Title

Cellular distribution and motion of essential magnetosome proteins expressed in mammalian cells

Authors

Qin Sun[1,2,3], Cecile Fradin[4,5], R. Terry Thompson[1,2,6,7], Frank S. Prato[1,2,3,6,7] and Donna E. Goldhawk[1,2,3]*

Affiliations

[1]Imaging, Lawson Health Research Institute, London, Ontario, Canada.

[2]Medical Biophysics, Western University, London, Ontario, Canada.

[3]Collaborative Graduate Program in Molecular Imaging, Western University, London, Ontario, Canada.

[4]Physics & Astronomy, McMaster University, Hamilton, Ontario, Canada.

[5]Biochemistry & Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada.

[6]Medical Imaging, Western University, London, Ontario, Canada.

[7]Physics & Astronomy, Western University, London, Ontario, Canada.

*** Corresponding Author**

Dr. Donna E. Goldhawk
Imaging Program, Lawson Health Research Institute
268 Grosvenor St., PO Box 5777, Station B
London, ON, Canada N6A 4V2
(519) 646-6100 ext. 64863
(519) 646-6110 FAX
dgoldhawk@lawsonimaging.ca

Running title

Distribution and motion of essential magnetosome proteins in mammalian cells

Keywords

Molecular imaging, microparticles, biosensor, nanotechnology, magnetosome, mobility,

mammalian cells

Abstract

Magnetosomes are organelle-like structures within magnetotactic bacteria that store iron biominerals in membrane-bound vesicles. In the bacteria, formation of these structures is highly regulated by approximately 30 genes which are conserved throughout different species. To compartmentalize iron in mammalian cells for magnetic resonance imaging using gene-based contrast, we are introducing key magnetosome proteins. We have previously expressed essential magnetosome genes *mamI* and *mamL* as fluorescent fusion proteins in the human melanoma cell line MDA-MB-435 and confirmed their co-localization and interaction. Here we investigate the expression of magnetosome genes *mamB* and *mamE* in MDA-MB-435 cells, using confocal fluorescence microscopy to observe expression patterns and to analyze particle mobility. Custom software was developed to characterize fluorescent particle trajectories. In mammalian cells, essential magnetosome proteins displayed different diffusive behaviours. However, all magnetosome proteins travelled at similar velocities when undergoing directed motion, suggesting that MamL, MamL+MamI, MamB, and MamE interact with similar mammalian mobile elements. These results confirm that localization and interaction of essential magnetosome proteins is tenable in the mammalian intracellular compartment.

Introduction

In magnetotactic bacteria (MTB), magnetosome formation allows the cell to compartmentalize and concentrate iron biominerals in membrane-enclosed vesicles (1). Magnetosome formation is a stepwise, protein-directed process that begins with vesicle formation and culminates with iron biomineralization. This entire process is regulated by approximately 30 genes, the majority of which are located on the magnetosome genomic island (2, 3). Many of these genes encode membrane proteins which are involved in different steps of

bacterial magnetosome formation: inner membrane invagination leading to vesicle formation, recruitment of proteins to the vesicle membrane, alignment of vesicles along a protein filament, concentration of iron inside the vesicle and finally magnetic crystal nucleation and growth. To stimulate the formation of a rudimentary magnetosome-like particle in mammalian cells, we are introducing select genes (alone or in combination) deemed essential for the initial stages of magnetosome formation, namely *mamI*, *mamL*, *mamB* and *mamE*. These four genes are clustered on the *mamAB* operon and are highly conserved in most species of MTB (4). The gene products of *mamI*, *mamL*, and *mamB* are believed to play an essential role in the first steps of magnetosome vesicle formation (2, 5, 6), including designation of the vesicle and recruitment of other magnetosome-associated proteins like MamE (7, 8). The latter plays a crucial role in the initiation of iron biomineralization and recruitment of additional magnetosome-associated proteins (9). We hypothesize that together MamI, MamL, MamB and MamE support the base structure upon which magnetosome formation relies in any cell. As such, we expect these essential magnetosome proteins (co)localize and interact in the intracellular membranous compartment of mammalian systems.

We have previously shown that the red fluorescent MamL fusion protein, Tomato-MamL, assembles into punctate structures that move intracellularly in a mammalian cell line. We have also shown that MamL recruits the green fluorescent MamI fusion protein, EGFP-MamI, to these structures via magnetosome protein-protein interactions. In contrast, EGFP-MamI expressed alone in mammalian cells has a very different cellular distribution (10). The novel and unexpected discovery of the mobility associated with these magnetosome proteins in a foreign cell environment prompted further investigation into particle trajectory analysis.

Here, we quantitatively compare the distribution and mobility of a larger set of magnetosome proteins (MamL, MamI, MamB, MamE), again expressed as fluorescent fusion proteins in a mammalian cell line. When expressed on their own, Tomato-MamL, Tomato-MamB and EGFP-MamE all form (or localize to) punctate structures, which are mostly mobile in the case of MamL and MamB, but immobile in the case of MamE. For EGFP-MamI, co-expression with MamL was necessary to ensure a similar localization to punctate structures so that a comparison in their mobility could be done. These punctate structures will herein be referred to as "particles". We track fluorescent particle motion and describe their mean-squared displacement (MSD) with a simple power-law analysis to distinguish between different types of mobility. With these tools, the trajectories of punctate intracellular structures are classified into three different categories corresponding to confined, diffusing and actively transported particles. The motions detected for MamL and MamB (and for MamI when interacting with MamL) show that a portion of these structures displays directed motion. This demonstrates the capacity of magnetosome proteins to spontaneously interact with elements of eukaryotic transport machinery.

Materials and Methods

Molecular Cloning

Magnetosome genes *mamI* and *mamL* were amplified by PCR from the genomic DNA of *Magnetospirillum magneticum* strain AMB-1 (ATCC 700264) using custom primers (10). The *mamI* and *mamL* amplicons were purified using a PCR clean-up kit (Invitrogen, Life Technologies, Burlington, Canada); digested with appropriate restriction enzymes; and purified once more, prior to insertion in the molecular cloning vectors pEGFP-C1 (Clontech) and

ptdTomato-C1 (Clontech), respectively. After propagation in *Escherichia coli* strain XL10GOLD, the vector-insert plasmid constructs were purified and used for mammalian cell transfection.

Sequence analysis of the MamL protein indicates the presence of a positively-charged peptide (2) in the C-terminal 15 amino acids. Cationic peptides, also referred to as cell-penetrating peptides, have been implicated in endocytotic pathways (11). For cloning of Tomato-$MamL_{trunc}$, the last 15 amino acids from the C-terminus of MamL were removed by PCR site-directed mutagenesis (10). Briefly, primers were designed that included a stop codon 45 nucleotides upstream from the end of *mamL*. These primers were used in PCR amplification of the $mamL_{trunc}$ gene, which was then inserted into the ptdTomato-C1 vector with restriction enzymes EcoRI and BglII.

For cloning of FLAG-MamL, primers were designed to flank the *mamL* gene, and a FLAG tag was included for immunodetection (10). The insert was amplified using PCR, purified using a PCR clean-up kit (Invitrogen, Life Technologies, Burlington, Canada), and restriction digested using SacI and EcoRI. *FLAG-mamL* was then inserted into the pSF-EMCV-*FLuc* vector and propagated in *Escherichia coli* strain XL10GOLD.

For cloning of Tomato-MamB, the *mamB* gene was amplified by PCR from AMB-1 genomic DNA using custom primers (Table 1). The amplicon was purified with a PCR clean-up kit, digested with restriction enzymes (Table 1), purified once more, and inserted into the ptdTomato-C1 vector (Clontech; Fig. 1). The vector-insert construct was propagated in *E. coli* strain XL10GOLD, purified, and then used for mammalian cell transfection. Sequencing information is provided in Supplementary Figure S1.

For cloning of EGFP-MamE, the *mamE* gene was amplified by PCR from the AMB-1 genomic DNA using custom primers (Table 1). The amplicon was purified with a PCR clean-up kit, digested with restriction enzymes (Table 1), purified once more, and inserted into the pEGFP-C1 vector (Clontech; Fig. 2). The vector-insert construct was propagated in *E. coli* strain XL10GOLD, purified, then used for mammalian cell transfection.

When sequencing EGFP-*mamE*, there were two mutations (G49S and T641S) present in the gene sequence (Supplementary Figure 2). These mutations were not found in column-purified EGFP-*mamE* miniprep DNA but arose spontaneously in maxiprep DNA. Therefore, mutated EGFP-MamE was used in some experiments (ie. confocal fluorescence). After another attempt at cloning *mamE*, the new EGFP-*mamE* was sequenced and there was one mutation found at amino acid 317 (Thr → Ala, Supplementary Figure 2). Constructs of *mamE* will be referred to as EGFP-*mamE* (G49S, T641S) and EGFP-*mamE* (T317A). Sequencing information of both constructs is provided in Supplementary Figure 2.

Cell Culture

MDA-MB-435 cells (ATCC HTB-129; derived from an adult female and characterized as a melanoma cell line) are a model of aggressive tumorigenesis (12). Cells were cultured in 100 mm cell culture dishes (CELLSTAR, VWR International, Mississauga, Canada) with Dulbecco's Modified Eagle Medium (DMEM) containing 1 g/L glucose (Gibco, Life Technologies, Burlington, Canada), 10% fetal bovine serum (FBS; Gibco), 4 U/mL penicillin, and 4 µg/mL streptomycin at 37°C with 5% CO_2.

To create cell lines expressing the red fluorescent protein fusions tdTomato (Tomato)-MamL, Tomato-$MamL_{trunc,}$ or Tomato-MamB, the enhanced green fluorescent protein (EGFP)

fusion protein EGFP-MamE, or the FLAG-tag (DYKDDDDK) fusion protein FLAG-MamL, cells were grown to 60-70% confluency on a 100 mm dish and transfected using Lipofectamine 2000 (Invitrogen), according to company protocol, using 8 µg of each construct. After 16 hours, cells were placed in full medium for 48 hours before commencing antibiotic selection. To select stable cell lines, cells were grown in the presence of 500 µg/mL Geneticin (G418; Gibco) for Tomato or EGFP expression systems and 500 ng/mL Puromycin (Gibco) for the FLAG expression system.

For co-expression of both pEGFP-*mamI* and ptdTomato-*mamL,* cells stably expressing Tomato-MamL were transfected with 8 µg of pEGFP-*mamI*. After 16 hours, cells were placed in full medium for 48 hours before commencing antibiotic selection. Cells were then sorted using Fluorescence Activated Cell Sorting (FACS; London Regional Flow Cytometry Facility, Robarts Research Institute, London, Ontario, Canada) to obtain a population of cells fluorescing both green and red (ie. expressing both EGFP-MamI and Tomato-MamL, respectively). For co-expression with the pSF-FLAG-*mamL*-EMCV-FLuc construct, MDA-MB-435/EGFP-MamI cells were grown to 60-70% confluency on a 100 mm dish and transfected using Lipofectamine 2000 (Invitrogen), according to company protocol, using 8 µg of pSF-FLAG-*mamL*-EMCV-FLuc. After 16 hours, cells were placed in full medium for 48 hours before commencing antibiotic selection. To select stable cell lines, cells were grown in the presence of 500 µg/mL Geneticin (G418; Gibco) and 500 ng/mL Puromycin (Gibco).

Protein Sample Preparation

Stably transfected cells were cultured to 70% confluency on a 100 mm dish, then washed twice using 10 mL phosphate buffered saline pH 7.4 (PBS, 137 mM NaCl/2.7 mM KCl/10 mM

Na_2HPO_4). Four to five dishes of cells were then collected into a 1 mL lysis solution containing 850 µL radioimmunoprecipitation assay buffer (RIPA, 10 mM Tris-HCl pH 7.5/140 mM NaCl/1% NP-40/1% sodium deoxycholate/0.1% sodium dodecyl sulfate (SDS)) and 150 µL of Complete Mini protease inhibitor cocktail (Roche Diagnostic Systems, Laval, Canada). Harvested cells were sonicated using three 12-second bursts of a Sonic Dismembrator (model 500, Thermo Fischer Scientific, Ottawa, Canada) at an amplitude of 30%. Total amount of protein was quantified using the BCA assay (13).

Western Blot

Protein samples of MDA-MB-435 cells stably expressing EGFP (40 µg), Tomato (40 µg), EGFP-MamE (40 µg), or Tomato-MamB (40 µg) were reduced with 100 mM dithiothreitol in sample preparation buffer (1 M Tris-HCl pH 6.8/10% SDS/0.1% Bromophenol Blue/43% glycerol) and heated at 85˚C for at least 5 min. Reduced samples were then subjected to discontinuous SDS polyacrylamide gel electrophoresis (SDS-PAGE) using a 10% running gel. Protein was transferred onto a nitrocellulose blot using the Original iBlot Gel Transfer Device (Life Technologies, Burlington, Canada).

For EGFP detection, nonspecific protein binding was blocked in 5% bovine serum albumin (BSA)/Tris-buffered saline pH 7.4 (TBS) for 3 h at room temperature. Blots were then incubated for 15 h in 1:1000 mouse α-GFP (Invitrogen)/3% BSA/TBS/0.02 % sodium azide (TBSA); then washed using TBS/0.1% Tween 20 (TBST; Sigma-Aldrich, Oakville, Canada) for 30 min with 4 changes of buffer; and incubated for 2 h in 1:20,000 horseradish peroxidase (HRP)-conjugated goat α-mouse IgG (Sigma-Aldrich)/1% BSA/TBS. All incubations were performed at room temperature. Blots were then washed with 0.1% TBST for 30 min with 4

changes of buffer and imaged using the Chemigenius Gel Doc (Syngene). A chemiluminescent signal was detected using SuperSignal West Pico Chemiluminescent Substrate (Thermo Fischer Scientific), according to the manufacturer's instructions.

For Tomato detection, blots were blocked in 3% BSA/TBSA for approximately 18 h at room temperature and then incubated for 18 h in 1:1000 primary goat α-tdTomato (MyBioSource, San Diego, USA)/3% BSA/TBSA at 4°C. After washing in 0.1% TBST as described above, blots were incubated for 1 h in 1:20,000 HRP-conjugated rabbit α-goat IgG (Sigma-Aldrich)/1% BSA/TBS at room temperature.

Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as a loading control. For GAPDH detection, blots were placed in stripping solution (1 M Tris-HCl pH 6.8/10% SDS/0.016% β-mercaptoethanol) and agitated in a 37°C water bath for 30 min prior to washing in 0.1% TBST and blocking in 5% BSA/TBS. The primary and secondary antibodies were 1:2000 rabbit α-GAPDH (Sigma-Aldrich)/3% BSA/TBSA and 1:20,000 HRP-conjugated goat α-rabbit IgG (Sigma-Aldrich)/1% BSA/TBS, respectively.

Confocal Imaging

Stably transfected cell lines were examined with confocal fluorescence microscopy (using a Nikon A1R confocal microscope) to confirm expression and characterize the intracellular localization of EGFP-MamI, Tomato-MamL, Tomato-MamL$_{trunc,}$ FLAG-MamL, EGFP-MamE (T317A), and Tomato-MamB fusion proteins. In preparation for confocal microscopy, approximately 100 thousand cells were cultured in a 35 mm glass-bottom dish (MatTek Corporation, Cedarlane, Burlington, Canada) for 48 hours. On the day of imaging, the

dish was placed in a stage-top incubator to maintain 37°C and 5% CO_2. Images and cines were captured using a Galvano scanner with NIS-Elements AR 5.11.01 (Nikon Instruments Inc.), and a 20x air objective with 0.75 numerical aperture. To capture images of cells expressing a single fluorophore, the FITC microscope filter (495 nm excitation/519 nm emission) was used for cells expressing the EGFP fluorophore and the TRITC microscope filter (557 nm excitation/576 nm emission) was used for cells expressing the Tomato fluorophore. To capture images of cells co-expressing both fluorophores (EGFP and Tomato), the FITC and TRITC filters allowing simultaneous 495 nm and 557 nm excitations was used. When needed, captured images of cells in both channels were then merged in Adobe Photoshop CS7.

Time-lapses were acquired with the time lapse function in NIS-Elements AR 5.11.01, recording an image every 1 s for a total of 60 s. Cines were captured in either channel or both channels simultaneously, as described above. The NIS-Elements software automatically generated a time lapse video in nd2 format with single or merged channels. This video was then edited in Adobe Photoshop CS7 by cropping and then pasting into a single file which could be exported in a graphics interchange format (GIF).

Particle Trajectory Analysis

Particle Tracking

Using a plugin called Mosaic Particle Tracker 2D/3D (version 1.0.1) (14) in the software ImageJ version 1.8.0 (15), the trajectories of particles observed in cells expressing either Tomato-MamL, Tomato-MamL/EGFP-MamI, Tomato-$MamL_{trunc}$, Tomato-MamB, or EGFP-MamE (T317A) were determined from confocal cines, and analyzed.

Cines converted to GIF files were loaded into ImageJ and prepared for analysis by converting them to greyscale; cropping to reduce their size and retaining only the portion of the movie with a single cell; then optimizing brightness and contrast. This last step does not affect the particle tracking but makes it easier for the user to visualize trajectories. For accurate estimation of diffusion coefficients and velocities, the pixel width $d = L/N$ (where L is the width of the field of view in micron and L in number of pixels) and time interval between two consecutive frames $\boldsymbol{\tau} = T/(F-1)$ (where T is the total duration of the video and F is the total number of frames) were calculated and added to the image properties in ImageJ.

After launching Mosaic Particle Tracker, the parameters used for particle detection were optimized for each cine: values of the radius (size of the tracked particles), cutoff (threshold intensity value below which detected particles are rejected), and percentile (range of intensities below the maximum intensity in the image for which fluorescent spots are considered to be particles) were manually adjusted to allow the software to detect the most manifest particles in the first image of the movie while not picking up lower-intensity noise speckle. For particle linking properties, the link range was set to 3 for all cines. In this way, the software would stop tracking a specific particle if it was absent for 3 consecutive frames. The displacement (maximum displacement allowed for a particle between two consecutive frames), which should be set to at least twice the average displacement of a particle between two frames, was set to 10 pixels. The software provided the total number of trajectories detected, a file containing the information relative to each detected trajectory (i.e., the position of the particle in each frame for which it was detected) and the MSD, $\langle r^2(t)\rangle$, of each detected particle.

Simple Trajectory Analysis

The mobility of each particle was first assessed by fitting the MSD for each trajectory (within Mosaic Particle Tracker) with a simple power-law function:

$$\langle r^2(t)\rangle = 4D_a\, t^{\alpha} \qquad \text{Eq. 1}$$

A particle undergoing free Brownian motion should have an MSD close to linear in time - that is, with an exponent α close to 1. Constrained Brownian motion would result in $\alpha < 1$. In contrast, a particle undergoing directed motion should be characterized by α close to 2. If the motion is mixed, with alternate periods of Brownian and directed motion, then one would expect α to be somewhere between 1 and 2. As explained below, results of the fit of MSD was among the information used to differentiate between direct and Brownian motion. For the first group of particles, D_a can be considered to be an apparent diffusion coefficient (note that when $\alpha \neq 1$, this quantity does not have the dimension of a diffusion coefficient, as it only represents the diffusion coefficient that would be estimated from the fitted value of the MSD at t = 1 s, which is why it is more accurate to speak of an apparent diffusion coefficient).

Refined Algorithm for Trajectory Analysis

Starting from the position and intensity of the tracked particle in each frame returned by Mosaic Particle Tracker, trajectories were further categorized and analyzed using an algorithm written for Mathematica. The main steps of this algorithm are as follows.

Trajectories were first evaluated for inclusion or exclusion based on the following criteria. Trajectories were rejected when either too short (less than 8 frames) or having a frame-to-frame displacement that was too large: that is, more than about 1 micron between successive

frames or more than about 3 pixels between non successive frames, if the particle was not detected for one frame (these numbers were adjusted for each cell depending on the imaging parameters). Trajectories were also rejected if the average intensity of the particle was too low (less than one standard deviation below the mean intensity for all particles) or was detected for less than 70% of the trajectory duration.

For each of the remaining trajectories, the MSD was calculated and fitted (for lag times between 1 and 10 s) using Eq. 1. An estimate of the particle maximum velocity was obtained by identifying the maximum value of the correlation between two successive measurements of the particle apparent velocity:

$$\sqrt{(v \rightarrow (t) \cdot v \rightarrow (t + \tau))} \qquad \text{Eq. 2}$$

in which τ is the time interval between consecutive frames.

For all trajectories in a given cell, the distribution of single-step displacements (i.e., displacements between two successive frames, corresponding to a time interval $\tau = 1$ s) was generated and fitted with the expression below, which accounts for the existence of two diffusive populations (i.e., the sum of two Rayleigh's distributions):

$$p(r) = \frac{f_1}{{s_1}^2} * r * e^{\frac{-r^2}{2{s_1}^2}} + \frac{f_2}{{s_2}^2} * r * e^{\frac{-r^2}{2{s_2}^2}} \qquad \text{Eq. 3}$$

In some cases, a third term (a Gaussian peak attributed to steps taken during directed motion) was needed in order to account for a small fraction of longer displacement. Each of these populations is characterized by an average step size. The values of s_1 and s_2 (in order of increasing length) became the basis for sorting each individual trajectory into either an immobile,

directed, or diffusive trajectory. The criteria used were as follows. For immobile trajectories, the total displacement and largest step are both less than $12 \times s_1$. For a directed trajectory, the total displacement is more than $1.5 \times s_2 \times \sqrt{(n)}$, where n is the number of steps in the trajectory, or the trajectory MSD is characterized by an exponent $\alpha > 1.1$, as explained above in the section on simple trajectory analysis. Diffusive trajectories are those that are neither immobile nor directed.

Results

The following subsections describe the localization of essential magnetosome proteins in a human melanoma cell line, as well as their intracellular mobility.

Cellular distribution of MamE and MamB

The magnetosome protein MamE consists of both conserved and variable sequences. We compared the expression of two EGFP fusion protein constructs to examine potential differences in their intracellular expression in MDA-MB-435 cells. Using confocal fluorescence microscopy, cells stably expressing EGFP-*mamE* (G49S, T641S) display intracellular fluorescence in a punctate pattern with a diffuse fluorescence background (Figure 3). These EGFP-MamE structures are numerous and cluster near the nucleus. Protein expression was confirmed by immunoblot (Figure 4 and Supplementary Figure 3) and reveals potential autoproteolysis of MamE as previously reported in MTB (16). Unlike the expression of MamL (10) or MamB (described below), MamE-expressing particles do not display mobility.

Cells expressing EGFP-*mamE* (T317A) confirm that both *mam*E constructs show similar expression patterns. Confocal fluorescence microscopy of cells expressing EGFP-*mamE* (T317A) shows distinct intracellular, punctate green fluorescence that clusters near the nucleus

(Figure 5). Similar to EGFP-MamE (G49S, T641S), the expression of EGFP-MamE (T317A) also exhibits a diffuse background and little or no mobility.

When Tomato-MamB is stably expressed in MDA-MB-435 cells, red fluorescence is displayed in a punctate pattern in the majority of the transfected cell population (~70%; Fig. 6A) and in a diffuse pattern in about 30% of the transfected cell population (Fig. 6B). Importantly, these punctate MamB structures are mobile. Protein expression was confirmed by immunoblot (Figure 7 and Supplementary Figure 4).

Analysis of EGFP-MamE trajectories

As no evident motion of the puncta associated with MamE protein was observed in confocal movies, the analysis of EGFP-MamE (T317A) cells were chosen to serve as an example of cells containing particles with low mobility. Using this reference point, the signature of these MamE particles is contrasted with that of the higher mobility particles observed in other expression systems. A visual representation of MamE trajectories and the distribution of their displacement is shown in Figures 8 and 9, respectively. It is apparent from this distribution that the motion of the MamE particles is very highly constrained, since almost no single displacement after t = 1s is larger than 0.2 µm, and since the displacements after 20 s are almost the same as after 1 s. Trajectory analysis of MamE fluorescent structures showed that 26% of these particles were immobile while 62% were undergoing Brownian motion and only 13% underwent directed motion at some point in their trajectories (Table 2). According to this analysis, the average apparent diffusion coefficient of MamE particles undergoing Brownian motion is $1.9 \pm 0.4 \times 10^{-3}$ $\mu m^2/s$. The average velocity of the very few MamE particles undergoing directed motion is 0.17 ± 0.03 µm/s.

Mobility of Tomato-MamB

Analyses of Tomato-MamB particle trajectories also reveal both directed and Brownian motion in each cell, with the occurrence of directed trajectories markedly greater than in MamE-expressing cells. A visual representation of the detected MamB particles is shown in Figure 10, and their displacement distribution is shown in Figure 11. The difference between MamE and MamB is clearly shown by the large number of displacements after t = 1 s between 0.2-0.4 µm and by the broad distribution of displacements after 20 s, as expected for mobile particles. Analysis of MamB trajectories showed that only 15% of MamB particles were immobile, while 57% were undergoing Brownian motion, and 28% underwent directed motion at some point in their trajectories (Table 2). The average apparent diffusion coefficient of MamB particles undergoing Brownian motion is $4.0 \pm 1.0 \times 10^{-3}$ $\mu m^2/s$. The average velocity of MamB particles undergoing directed motion is 0.24 ± 0.01 µm/s.

Mobility of Tomato-MamL

Previous analyses of Tomato-MamL particles (10) reveal a similar pattern to those of Tomato-MamB with respect to number of trajectories with directed and Brownian motion. A visual representation of detected Tomato-MamL particles is shown in Figure 12, and their displacement distribution is shown in Figure 13, where a broad distribution after 20 s is apparent. Analysis of these results showed that 18% of MamL particles were immobile, 56% were undergoing Brownian motion, and 25% underwent directed motion at some point in their trajectories (Table 2). The average apparent diffusion coefficient of MamL particles undergoing Brownian motion is $5.1 \pm 1.7 \times 10^{-3}$ $\mu m^2/s$. The average velocity of MamL particles undergoing directed motion is 0.19 ± 0.05 µm/s.

Mobility of Tomato-MamL$_{trunc}$

To better understand the mobility of MamL, we introduced an early stop codon to remove 15 amino acids from the C terminus of the MamL protein and examined this truncated form. The C-terminal peptide that was removed may interact with mobile cytoskeletal elements based on the similarity to cell-penetrating peptides (2). A visual representation of the detected MamL$_{trunc}$ particles is shown in Figure 14, and their displacement distribution is shown in Figure 15. Analysis of MamL$_{trunc}$ trajectories showed that 14% of MamL$_{trunc}$ particles were immobile, 60% were undergoing Brownian motion, and 26% underwent directed motion at some point in their trajectories (Table 2). While the distribution of particles is similar to native MamL, the average apparent diffusion coefficient of MamL$_{trunc}$ particles undergoing Brownian motion has been reduced more than 2-fold to $1.9 \pm 1.1 \times 10^{-3}$ $\mu m^2/s$. The average velocity of MamL$_{trunc}$ particles undergoing directed motion is 0.14 ± 0.06 µm/s and comparable to MamL particles.

Mobility of Tomato-MamL/EGFP-MamI

Tomato-MamL particles not only retain their mobility when co-expressed with EGFP-MamI, the latter appear to be recruited to the same mobile complex (10). These mobile particles, consisting of co-localized protein, also display both directed and Brownian motion. A visual representation of the detected MamL+MamI particles is shown in Figure 16 and their displacement distribution in Figure 17. Analysis of the MamL+MamI trajectories revealed that 21% of MamL+MamI particles were immobile, 45% were undergoing Brownian motion, and 32% underwent directed motion at some point in their trajectories (Table 2). The average apparent diffusion coefficient of MamL+MamI particles undergoing Brownian motion is 3.2 ±

2.5×10^{-3} μm^2/s while the average velocity of MamL+MamI particles undergoing directed motion is 0.23 ± 0.09 μm/s, consistent with the other expression systems examined.

Mobility of FLAG-MamL/EGFP-MamI

A FLAG-MamL/EGFP-MamI co-expression system was created to simplify confocal image analysis using a single chromophore (EGFP) and to remove bulky fluorescent proteins (i.e. tandem Tomato fluorescence) attached to MamL that may compromise magnetosome protein interactions (10). The trajectories of particles from stable FLAG-MamL/EGFP-MamI co-expression were analyzed and compared to those of the dual fluorescent Tomato-MamL/EGFP-MamI particles. A visual representation of detected FLAG-MamL+MamI particles is shown in Figure 18, and their displacement distribution is shown in Figure 19. Similar to Tomato-MamL/EGFP-MamI, analysis of FLAG-MamL/EGFP-MamI trajectories showed that a minority of the particles were immobile (21%) while the majority of particles were undergoing Brownian motion (45%) and directed motion (32%) at some point in their trajectories (Table 2). The average apparent diffusion coefficient of FLAG-MamL+MamI particles undergoing Brownian motion is $5.0 \pm 0.9 \times 10^{-3}$ μm^2/s. The average velocity of FLAG-MamL+MamI particles undergoing directed motion is 0.23 ± 0.09 μm/s.

Discussion

This is the first report to describe the stable expression of MamE and MamB in mammalian cells and analyze the intracellular mobility of punctate structures formed by magnetosome proteins expressed in this eukaryotic environment. For each fluorescent fusion protein analyzed (MamL, $MamL_{trunc}$, MamL+MamI, MamB, or MamE), three types of particle

movement were observed for the puncta: confined trajectories, Brownian trajectories, and trajectories with stretches of directed motion.

Both MamE and MamB fusion proteins displayed intracellular, punctate fluorescence similar to the pattern of MamL expression (10). However, EGFP-MamE-expressing cells often displayed punctate fluorescence in the perinuclear region while approximately 30% of Tomato-MamB-expressing cells had no discernable punctate fluorescence. Also unlike MamL expression, both MamB- and MamE-expressing cells displayed diffuse fluorescence.

Trajectories undergoing Brownian motion have apparent diffusion coefficients between 2 x 10^{-3} and 5 x 10^{-3} $\mu m^2/s$. This diffusion coefficient is very slow compared to that of a soluble protein in the cytoplasm of mammalian cells, which is on the order of 10 $\mu m^2/s$ (17). Assuming that MamI, MamL, MamB, and MamE are all membrane associated proteins when expressed in mammalian cells, as in MTB, the measured diffusion coefficients are consistent with the notion that these magnetosome proteins are part of a rather large (membranous) structure and not simply part of the soluble protein fraction. Since these punctate structures cannot be resolved by confocal microscopy, the radius of these fluorescent particles is expected to be approximately less than or equal to 100 nm (i.e. smaller than the resolution of the microscope). Hence, the possibilities for particles of this diameter include localization in the membrane of pre-existing lipid vesicles (18) or a role in the formation of new ones. Confined trajectories may thus correspond to particles localized in immobilized vesicles.

Particles that undergo active directed motion have a maximum velocity from 0.1 to 0.3 µm/s. These particles usually also undergo stretches of Brownian motion or immobility. The fluorescent magnetosome fusion proteins may be localized in vesicles that are attached to

molecular motors and/or may be directly interacting with molecular motors as they transiently walk along protein filaments.

Brownian motion of magnetosome proteins

MamE particles are either immobile (26%) or undergoing some form of Brownian motion (62%); however, several features of this analysis point toward a restricted type of Brownian motion. First, Brownian MamE particles have an apparent diffusion coefficient of $0.8 \pm 0.1 \times 10^{-3}$ $\mu m^2/s$, which is significantly lower than the diffusion of MamL and MamB particles (Table 3). This suggests that MamE is localized in different structures than MamL and MamB, with more restricted mobility. Additionally, Brownian MamE particles have trajectories characterized by an extremely low anomalous exponent of 0.15 ± 0.04, the lowest α value of all the magnetosome protein expression systems examined. An α value below 1 points to constrained or restricted diffusion (i.e. diffusion in the presence of obstacles), and a value below 0.5 further points towards caged diffusion (19, 20). Taken together, the low diffusion coefficient of MamE particles, extremely low value of the anomalous exponent α, and the high occurrence of stationary MamE particles suggest that this protein is the most constrained out of the magnetosome proteins studied in mammalian cells.

In contrast to MamE, Brownian MamL and MamB particles have a noticeably larger diffusion coefficient (on the order of $4\text{-}5 \times 10^{-3}$ $\mu m^2/s$), and a larger anomalous exponent α (on the order of 0.4 to 0.5). These values are consistent with magnetosome proteins localizing on particles (e.g. lipid vesicles) that are tens to hundreds of nm in size and undergoing restricted diffusion (18). $MamL_{trunc}$ particles, which diffuse at $1.9 \pm 1.1 \times 10^{-3}$ $\mu m^2/s$, diffuse significantly slower than MamL particles ($p < 0.001$). Removal of the MamL C-terminal peptide may cause the protein to localize in different (more confined) structures than full-length MamL.

Alternatively, the MamL C-terminal may confer mobility, either through its association with the particle structure or by binding directly to molecular motors.

When MamL is co-expressed with MamI, the diffusion coefficient of Tomato-MamL/EGFP-MamI particles ($3.2 \pm 2.5 \times 10^{-3}$ $\mu m^2/s$) is no different than the diffusion of structures with MamL alone ($5.1 \pm 1.7 \times 10^{-3}$ $\mu m^2/s$). This is consistent with the conclusion that MamL and MamI are localized within the same structures/particles, and that MamI is recruited to the same structure as MamL when they interact. This is further supported by the FLAG-MamL/EGFP-MamI expression system. These particles diffuse at $5.0 \pm 0.9 \times 10^{-3}$ $\mu m^2/s$, which is not significantly different from the diffusion coefficient of particles containing either Tomato-MamL/EGFP-MamI or Tomato-MamL alone. Furthermore, the α value of Tomato-MamL alone, Tomato-MamL/EGFP-MamI, and FLAG-MamL/EGFP-MamI particles are all comparable.

Directed motion of magnetosome proteins

Although the magnetosome proteins studied undergo directed motion to a different extent (i.e. variability in the percentage of detected particles that display directed motion and in the average value of the α exponent for these directed trajectories), similar velocities (around 0.2 $\mu m/s$) were measured for all particles (Table 4), and except for MamE particles, the aspect of these directed trajectories (which very often are accompanied by period of intermittent Brownian motion) is similar for all studied proteins (Figure 20).

The directed motion observed suggests that detected particles interact with some kind of linear molecular motor present in the cell, for example a type of myosin motor. In mammalian cells, myosin motors are a group of molecular motor proteins responsible for directed movement of intracellular cargo. These structures, specifically myosin IB, myosin II, and myosin V, all

move at a velocity of 0.2 μm/s (21). The speeds observed for the detected particles are therefore consistent with interactions with myosin. Molecular motors with a negative charge, or that associate with cargo adaptor proteins that are negatively charged, may interact with positively-charged proteins in the cell (11).

MamE stands out from other magnetosome proteins studied here in that many fewer MamE particles are undergoing directed motion and/or they spend less time undergoing active directed motion (Fig. 20). This is further indicated by the low α exponent of 0.26 ± 0.32. Particles undergoing directed motion should have α exponents close to 2; however, a lower exponent can also indicate that the directed motion alternates with either Brownian motion or immobilized motion, as is the case with all particles detected and especially with MamE. When undergoing directed motion, MamE particles move at a velocity of 0.17 μm/s. From protein structure analysis, the cytoplasmic C-terminal polypeptide of MamE is short and has only 2 positively-charged amino acids; therefore, MamE most likely does not bind to molecular motors strongly or frequently due to the small positive charge. Alternatively, MamE may be located in a vesicle that does not interact often with molecular motors. This may explain why MamE is not undergoing directed motion as often and why the protein seems mostly immobile while observed under a confocal fluorescence microscope.

MamB particles undergoing active directed motion travel at 0.24 μm/s. From protein structure analysis, the cytoplasmic domain of MamB is quite large and almost 20% of its amino acids have positive charges. These positive charges, therefore, have opportunities to interact with negatively-charged molecular motors in the mammalian cytoplasm. We can thus infer that MamB either interacts directly with molecular motors or is located in a vesicle that interacts with molecular motors.

MamL particles undergoing active directed motion move at approximately 0.2 µm/s, and are thus expected to also be interacting with molecular motors (i.e. myosin). Although the cationic amino acid residues of the MamB and MamL cytoplasmic domains are not arranged in a similar pattern, we can speculate that these two proteins may be interacting with similar molecular motors since they move at similar velocities. $MamL_{trunc}$ particles that are undergoing active directed motion travel at 0.15 µm/s, which is not significantly different from the velocity of full-length MamL particles. It is notable to point out that the α exponent of $MamL_{trunc}$ particles is smaller than that of MamL particles (0.88 ± 0.31 vs. 1.01 ± 0.18). This result aligns with what is seen with the confocal microscope: that $MamL_{trunc}$ particles do not undergo directed motion as frequently as full-length MamL particles (10).

Tomato-MamL/EGFP-MamI and FLAG-MamL/EGFP-MamI particles undergoing active directed motion move at similar velocities as Tomato-MamL alone (0.23 µm/s vs. 0.14 µm/s vs. 0.20 µm/s, respectively). This suggests that Tomato-MamL alone, Tomato-MamL/EGFP-MamI, and FLAG-MamL/EGFP-MamI particles interact with the same molecular motors. Unlike the predicted protein structures of MamL, MamI has no cytoplasmic domain and thus has no opportunity to interact with molecular motors (22). Whether expressed alone or together with MamI, MamL is the protein that likely interacts with molecular motors.

Conclusion

The mobility of magnetosome proteins MamL, MamB, and MamE in a mammalian system indicates their possible interaction with mammalian molecular motors. The pattern of motion of the structures to which these proteins localize is consistent with the motion of lipid vesicles (endosomes, lysosomes) observed in mammalian cells, namely a mixture of Brownian

motion (with very small diffusion coefficient on the order of 10^{-3} to 10^{-2} $\mu m^2/s$) and directed motion (with velocity on the order of 0.1 to 1 μm/s) (23, 24). Moreover MamL, MamB, and MamE all move at similar velocities when undergoing active directed motion. This suggests that, when expressed in mammalian cells, these magnetosome proteins may interact with the same type of molecular motor or localize to a structure that interacts with a common type of molecular motor.

When co-expressed, MamI and MamL localize to the same intracellular compartment and form punctate, mobile structures. Importantly, MamL+MamI particles undergoing directed motion move at the same velocity as particles in cells expressing MamL alone, suggesting they interact with the same molecular motor. In addition, MamL+MamI particles diffuse slower than particles with MamL alone, consistent with the expected slower diffusion of larger structures relative to smaller ones.

Taken together, this trajectory analysis supports mounting evidence that magnetosome proteins are compatible with mammalian cell systems, and localize and interact with mammalian membranous compartments. The diffusion coefficients and similar velocities at which these essential magnetosome proteins travel inside the cell are consistent with their proposed localization on an intracellular membrane. Finally, the evidence for co-localization and interaction of MamI and MamL is further substantiated through analysis and comparison of the mobility of particles with MamL alone versus MamL+MamI. This study is thus consistent with our hypothesis that essential magnetosome proteins MamI, MamL, MamB, and MamE interact in a membranous intracellular compartment in mammalian cells.

References

1. Goldhawk DE, Gelman N, Thompson RT, Prato FS. Forming Magnetosome-Like Nanoparticles in Mammalian Cells for Molecular MRI. In: Bulte JWM, Modo MMJ, editors. Design and Applications of Nanoparticles in Biomedical Imaging. Cham: Springer International Publishing; 2017. p. 187-203.
2. Komeili A. Molecular mechanisms of compartmentalization and biomineralization in magnetotactic bacteria. FEMS Microbiol Rev. 2012;36(1):232-55.
3. Uebe R, Schüler D. Magnetosome biogenesis in magnetotactic bacteria. Nat Rev Microbiol. 2016;14(10):621-37.
4. Lefèvre CT, Trubitsyn D, Abreu F, Kolinko S, Jogler C, de Almeida LG, et al. Comparative genomic analysis of magnetotactic bacteria from the Deltaproteobacteria provides new insights into magnetite and greigite magnetosome genes required for magnetotaxis. Environ Microbiol. 2013;15(10):2712-35.
5. Murat D, Quinlan A, Vali H, Komeili A. Comprehensive genetic dissection of the magnetosome gene island reveals the step-wise assembly of a prokaryotic organelle. Proc Natl Acad Sci U S A. 2010;107(12):5593-8.
6. Raschdorf O, Forstner Y, Kolinko I, Uebe R, Plitzko JM, Schüler D. Genetic and Ultrastructural Analysis Reveals the Key Players and Initial Steps of Bacterial Magnetosome Membrane Biogenesis. PLoS Genet. 2016;12(6):e1006101.
7. Uebe R, Junge K, Henn V, Poxleitner G, Katzmann E, Plitzko JM, 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. Mol Microbiol. 2011;82(4):818-35.
8. Uebe R, Keren-Khadmy N, Zeytuni N, Katzmann E, Navon Y, Davidov G, et al. The dual role of MamB in magnetosome membrane assembly and magnetite biomineralization. Molecular Microbiology. 2018;107(4):542-57.
9. Quinlan A, Murat D, Vali H, Komeili A. The HtrA/DegP family protease MamE is a bifunctional protein with roles in magnetosome protein localization and magnetite biomineralization. Mol Microbiol. 2011;80(4):1075-87.
10. Sun, Q, Yu, L, Donnelly, SC, Fradin, C, Thompson, RT, Prato, FS and Goldhawk, DE Essential magnetosome proteins MamI and MamL from magnetotatic bacteria interact in mammalian cells. 2023. BIORXIV/2023/572379.
11. Schmidt N, Mishra A, Lai GH, Wong GC. Arginine-rich cell-penetrating peptides. FEBS Lett. 2010;584(9):1806-13.
12. Rohani R, Figueredo R, Bureau Y, Koropatnick J, Foster P, Thompson RT, et al. Imaging tumor growth non-invasively using expression of MagA or modified ferritin subunits to augment intracellular contrast for repetitive MRI. Mol Imaging Biol. 2014;16(1):63-73.
13. Smith P, Krohn R, Hermanson G, Mallia A, Gartner F, Provenzano M, et al. Measurement of protein using bicinchoninic acid. Anal Biochem. 1985;150:76-85.
14. Sbalzarini IF, Koumoutsakos P. Feature point tracking and trajectory analysis for video imaging in cell biology. J Struct Biol. 2005;151(2):182-95.
15. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5.
16. Hershey DM, Browne PJ, Iavarone AT, Teyra J, Lee EH, Sidhu SS, et al. Magnetite

Biomineralization in Magnetospirillum magneticum Is Regulated by a Switch-like Behavior in the HtrA Protease MamE. J Biol Chem. 2016;291(34):17941-52.
17. Wachsmuth M, Waldeck W, Langowski J. Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. J Mol Biol. 2000;298(4):677-89.
18. Bandyopadhyay D, Cyphersmith A, Zapata JA, Kim YJ, Payne CK. Lysosome transport as a function of lysosome diameter. PLoS One. 2014;9(1):e86847.
19. Weeks ER, Weitz DA. Subdiffusion and the cage effect studied near the colloidal glass transition. Chemical Physics. 2002;284(1):361-7.
20. Kapanidis AN, Uphoff S, Stracy M. Understanding Protein Mobility in Bacteria by Tracking Single Molecules. J Mol Biol. 2018;430(22):4443-55.
21. Howard J. Mechanics of Motor Proteins and the Cytoskeleton Sunderland. 552001.
22. Nudelman H, Zarivach R. Structure prediction of magnetosome-associated proteins. Front Microbiol. 2014;5:9.
23. Matteoni R, Kreis TE. Translocation and clustering of endosomes and lysosomes depends on microtubules. J Cell Biol. 1987;105(3):1253-65.
24. Cordonnier MN, Dauzonne D, Louvard D, Coudrier E. Actin filaments and myosin I alpha cooperate with microtubules for the movement of lysosomes. Mol Biol Cell. 2001;12(12):4013-29.

Tables

Table 1. Primer design for the cloning of MTB genes *mamB* and *mamE* into ptdTomato-C1 and pEGFP-C1, respectively.

Gene	Primer (5' – 3')		Restriction Site*	Vector
mamB	Forward	CAATCTTGTG**GAATTC**AGAACCG	*EcoRI*	ptdTomato-C1
	Reverse	ACGCTCTGG**CCCGGG**ATGTCC	*SmaI*	
mamE	Forward	ACCCTG**AGATCT**GGATGGTTG	*BglII*	pEGFP-C1
	Reverse	GCCATTATCC**GAGCTC**CACCA	*SacI*	

*Restriction enzyme sites appear in bold in the primer sequence.

Table 2. Summary of trajectory analysis in mammalian cells expression magnetosome proteins.

Fusion Protein	Number of cells	Number of analyzed trajectories	Percent immobile	Percent Brownian	Percent directed	*Apparent diffusion coefficient (10^{-3} μm^2/s)	*Velocity (μm/s)
EGFP-MamE	4	47	26	62	13	1.9 ± 0.4	0.17 ± 0.03
Tomato-MamB	5	571	15	57	28	4.0 ± 1.0	0.24 ± 0.01
Tomato-MamL	7	461	18	56	25	5.1 ± 1.7	0.19 ± 0.05
Tomato-MamL$_{trunc}$	7	516	14	60	26	1.9 ± 1.1	0.14 ± 0.06
Tomato-MamL + EGFP-MamI	4	225	21	45	32	3.2 ± 2.5	0.23 ± 0.09
FLAG-MamL + EGFP-MamI	5	520	8	60	22	5.0 ± 0.9	0.14 ± 0.05

*The apparent diffusion coefficient is measured from particles undergoing Brownian motion. The velocity is measured from particles undergoing directed motion.

Table 3. Summary of apparent diffusion coefficient and anomalous exponent values for magnetosome proteins particles undergoing Brownian motion.

Magnetosome protein	% Particles undergoing Brownian motion	*Apparent diffusion coefficient (10^{-3} µm^2/s)	α value
EGFP-MamE	62	0.8 ± 0.1 †, §	0.15 ± 0.04
Tomato-MamB	57	4.0 ± 1.0 §	0.54 ± 0.08
Tomato-MamL	56	5.1 ± 1.7 †, ‡	0.43 ± 0.17
Tomato-MamL$_{trunc}$	60	1.9 ± 1.1 ‡	0.52 ± 0.11
Tomato-MamL + EGFP-MamI	45	3.2 ± 2.5	0.42 ± 0.11
FLAG-MamL + EGFP-MamI	60	5.0 ± 0.9	0.41 ± 0.11

* Data are the mean ± standard deviation †, ‡ $p < 0.001$ § $p < 0.05$

Table 4. Summary of velocity of magnetosome proteins undergoing directed motion.

Magnetosome protein	% Particles undergoing directed motion	* Velocity (μm/s)	α value
EGFP-MamE	13	0.17 ± 0.03	0.26 ± 0.32
Tomato-MamB	28	0.24 ± 0.01	1.24 ± 0.04
Tomato-MamL	25	0.19 ± 0.05	1.01 ± 0.18
Tomato-MamL$_{trunc}$	26	0.14 ± 0.06	0.88 ± 0.31
Tomato-MamL + EGFP-MamI	32	0.23 ± 0.09	0.66 ± 0.10
FLAG-MamL + EGFP-MamI	22	0.14 ± 0.05	0.95 ± 0.33

* Data are the mean ± standard deviation.

Figures

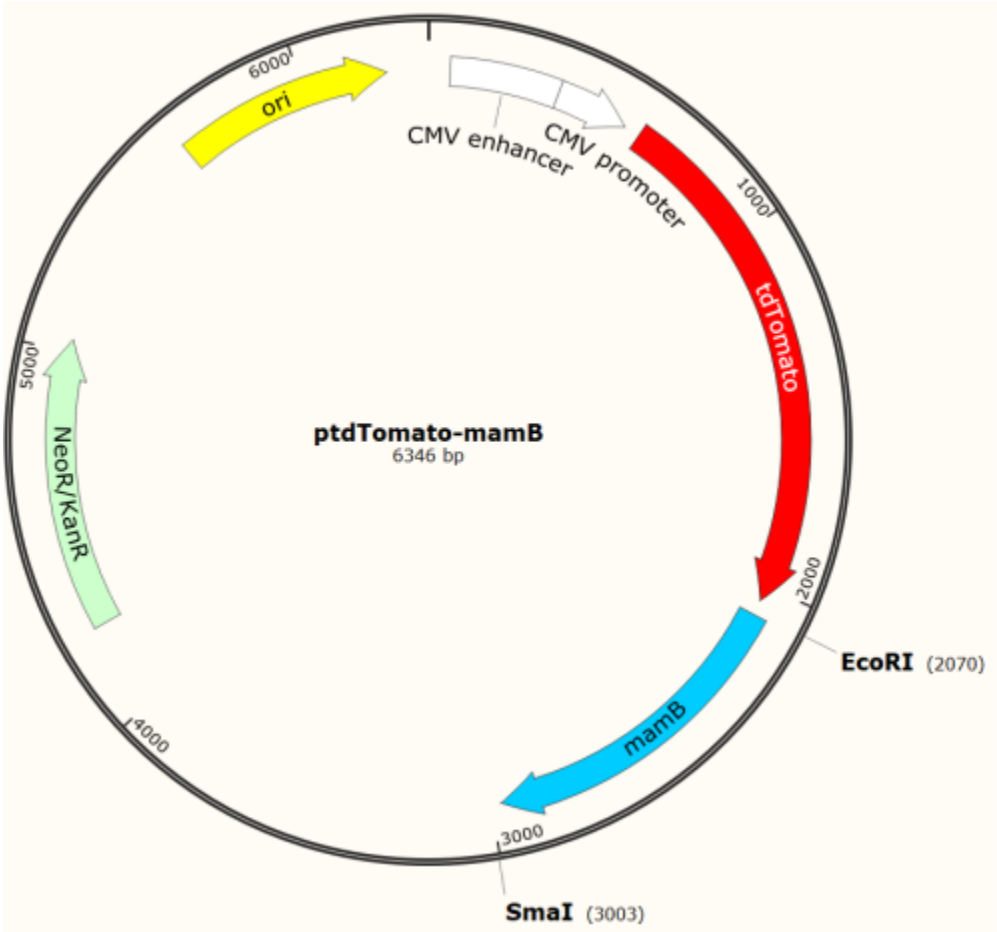


Figure 1. Vector-insert map of the cloning strategy of magnetosome gene *mamB*.

The vector-insert map of *ptdTomato-mamB* shows *mamB* (in blue) cloned in downstream of tdTomato (red) between restriction sites EcoRI and SmaI. The bacterial origin of replication (ori) is shown in yellow. The mammalian CMV enhancer and promoter, which allow for constitutive gene expression, is shown in white. The NeoR/KanR/G418R gene (light green) provides kanamycin antibiotic resistance in bacterial cells and G418 antibiotic resistance in mammalian cells.

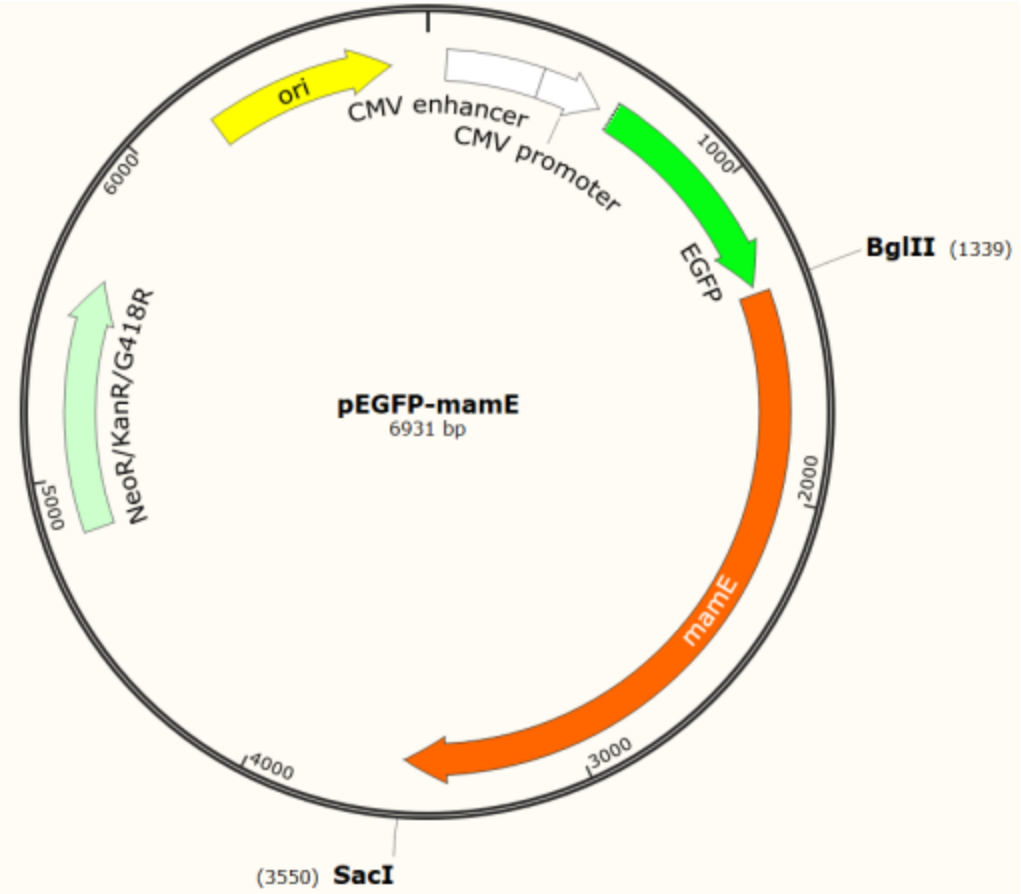


Figure 2. Vector-insert map of the cloning strategy of magnetosome gene *mamE*.

The vector-insert map of *pEGFP-mamE* shows *mamE* (orange) cloned in downstream of EGFP (bright green) between restriction sites BglII and SacI. The bacterial origin of replication (ori) is shown in yellow. The mammalian CMV enhancer and promoter, which allows for constitutive gene expression, is shown in white. The NeoR/KanR/G418R gene (light green) provides kanamycin antibiotic resistance in bacterial cells and G418 antibiotic resistance in mammalian cells.

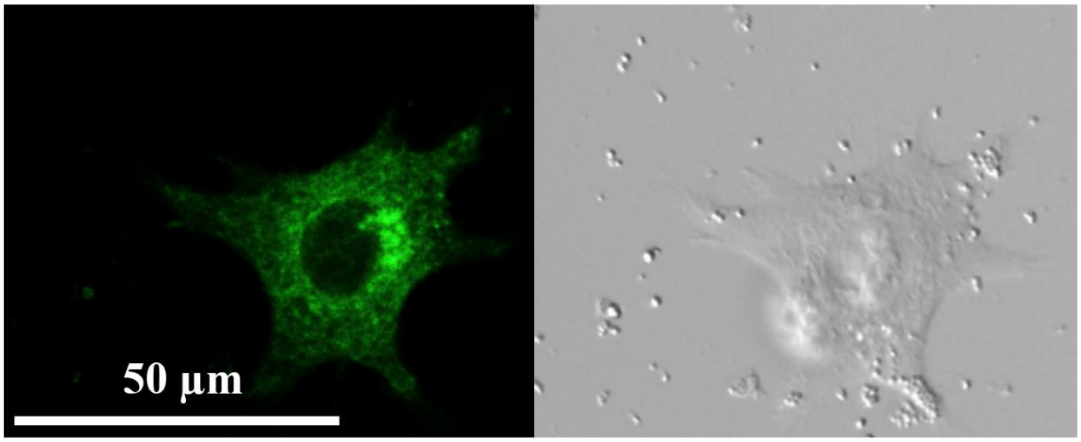


Figure 3. Confocal fluorescence microscopy of a mammalian cell stably expressing EGFP-MamE (G49S, T641S).

The EGFP-MamE fusion protein displays a punctate intracellular fluorescence pattern. These structures are more numerous around the nucleus.

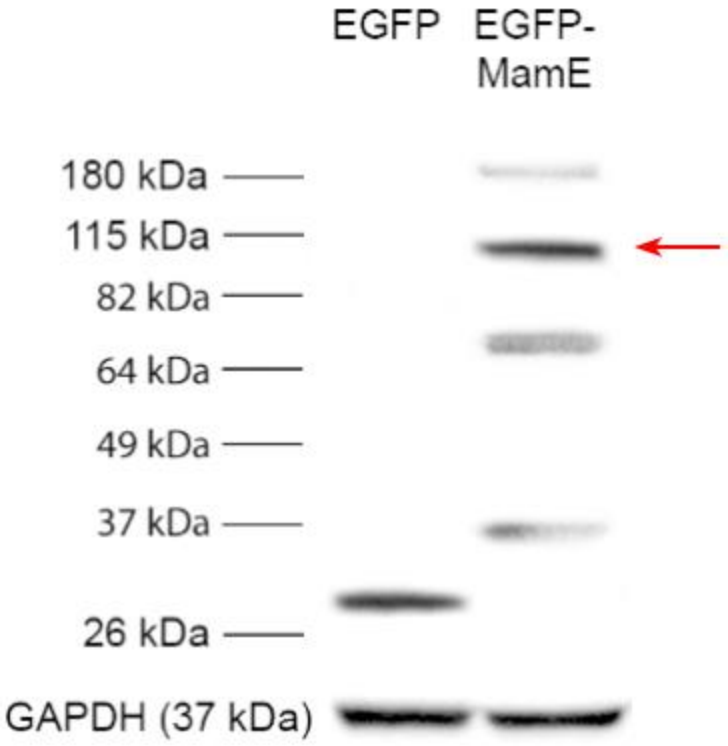


Figure 4. Expression of fluorescent EGFP-MamE fusion protein in mammalian cells.

Total cellular protein from MDA-MB-435 cells stably expressing EGFP or EGFP-MamE (G49S, T641S) were examined by western blot using mouse α-EGFP as the primary antibody. The approximate size of EGFP-MamE is 110 kDa (red arrow). Bands at lower molecular weights are products of MamE proteolysis (15). Approximate MW is shown in the left margin. The loading control was GAPDH and appears in the bottom panels.

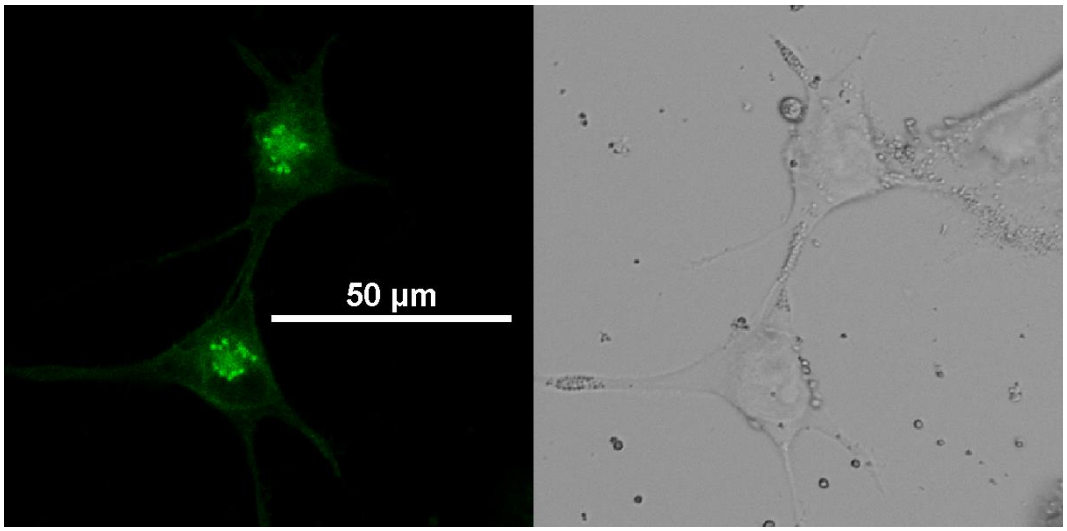


Figure 5. Confocal fluorescence microscopy of mammalian cells stably expressing EGFP-MamE (T317A).

Transfected cells were placed under selection to obtain populations expressing EGFP fused to MamE. EGFP-MamE (T317A) fusion protein displays a punctate intracellular fluorescence pattern.

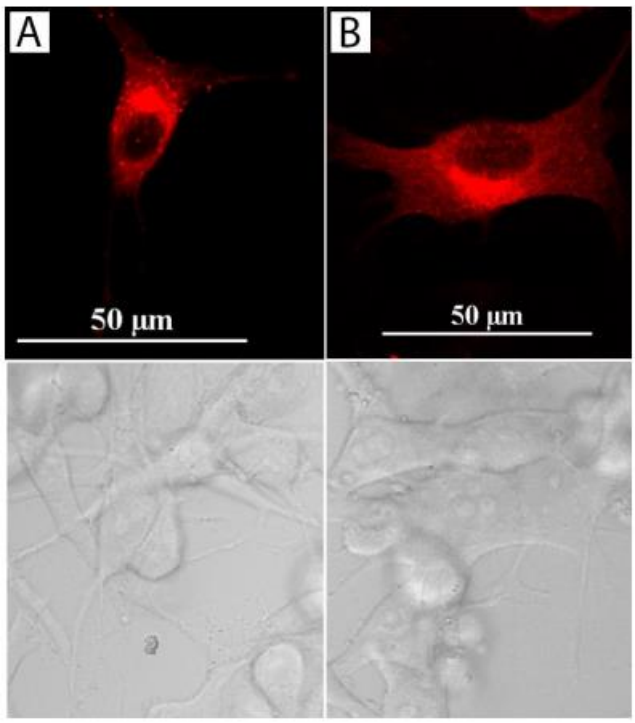


Figure 6. Confocal fluorescence microscopy of mammalian cells stably expressing Tomato-MamB fusion protein.

Transfected cells were placed under selection to obtain populations expressing Tomato fused to MamB. Tomato-MamB fusion protein displays either a punctate intracellular fluorescence pattern (A) or a diffuse pattern (B). The punctate structures in (A) are dispersed throughout the cell.

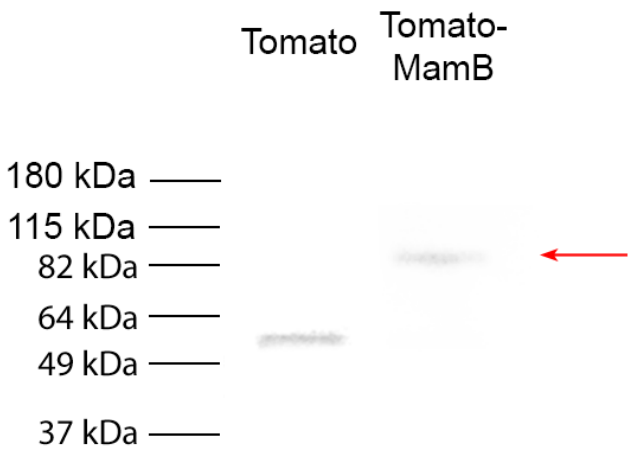


Figure 7. Expression of fluorescent Tomato-MamB fusion protein in mammalian cells.

Total cellular protein content from MDA-MB-435 cells stably expressing Tomato or Tomato-MamB were examined by western blot using rabbit α-Tomato as the primary antibody. The approximate size of Tomato-MamB is 91 kDa (red arrow). Approximate MW is shown in the left margin.

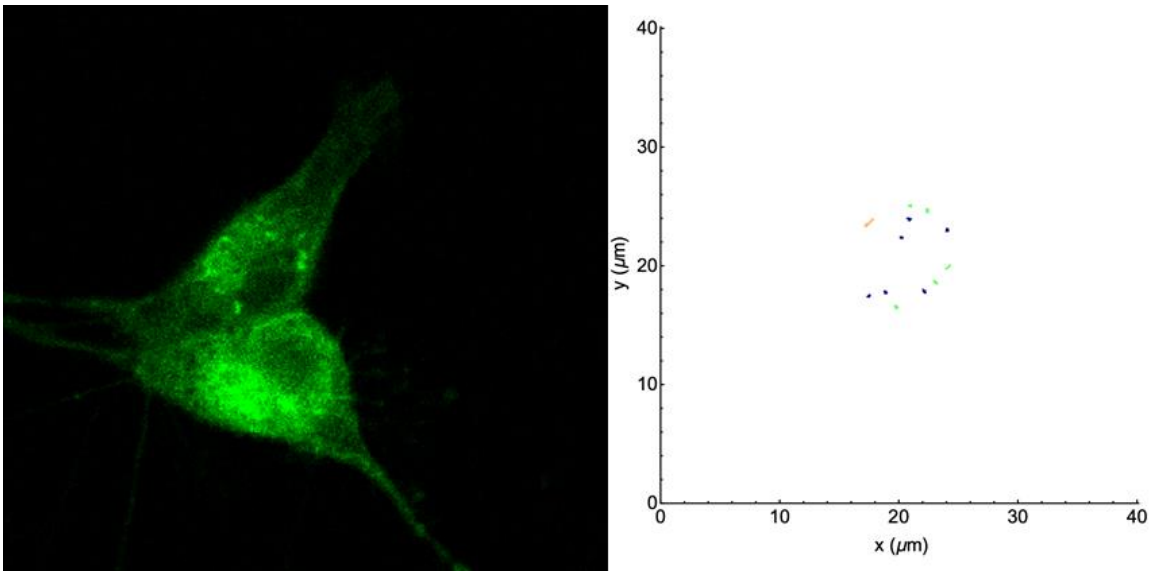


Figure 8. Representative cell expressing MamE particle trajectories.

A confocal image of the analyzed cell is shown on the left. Analyzed trajectories are shown on the right, with stationary particles shown in dark blue, and Brownian motion and directed motion trajectories represented in green and orange, respectively.

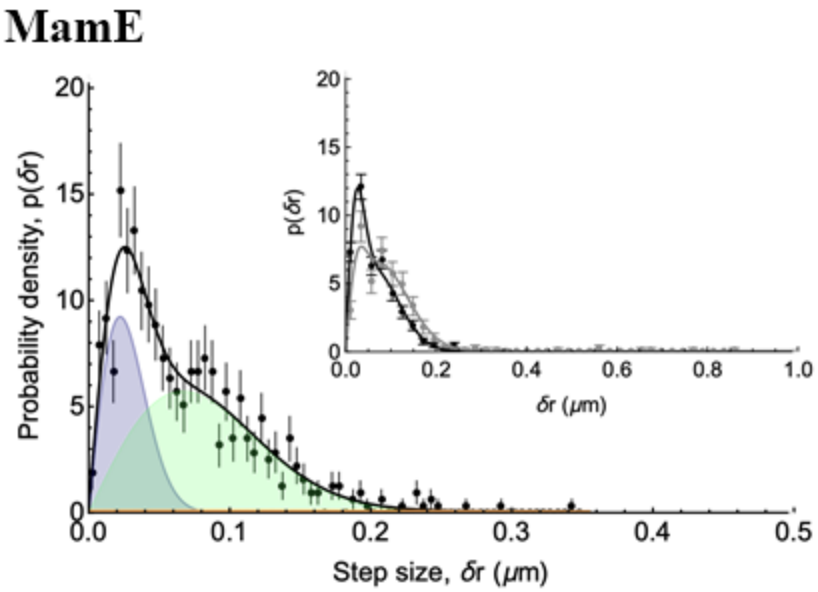


Figure 9. Distribution of displacements of punctate structures in a representative EGFP-MamE-expressing cell.

In the main panel the probability density of displacements after t = 1s is shown (black symbols) and fitted assuming two diffusive populations and one directed population. The blue and green shaded peaks represent the contributions of the immobile and Brownian diffusing populations, the orange shaded peak - barely visible in this case – is that of the directed population and the black curve shows the sum of these three contributions. The inset shows that there is very little difference in the distribution of displacements after 1 s (black symbols) and 20 s (grey symbols).

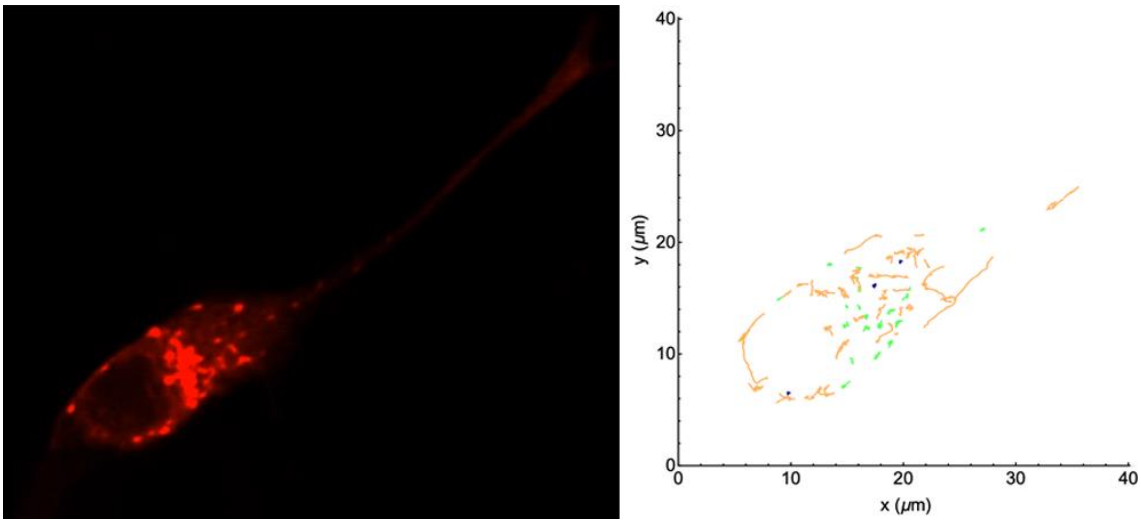


Figure 10. Representative cell expressing MamB particle trajectories.

A confocal image of the analyzed cell is shown on the left. Analyzed trajectories are shown on the right, with stationary particles shown in black, and Brownian motion and directed motion trajectories represented in green and orange, respectively.

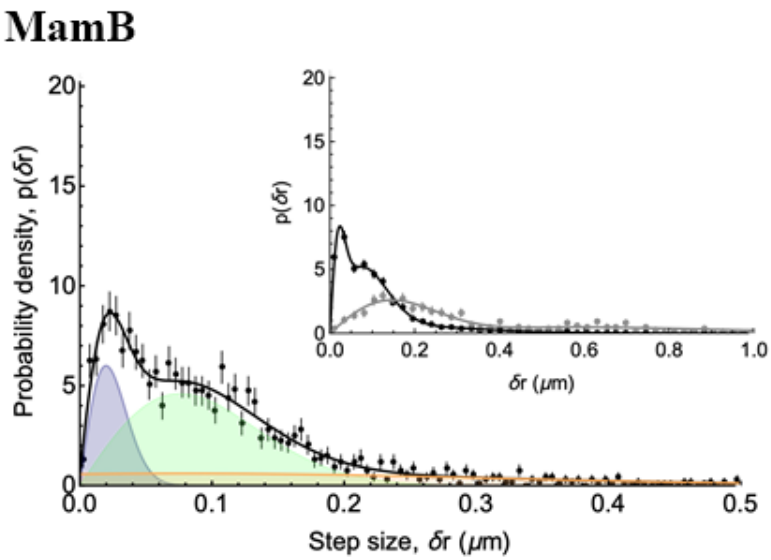


Figure 11. Distribution of displacements in a representative Tomato-MamB-expressing cell. In the main panel the probability density of displacements after t = 1s is shown (black symbols) and fitted assuming two diffusive populations and one directed population. The blue and green shaded peaks represent the contributions of the slow and fast diffusing populations, the orange shaded peak that of the directed population and the black curve shows the sum of these three contributions. The inset shows a comparison between the distribution of displacements after t = 1s (black symbols) and after t = 20 s (grey symbols).

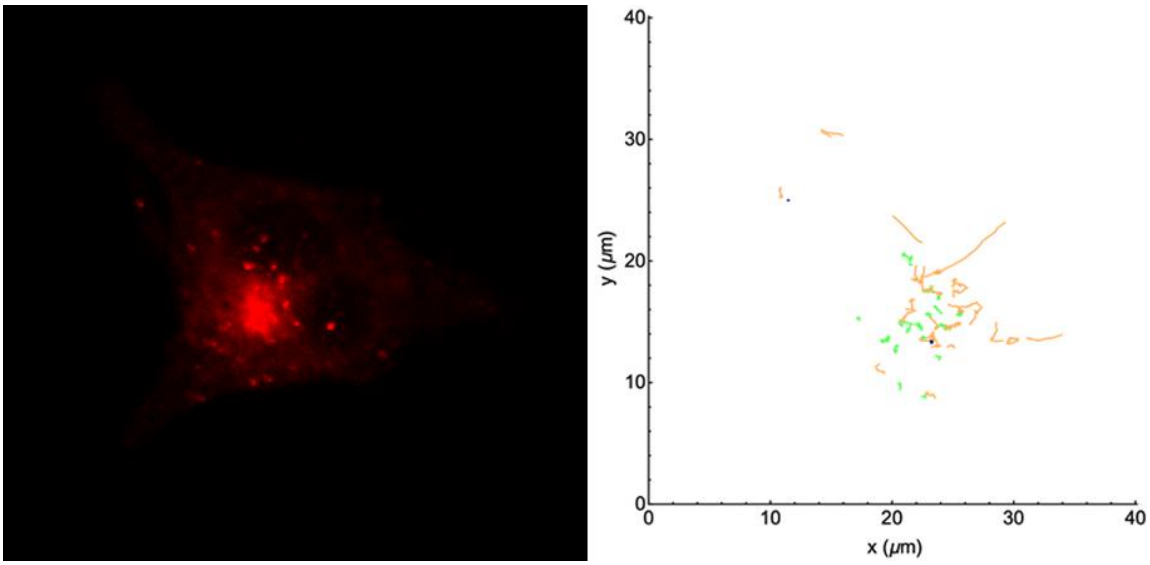


Figure 12. Representative cell expressing MamL particle trajectories.

A confocal image of the analyzed cell is shown on the left. Analyzed trajectories are shown on the right, with stationary particles shown in dark blue, and Brownian motion and directed motion trajectories represented in green and orange, respectively.

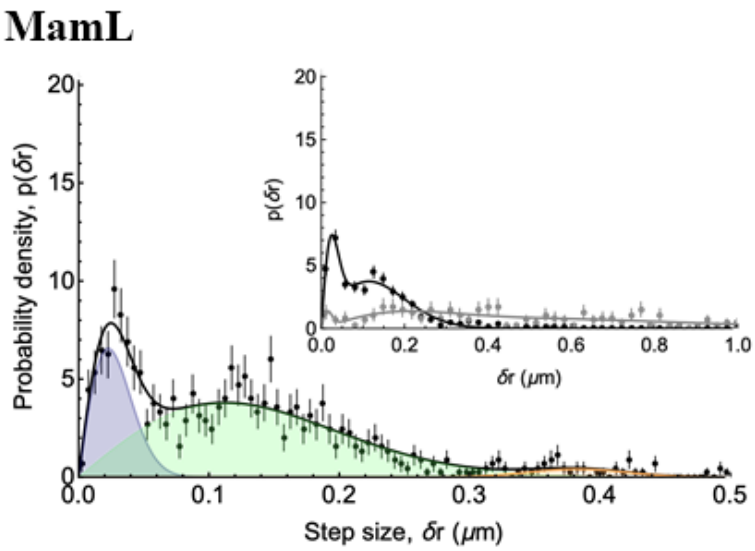


Figure 13. Distribution of displacements in a representative MamL-expressing cell.

In the main panel the probability density of displacements after t = 1s is shown (black symbols) and fitted assuming two diffusive populations and one directed population. The blue and green shaded peaks represent the contributions of the slow and fast diffusing populations, the orange shaded peak that of the directed population and the black curve shows the sum of these three contributions. The inset shows a comparison between the distribution of displacements after t = 1s (black symbols) and after t = 20 s (grey symbols).

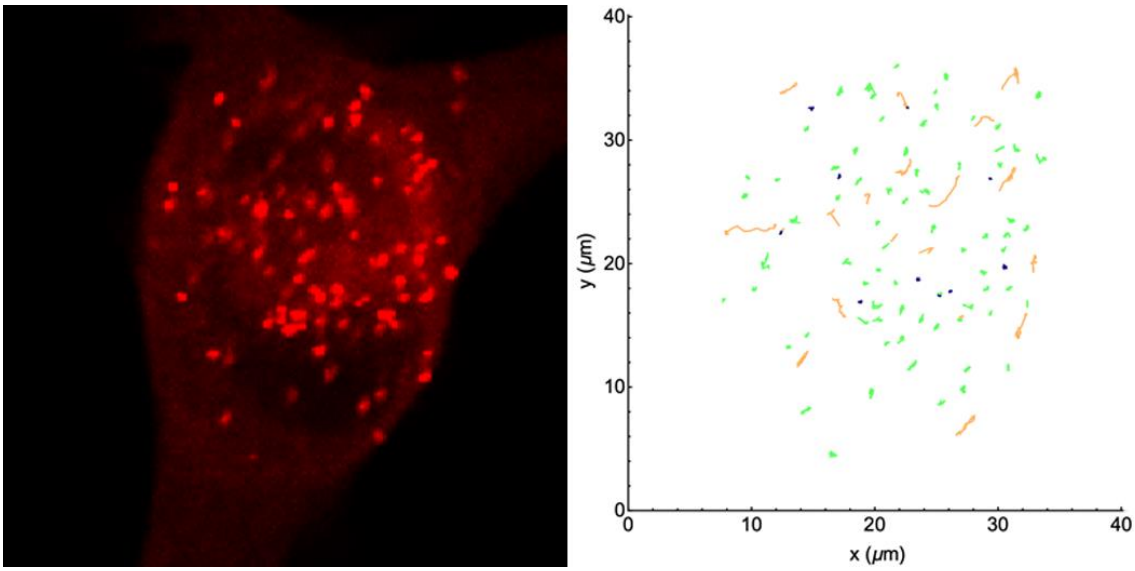


Figure 14. Representative cell expressing MamL$_{trunc}$ particle trajectories.

A confocal image of the analyzed cell is shown on the left. Analyzed trajectories are shown on the right, with stationary particles shown in dark blue, and Brownian motion and directed motion trajectories represented in green and orange, respectively.

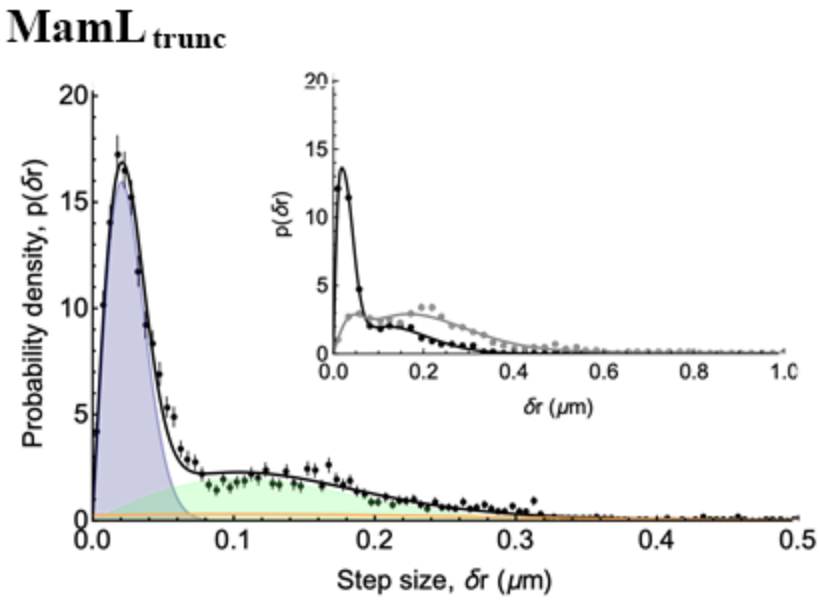


Figure 15. Distribution of displacements in a representative MamL$_{trunc}$-expressing cell. In the main panel the probability density of displacements after t = 1s is shown (black symbols) and fitted assuming two diffusive populations and one directed population. The blue and green shaded peaks represent the contributions of the slow and fast diffusing populations, the orange shaded peak that of the directed population and the black curve shows the sum of these three contributions. The inset shows a comparison between the distribution of displacements after t = 1s (black symbols) and after t = 20 s (grey symbols).

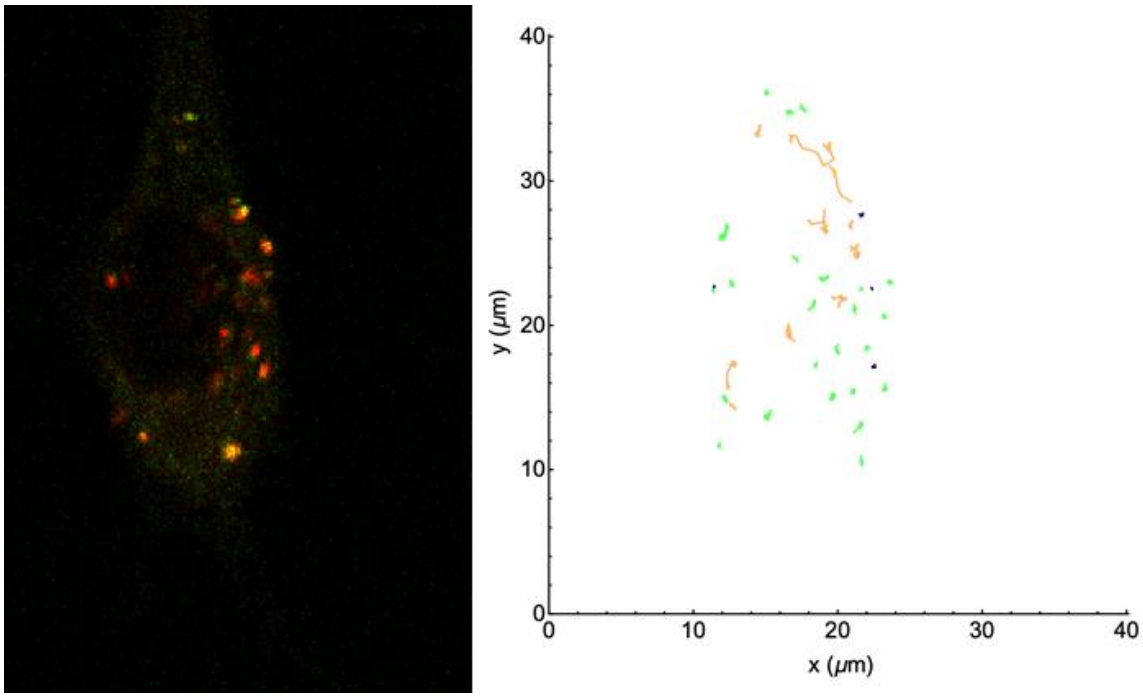


Figure 16. Representative cell expressing MamL+MamI particle trajectories.

A confocal image of the analyzed cell is shown on the left. Analyzed trajectories are shown on the right, with stationary particles shown in dark blue, and Brownian motion and directed motion trajectories represented in green and orange, respectively.

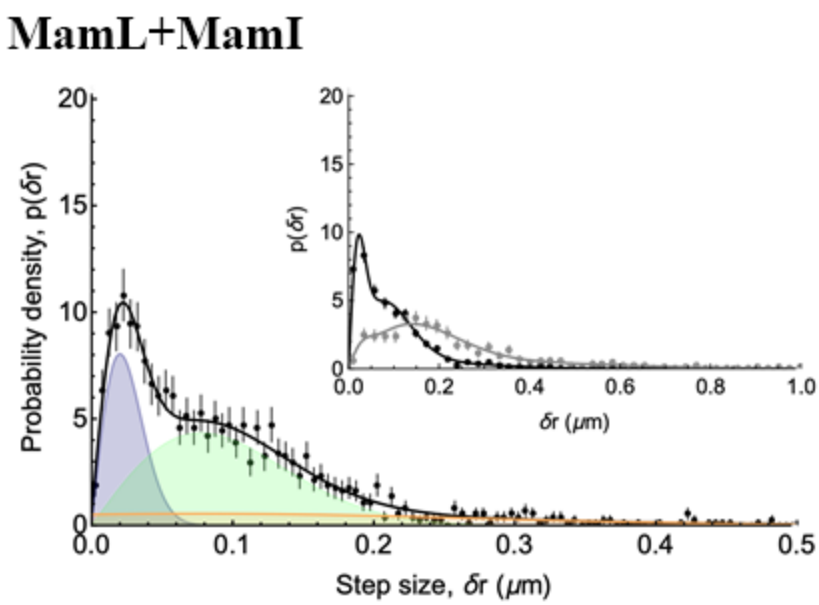


Figure 17. Distribution of displacements in a representative MamL+MamI-expressing cell. In the main panel the probability density of displacements after t = 1s is shown (black symbols) and fitted assuming two diffusive populations and one directed population. The blue and green shaded peaks represent the contributions of the slow and fast diffusing populations, the orange shaded peak that of the directed population and the black curve shows the sum of these three contributions. The inset shows a comparison between the distribution of displacements after t = 1s (black symbols) and after t = 20 s (grey symbols).

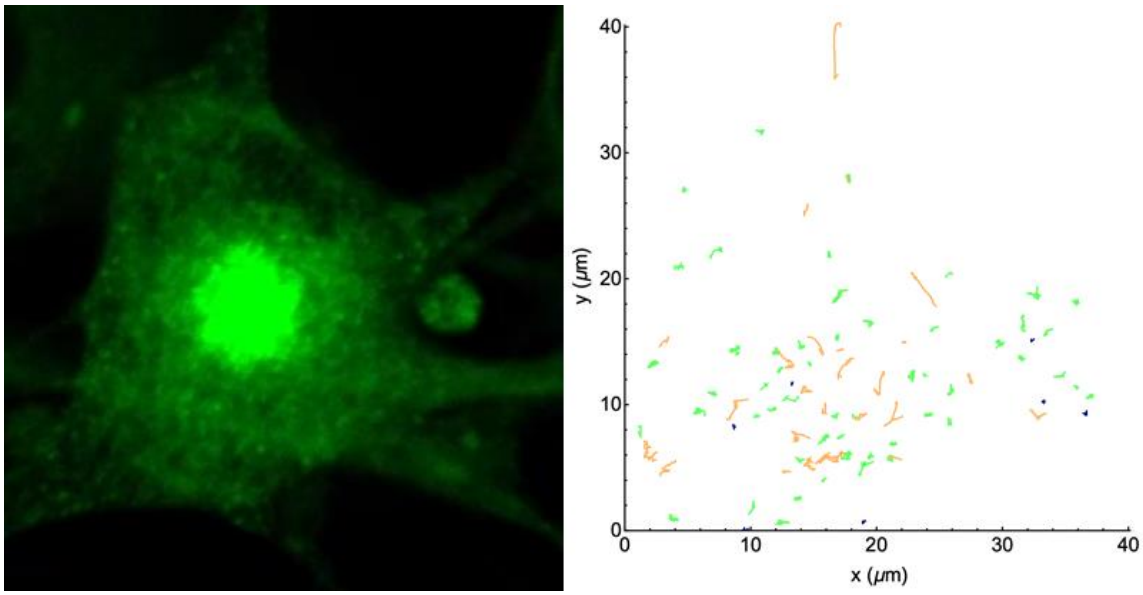


Figure 18. Representative cell expressing FLAG-MamL+MamI particle trajectories.

A confocal image of the analyzed cell is shown on the left. Analyzed trajectories are shown on the right, with stationary particles shown in dark blue, and Brownian motion and directed motion trajectories represented in green and orange, respectively.

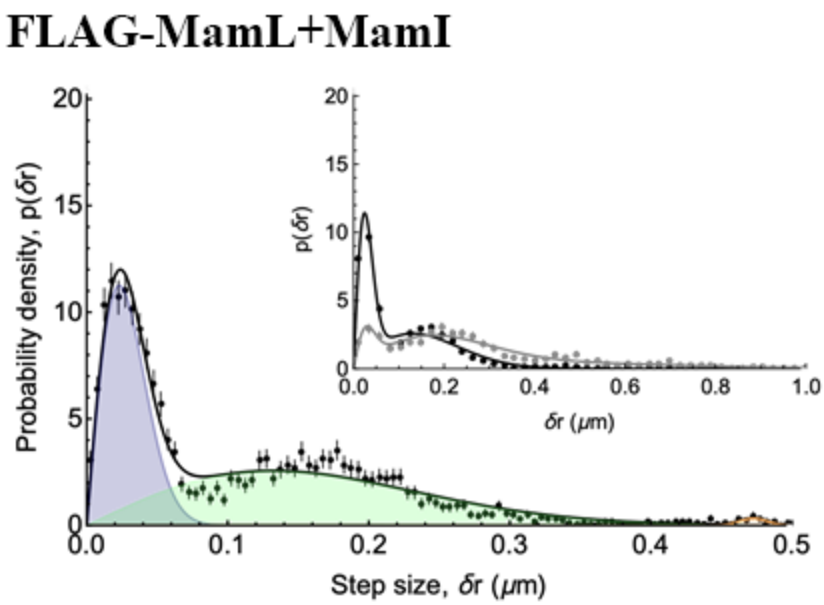


Figure 19. Distribution of displacements in a representative FLAG-MamL+MamI-expressing cell.

In the main panel the probability density of displacements after t = 1s is shown (black symbols) and fitted assuming two diffusive populations and one directed population. The blue and green shaded peaks represent the contributions of the slow and fast diffusing populations, the orange shaded peak that of the directed population and the black curve shows the sum of these three contributions. The inset shows a comparison between the distribution of displacements after t = 1s (black symbols) and after t = 20 s (grey symbols).

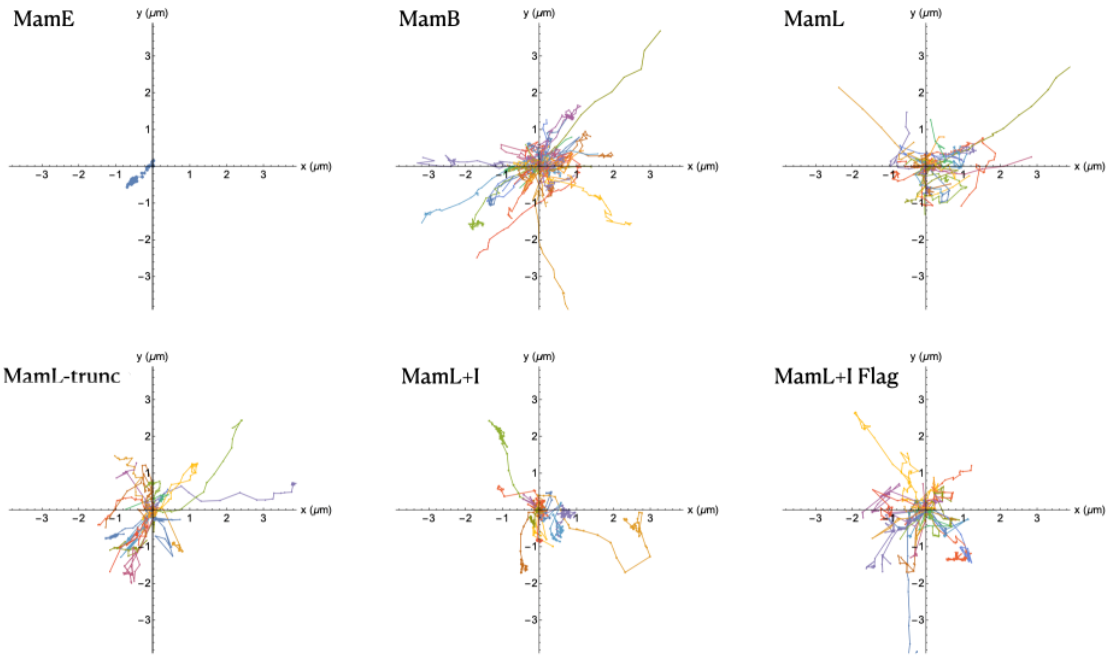


Figure 20. Representative directed trajectories of magnetosome proteins.

Each panel shows all the directed trajectories detected in a single representative cell, for each of the different expression systems under study. All trajectories have been plotted by placing the first point of the trajectory at the point with coordinates x = y = 0.

Supplementary Figures

Supplementary Figure S1. *mamB* sequence information. The *mamB* gene (cyan) is shown downstream and in frame with Tomato (red). Restriction sites used in the cloning of *mamB* are labeled and boxed in green.

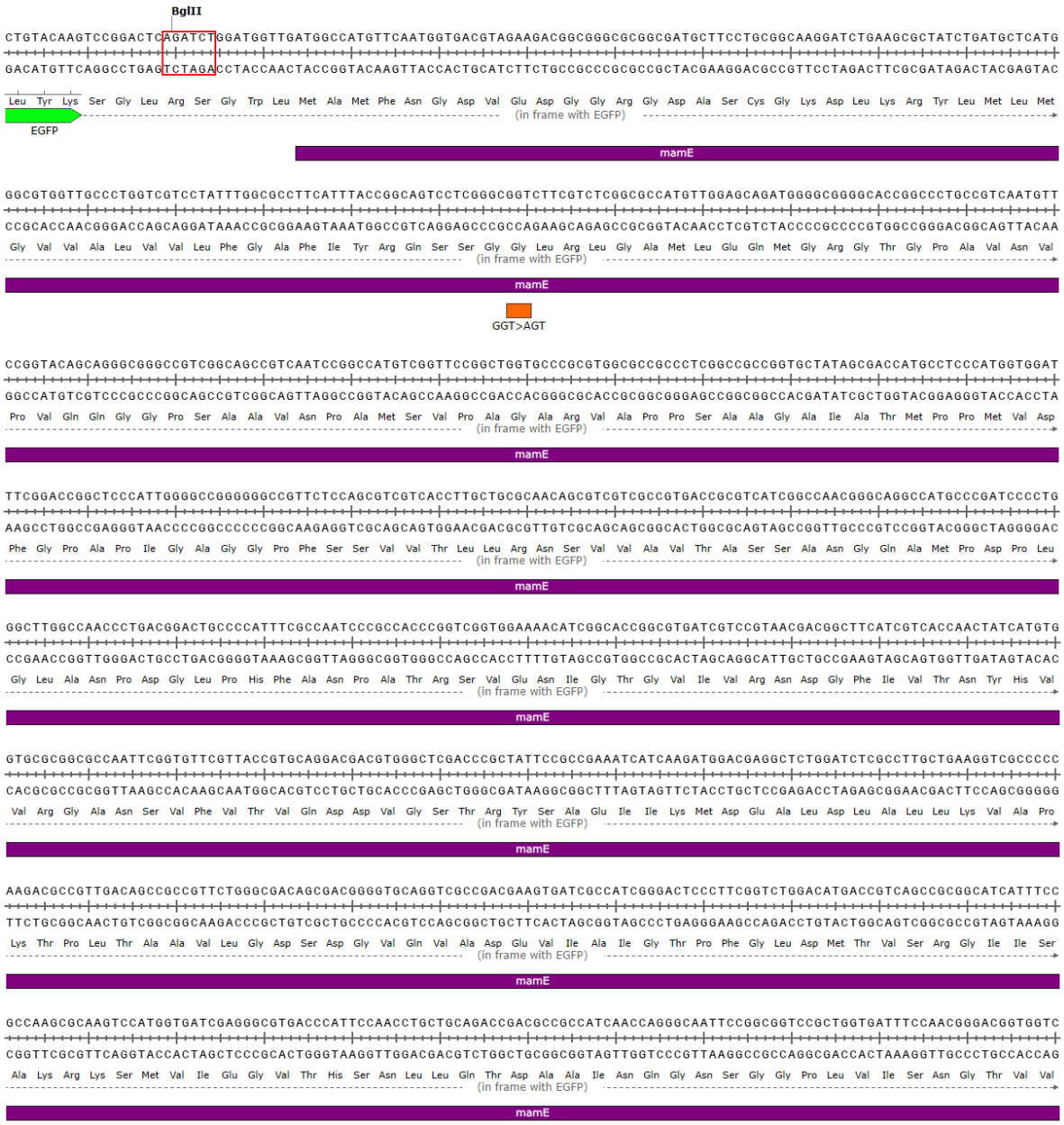
```
GGAGGAAATCCCCCACGCCCGCGACATCCGCCTGTTCGTCAGCCCGGCACCCGCCGCCGCGGCACGGGCCTGAGCGAGAGAGGGGCGGCGTCATGGACAT
CCTCCTTTAGGGGGTGCGGGCGCTGTAGGCGGACAAGCAGTCGGGCCGTGGGCGGCGGCGCCGTGCCCGGACTCGCTCTCTCCCCGCCGCAGTACCTGTA
```

```
     275                 280                 285                 290                 295
Glu  Glu  Ile  Pro  His  Ala  Arg  Asp  Ile  Arg  Leu  Phe  Val  Ser  Pro  Ala  Pro  Ala  Ala  Ala  Ala  Arg  Ala  *
                                                  mamB
```

SmaI

```
CCCGGGATCCACCGGATCTAGATAACTGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTGAAC
GGGCCCTAGGTGGCCTAGATCTATTGACTAGTATTAGTCGGTATGGTGTAAACATCTCCAAAATGAACGAAATTTTTTGGAGGGTGTGGAGGGGGACTTG
```

Supplementary Figure S2. *mamE* sequence information. The *mamE* gene (purple) is shown downstream and in frame with EGFP (green). MamE (G49S, T641S) mutations are labeled in orange and MamE (T317A) mutation is labeled in pink. Restriction sites used in the cloning of *mamE* are labeled and boxed in red.



GGCATCAATACCGCCATCTACACCCCCAACGGCGCCTTTGCCGGAATAGGCTTCGCCGTGCCCAGCAATCAGGCCCGCCTGTTCATCTTGGATGAAGTGGGCTGGCTGCCGACCTCCACC
CCGTAGTTATGGCGGTAGATGTGGGGGTTGCCGCGGAAACGGCCTTATCCGAAGCGGCACGGGTCGTTAGTCCGGGCGGACAAGTAGAACCTACTTCACCCGACCGACGGCTGGAGGTGG

Gly Ile Asn Thr Ala Ile Tyr Thr Pro Asn Gly Ala Phe Ala Gly Ile Gly Phe Ala Val Pro Ser Asn Gln Ala Arg Leu Phe Ile Leu Asp Glu Val Gly Trp Leu Pro Thr Ser Thr
(in frame with EGFP)

mamE

ACC>GCC

GCCGAGGGGGCCTCCATGGGCCTGGTGGCCATGCAGCGCCCCATGGGCGGCGGCGTGGGCGCGGCGGGTCCGGCCATTTTCGCCGGGACTAGGGCGCCCCACACCGATGGCCGCCAGAAC
CGGCTCCCCCGGAGGTACCCGGACCACCGGTACGTCGCGGGGTACCCGCCGCCGCACCCGCGCCGCCCAGGCCGGTAAAAGCGGCCCTGATCCCGCGGGGTGTGGCTACCGGCGGTCTTG

Ala Glu Gly Ala Ser Met Gly Leu Val Ala Met Gln Arg Pro Met Gly Gly Gly Val Gly Ala Ala Gly Pro Ala Ile Phe Ala Gly Thr Arg Ala Pro His Thr Asp Gly Arg Gln Asn
(in frame with EGFP)

mamE

ATGGATTGCACCACCTGCCACGATCTGATTCCGGCCGGAAACGGGCGGCCCGCGCCGATGATGCCCATCGCCGCCCCGATTCCGCCGCCGCCGATTCCCATGGGCGCCGTCTCGCCCCAT
TACCTAACGTGGTGGACGGTGCTAGACTAAGGCCGGCCTTTGCCCGCCGGGCGCGGCTACTACGGGTAGCGGCGGGGCTAAGGCGGCGGCGGCTAAGGGTACCCGCGGCAGAGCGGGGTA

Met Asp Cys Thr Thr Cys His Asp Leu Ile Pro Ala Gly Asn Gly Arg Pro Ala Pro Met Met Pro Ile Ala Ala Pro Ile Pro Pro Pro Pro Ile Pro Met Gly Ala Val Ser Pro His
(in frame with EGFP)

mamE

ACCGACGGCCGTCAGAACATGAACTGCGCCAATTGCCATCAGATGCTGGGCGGTGCCGCGCCCATCGCCGCACCGGGACTCGGCGGGGGGGCTTACCGCTTCGCGCAGCCGCCGGGAAGC
TGGCTGCCGGCAGTCTTGTACTTGACGCGGTTAACGGTAGTCTACGACCCGCCACGGCGCGGGTAGCGGCGTGGCCCTGAGCCGCCCCCCCGAATGGCGAAGCGCGTCGGCGGCCCTTCG

Thr Asp Gly Arg Gln Asn Met Asn Cys Ala Asn Cys His Gln Met Leu Gly Gly Ala Ala Pro Ile Ala Ala Pro Gly Leu Gly Gly Gly Ala Tyr Arg Phe Ala Gln Pro Pro Gly Ser
(in frame with EGFP)

mamE

CTGGCCATCAATATTCAGGGGCCGCGCGGCGGGCAGTCGACCGCCGCCGGAACCGGCCGGGTCACGCTTTTGGGCGCCGCCTTGACTCCCATGTCGCAGCGCCTTGGCGCCCAGACCGGC
GACCGGTAGTTATAAGTCCCCGGCGCGCCGCCCGTCAGCTGGCGGCGGCCTTGGCCGGCCCAGTGCGAAAACCCGCGGCGGAACTGAGGGTACAGCGTCGCGGAACCGCGGGTCTGGCCG

Leu Ala Ile Asn Ile Gln Gly Pro Arg Gly Gly Gln Ser Thr Ala Ala Gly Thr Gly Arg Val Thr Leu Leu Gly Ala Ala Leu Thr Pro Met Ser Gln Arg Leu Gly Ala Gln Thr Gly
(in frame with EGFP)

mamE

GAAGAAGTCTCCGCCATCATGGTCGAGATGCATGCCGGACGGTCGGTGCGGCTCGGCGTCCTGCGCGATGGCGATGTGCGCAACATGACCCTGGTCGCGGGACCCGCCGGTCTGGCCGCC
CTTCTTCAGAGGCGGTAGTACCAGCTCTACGTACGGCCTGCCAGCCACGCCGAGCCGCAGGACGCGCTACCGCTACACGCGTTGTACTGGGACCAGCGCCCTGGGCGGCCAGACCGGCGG

Glu Glu Val Ser Ala Ile Met Val Glu Met His Ala Gly Arg Ser Val Arg Leu Gly Val Leu Arg Asp Gly Asp Val Arg Asn Met Thr Leu Val Ala Gly Pro Ala Gly Leu Ala Ala
(in frame with EGFP)

mamE

GCAGCGGTCCAGGCTCCGGCCATTGCCGACATGGCGCAGCCGCCCATGGGCGGCATGGCCCCCACTGCACCCGGCATGGTCGCGGTTCCCGGTGGTCCGGCCGTCATGCCGAAGCCTCCG
CGTCGCCAGGTCCGAGGCCGGTAACGGCTGTACCGCGTCGGCGGGTACCCGCCGTACCGGGGGTGACGTGGGCCGTACCAGCGCCAAGGGCCACCAGGCCGGCAGTACGGCTTCGGAGGC

Ala Ala Val Gln Ala Pro Ala Ile Ala Asp Met Ala Gln Pro Pro Met Gly Gly Met Ala Pro Thr Ala Pro Gly Met Val Ala Val Pro Gly Gly Pro Ala Val Met Pro Lys Pro Pro
(in frame with EGFP)

mamE

ACCGAGTTCAACTGGCTGGGCATGGAGATCGAAACCTTCCAGGCGCCGCGTCCGATCACCGGTGTGCCGGGCGCCGTTCCGGTTCCCGGAGCCAAGGGTGCCCAGGTGGCCGAGGTTCTG
TGGCTCAAGTTGACCGACCCGTACCTCTAGCTTTGGAAGGTCCGCGGCGCAGGCTAGTGGCCACACGGCCCGCGGCAAGGCCAAGGGCCTCGGTTCCCACGGGTCCACCGGCTCCAAGAC

Thr Glu Phe Asn Trp Leu Gly Met Glu Ile Glu Thr Phe Gln Ala Pro Arg Pro Ile Thr Gly Val Pro Gly Ala Val Pro Val Pro Gly Ala Lys Gly Ala Gln Val Ala Glu Val Leu
(in frame with EGFP)

mamE

ACC>TCC

GTCGGGTCCCGGGCCGCCGTGGCGGGGCTCCAGGCCAATGACCTCATCCTCGAGGTCAATAATCGGCCGGTGGCCGGTCCGGCTCGTCTCGACGCCGCCATTAAGGGGGCGACGAATGCC
CAGCCCAGGGCCCGGCGGCACCGCCCCGAGGTCCGGTTACTGGAGTAGGAGCTCCAGTTATTAGCCGGCCACCGGCCAGGCCGAGCAGAGCTGCGGCGGTAATTCCCCCGCTGCTTACGG

Val Gly Ser Arg Ala Ala Val Ala Gly Leu Gln Ala Asn Asp Leu Ile Leu Glu Val Asn Asn Arg Pro Val Ala Gly Pro Ala Arg Leu Asp Ala Ala Ile Lys Gly Ala Thr Asn Ala
(in frame with EGFP)

mamE

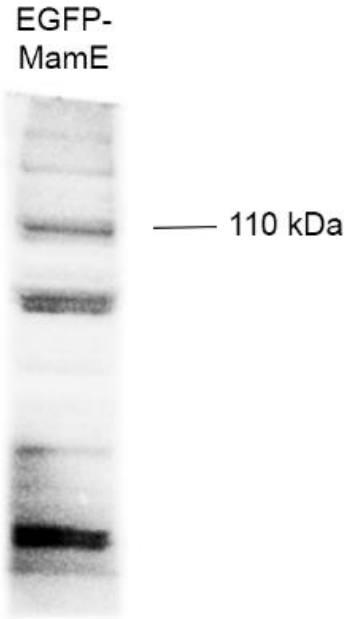
SacI

GGGCAGCAGATTTTGCTGAAGGTCAACCGCAATGGCCAAGAGTTCTGGATTGTCCTTTGATGGTGGAGCTCAAGCTTCGAATTCTGCAGTCGACGGTACCGCGGGCCCGGGATCCACCGG
CCCGTCGTCTAAAACGACTTCCAGTTGGCGTTACCGGTTCTCAAGACCTAACAGGAAACTACCACCTCGAGTTCGAAGCTTAAGACGTCAGCTGCCATGGCGCCCGGGCCCTAGGTGGCC

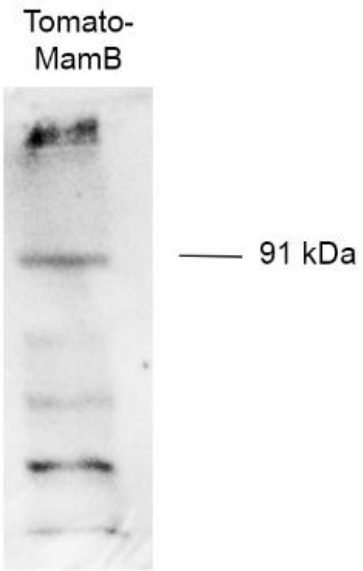
Gly Gln Gln Ile Leu Leu Lys Val Asn Arg Asn Gly Gln Glu Phe Trp Ile Val Leu
(in frame with EGFP)

mamE

Supplementary Figure S3. Immunoblot of total cellular protein from mammalian cells expressing EGFP-MamE. Total cellular protein from MDA-MB-435 cells stably expressing EGFP-MamE was examined with mouse α-EGFP. The full-length western blot is displayed.

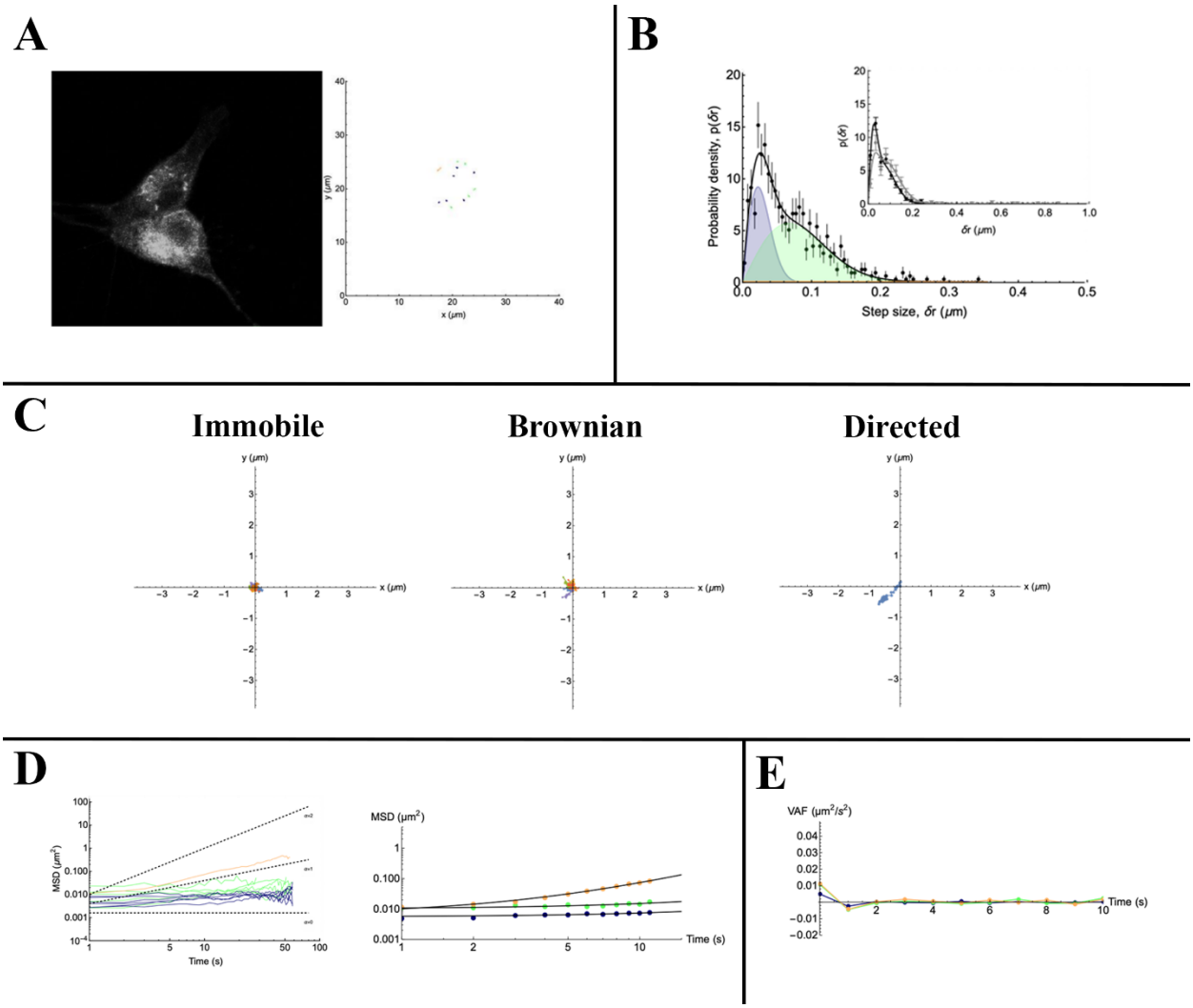


Supplementary Figure S4. Immunoblot of total cellular protein from mammalian cells expressing Tomato-MamB. Total cellular protein from MDA-MB-435 cells stably expressing Tomato-MamB was examined with rabbit α-Tomato. The full-length western blot is displayed.



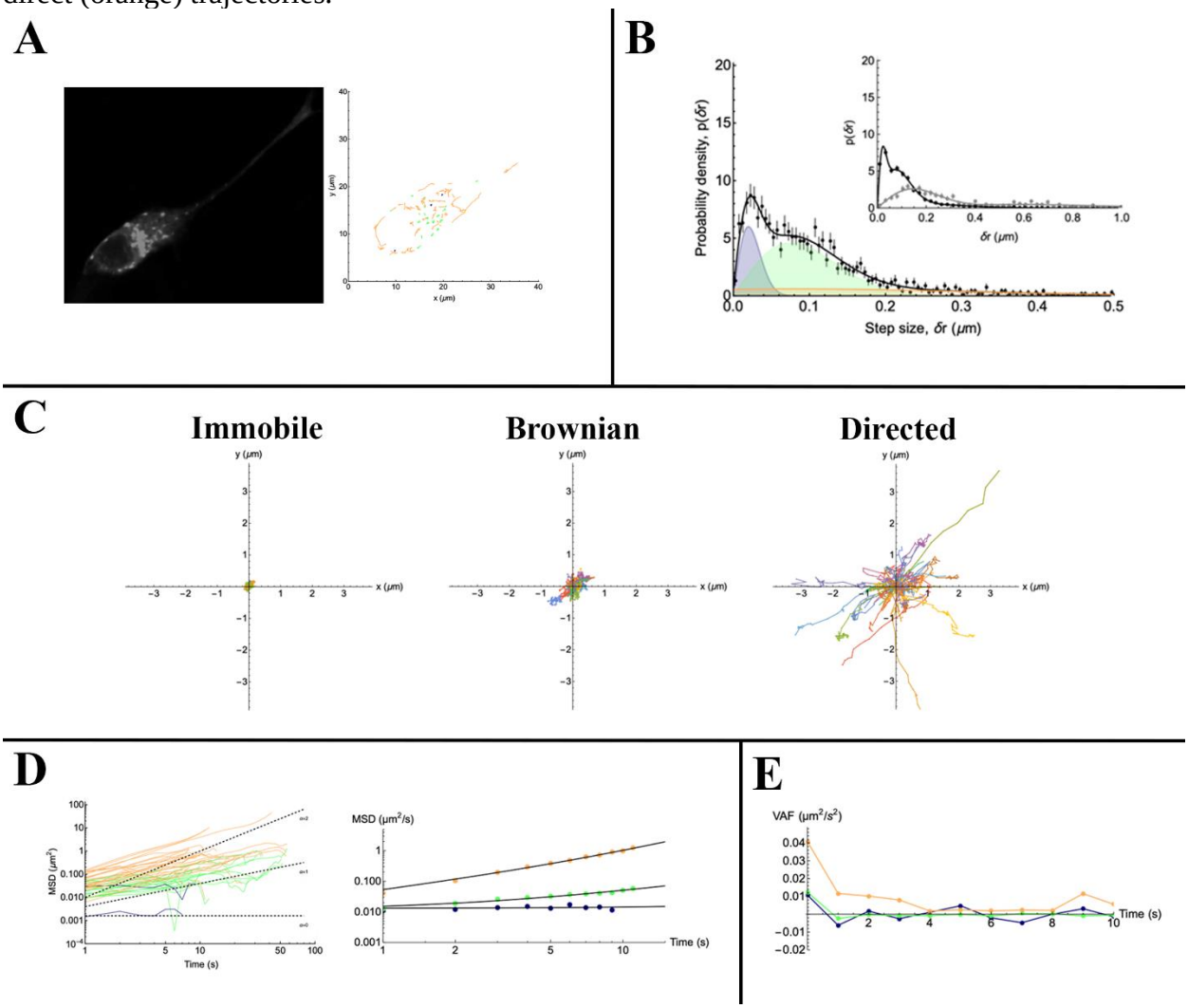
Supplementary Figure S5: Full panel of trajectory analysis figures for EGFP-MamE particles.

A. Representative cell and trajectories detected within the cell. Immobile, Brownian, and directed trajectories are represented in black, green, and orange, respectively. **B.** Representative distribution of displacements, where the immobile, Brownian, and direct populations are represented by the blue, green, and orange peaks, respectively. **C.** Immobile, Brownian, and directed trajectories plotted on a cartesian plane. **D.** Immobile (black), Brownian (green), and direct (orange) trajectories represented as mean-squared displacement (MSD) as a function of lag time. Dashed lines indicate the expected slope of the MSD, and the figure on the right shows average MSD for all trajectories. **E.** Average VAF of immobile (black), Brownian (green), and direct (orange) trajectories.



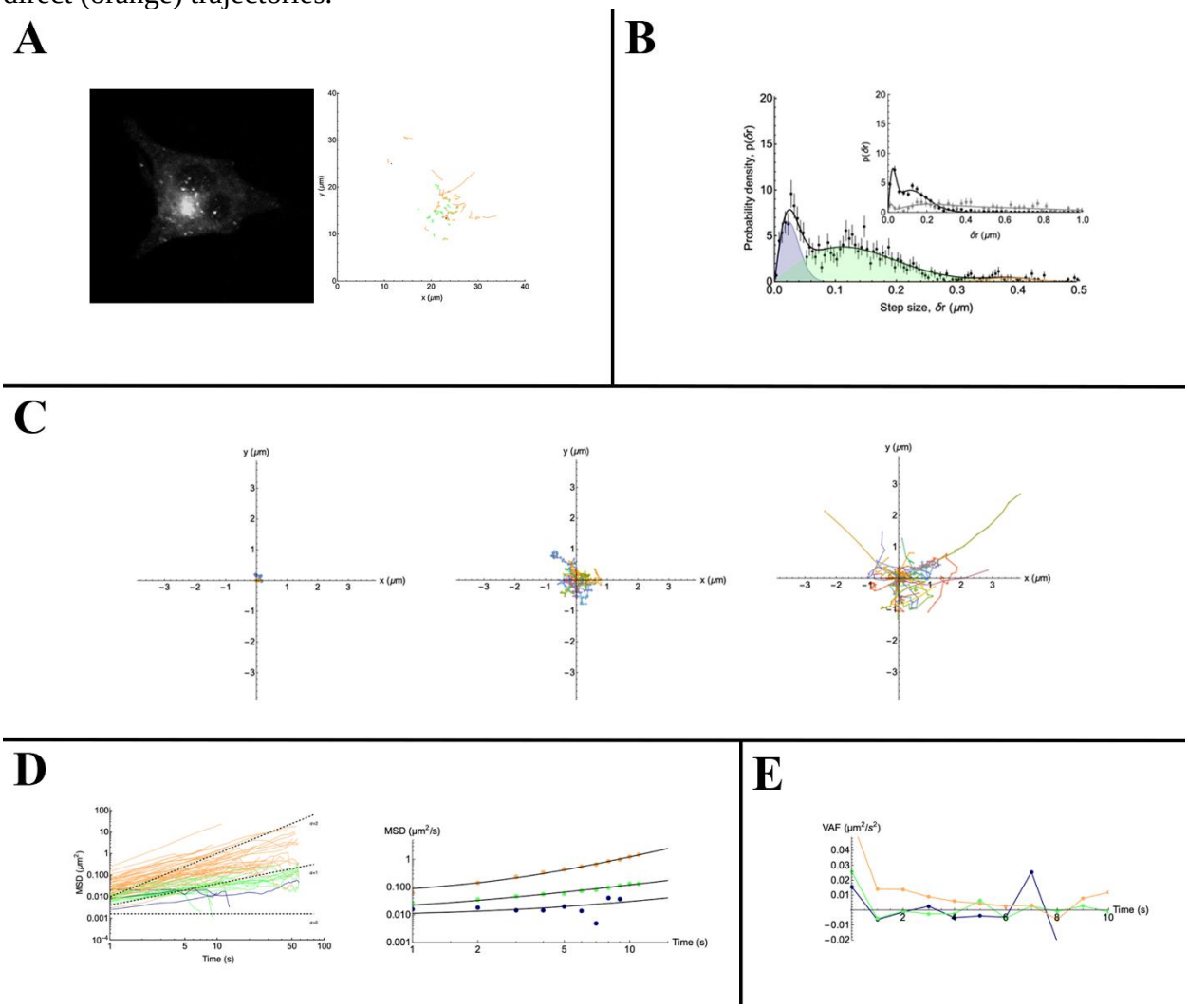
Supplementary Figure S6: Full panel of trajectory analysis figures for Tomato-MamB particles.

A. Representative cell and trajectories detected within the cell. Immobile, Brownian, and directed trajectories are represented in black, green, and orange, respectively. **B.** Representative distribution of displacements, where the immobile, Brownian, and direct populations are represented by the blue, green, and orange peaks, respectively. **C.** Immobile, Brownian, and directed trajectories plotted on a cartesian plane. **D.** Immobile (black), Brownian (green), and direct (orange) trajectories represented as mean-squared displacement (MSD) as a function of lag time. Dashed lines indicate the expected slope of the MSD, and the figure on the right shows average MSD for all trajectories. **E.** Average VAF of immobile (black), Brownian (green), and direct (orange) trajectories.



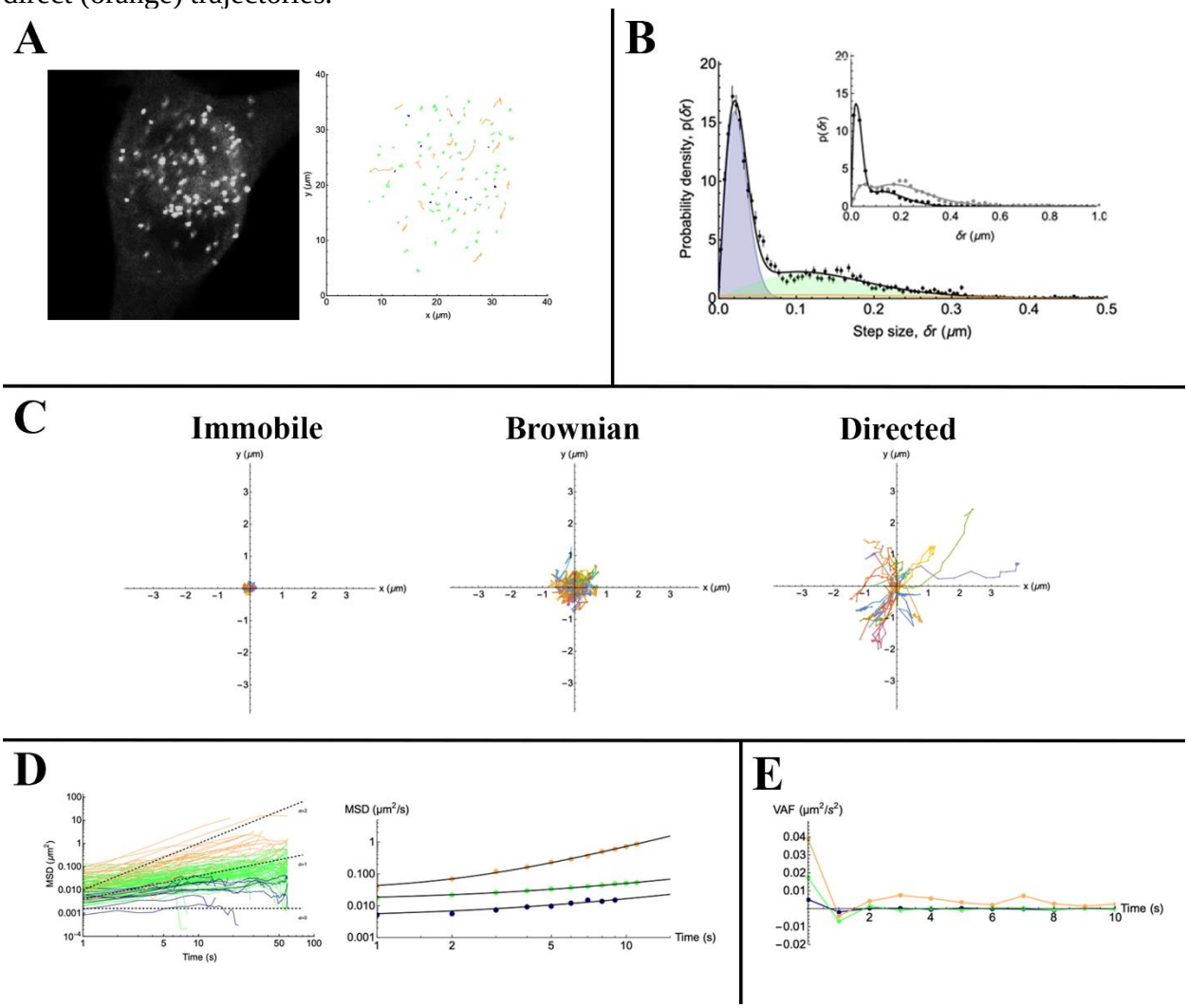
Supplementary Figure S7: Full panel of trajectory analysis figures for Tomato-MamL particles.

A. Representative cell and trajectories detected within the cell. Immobile, Brownian, and directed trajectories are represented in black, green, and orange, respectively. **B.** Representative distribution of displacements, where the immobile, Brownian, and direct populations are represented by the blue, green, and orange peaks, respectively. **C.** Immobile, Brownian, and directed trajectories plotted on a cartesian plane. **D.** Immobile (black), Brownian (green), and direct (orange) trajectories represented as mean-squared displacement (MSD) as a function of lag time. Dashed lines indicate the expected slope of the MSD, and the figure on the right shows average MSD for all trajectories. **E.** Average VAF of immobile (black), Brownian (green), and direct (orange) trajectories.



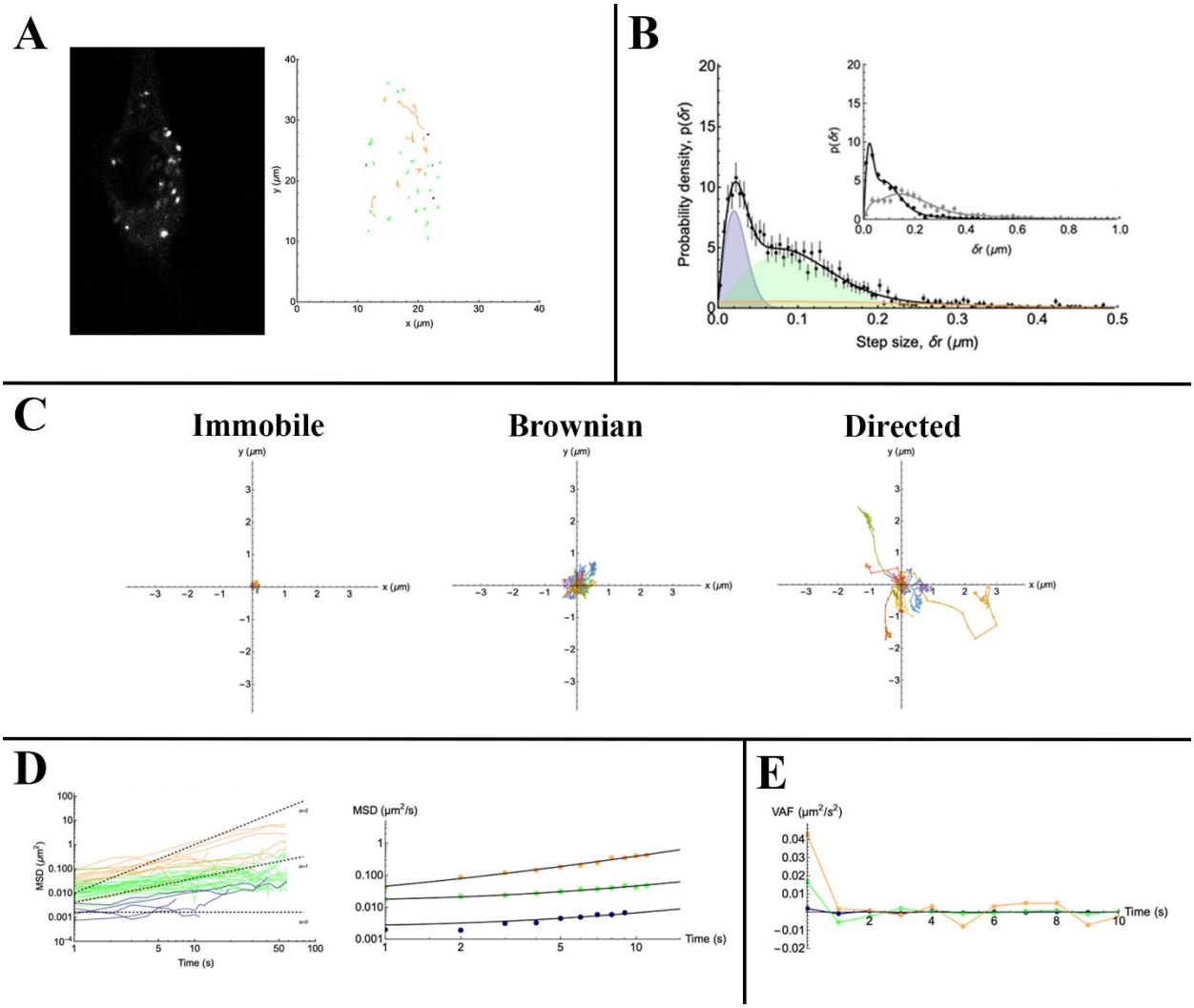
Supplementary Figure S8: Full panel of trajectory analysis figures for Tomato-MamL$_{trunc}$ particles.

A. Representative cell and trajectories detected within the cell. Immobile, Brownian, and directed trajectories are represented in black, green, and orange, respectively. **B.** Representative distribution of displacements, where the immobile, Brownian, and direct populations are represented by the blue, green, and orange peaks, respectively. **C.** Immobile, Brownian, and directed trajectories plotted on a cartesian plane. **D.** Immobile (black), Brownian (green), and direct (orange) trajectories represented as mean-squared displacement (MSD) as a function of lag time. Dashed lines indicate the expected slope of the MSD, and the figure on the right shows average MSD for all trajectories. **E.** Average VAF of immobile (black), Brownian (green), and direct (orange) trajectories.



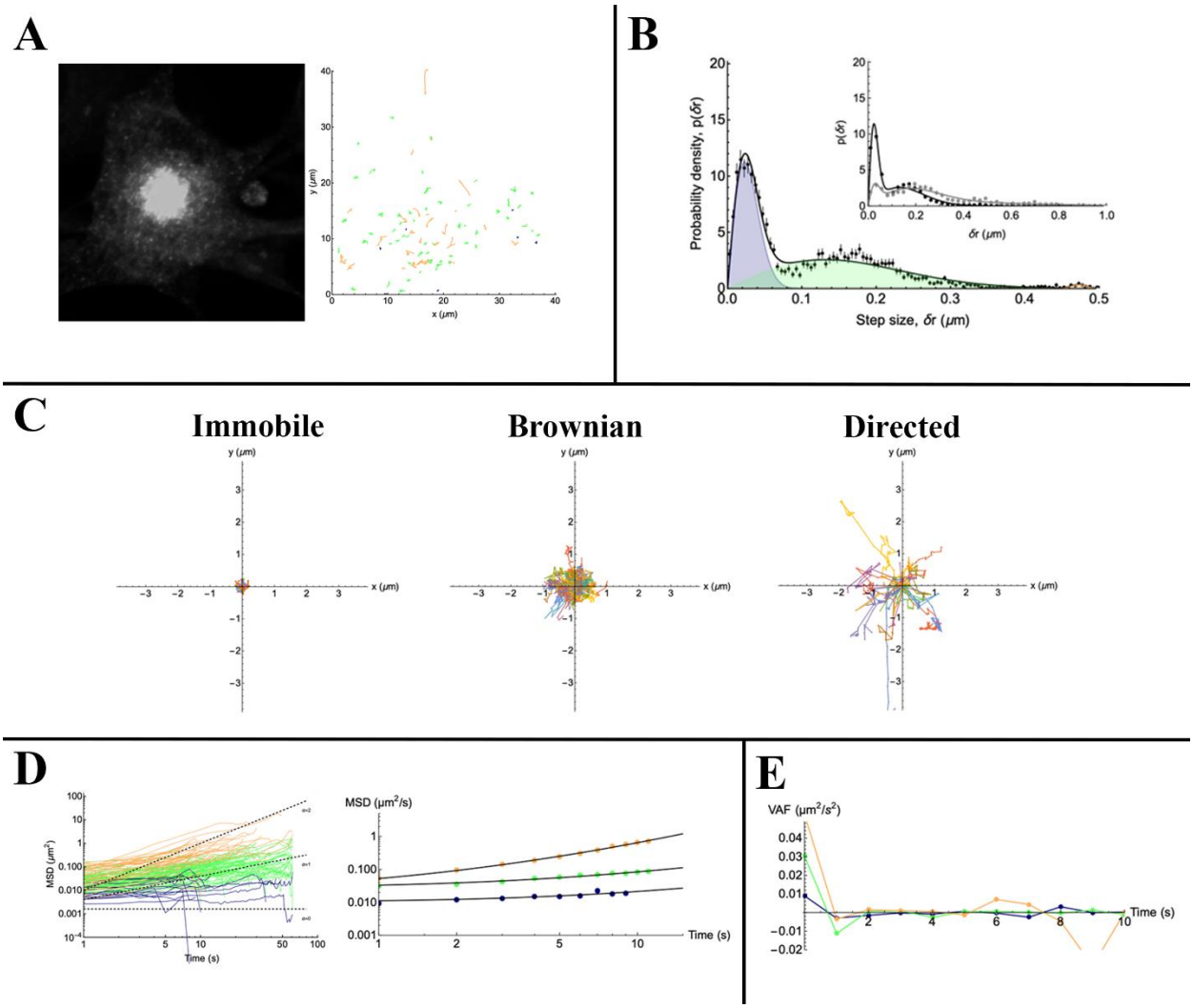
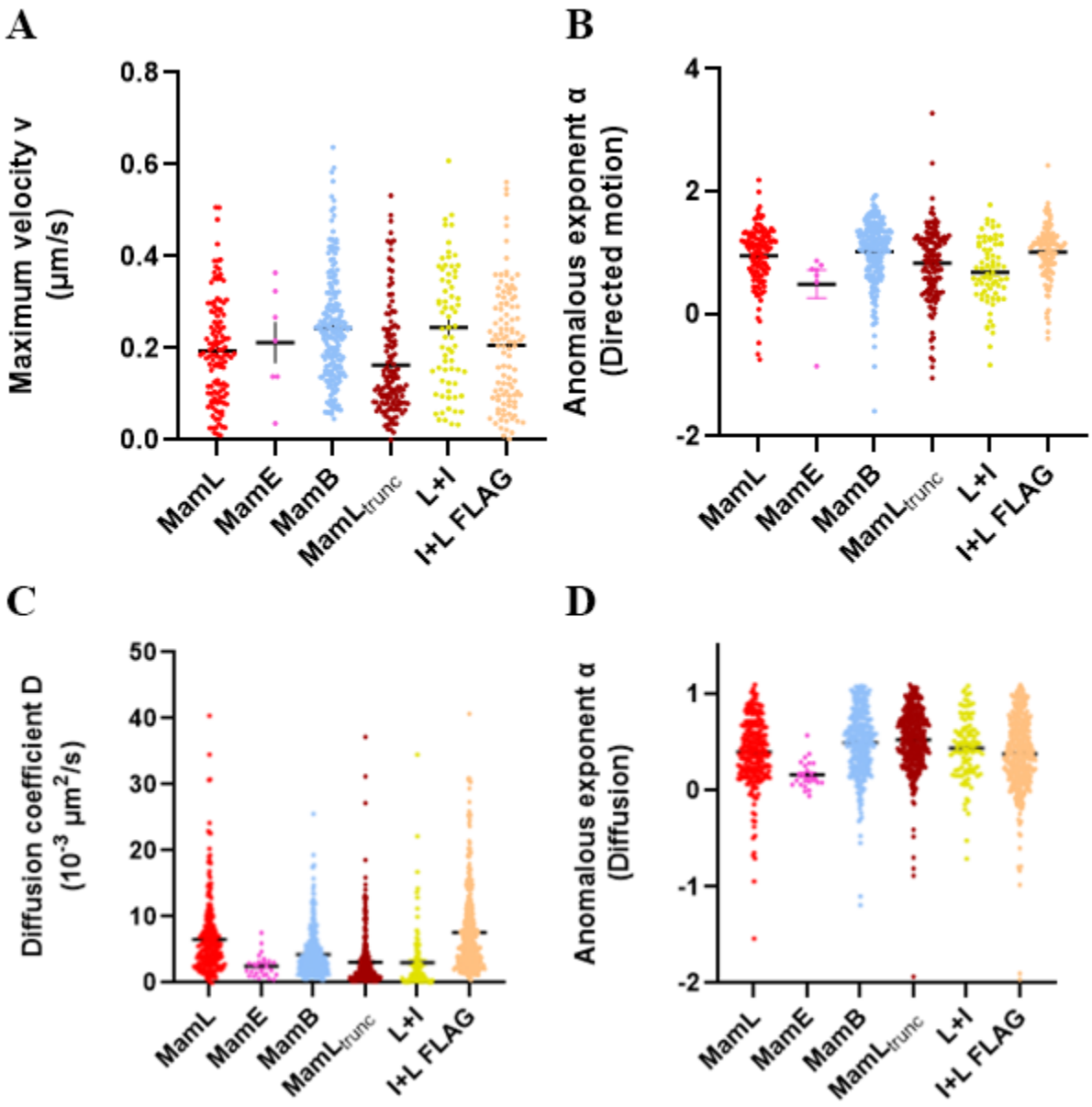
Supplementary Figure S9: Full panel of trajectory analysis figures for Tomato-MamL/EGFP-MamI particles.

A. Representative cell and trajectories detected within the cell. Immobile, Brownian, and directed trajectories are represented in black, green, and orange, respectively. **B.** Representative distribution of displacements, where the immobile, Brownian, and direct populations are represented by the blue, green, and orange peaks, respectively. **C.** Immobile, Brownian, and directed trajectories plotted on a cartesian plane. **D.** Immobile (black), Brownian (green), and direct (orange) trajectories represented as mean-squared displacement (MSD) as a function of lag time. Dashed lines indicate the expected slope of the MSD, and the figure on the right shows average MSD for all trajectories. **E.** Average VAF of immobile (black), Brownian (green), and direct (orange) trajectories.



Supplementary Figure S10: Full panel of trajectory analysis figures for FLAG-MamL/EGFP-MamI particles.

A. Representative cell and trajectories detected within the cell. Immobile, Brownian, and directed trajectories are represented in black, green, and orange, respectively. **B.** Representative distribution of displacements, where the immobile, Brownian, and direct populations are represented by the blue, green, and orange peaks, respectively. **C.** Immobile, Brownian, and directed trajectories plotted on a cartesian plane. **D.** Immobile (black), Brownian (green), and direct (orange) trajectories represented as mean-squared displacement (MSD) as a function of lag time. Dashed lines indicate the expected slope of the MSD, and the figure on the right shows average MSD for all trajectories. **E.** Average VAF of immobile (black), Brownian (green), and direct (orange) trajectories.



Supplementary Figure S11: Trajectory parameters measured for individual particles detected in cells expressing different magnetosome proteins. Trajectories classified as directed motion trajectories (A,B) for which both a maximum velocity (A) and an anomalous exponent (B), were extracted. Trajectories classified as entirely Brownian (C,D), for which both a diffusion coefficient (C) and an anomalous exponent (D), were extracted.