membrane proteins and only contains a smaller number of “free” lipids. For comparison, the hepta-spanning archaeal bacteriorhodopsin is one of the most tightly clustered transmembrane proteins and might be present in up to 600,000 molecules on a surface of 15 μm^2 [42]. Assuming a trimer configuration of the molecule with an approximate TMD-diameter of 5.2 nm [PDB] (including 3 enclosed lipids), this would indicate a total membrane coverage of 28%, which is somewhat higher but in the same range as our conservatively estimated coverage of the magnetosome membrane.

Conclusions

In summary, our proteomic data and modelling estimated the composition of the integral magnetosome-membrane proteome of MSR-1. Although the prediction of relative protein abundancies from label-free mass spectroscopy data is inherently error-prone and might lead to under- or overestimations for individual proteins, the cautious interpretation of our data allowed us to approximate the protein composition and coverage of the magnetosome membrane in a novel approach. Absolute quantifications of other individual magnetosomes-membrane proteins in the future

will further refine our model, which will be the basis for more precise determination of the structure of this unique bacterial organelle. By directly comparing protein abundances of magnetosomes purified from single magnetosome gene deletion mutants with those of the wild-type, our approach could be utilized to systematically assay the interdependency on protein localization to the organelle. This might prove as a powerful tool to further investigate the complex interaction-network of magnetosome proteins. Finally, our prediction of an unusually crowded protein composition within the membrane of the organelle also might substantially contribute to the assumption that a lipid raft like association of magnetosome-membrane proteins takes place prior to the magnetosome invagination [3,5]. Altogether our results will help to elucidate the processes involved in biogenesis of magnetosomes.

Acknowledgments

This work was supported by grants from the Deutsche Forschungsgemeinschaft (Schu1080/9-2 and 15-3) and the European Research Council (Proposal N° 692637 Syntomagx) to Dirk Schüler.

References

[1] F. Popp, J.P. Armitage, D. Schüler, Polarity of bacterial magnetotaxis is controlled by aerotaxis through a common sensory pathway, Nat. Commun. 5 (2014). doi:10.1038/ncomms6398.

[2] E. Cornejo, P. Subramanian, Z. Li, G.J. Jensen, Dynamic Remodeling of the Magnetosome Membrane Is Triggered by the Initiation of Biomineralization, 7 (2016) 1–9. doi:10.1128/mBio.01898-15.Editor.

[3] O. Raschdorf, Y. Forstner, I. Kolinko, R. Uebe, J.M. Plitzko, D. Schüler,

Genetic and Ultrastructural Analysis Reveals the Key Players and Initial Steps of Bacterial Magnetosome Membrane Biogenesis, PLoS Genet. 12 (2016) 1–23. doi:10.1371/journal.pgen.1006101.

[4] R. Uebe, D. Schüler, Magnetosome biogenesis in magnetotactic bacteria, Nat. Rev. Microbiol. 14 (2016) 621–637. doi:10.1038/nrmicro.2016.99.

[5] H. Nudelman, R. Zarivach, Structure prediction of magnetosome-associated proteins, Front. Microbiol. 5 (2014) 1–17. doi:10.3389/fmicb.2014.00009.

[6] Y.A. Gorby, T.J. Beveridge, R.P. Blakemore, Characterization of the bacterial magnetosome membrane., J. Bacteriol. 170 (1988) 834–41. http://jb.asm.org/content/170/2/834.

[7] K. Grünberg, C. Wawer, B.M. Tebo, D. Schüler, A Large Gene Cluster Encoding Several Magnetosome Proteins Is Conserved in Different Species of Magnetotactic Bacteria, Appl. Environ. Microbiol. 67 (2001) 4573–4582. doi:10.1128/AEM.67.10.4573-4582.2001.

[8] M. Tanaka, Y. Okamura, A. Arakaki, T. Tanaka, H. Takeyama, T. Matsunaga, Origin of magnetosome membrane: proteomic analysis of magnetosome membrane and comparison with cytoplasmic membrane., Proteomics. 6 (2006) 5234–47. doi:10.1002/pmic.200500887.

[9] K. Grünberg, E.-C. Müller, A. Otto, R. Reszka, D. Linder, M. Kube, R. Reinhardt, D. Schüler, Biochemical and Proteomic Analysis of the Magnetosome Membrane in Magnetospirillum gryphiswaldense, Appl. Environ. Microbiol. 70 (2004) 1040–1050. doi:10.1128/AEM.70.2.1040-1050.2004.

[10] M. Richter, M. Kube, D.A. Bazylinski, T. Lombardot, F.O. Glöckner, R. Reinhardt, D. Schüler, Comparative genome analysis of four magnetotactic bacteria reveals a complex set of group-specific genes implicated in magnetosome biomineralization and function., J. Bacteriol. 189 (2007) 4899–910. doi:10.1128/JB.00119-07.

[11] S. Ullrich, M. Kube, S. Schübbe, R. Reinhardt, D. Schüler, A Hypervariable 130-Kilobase Genomic Region of Magnetospirillum gryphiswaldense Comprises a Magnetosome Island Which Undergoes Frequent Rearrangements during Stationary Growth, J. Bacteriol. 187 (2005) 7176–7184. doi:10.1128/JB.187.21.7176.

[12] T. Matsunaga, M. Nemoto, A. Arakaki, M. Tanaka, Proteomic analysis of irregular, bullet-shaped magnetosomes in the sulphate-reducing magnetotactic bacterium Desulfovibrio magneticus RS-1., Proteomics. 9 (2009) 3341–3352. doi:10.1002/pmic.200800881.

[13] A. Lohße, S. Ullrich, E. Katzmann, S. Borg, G. Wanner, M. Richter, B. Voigt, T. Schweder, D. Schüler, Functional Analysis of the Magnetosome Island in Magnetospirillum gryphiswaldense: The mamAB Operon Is Sufficient for Magnetite Biomineralization, PLoS One. 6 (2011) e25561. doi:10.1371/journal.pone.0025561.

[14] U. Heyen, D. Schüler, Growth and magnetosome formation by microaerophilic Magnetospirillum strains in an oxygen-controlled fermentor., Appl. Microbiol. Biotechnol. 61 (2003) 536–44. doi:10.1007/s00253-002-1219-x.

[15] D. Schüler, R. Uhl, E. Bäuerlein, A simple light scattering method to assay

magnetism in Magnetospirillum gryphiswaldense, FEMS Microbiol. Lett. 132 (1995) 139–145. doi:10.1016/0378-1097(95)00300-T.

[16] D. Schultheiss, D. Schüler, Development of a genetic system for Magnetospirillum gryphiswaldense., Arch. Microbiol. 179 (2003) 89–94. doi:10.1007/s00203-002-0498-z.

[17] S. Ullrich, D. Schüler, Cre-lox-based method for generation of large deletions within the genomic magnetosome island of Magnetospirillum gryphiswaldense., Appl. Environ. Microbiol. 76 (2010) 2439–44. doi:10.1128/AEM.02805-09.

[18] F. Bonn, J. Bartel, K. Büttner, M. Hecker, A. Otto, D. Becher, Picking vanished proteins from the void: How to collect and ship/share extremely dilute proteins in a reproducible and highly efficient manner, Anal. Chem. 86 (2014) 7421–7427. doi:10.1021/ac501189j.

[19] S. Wolff, H. Hahne, M. Hecker, D. Becher, Complementary analysis of the vegetative membrane proteome of the human pathogen Staphylococcus aureus., Mol. Cell. Proteomics. 7 (2008) 1460–8. doi:10.1074/mcp.M700554-MCP200.

[20] J.A. Vizcaíno, A. Csordas, N. del-Toro, J.A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q.-W. Xu, R. Wang, H. Hermjakob, 2016 update of the PRIDE database and its related tools, Nucleic Acids Res. 44 (2016) D447–D456. doi:10.1093/nar/gkv1145.

[21] S. Borg, J. Hofmann, A. Pollithy, C. Lang, D. Schüler, New Vectors for Chromosomal Integration Enable High-Level Constitutive or Inducible

Magnetosome Expression of Fusion Proteins in Magnetospirillum gryphiswaldense., Appl. Environ. Microbiol. 80 (2014) 2609–16. doi:10.1128/AEM.00192-14.

[22] K. Jacobson, O.G. Mouritsen, R.G.W. Anderson, Lipid rafts: at a crossroad between cell biology and physics., Nat. Cell Biol. 9 (2007) 7–14. doi:10.1038/ncb0107-7.

[23] F.D. Gunstone, J.L. Harwood, F.B. Padley, The Lipid Handbook, second edi, Chapman & Hall/CRC, London, 1994.

[24] J. Rappsilber, U. Ryder, A.I. Lamond, M. Mann, Large-Scale Proteomic Analysis of the Human Spliceosome, Genome Res. 12 (2002) 1231–1245. doi:10.1101/gr.473902.

[25] O. Raschdorf, J.M. Plitzko, D. Schüler, F.D. Müller, A tailored galK counterselection system for efficient markerless gene deletion and chromosomal tagging in Magnetospirillum gryphiswaldense, Appl. Environ. Microbiol. 80 (2014) 4323–4330. doi:10.1128/AEM.00588-14.

[26] R. Arai, H. Ueda, A. Kitayama, N. Kamiya, T. Nagamune, Design of the linkers which effectively separate domains of a bifunctional fusion protein., Protein Eng. 14 (2001) 529–32. http://www.ncbi.nlm.nih.gov/pubmed/11579220.

[27] R. Uebe, K. Junge, V. Henn, G. Poxleitner, E. Katzmann, J.M. Plitzko, R. Zarivach, T. Kasama, G. Wanner, M. Pósfai, L. Böttger, B.F. Matzanke, D. Schüler, 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., Mol. Microbiol. 84 (2011) 818–835. doi:10.1111/j.1365-2958.2011.07863.x.

[28] X. Wang, Q. Wang, W. Zhang, Y. Wang, L. Li, T. Wen, T. Zhang, Y. Zhang, J. Xu, J. Hu, S. Li, L. Liu, J. Liu, W. Jiang, J. Tian, Y. Li, L. Wang, J. Li, Complete Genome Sequence of Magnetospirillum gryphiswaldense, Genome Announc. 2 (2014) 2–3. doi:10.1128/genomeA.00171-14.Copyright.

[29] A. Scheffel, A. Gärdes, K. Grünberg, G. Wanner, D. Schüler, The major magnetosome proteins MamGFDC are not essential for magnetite biomineralization in Magnetospirillum gryphiswaldense but regulate the size of magnetosome crystals., J. Bacteriol. 190 (2008) 377–86. doi:10.1128/JB.01371-07.

[30] D. Murat, A. Quinlan, H. Vali, A. Komeili, 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 (2010) 5593–5598. doi:10.1073/pnas.0914439107.

[31] A. Taoka, R. Asada, H. Sasaki, K. Anzawa, L.-F. Wu, Y. Fukumori, Spatial localizations of Mam22 and Mam12 in the magnetosomes of Magnetospirillum magnetotacticum., J. Bacteriol. 188 (2006) 3805–12. doi:10.1128/JB.00020-06.

[32] N. Zeytuni, E. Ozyamak, K. Ben Harush, G. Davidov, M. Levin, Y. Gat, T. Moyal, A. Brik, A. Komeili, R. Zarivach, Self-recognition mechanism of MamA, a magnetosome-associated TPR-containing protein, promotes complex assembly., Proc. Natl. Acad. Sci. U. S. A. (2011).

doi:10.1073/pnas.1103367108.

[33] D. Yamamoto, A. Taoka, T. Uchihashi, H. Sasaki, H. Watanabe, T. Ando, Y. Fukumori, Visualization and structural analysis of the bacterial magnetic organelle magnetosome using atomic force microscopy., Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 9382–7. doi:10.1073/pnas.1001870107.

[34] A. Lohße, S. Borg, O. Raschdorf, I. Kolinko, É. Tompa, M. Pósfai, D. Faivre, J. Baumgartner, D. Schülera, Genetic dissection of the mamAB and mms6 operons reveals a gene set essential for magnetosome biogenesis in magnetospirillum gryphiswaldense, J. Bacteriol. 196 (2014) 2658–2669. doi:10.1128/JB.01716-14.

[35] C. Rong, C. Zhang, Y. Zhang, L. Qi, J. Yang, G. Guan, Y. Li, J. Li, FeoB2 Functions in Magnetosome Formation and Oxidative Stress Protection in Magnetospirillum gryphiswaldense Strain MSR-1., J. Bacteriol. 194 (2012) 3972–6. doi:10.1128/JB.00382-12.

[36] C. Lang, D. Schüler, Expression of green fluorescent protein fused to magnetosome proteins in microaerophilic magnetotactic bacteria., Appl. Environ. Microbiol. 74 (2008) 4944–53. doi:10.1128/AEM.00231-08.

[37] D. Murat, A. Quinlan, H. Vali, A. Komeili, Supporting Information - 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 (2010).

[38] M. Tanaka, A. Arakaki, S.S. Staniland, T. Matsunaga, Simultaneously discrete biomineralization of magnetite and tellurium nanocrystals in magnetotactic

bacteria., Appl. Environ. Microbiol. 76 (2010) 5526–32. doi:10.1128/AEM.00589-10.

[39] A. Poetsch, D. Wolters, Bacterial membrane proteomics, Proteomics. 8 (2008) 4100–4122. doi:10.1002/pmic.200800273.

[40] Y. Okamura, H. Takeyama, T. Matsunaga, A Magnetosome-specific GTPase from the Magnetic Bacterium Magnetospirillum magneticum AMB-1, J. Biol. Chem. 276 (2001) 48183–48188. doi:10.1074/jbc.M106408200.

[41] D. Schultheiss, R. Handrick, D. Jendrossek, M. Hanzlik, D. Schüler, The Presumptive Magnetosome Protein Mms16 Is a Poly(3-Hydroxybutyrate) Granule-Bound Protein (Phasin) in Magnetospirillum gryphiswaldense, J. Bacteriol. 187 (2005) 2416–2425. doi:10.1128/JB.187.7.2416.

[42] R.C.H. del Rosario, C. Oppawsky, J. Tittor, D. Oesterhelt, Modeling the membrane potential generation of bacteriorhodopsin, Math. Biosci. 225 (2010) 68–80. doi:10.1016/j.mbs.2010.02.002.

Figures and tables

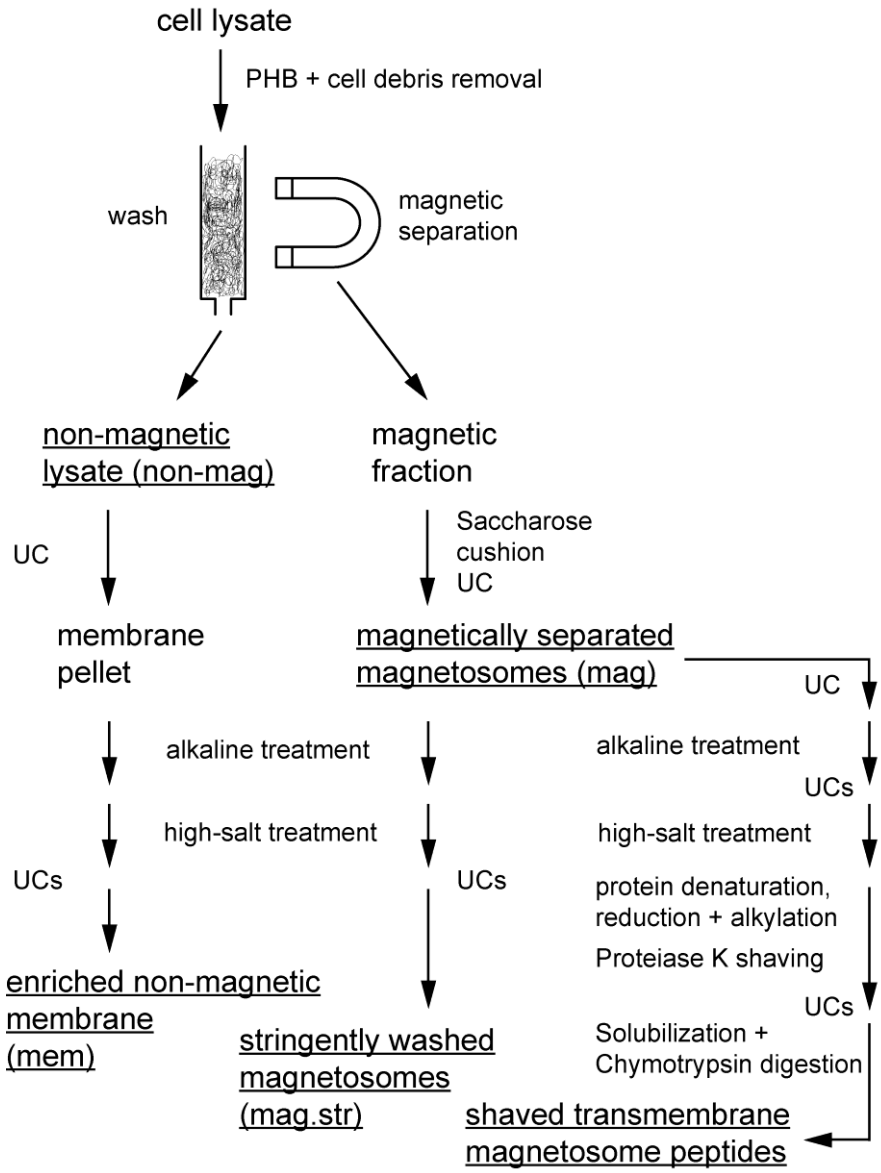

Figure 1: Fractionation workflow of magnetically separated magnetosomes, stringently washed magnetosomes, non-magnetic cell lysate and enriched non-magnetic membrane fraction. Additionally, workflow to obtain transmembrane peptides from magnetosome proteins by Proteinase K membrane shaving is outlined**.** UC: Ultracentrifugation

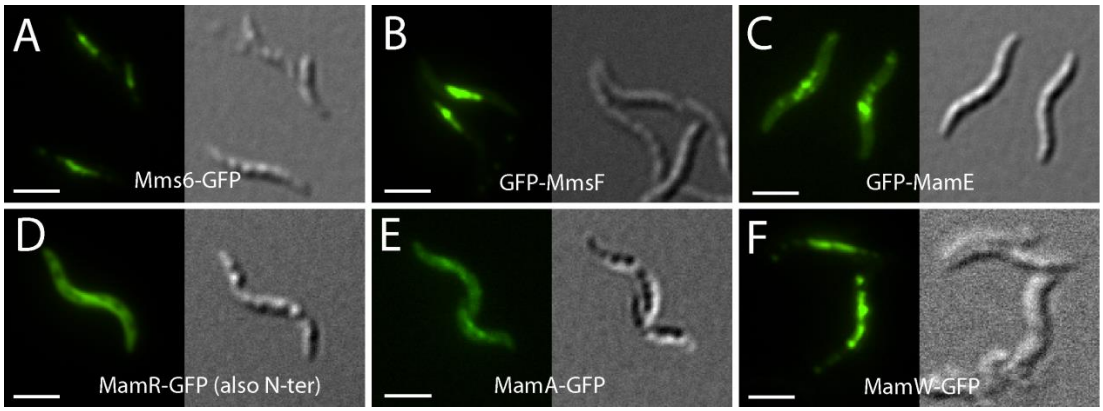

Figure 2: Cellular localization of six Mam/Mms proteins in *M. gryphiswaldense*. All proteins were C- or N-terminally fused to EGFP and expressed in the wild type. A 30 amino acid alpha-helical linker (HL) was placed in between the fusion proteins. A) C-terminal P_{mamDC}-*mms6*-HL-*egfp* construct, expressed from plasmid. B) N-terminal P_{mamDC}-*mmsF*-HL-*egfp* construct, expressed from plasmid. C) N-terminal P_{mamDC}-*mamE*-HL-*egfp* construct, expressed from ectopic chromosomal locus. D) C-terminal P_{mamDC}-*mamR*-HL-*egfp* construct, expressed from plasmid. N-terminal fusion exhibited comparable localization pattern. E) C-terminal *mamA*-HL-*egfp* construct, expressed from native chromosomal locus (in-frame gene fusion) F) C-terminal P_{mamDC}-*mamW*-HL-*egfp* construct, expressed from plasmid. Fluorescence (left) and corresponding differential interference contrast (right) images are shown. Scale bar: 2µm

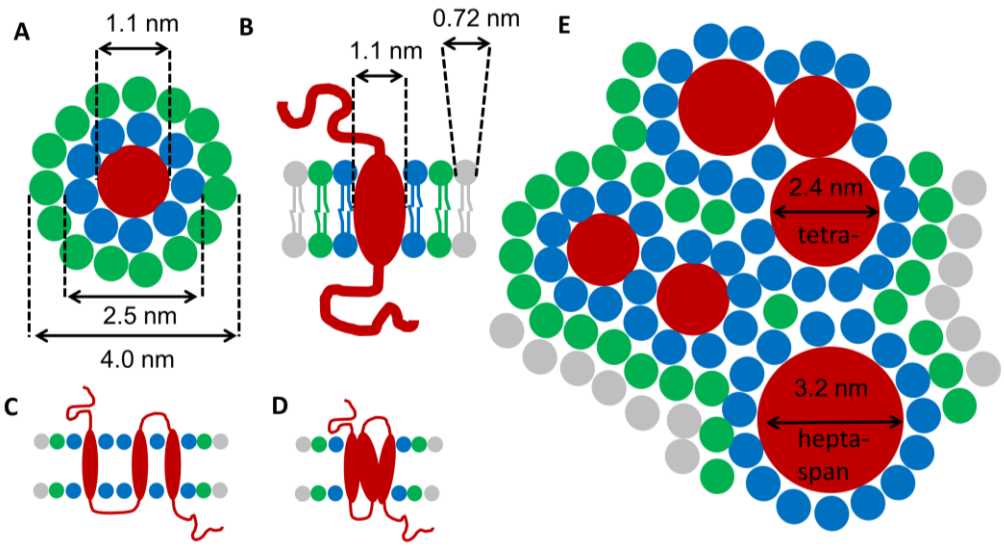

Figure 3: Model of transmembrane helices (domains) embedded in a membrane. Transmembrane helices (domains) are colored in red, primary boundary lipids in blue, secondary boundary lipids in green and non-interacting ('free') lipids in grey. Approximate cross section diameters of several entities are indicated within the figure. A) Top view of single transmembrane helix with two layers of boundary lipids. B) Side view of transmembrane helix with two layers of boundary lipids and one layer of free lipids. C) Side view of three transmembrane helices of a single protein all embedded by boundary lipids. D) Side view of three transmembrane helices of a single protein packed in one domain embedded by boundary lipids but without internal lipids. E) Top view of an array of transmembrane domains of several proteins surrounded by variable numbers of boundary lipids. One of these proteins is a packed protein with transmembrane domains consisting of three transmembrane helices (tetra-span) and another protein with seven (hepta-span) helices.

Table 1: Number of assigned genuine magnetosome proteins under different parameter conditions

Parameters									
$A_{mem}/A_{non\text{-}mag} \geq$	**1**	0	1	1	0	**1**	1	1	0
$A_{mag}/A_{mem} >$	**1**	1	0	1	1	**4**	1	4	4
$A_{mag.str}/A_{mag} \geq$	**1**	1	1	0	0	**1**	1	1	1
$A_{mag.str}/A_{mag.str(MamC)} >$	**-**	-	-	-	-	-	0.01	0.01	0.01
Total # of assigned proteins [A]	**81**	151	556	155	651	**30**	43	26	30
# assigned predicted MMP [B]	**23**	24	24	24	27	**21**	23	21	22
# non-assigned predicted MMP [C]	**9**	8	8	8	5	**11**	9	11	10
# assigned predicted non-MMP [D]	**58**	127	532	131	624	**9**	20	5	8

[A] total number of proteins that meet the applied parameter filter set (= assigned)

[B] number of assigned proteins that are predicted magnetosome membrane proteins (MMPs) [encoded within *mam* or *mms* operons (including *mamW*, *mamF2*, *mamD2* and *ftsZm*]

[C] number of proteins that are predicted MMPs, but do not meet parameter criteria

[D] number of assigned proteins that are not predicted MMPs

Table 2: Top20-abundant and all assigned genuine magnetosome proteins. The following parameters were applied: $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} > 4$, $A_{mag.str}/A_{mag} \geq 1$. Abundance relative to MamC was calculated by $A_{mag.str}/A_{mag.str(MamC)}$ value. The number of predicted protein copies is estimated by assuming a MamC copy number of 100 per magnetosome [21]. Genuine magnetosome-membrane proteins meeting the parameter criteria are marked by grey columns; other proteins only show high abundance values compared to MamC copy number.

Rank Abundance ($A_{mag.str}$)	Abundance relative to MamC [predicted copy number]	Protein name [main putative function[4]]	Molecular weight
1	1.21 [121]	**Mms6** [magnetite crystal nucleation and growth]	13 kDa
2	1.00 [100]	**MamC** [magnetite crystal growth]	12 kDa
3	0.51 [51]	**MamD** [magnetite crystal growth]	30 kDa
4	0.48	**YajC** [preprotein translocase subunit]	13 kDa
5	0.31 [31]	**MamF2** [unknown]	12 kDa
6	0.31 [31]	**MamE** [magnetosome maturation, magnetite crystal nucleation]	78 kDa
7	0.26	**MGR_0659** (Mms16) [Phasin]	16 kDa
8	0.25 [25]	**MmsF** [magnetite crystal growth]	14 kDa
9	0.24	**MGR_3650** [Outer membrane protein (porin)]	41 kDa
10	0.23 [23]	**MamB** [magnetosomal iron transport, magnetosome membrane formation]	32 kDa
11	0.21 [21]	**MamM** [magnetosomal iron transport]	34 kDa
12	0.21 [21]	**MamF** [magnetite crystal growth]	12 kDa
13	0.20	**MamR** [magnetite crystal growth]	8 kDa
14	0.18	**AtpF** [ATP synthase B chain precursor]	19 kDa
15	0.18	**MGR_1798** [Outer membrane protein]	17 kDa
16	0.17	**MamA** [magnetosome maturation]	24 kDa
17	0.17 [17]	**MamY** [magnetosome membrane maturation]	41 kDa
18	0.14 [14]	**MamO** [magnetite crystal nucleation]	65 kDa
19	0.13	**AtpG** [ATP synthase B' chain]	18 kDa
20	0.13	**MamJ** [magnetosome chain formation]	44 kDa
24	0.11 [11]	**MamP** [magnetite crystal nucleation and growth, redox control]	28 kDa
25	0.10 [10]	**MamT** [magnetosomal redox control]	19 kDa
28	0.09 [9]	**MamS** [magnetite crystal growth]	19 kDa
36	0.06 [6]	**MamG** [magnetite crystal growth]	8 kDa
40	0.06 [6]	**MamI** [magnetite crystal nucleation and growth]	8 kDa
43	0.05 [5]	**MamW** [unknown]	15 kDa
47	0.05 [5]	**MGR_4114** [unknown]	7 kDa

58	0.04 [4]	**MamH** [unknown, putative iron importer]	46 kDa
60	0.04 [4]	**MamN** [magnetite crystal growth]	46 kDa
95	0.02 [2]	**MGR_0622** [ATP-binding region, Histidine kinase]	50 kDa
100	0.02 [2]	**MGR_2552** [Cytochrome oxidase maturation cbb3-type]	7 kDa
102	0.02 [2]	**MGR_2833** [unknown]	26 kDa
111	0.02 [2]	**MamX** [magnetosomal redox control]	28 kDa
127	0.02 [2]	**MamZ** [magnetosomal redox control, putative iron importer]	72 kDa
187	0.01 [1]	**MGR_2730** [unknown]	20 kDa
226	0.01 [1]	**MGR_0916** [unknown]	9 kDa
444	0.00 [0]	**MGR_0581** [unknown]	10 kDa
535	0.00 [0]	**MGR_2491** [unknown]	8 kDa
1040	0.00 [0]	**MGR_3321** [two-comp. sensor histidine kinase]	47 kDa

Table 3: Additional candidate genuine magnetosome membrane proteins predicted by less stringent filter parameters. Proteins here are only listed if not already mentioned in Table 2.

$A_{mem}/A_{non\text{-}mag}$	A_{mag}/A_{mem}	$A_{mag.str}/A_{mag}$	$A_{mag.str}/A_{mag.str(MamC)}$	Rank ($A_{mag.str}$)	**Set parameters: $A_{mem}/A_{non\text{-}mag} \geq 1$, $A_{mag}/A_{mem} > 1$, $A_{mag.str}/A_{mag} \geq 1$, $A_{mag.str(N)}/A_{mag.str(MamC)} \geq 0.02$**
4.3	2.0	1.3	0.07	32	MamQ [magnetosome membrane maturation] (30 kDa)
3.4	1.3	1.5	0.07	34	MGR_3120 Bacterial sec-independent translocation protein mttA/Hcf106 (8 kDa)
2.9	1.3	1.3	0.07	35	MGR_1712 translocase, subunit Tim44 (26 kDa)
1.2	1.4	1.0	0.05	46	MGR_0255 conserved hypothetical protein (11 kDa)
5.0	3.7	1.3	0.03	66	MGR_1199 Peptidase M48, Ste24p (33 kDa)
1.1	1.2	1.2	0.02	87	MGR_4238 regulatory protein (22 kDa)
3.9	1.1	1.1	0.02	93	MGR_0007 glycosyl transferase, group 2 family (27 kDa)
6.6	1.3	1.9	0.02	114	MGR_0417 serine O-acetyltransferase (27 kDa)
4.3	1.9	1.1	0.02	129	MGR_3354 phos.lipid/glycerol acyltransferase (30 kDa)
$A_{mem}/A_{non\text{-}mag}$	A_{mag}/A_{mem}	$A_{mag.str}/A_{mag}$	$A_{mag.str}/A_{mag.str(MamC)}$	Rank ($A_{mag.str}$)	**Set parameters: $A_{mem}/A_{non\text{-}mag} \geq 0$ $A_{mag}/A_{mem} > 4$, $A_{mag.str}/A_{mag} \geq 1$, $A_{mag.str(N)}/A_{mag.str(MamC)} > 0.01$**
0.0	∞	1.1	0.20	13	MamR [magnetite crystal growth] (8 kDa)
0.4	6.5	1.5	0.10	27	MGR_2633 Phasin (12 kDa)
0.1	6.8	1.2	0.02	116	MGR_2416 cytochrome c (12 kDa)
0.8	5.0	1.1	0.01	142	MGR_1351 CreA (17 kDa)

Table 4: Coverage of the magnetosome membrane surface by transmembrane domains of integral magnetosome proteins

Integral magnetosome membrane coverage by TMDs [A]	All TMHs [B] are isolated from each other	TMHs of individual proteins are tightly packed
No boundary lipid	**18 %**	**20 %**
One boundary lipid	**97 %**	**63 %**
Two boundary lipids	**238 %**	**132 %**

A: Transmembrane domain
B: Transmembrane helix

Supplementary information

suppl. File 1: Calculations and interactive table to analyze magnetosome membrane proteom data

Fig S 1: Organization of genes associated with magnetosome formation within the genomic magnetosome island (MAI) of *M. gryphiswaldense*.

Fig S 2: 2D SDS-PAGE of fractions employed for mass spectrometry analysis:

Fig S 3: Representative Transmission Electron Micrographs of *M. gryphiswaldense* wild type (A), Δ*mgr3691* (Δ*MM22*) (B) and Δ*mgr4114op* (C). Scale bar represents 500 nm

suppl. Table S 1. Number of assigned genuine magnetosome proteins under different parameter conditions

suppl. Table S 2: Number of genuine magnetosome proteins assigned by increasing A_{mag}/A_{mem} ratios

suppl. Table S 3: Introduction of MamC abundance threshold to regulate the number of assigned genuine magnetosome membrane proteins

suppl. Table S 4: Proteins with identified peptides from shaving assay.

suppl. Table S 5: Estimated number of protein copies for membrane integral Mam/Mms Proteins according to different quantification methods.

suppl. Table S 6: Coverage of the magnetosome membrane surface by transmembrane domains of integral magnetosome proteins calculated according to different quantification methods

suppl. Table S 7: Strains and plasmids used in this study

suppl. Table S 8: Oligonucleotides used in this study