Uncovering diffusive states of the yeast membrane protein, Pma1, and how labeling method can change diffusive behavior

Mary Lou P. Bailey, Susan E. Pratt, Michael Hinrichsen, Yongdeng Zhang, Joerg Bewersdorf, Lynne J. Regan, and Simon G. J. Mochrie

Abstract

We present and analyze video-microscopy-based single-particle-tracking measurements of the budding yeast (Saccharomyces cerevisiae) membrane protein, Pma1, fluorescently labeled either by direct fusion to the switchable fluorescent protein, mEos3.2, or by a novel, light-touch, labeling scheme, in which a 5 amino acid tag is directly fused to the C-terminus of Pma1, which then binds mEos3.2. The track diffusivity distributions of these two populations of single-particle tracks differ significantly, demonstrating that labeling method can be an important determinant of diffusive behavior. We also applied perturbation expectation maximization (pEMv2) (Koo and Mochrie in Phys Rev E 94(5):052412, 2016), which sorts trajectories into the statistically optimum number of diffusive states. For both TRAP-labeled Pma1 and Pma1-mEos3.2, pEMv2 sorts the tracks into two diffusive states: an essentially immobile state and a more mobile state. However, the mobile fraction of Pma1-mEos3.2 tracks is much smaller (∼0.16) than the mobile fraction of TRAP-labeled Pma1 tracks (∼0.43). In addition, the diffusivity of Pma1-mEos3.2’s mobile state is several times smaller than the diffusivity of TRAP-labeled Pma1’s mobile state. Thus, the two different labeling methods give rise to very different overall diffusive behaviors. To critically assess pEMv2’s performance, we compare the diffusivity and covariance distributions of the experimental pEMv2-sorted populations to corresponding theoretical distributions, assuming that Pma1 displacements realize a Gaussian random process. The experiment–theory comparisons for both the TRAP-labeled Pma1 and Pma1-mEos3.2 reveal good agreement, bolstering the pEMv2 approach.

1 Introduction

The overarching goal of this work is to carefully examine the extent to which the measured diffusive behavior of a protein of interest (POI) in a heterogeneous, biological environment can be convincingly described in terms of a limited number of discrete diffusive states, each with its own diffusive properties. Such diffusive states might correspond to the POI being bound to different binding partners or being located in different local environments. To determine the number of diffusive states, we employ the perturbation expectation maximization (pEMv2) software, an unsupervised, systems-level, machine-learning-based data analysis and classification method that sorts a heterogeneous population of single-particle tracks. We then developed a method to validate the number of states found by pEMv2, based on the theoretical displacement covariance distributions. In this paper, we studied the budding yeast (Saccharomyces cerevisiae) membrane protein, Pma1 [1–12]. For Pma1, we find that the observed population of single-molecule trajectories can be well described in terms of just two diffusive states. One of these states corresponds to simple diffusion. The other is essentially immobile.

Spurred by the observation of two diffusive states for Pma1, our second goal is to quantify the differences between the diffusivities and displacement covariances exhibited by Pma1, fluorescently labeled in two different ways. The first labeling method—which we refer

Mary Lou P. Bailey and Susan E. Pratt have contributed equally to this work.

*Corresponding author

E-mail: simon.mochrie@yale.edu
to as “TRAP labeling”—is a light-touch method in which wild-type Pma1 is replaced with a version of Pma1, in which a 5 amino acid tag is directly fused to the C-terminus of Pma1 [13–17]. This is done in cells expressing a version of a fluorescent protein engineered to bind the tag, thus labeling Pma1. Second is the commonly used direct-labeling method, which replaces wild-type Pma1 with a Pma1-fluorescent protein (FP) direct fusion. In many cases, the assumption that the POI’s intrinsic biological function will be unaffected by the direct addition of the FP is surely correct. However, in some cases, adding an FP to a POI can cause the modified protein to mislocalize [18] or malfunction [19–22]. For example, it was recently found that tagging dynamin-related protein 1 with GFP alters oligomerization dynamics, causing impaired oligomerization compared to native protein [23]. In another study, overexpressed, membrane-targeted GFP fusion proteins were found to form organelle aggregates, therefore changing the motion of the protein being studied [24]. Similarly, Pma1 provides an example of a protein that is sensitive to labeling method: Yeast strains expressing direct fusions of Pma1 and the FPs, mCherry or EGFP, both exhibit compromised growth and mislocalization of Pma1-mCherry or Pma1-EGFP, respectively, to the vacuole [15]. The possibility of mislocalization or misfunction of direct-fusion proteins has motivated efforts to develop and deploy alternative in vivo labeling methods that are less disruptive of the protein’s localization or function [13–17]. The different diffusive behaviors that we observe for TRAP-labeled and direct-fusion Pma1 demonstrate that it may be necessary to broadly employ minimally perturbing labeling schemes in order to fully realize and study intrinsic biological behavior in live cells.

This paper is organized as follows. In Sect. 2, we present necessary background material. Sect. 2.1 briefly summarizes what is known about Pma1. In Sect. 2.2, we review the perturbation expectation maximization (pEMv2) algorithm, which determines the number of unique diffusive states within a population of single-particle tracks and sorts individual tracks into those states. In Sect. 2.3, we review theoretical results for probability distributions of two-dimensional displacement covariance matrix elements [25], which can be compared to the corresponding experimental distributions.

In Sect. 3.1, we describe the yeast strains employed in this study and how samples were prepared. We describe the microscopy setup used to collect the data in Sect. 3.2, and the methods we used for single-particle tracking in Sect. 3.3.

Section 4 presents and discusses our results. In Sect. 4.1, we present microscopy images of the strains studied. In contrast to cells where Pma1 is expressed as a fusion protein with the fluorescent proteins, mCherry and EGFP, which show defective growth and behavior, we find that cells expressing Pma1-mEos3.2 do not show a growth defect, nor does Pma1-mEos3.2 mislocalize to the vacuole. However, these gross observations do not rule out more subtle differences in behavior between Pma1-mEos3.2 and TRAP-labeled Pma1.

In Sect. 4.2, we present the experimental track lifetime distributions of TRAP-labeled Pma1 and Pma1-mEos3.2. In both cases, we find that the track lifetime distribution decreases monotonically in time, indicating that Pma1 is a monomer in the cell membrane. We also show that the lifetime of TRAP-labeled Pma1 tracks is shorter than that of Pma1-mEos3.2. We interpret this difference in lifetime in terms of a nonzero TRAP-peptide unbinding rate, which allows us to estimate this rate to be about $6 \text{s}^{-1}$.

To examine the differences between the diffusive behavior of TRAP-labeled Pma1 and Pma1-mEos3.2, we subjected both of these populations of trajectories to pEMv2 analysis [26–28]. In Sect. 4.4, we describe the application of pEMv2 to the population of TRAP-labeled Pma1 single-particle trajectories. pEMv2 sorts TRAP-labeled Pma1 trajectories into two states. We test the sorted populations by comparing their covariance and diffusivity distributions to the corresponding theoretical distributions. We find good agreement between pEMv2-sorted covariance distributions and theory (Sect. 4.5), bolstering the pEMv2-based approach. To quantitatively compare the pEMv2-sorted experimental distributions and theory, we explore the Kolmogorov–Smirnov (KS) statistic in Sect. 4.5.1. Next, in Sect. 4.6, we analyze the population of Pma1-mEos3.2 direct-fusion tracks with pEMv2, and find that it also shows two states. However, in contrast to TRAP-labeled Pma1, where the mobile fraction is about 0.43, the fraction of Pma1-mEos3.2 direct-fusion tracks in the mobile state that is 0.16. In addition to the much reduced mobile fraction in Pma1-mEos3.2 compared to TRAP-labeled Pma1, the diffusivities of the mobile subpopulations alone appear to be different for the two labeling strategies with the mean diffusivity of the mobile subpopulation of TRAP-labeled Pma1 ($D_2 \approx 0.25 \mu \text{m}^2\text{s}^{-1}$) being several-fold larger than the mean diffusivity of the mobile state (state 2) of Pma1-mEos3.2 ($D_2 \approx 0.08 \mu \text{m}^2\text{s}^{-1}$). Direct comparison between the sorted experimental diffusivity and covariance distributions with the theoretical diffusivity and covariance distributions from Refs. [25, 29] reveals good agreement between experiment and theory in this case too. Next we present analysis of the mean squared displacements of the sorted tracks in Sect. 4.7, highlighting that the state 1 MSDs are essentially flat, or immobile, compared to state 2. Finally, in Sect. 5, we summarize and conclude.

# 2 Background

## 2.1 Pma1

Pma1 is the most abundant protein in the plasma membrane of budding yeast (Saccharomyces cerevisiae). It is a transmembrane protein which pumps protons out of the cell and thus plays a role in regulating...
the pH of the cytoplasm. Pma1 is also a marker of cell aging because, interestingly, there is less Pma1 in the plasma membranes of newly budded daughter cells, than in the membranes of their mother cell [9]. The yeast plasma membrane is laterally organized into several different membrane “compartments” or domains. As the name implies, Pma1 is the majority protein component of the membrane compartment of Pma1 (MCP). Membrane compartments, including MCP, show characteristic linear dimensions of about 0.1 μm [5, 7], and differ from each other in their composition, size, shape, etc. [10]—MCPs are enriched in sphingolipids, as well as Pma1, and show non-compact, “network-like” shapes [7]. Pma1 contains 918 amino acids, comprising four domains: a membrane domain which includes ten transmembrane α-helices, a phosphorylation domain, a nucleotide-binding domain, and an actuator-domain, which experiences significant rearrangements when Pma1 cycles between the two allosteric states, activated and inhibited, involved in its enzymatic cycle. The change from Pma1’s inhibited to activated state has been proposed to be a consequence of phosphorylation of a specific Ser residue (Ser 899) and the tandem phosphorylation of a Ser/Thr pair (Ser911 and Thr912) [1, 2]. Recently, two cryo-electron microscopy studies of detergent-extracted, lipid-reconstituted, hexamerically associated Pma1 (from S. cerevisiae and Neurospora crassa) provided microscopic details of the inhibited and activated molecular structures [11, 12]. Other studies involving in vitro reconstitution into liposomes [3] or nanodiscs [30] report that Pma1 monomers are active in proton pumping. As far as we are aware, no study has yet definitively identified the in vivo association state of Pma1.

### 2.2 Perturbation expectation maximization

A number of methods have been introduced for addressing particle tracks exhibiting biological heterogeneity [26, 27, 31–37]. In particular, Refs. [26, 27] describe perturbation expectation maximization, pEMv2, which simultaneously analyzes a population of particle trajectories and sorts the trajectories into distinct diffusive states, each with its own diffusion properties. pEMv2 is a machine-learning approach that makes no a priori assumptions concerning the character of a diffusive state—e.g., whether it corresponds to simple diffusion or not—but rather determines each state’s diffusive properties directly from the sorted tracks.

Although pEMv2 performs well on simulations, to date it has been applied to relatively few experimental datasets [27, 28]. To further explore pEMv2’s performance, we apply it to our experimental dataset of Pma1 single-particle tracks. In particular, we compare the statistical properties of pEMv2-sorted populations, which pEMv2 asserts are homogeneous, to theoretical expectations for a population of tracks with a single set of diffusion parameters [25, 29].

pEMv2 is described in detail in Refs. [26, 27]. In brief, it is an unsupervised, systems-level, machine-learning-based data analysis and classification method, that takes as input a heterogeneous population of single-particle trajectories. It hypothesizes the existence of several different diffusive states within the population; it then determines the most likely diffusive properties of each diffusive state, while sorting each trajectory into the most likely of these diffusive states; it follows this procedure for different numbers of diffusive states, and finally picks the optimum number of diffusive states.

Specifically, for K diffusive states and M tracks, pEMv2 maximizes the log-likelihood of obtaining the measured tracks:

$$\log L = \sum_{m=1}^{M} \log \left( \sum_{k=1}^{K} \pi_k P(\Delta x_m, \Delta y_m | \Sigma_k) \right),$$  \hspace{1cm} (1)

where $\Delta x_m$ is the vector of displacements along x for track m, $\Delta y_m$ is the vector of displacements along y for track m, $\pi_k$ is the fraction of tracks in diffusive state k, $\Sigma_k$ is the displacement covariance matrix of diffusive state k (assumed the same for x and y), and

$$P(\Delta x_m, \Delta y_m | \Sigma_k) = \frac{e^{-\frac{1}{2} \Delta x_m^T \Sigma_k^{-1} \Delta x_m - \frac{1}{2} \Delta y_m^T \Sigma_k^{-1} \Delta y_m} }{(2\pi)^N |\Sigma_k|},$$  \hspace{1cm} (2)

is the probability of realizing trajectory m, given that trajectory m corresponds to diffusive state k, with $|\Sigma_k|$ and $\Sigma_k^{-1}$ the determinant and inverse of $\Sigma_k$, respectively. For trajectories comprising N displacements in x and y, each of $\Delta x_m$ and $\Delta y_m$ is an N-component vector and $\Sigma_k$ is an $N \times N$ symmetric Toeplitz matrix. That is, pEMv2 assumes that particle displacements are multivariate Gaussian random variables. To apply pEMv2 to experimental tracks, we subdivide longer tracks into N step tracks. pEMv2 maximizes log $L$ iteratively by appropriately picking $\pi_k$ and the matrix elements of $\Sigma_k$ for each diffusive state k, and by assigning each track to the most likely diffusive state.

Model selection in pEMv2—that is, picking the appropriate value of K—is implemented by picking the state with the largest Bayesian Information Criterion (BIC), defined here as

$$\text{BIC} = \log L - \frac{1}{2} N_p \log N_D,$$  \hspace{1cm} (3)

where $N_D = 2NM$ is the number of data points, and $N_p = KN + K - 1$ is the number of model parameters, equal to the sum of the number of independent covariance matrix elements, KN, plus the number of independent population fractions, K−1. The log-likelihood always increases as the number of parameters, and therefore the number of states, increases. Counteracting this behavior, the second term on the right hand side of Eq. 3 penalizes a larger number of parameters, and therefore a larger number of states. Together these two contributions lead to an optimum value of K.
2.3 Covariance and diffusivity distributions

According to pEMv2, its sorted populations each correspond to a single diffusive state with well-defined diffusive properties and parameters. Assuming that a given diffusive state’s displacements are zero-mean Gaussian random variable (Eq. 2), all statistical properties of the displacements are determined solely by the mean covariance matrix. Therefore, to test pEMv2’s performance, we sought to compare the pEMv2-sorted covariance- and diffusivity distributions to the corresponding theoretical expectations, given the experimental mean covariances.

The development of theoretical predictions for covariance- and diffusivity distributions is described in detail in Refs. [25,29]. In brief, for a population of two-dimensional, single-particle trajectories, each of length \( N \), and each corresponding to the same diffusive state, the probability density for a track to yield an estimate of the covariance matrix element, \( n \) steps away from the diagonal, equal to \( S_n \), is:

\[
P(S_n|\Sigma) = \int d(\Delta x_1)\ldots d(\Delta y_1)\ldots d(\Delta x_n)\ldots d(\Delta y_n)\]

\[
P(\Delta x, \Delta y|\Sigma) \propto \left( S_n - \frac{1}{2} \Delta x^T C_n \Delta x - \frac{1}{2} \Delta y^T C_n \Delta y \right)
\]

\[
= \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{1}{|I + \frac{1}{\Delta t} \omega \Sigma|} e^{i\omega S_n}, \quad (4)
\]

where \([C_0]_{jk} = \frac{2}{N} I\) and \([C_n]_{jk} = \frac{1}{(N-n)} \delta_{j,k+1} + \delta_{j,k-1}\) for \( n > 0 \), where \( j = 1 \) through \( N \).

Many experimental systems, including TRAP-labeled Pma1 and Pma1-mEos3.2, show diffusive behavior consistent with simple diffusion with experimental errors, corresponding to a symmetric, tridiagonal covariance matrix, where the only nonzero mean covariance matrix elements are on the diagonal, namely \( S_0 \), and one away from the diagonal, namely \( S_1 \). Each individual track yields a measurement of \( S_1 \) and \( S_0 \), which are related to a measurement of the diffusivity, \( D \), and the static localization noise, \( \sigma^2 \), for that track via [29]

\[
S_0 = \sum_{j=1}^{N} (\Delta x_j^2 + \Delta y_j^2) = 4D \Delta t - \frac{4}{3} D \Delta t_E + 2\sigma^2 \quad (5)
\]

and

\[
S_1 = \sum_{j=1}^{N-1} (\Delta x_j \Delta x_{j+1} + \Delta y_j \Delta y_{j+1}) = \frac{2}{3} D \Delta t_E - \sigma^2 , \quad (6)
\]

where \( \Delta t \) is the time between camera exposures and \( \Delta t_E \) is the exposure time. The terms involving the exposure time, \( \Delta t_E \), correspond to motion blur, because measurement of the particle position is integrated while the shutter is open. To be clear, \( S_0 \), \( S_1 \), \( D \), and \( \sigma^2 \) are random variables. Their respective means are \( \bar{S}_0 \), \( \bar{S}_1 \), \( \bar{D} \), and \( \bar{\sigma}^2 \).

Solving Eqs. 5 and 6 for \( D \) and rewriting in terms of \( C_0 \), \( C_1 \), \( \Delta x \), and \( \Delta y \), [25,29], we find

\[
D = \frac{S_0}{4\Delta t} + \frac{S_1}{2\Delta t} = \frac{1}{4\Delta t} (\Delta x^T C_0 \Delta x + \Delta y^T C_0 \Delta y) + \frac{1}{2\Delta t} (\Delta x^T C_1 \Delta x + \Delta y^T C_1 \Delta y) . \quad (7)
\]

Similarly,

\[
\sigma^2 = \frac{\Delta t_E}{6\Delta t} S_0 + \left( \frac{\Delta t_E}{3\Delta t} - 1 \right) S_1 = \frac{\Delta t_E}{6\Delta t} (\Delta x^T C_0 \Delta x + \Delta y^T C_0 \Delta y) + \left( \frac{\Delta t_E}{3\Delta t} - 1 \right) (\Delta x^T C_1 \Delta x + \Delta y^T C_1 \Delta y) . \quad (8)
\]

It follows that the probability densities of the diffusivity and the static localization noise can be expressed as

\[
P(D|\Sigma) = \int \frac{d\omega}{2\pi} \frac{1}{|I + \frac{1}{\Delta t} \omega \Sigma(C_0 + 2C_1)|} e^{i\omega D} . \quad (9)
\]

and

\[
P(\sigma^2|\Sigma) = \int \frac{d\omega}{2\pi} \frac{1}{|I + \frac{1}{\Delta t} \omega \Sigma(C_0 + (\frac{\Delta t_E}{3\Delta t} - 1) C_1)|} e^{i\omega D}, \quad (10)
\]

respectively. For each of Eqs. 4, 9, and 10, to provide explicit results, we calculate the determinant in the integrand as a function of \( \omega \) and then carry out each integral over \( \omega \) numerically. An important aspect of Eqs. 7 and 9 is that there is a nonzero probability that the diffusivity estimated from an individual track (via Eq. 7) is negative, notwithstanding that the underlying diffusivity that governs the track’s motion is positive. In this paper, we will generally refer to the diffusivity, estimated from a single track, as the “track diffusivity” A population average of the track diffusivity equals the diffusivity.

Equations 4, 9, and 10 are applicable when there is one diffusive state. The generalization to \( K \) states, with population fractions specified by \( \{\pi_k\} \) and diffusion properties specified by \( \{\Sigma_k\} \), is straightforward:

\[
P(S_n|\{\pi_k\},\{\Sigma_k\}) = \sum_{k=1}^{K} \pi_k P(S_n|\Sigma_k), \quad (11)
\]

\[
P(D|\{\pi_k\},\{\Sigma_k\}) = \sum_{k=1}^{K} \pi_k P(D|\Sigma_k), \quad (12)
\]
and

\[ P(\sigma^2|\{\pi_k\}, \{\Sigma_k\}) = \sum_{k=1}^{K} \pi_k P(\sigma^2|\Sigma_k). \] (13)

where \{..\} indicates “the set of ...”.

3 Materials and methods

3.1 Sample preparation

All measurements described in this paper employed strains of the budding yeast, *S. cerevisiae*. The construction of these strains is described in detail in [15,38]. For strains expressing a modified version of Pma1, we replaced the native chromosomal PMA1 gene with a gene encoding modified Pma1. Modified Pma1 was expressed from the endogenous PMA1 promoter, allowing us to study behavior at endogenous expression levels.

For our microscopy experiments, overnight cultures were grown in synthetic defined (SD) media, with 2% sucrose and 1% raffinose. These starter cultures were diluted into fresh media, containing 2% galactose, to obtain a final OD\textsubscript{600} ≈ 0.05. Growth was then continued at 30°C for a further 8 hrs. Cells from these cultures were then imaged as follows: 1 mg/mL of concanavalin A (conA) was applied to a clean cover slip, and incubated at room temperature for 10 minutes. Then, 0.5–1.0 mL of fresh milli-Q water was used to rinse off the excess conA. Next, the yeast culture, previously vortexed for approximately 30 s to separate any cell clusters, was added to the conA-coated cover slip and incubated at room temperature for an additional 10 minutes. Excess, unbound cells were rinsed from the cover slip, which was then sealed to a microscope slide using a 1:1:1 ratio mixture of vaseline, lanolin, and paraffin wax (VALAP).

3.2 Microscopy

Microscopy measurements to track the motions of individual molecules of TRAP-labeled Pma1 and Pma1-mEos3.2 were carried out using the custom-built microscope, described in Ref. [39], which has both total-internal reflection fluorescence (TIRF) and photoactivation localization microscopy (PALM) capabilities. We employed a 405 nm-wavelength laser to switch mEos3.2 into its red state and a 560 nm-wavelength laser for imaging of switched mEos3.2. As noted previously, Pma1 is the most abundant yeast membrane protein. Therefore, switching a subset of the population into mEos3.2’s red fluorescent state ensures sufficiently isolated, and therefore resolvable, individual trajectories for Pma1-mEos3.2, that are suitable for unambiguous single-particle tracking. The trajectories reported and analyzed in this paper were collected under conditions of TIRF illumination, which restricted switching and fluorescence excitation to the portion of a cell’s membrane in close proximity to the coverslip. In addition, however, a limited quantity of wide-field PALM data were collected to visually assess the extent to which PALM signal is associated with the cell membrane. The intensity of the 405 nm-wavelength laser was manually adjusted during the experiments to ensure a sufficient signal rate throughout the duration of data acquisition. The fields-of-view imaged were 256 × 256 pixels, with square pixels, each spanning 103 nm on a side. Images were collected at a rate of 100 frames per second (fps), corresponding to Δt = 0.01 s. The exposure time was also Δt\textsubscript{E} = 0.01 s. A custom, reflection-based autofocusing system was deployed during data acquisition to maintain the microscope focus.

3.3 Single-particle tracking

Single-particle tracking was accomplished using a locally customized MATLAB version (The MathWorks, Inc.) of the software, described in Ref. [40], initially resulting in a number of candidate localizations in each movie frame. To achieve our final set of localizations, we manually segmented brightfield images to find cells and excluded any (apparent) localizations outside cells. We also excluded localizations with standard deviations less than 0.5 pixels or more than 3.5 times the mean standard deviation of the population of localizations. Finally, to construct single-particle trajectories, spanning multiple movie frames—i.e., spanning time—we insisted that the maximum number of pixels that a particle can move between successive frames is 2 pixels, and that the maximum number of frames for which a particle can transiently disappear and still be considered part of a specific trajectory is 1 frame.

4 Results and discussion

4.1 Microscopy

Figure 1a shows wide-field fluorescence microscopy images of cells expressing a Pma1-mCherry direct fusion. Apparent in this image is a bright halo of fluorescence intensity at the cell periphery, correspond-
each case, the scale bars correspond to 3 µm.

Cells expressing untagged Pma1 and TRAP-mEos3.2. In the case of the Pma1-MEEVF/TRAP-mEos3.2 TRAP-peptide system; and the dissociation constant to be $K_D \simeq 300$ nM. In Fig. 1c, fluorescence intensity also originates within the cell, corresponding to free TRAP-mEos3.2 in solution. Finally, Fig. 1d shows cells expressing TRAP-mEos3.2 but with unmodified Pma1. As expected, there is no longer an intensity halo at the cells’ periphery, because the TRAP-mEos3.2’s MEEVF binding partner is absent from Pma1 or anywhere else.

For the images in Fig. 1b–d, mEos3.2 was visualized through its unswitched green fluorescence. However, mEos3.2 can be switched by exposure to 405 nm-wavelength light into a red fluorescent state. Figure 2a, b depicts trajectories of switched, red-emitting proteins containing four or more steps obtained from movies collected under 565-nm-illumination from cells with (a) Pma1-mEos3.2 and (b) TRAP-labeled Pma1. In both cases, the spatial distribution of trajectories clearly outlines the periphery of the cell that intersects the focal plane. There is no obvious difference in the spatial distribution of the tracks obtained for Pma1-mEos3.2, which we expect to be strictly confined to the cell membrane, and tracks obtained for TRAP-labeled Pma1. Initially, it may seem surprising that there are no trajectories corresponding to the cytosolic fluorescence apparent in Fig. 1c, which depicts the same yeast strain as Fig. 2b. However, we expect that an unbound TRAP-mEos3.2, freely and rapidly diffusing in the cytosol, would yield either a localization in just a single movie frame or a very short track at most. We infer, therefore, that trajectories with four or more steps in cells with TRAP-peptide labeling overwhelmingly must correspond to labeled Pma1 in the cell membrane, and not to unbound MEEVF-mEos3.2 in the cytosol.

More generally, the observation that unbound TRAP-mEos3.2 may be eliminated from consideration by insisting on sufficiently long tracks suggests a possible general strategy for imaging TRAP-labeled POIs [14,41], namely to insist that tracks endure for some minimum number of time steps. In this case, the fluorescence from unbound TRAP-FPs may be less of a problem for imaging than it might originally have seemed.

Figure 2c, d shows trajectories also containing four or more steps, obtained from PALM movies collected with TIRF illumination, from cells with (c) Pma1-mEos3.2 and (d) TRAP-labeled Pma1. For both labeling methods, tracks appear largely confined to roughly circular regions. Because the tracks are derived from TIRF data, so that we expect to visualize only tracks close to the cover slide, we identify each roughly circular region as the area of the cell in contact with the cover slide. Apparent tracks outside these regions were eliminated from further consideration. Because all of these trajectories involve four or more steps and because they were acquired through TIRF, we are confident that overwhelmingly they correspond to TRAP-labeled Pma1 and not freely diffusing TRAP-mEos3.2. Although wide-field PALM measurements were important for establishing the conditions necessary to examine TRAP-labeled Pma1, in the remainder of the paper, we focus solely on the two-dimensional trajectories collected with TIRF illumination, in order to avoid any ambiguities associated with the sideways view of membrane-based trajectories acquired in wide-field.

4.2 Distribution of track lifetimes

Figure 3 displays the number of tracks versus track length on linear-logarithmic axes for both TRAP-labeled Pma1 (blue) and Pma1-mEos3.2 (pink). Overlaps in the histograms are represented in bordeaux. The track length distributions depend on the FP flu-
Fig. 2 Trajectories comprising four or more steps obtained from 20000 frames of wide-field PALM movies from cells with a Pma1-mEos3.2 and b TRAP-labeled Pma1. Trajectories comprising four or more steps obtained from 20000 frames of PALM movies, under TIRF illumination, from cells with c Pma1-mEos3.2 and d TRAP-labeled Pma1. In each panel, the scale bar represents 3 µm.

Fluorescent state lifetime and on the FP-TRAP dissociation rate, in the case of TRAP-labeled Pma1. Evidently, for both labeling methods, the number of tracks is a monotonically decreasing function of the number of steps in the track. In contrast to this behavior, the distribution of track lifetimes—defined as the time at which all fluorescence finally disappears—for a protein complex, which includes several FPs moving together, is predicted to start at zero, then increase to a peak at nonzero time, before decreasing to zero again at large times. A proof of this result is presented in Appendix A. Thus, the observed lifetime distributions indicate that the observed tracks correspond to single FPs, i.e., Pma1 is a monomer in vivo. (In principle, the observed track distribution also depends on the rate at which tracks leave the illuminated volume, which would end multiple-FP tracks at the same time, but tracks that disappear by leaving the illuminated volume constitute a small fraction of the total number of tracks.)

Although the tracks of both TRAP-labeled Pma1 and Pma1-mEos3.2 correspond to monomers, there is nevertheless a clear difference in the track length distribution between the two methods with the direct fusion showing more long tracks than TRAP-labeled Pma1. This observation is consistent with the interpretation that when Pma1 is directly bound to the FP, the lifetime of the signal is limited by photobleaching, whereas when Pma1 is labeled by the TRAP-peptide pair, the lifetime is determined by a combination of photobleaching and the lifetime of TRAP-MEEVF binding. Thus, we expect shorter lifetimes for TRAP-labeled Pma1 than for Pma1-mEos3.2. In fact, the normalized distribution of TRAP-labeled Pma1 track lengths (Fig. 3b) falls a factor of $e$ below that of Pma1-mEos3.2 for a track length of about 15, suggesting that the mean lifetime of the MEEVF-TRAP bond is about 0.15 s. Equivalently, the MEEVF-TRAP unbinding rate is about 6 s$^{-1}$.

4.3 Population-averaged track diffusivity distributions

Figure 3 makes clear that the number of observed tracks is a rapidly decreasing function of the track length. Therefore, to maximize the available data, while being sure to exclude tracks corresponding to unbound mEos3.2 in the cytosol, we have chosen to focus on tracks consisting of four steps in order to analyze the diffusive behavior of TRAP-labeled and direct-fusion Pma1. Specifically, we use all tracks of length four steps.
Fig. 3  

(a) Number of tracks and (b) probability versus track length plotted on logarithmic-linear axes for TRAP-labeled Pma1, shown in blue, and Pma1-mEos3.2, shown in pink. Overlap in the histograms is represented in bordeaux. Two is the minimum number of connected steps to be considered a track.

Fig. 4  

Comparison between the distributions of track diffusivities, $D$, for (a) experimental TRAP-labeled Pma1 and (b) experimental direct fusion data. In both cases, the track length equals four steps. The overall, unsorted, population-averaged distributions, shown as the light gray histograms, are plotted with the theoretical two-component curve, shown as the red curve, calculated using Eq. 12 and the pEMv2-found covariance values. The sorted track diffusivity distributions (dark and darker gray histograms) are shown with their corresponding single state theory curves (blue and cyan curves), given by Eq. 9. Overlap between histograms is denoted by a deep gray color.

or longer. Thus, included in the analysis are tracks that are split or truncated versions of longer tracks. For example, if we consider a track that is 22 steps long, the way that this track enters our analysis of 4-step tracks is that it is split into five tracks of four steps each, and the remaining two steps at the end are discarded. This approach of splitting longer tracks into $n$-step (here, 4-step) segments is similar to that employed in Ref. [27], where the application of pEMv2 to simulated data, partitioned in this fashion, was shown to accurately identify diffusive states that transition among themselves, provided that the time scales of the transitions were longer than the duration of the segments.

Figure 4 compares population-averaged, track diffusivity distributions, calculated on the basis of Eq. 7, for (a) the population of TRAP-labeled Pma1 4-step tracks and (b) the population of Pma1-mEos3.2 4-step tracks. In the figure, the overall diffusivity distributions of these populations are shown as the light gray histograms. Although both protein variants’ overall track diffusivity distributions are peaked near zero, the Pma1-mEos3.2’s diffusivity distribution is largely confined within $\pm0.1 \mu m^2/s^{-1}$. By contrast, TRAP-labeled Pma1’s track diffusivity distribution shows a large tail, extending beyond $0.5 \mu m^2/s^{-1}$. Evidently, there is a substantial, qualitative difference between the overall track diffusivity distribution of Pma1-mEos3.2 and the track diffusivity distribution of TRAP-labeled Pma1. This result shows directly, without any further detailed analysis, that labeling strategy significantly affects the dynamics of membrane-bound Pma1, with unknown consequences to Pma1’s biological roles. Because TRAP labeling minimally modifies the protein of interest, and direct fusions with other fluorescent proteins cause growth defects and Pma1 mislocalization, our assumption is that TRAP-labeled Pma1 more closely represents the intrinsic biological behavior of wild-type Pma1, than does Pma1-mEos3.2.
4.4 pEMv2 sorts TRAP-labeled Pma1 trajectories into two diffusive states

For the entire population of tracks, we find that the mean covariances, $\bar{S}_n$, are very close to zero for all $n > 1$. Therefore, for simplicity we chose to set these covariances identically equal to zero when running pEMv2. The statement that $\bar{S}_n = 0$ for $n > 1$ is equivalent to the statement that labeled Pma1 undergoes simple diffusion in the presence of experimental errors.

Figure 5 plots the values of $\text{BIC}_K - \text{BIC}_1$ returned by pEMv2 versus the number of diffusive states, $K$, for TRAP-labeled Pma1 (blue). The BIC achieves a maximum for $K = 2$, indicating that pEMv2 robustly supports two subpopulations, namely state 1 and state 2—with distinct diffusive properties. Table 2 reports the mean covariances, $\bar{S}_0$ and $\bar{S}_1$, the mean track diffusivity, $\bar{D}$ and the mean localization noise for these two diffusive states, as well as the fraction of the population corresponding to state 1, $\phi_1$, the fraction of the population corresponding to state 2, $\phi_2$. Evidently, pEMv2 indicates the existence of a relatively low diffusivity state (state 1) with $D_1 \approx 0.0013 \, \mu\text{m}^2\text{s}^{-1}$ and a relatively high diffusivity state (state 2) with $D_2 \approx 0.25 \, \mu\text{m}^2\text{s}^{-1}$. Both states have roughly equal representation in the overall population with $\phi_1 \approx 0.57$ and $\phi_2 \approx 0.43$.

4.5 Covariance distributions of pEMv2-sorted TRAP-labeled Pma1 trajectories agree with theory

To further assess the performance of pEMv2, we first sought to compare the theoretical covariance distributions for each state, conditioned on their respective experimental mean covariances, $\bar{S}_0$ and $\bar{S}_1$, to the corresponding experimental distributions. This is shown in Fig. 6, which plots both experimental and theoretical covariance distributions sorted into states 1 and 2 and for the overall population of unsorted trajectories. Stepped histograms represent the experimental covariance distributions. Smooth lines are the theoretical expectations according to pEMv2.

Table 2. Results from applying pEMv2 to TRAP-labeled Pma1 and direct-fusion data: covariances ($\bar{S}_0$ and $\bar{S}_1$, $\mu\text{m}^2$), diffusivities ($D$, $\mu\text{m}^2\text{s}^{-1}$), localization errors ($\sigma_2$, $\mu\text{m}^2$), and volume fractions ($\phi$, $\%$) for the two states found by pEMv2

| Label | Method | $\bar{S}_0$ | $\bar{S}_1$ | $\bar{D}$ | $\sigma_2$ | $\phi_1$ | $\phi_2$ |
|-------|--------|------------|-------------|----------|-----------|--------|--------|
| TRAP  | $0.0065 \times 10^{-5}$ | $0.0032 \times 10^{-5}$ | $0.0013 \times 10^{-5}$ | $0.0031 \times 10^{-5}$ | $0.0028 \times 10^{-5}$ | $0.0128 \times 10^{-5}$ | $0.0063 \times 10^{-5}$ | $0.13$ |
| Direct fusion | $0.0066 \times 10^{-5}$ | $0.0029 \times 10^{-5}$ | $0.0031 \times 10^{-5}$ | $0.0028 \times 10^{-5}$ | $0.0045 \times 10^{-5}$ | $0.0063 \times 10^{-5}$ | $0.13$ |

Fig. 5 Relative Bayesian Information Criterion ($\text{BIC}_K - \text{BIC}_1$) values versus number of diffusive states for TRAP-labeled Pma1 (blue) and Pma1-mEos3.2 trajectories (red). The highest relative BIC value occurs for two states.
Overlap between histograms is denoted by a deep gray color given by Eq. 4. Overlap between histograms is denoted with their single theory curves (blue and cyan curves), and plot-ted with the theoretical two-component curve (red curve) given by Eq. 11, and the fitted two-component curve (solid black and dashed black lines), and plotted with the theoretical two-component curve (red curve), and the fitted two-component curve (solid black and dashed black lines). The tracks are sorted into two distributions representing the two distinct diffusive states found by pEMv2 (dark gray and darker gray histograms), and plotted with their single theory curves (blue and cyan curves), given by Eq. 4.

4.5.1 Kolomogorov–Smirnov test of pEMv2-sorted distributions

To quantitatively compare the pEMv2-sorted experimental distributions and theory, we adopted a simulation-based approach to estimate the p values appropriate to our measured distributions. Our p values are defined to be the probability of realizing a certain test statistic that is greater than or equal to the value of that test statistic measured experimentally. The test statistic in question is the Kolmogorov–Smirnov statistic (KS statistic), equal to the maximum absolute difference between experimental and simulated cumulative distribution functions (CDFs) of the covariances. We first simulated 40000 trajectories, far more than the number of experimental trajectories. The resultant simulated covariance distributions closely match the corresponding theoretical distributions with relatively small statistical errors. We used these 40000-trajectory-based distributions to determine an experimental KS statistic, equal to the maximum absolute difference between the CDF of the experimental covariance and the CDF of the covariance from the 40000-trajectory simulation. A single simulation of a number of trajectories equal to the number of experimental trajectories, together with the 40000-trajectory-based distribution, determines one simulated value of this KS statistic. To determine the cumulative distribution function of this KS statistic, we repeated such a simulation 2000 times, yielding 2000 values of the KS statistic. Then, the probability of realizing a KS statistic, greater than or equal to the experimental value—i.e., the p value of the experimental KS statistic—can then be read from a plot of one minus the CDF of the KS statistic versus KS statistic. Figure 7 displays the p value so-obtained versus KS statistic for each diffusive state and for each $S_0$ and $S_1$. In addition, highlighted in this figure are the corresponding experimental KS statistics and their corresponding p values, shown as black circles. Figure 8 replots the p value curves of Fig. 7 as a function of $\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \times$ KS statistic, where $n_1 = 40000$ and $n_2$ is the number of experimental tracks. Evidently, plotted in this way, all of the curves more-or-less collapse onto each other. In fact, according to the Kolmogorov theorem, in the large- $n_1$, large- $n_2$ limit, the p value plotted in this way for a one-dimensional probability distribution is a universal function, namely the Kolmogorov distribution, independent of the distribution of the underlying random variables. This prediction appears to be closely followed for our values of $n_1$ and $n_2$.

The KS statistics and p values are presented in Table 3. For both $S_1$-distributions, the estimated p value is larger than 0.05, indicating that both experimental $S_1$-distributions closely match the expected distribution. However, in spite of apparent good theory-experiment agreement in Fig. 6, for the state-2 $S_0$-distribution the experimental KS statistics fall outside of the range of simulated KS statistics, indicating a small p value, and suggesting the existence of systematic errors, that we have not accounted for. Figure 9 shows the CDF for the experimental covariance distributions (dashed black lines), compared to the 40000-trajectory simulated distributions, using the mean covariances determined by pEMv2 (solid blue and cyan lines for $S_0$, and solid red and magenta lines for $S_1$).
Fig. 7 $p$ values versus KS statistic for TRAP-labeled Pma1 tracks, determined as described in the text by simulation, for state 1 $S_0$ (red) and $S_1$ (orange), and state 2 $S_0$ (blue) and $S_1$ (cyan). Also shown is the analytic $p$ value according to the Kolmogorov distribution. The black circles correspond to the experimental KS statistics for each track length and their corresponding $p$ values.

Fig. 8 $p$ values versus $\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \times$ KS statistic, for TRAP-labeled Pma1 tracks, determined as described in the text by simulation, for state 1 $S_0$ (red) and $S_1$ (orange) and state 2 $S_0$ (blue) and $S_1$ (cyan).

Fig. 9 Sorted CDFs for TRAP-labeled tracks. Shown are pEMv2-sorted experimental $S_0$ distributions (black dashed lines) for state 1 and state 2, compared to 40000 simulated tracks for state 1 (blue) and state 2 (cyan), plotted with pEMv2-sorted experimental $S_1$ distributions (black dashed lines) for state 1 and state 2, compared to 40000 simulated tracks for state 1 (red) and state 2 (magenta).

4.6 pEMv2 finds two diffusive states for Pma1-mEos3.2 direct-fusion trajectories

Next, we applied pEMv2 and the theory of Sect. 2.3 to Pma1-mEos3.2 direct-fusion trajectories. As shown in Fig. 5 (red), the BIC was maximized for two states. However, the increase in the BIC, going from one state to two states, is much less in this case ($\text{BIC}_2 - \text{BIC}_1 = 200$), than for TRAP-labeled Pma1 trajectories ($\text{BIC}_2 - \text{BIC}_1 = 800$, Fig. 5), indicating that pEMv2’s preference for a two-state description over a one-state description is much less for Pma1-mEos3.2 than for TRAP-labeled Pma1.

Table 2 reports the mean covariances, $\bar{S}_0$ and $\bar{S}_1$, the mean track diffusivity, $D$ and the mean localization noise for the two diffusive states, found in this case, as well as the fraction of the population corresponding to state 1, $\phi_1$, the fraction of the population corresponding to state 2, $\phi_2$. From this table, it is clear that pEMv2 finds that Pma1-mEos3.2 is overall significantly less mobile than TRAP-labeled Pma1. First, the immobile state (state 1) constitutes about 84% of the total population. Second, the diffusivity of the higher diffusivity state (state 2) is just $D_2 \approx 0.08 \mu m^2 s^{-1}$, several times smaller than the diffusivity of TRAP-labeled Pma1’s higher diffusivity state, for which $D_2 \approx 0.25 \mu m^2 s^{-1}$. Hence, the more mobile second state in this labeling method is still far less mobile than the mobile state for the indirect labeling.

Figure 10 compares the experimental covariance distributions for Pma1-mEos3.2 trajectories to the corresponding theoretical predictions. There appears to be good agreement between theory and data, even though no fitting is involved. The fact that the distributions...
Table 3  KS test results, comparing the experimental pEMv2-sorted TRAP-labeled Pma1 covariance distributions to 40000 simulated tracks based on pEMv2-found values

| Track length | \( p_{S_{01}} \) | \( p_{S_{02}} \) | \( KS_{S_{01}} \) | \( KS_{S_{02}} \) | \( p_{S_{11}} \) | \( p_{S_{12}} \) | \( KS_{S_{11}} \) | \( KS_{S_{12}} \) | State 1 tracks | State 2 tracks |
|--------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|----------------|----------------|
| 4            | 0.378           | <0.0005         | 0.0146          | 0.0642          | 0.646           | 0.285           | 0.0120          | 0.0182          | 4406           | 3324           |

Shown are the \( p \) values and KS statistics for the diagonal and one-off diagonal covariances for state 1 (\( S_{01} \) and \( S_{11} \), respectively) and state 2 (\( S_{02} \) and \( S_{12} \), respectively). \( p \) values of < 0.0005 correspond to experimental KS statistics outside of the range of simulated KS statistics.

Fig. 10  Distribution of covariance elements, \( S_0 \) (a) and \( S_1 \) (b), for experimental direct-fusion tracks. The unsorted covariance distribution (light gray histogram) is plotted with the theoretical two-component curve (red curve) given by Eq. 11, and the fitted two-component curve (black dashed line). The tracks are sorted into two distributions representing the two distinct diffusive states found by pEMv2 (dark gray and darker gray histograms), and plotted with their single theory curves (blue and cyan curves), given by Eq. 4. Overlap between histograms is denoted by a deep gray color.

4.7 Mean-square displacements (MSDs) versus delay time of pEMv2-sorted trajectories

In Sect. 4.4, we used the pEMv2-determined mean covariances to estimate the diffusivity of each diffusive state. Reference [29] showed that such covariance-based estimators are statistically optimal. However, another widely used means of characterizing diffusivity is via the slope of the mean-square displacement (MSD) versus time delay, which is a linear function for simple diffusion. Figure 14 plots the MSDs for TRAP-labeled Pma1 trajectories for each state 1 (blue) and state 2 (cyan) versus the number of steps. Evidently, the MSD for state 1 is conspicuously constant versus time. This observation emphasizes that state 1 really is essentially immobile. By contrast, it is clear that the state-2 MSD increases approximately linearly versus time, consistent with what is expected for simple diffusion. Also shown are the MSDs for direct-fusion trajectories for state 1 (red) and state 2 (magenta). We observe a similar trend, where state 1 is essentially immobile, and state Pma1-mEos3.2's \( S_0 \) are smaller than for TRAP-labeled Pma1, indicating less overall agreement between experiment and theory in this case. Because the mobility of Pma1-mEos3.2 is smaller than for TRAP-labeled Pma1, we may expect that localization noise is a more dominant contributor to \( S_0 \) and \( S_1 \) for Pma1-mEos3.2 than for TRAP-labeled Pma1. Therefore, because the theory assumes Gaussian localization noise, any deviation of the localization noise from Gaussian behavior may manifest as poorer experiment–theory agreement for Pma1-mEos3.2 than for TRAP-labeled Pma1. Previously, Ref. [25] showed that heterogeneity in the localization noise could account for discrepancies between experimental and predicted \( S_0 \) and \( S_1 \) distributions.

Figure 4b shows the track diffusivity distributions for Pma1-mEos3.2 trajectories. Two features stand out from this figure in comparison to Fig. 4a. First, the population fraction of the higher diffusivity state (state 2) is significantly less for the direct fusion than for TRAP-labeled Pma1. Second, the mean track diffusivity of the higher diffusivity state (state 2) is significantly less for the direct fusion (\( \sim 0.05 \) \( \mu m^2 s^{-1} \)) than for TRAP-labeled Pma1 (\( \sim 0.16 \) \( \mu m^2 s^{-1} \)). These two features make the overall mean track diffusivity of Pma1-mEos3.2 significantly smaller than that of TRAP-labeled Pma1.
Table 4  KS test results, comparing the experimental pEMv2-sorted direct-fusion covariance distributions to 40000 simulated tracks based on pEMv2-found values

| Track length | $p_{S_{01}}$ | $p_{S_{02}}$ | $KS_{S_{01}}$ | $KS_{S_{02}}$ | $p_{S_{11}}$ | $p_{S_{12}}$ | $KS_{S_{11}}$ | $KS_{S_{12}}$ | Tracks in state 1 | Tracks in state 2 |
|--------------|-------------|-------------|--------------|--------------|-------------|-------------|--------------|--------------|-----------------|------------------|
| 4            | 0.0085      | < 0.0005    | 0.0258       | 0.106        | 0.634       | 0.0695      | 0.0108       | 0.0426       | 4904            | 910              |

Shown are the $p$ values and KS statistics for the diagonal and one-off diagonal covariances for state 1 ($S_{01}$ and $S_{11}$, respectively) and state 2 ($S_{02}$ and $S_{12}$, respectively). $p$ values of < 0.0005 correspond to experimental KS statistics outside of the range of simulated KS statistics.

![Fig. 11](image1.png)  
**Fig. 11** $p$ values versus KS statistic for direct-fusion tracks, determined as described in the text by simulation, for state 1 $S_0$ (red) and $S_1$ (orange), and state 2 $S_0$ (blue) and $S_1$ (cyan). Also shown is the analytic $p$ value according to the Kolmogorov distribution. The black circles correspond to the experimental KS statistics for each track length and their corresponding $p$ values.

![Fig. 12](image2.png)  
**Fig. 12** $p$ values versus $\sqrt{\frac{n_1 n_2}{n_1 + n_2}} \times$ KS statistic, for direct-fusion tracks, determined as described in the text by simulation, for state 1 $S_0$ (red) and $S_1$ (orange) and state 2 $S_0$ (blue) and $S_1$ (cyan).

2 increases linearly with time, however with a smaller slope than state 2 of the TRAP-labeled tracks, signifying less motion.

5 Conclusion

Two primary conclusions are to be drawn from this work. First, the diffusive behavior of Pma1 can be described convincingly in terms of two discrete diffusive states, each with its own diffusive properties, sorted by the pEMv2 software, and validated using our theoretical covariance distributions. Specifically, we investigated two differently labeled versions of Pma1 at the single-molecule level using single-particle tracking under TIRF illumination. Using a machine-learning-based approach that identifies and sorts a population of trajectories into a discrete number of diffusive states, pEMv2, we found that a minimally modified version of Pma1, with a C-terminal five-amino-acid tag that reversibly binds to a TRAP-mEos3.2 fusion, shows comparable population fractions of a mobile state, corresponding to simple diffusion with a diffusion constant consistent with what is expected for a membrane protein, and an essentially immobile state. By contrast,
we found that a Pma1-mEos3.2 direct fusion is overwhelmingly in an immobile state, and the small population fraction, assigned to a more mobile state by pEMv2, has a diffusion coefficient several times smaller than the diffusion coefficient of mobile TRAP-labeled Pma1. A comparison between the experimental pEMv2-sorted covariance and track diffusivity distributions, conditioned on the experimental mean covariances of the sorted tracks, and the corresponding theoretical distributions that we derived, shows overall good agreement without any fitting, providing strong support for pEMv2-based sorting. Additionally, we note that the method used to label a protein for microscopy visualization in living cells can affect the protein’s diffusive behavior in a substantial fashion with unknown consequences for its biological function. Further studies are required to determine which labeling method more closely represents the motion of the unlabeled protein. Second, we found that both variants of Pma1 studied in this work, TRAP-labeled Pma1 and Pma1-mEos3.2, are monomers in vivo. This result is noteworthy because in vitro preparations for electron microscopy show Pma1 forms hexamers. We estimated the lifetime of the TRAP-peptide bound-state to be 0.15 s. An open question remains as to whether Pma1’s two diffusive states are significant for Pma1’s biological function, and if so what their roles might be.

Acknowledgements This research was supported by the National Institutes of Health via NIH R01 GM118528 and NIH P30 DK045735. M.L.P.B. was supported by NIH T32EB019941 and the NSF GRFP. S.E.P. was supported by the NSF GRFP and NSF PHYS 1522467. We thank our reviewer for the proof in Appendix A.

Author contribution statement
MLPB and SEP contributed equally to this work. MLPB conceived the project, analyzed data, and wrote the paper; SEP conceived the project, collected data, analyzed data, and wrote the paper; YZ conceived the project, collected data, and wrote the paper; MH conceived the project, constructed the yeast strains, and wrote the paper; JB conceived the project and wrote the paper; LR conceived the project and wrote the paper; SGJM conceived the project and wrote the paper.

Data availability statement The datasets analyzed during this study are available from the corresponding author on reasonable request.

Declarations
Conflict of interest J. B. discloses significant financial interest in Bruker Corp., Hamamatsu Photonics, and Pan-luminate Inc.

Appendix A The lifetime distribution of multi-fluorophore tracks
Let \( p(t) \) denote the probability density of the lifetime until bleaching \( t \), for a single fluorophore. Then, the cumulative probability that a single fluorophore has a lifetime less than \( t \) is

\[
P(t) = \int_0^t p(t') \, dt'.
\]

(A1)

Now suppose that a protein complex is labeled with \( n \) fluorophores, and assume that the fluorophores do not affect each other’s lifetimes. Then, the probability that all \( n \) in a complex have bleached by time \( t \) is:

\[
P(t)^n.
\]

(A2)

Thus, the rate at which protein complexes labeled in this manner become invisible is

\[
\frac{d}{dt}(P(t)^n) = np(t)P(t)^{n-1}.
\]

(A3)

We can safely assume that the rate of bleaching of an individual fluorophore is constant in time or approximately so. It is certainly not singular at any time, since that would require an exceptional physical and/or chemical cause of bleaching to occur at that time. It follows that \( p(t) \) is non-singular at \( t = 0 \), and hence

\[
p(t) \approx p(0) + \dot{p}(0)t + ...
\]

(A4)

for small \( t \). Consequently, the fluorescent lifetime distribution of the \( n \) fluorophores protein complex is

\[
\frac{d}{dt}(P(t)^n) = np(t)P(t)^{n-1} \approx np(0)^n t^{n-1}.
\]

(A5)

which vanishes as \( t \to 0 \) unless \( n = 1 \). Thus, Fig. 3b demonstrates that \( n = 1 \) because it shows that the distribution, described by Eq. A5, is nonzero as \( t \to 0 \).
Although the lifetime distributions we observe are more complicated than simple exponentials, it is illuminating to consider the simple, concrete case of a fluorophore lifetime distribution given by

\[ p(t) = \lambda e^{-\lambda t}. \]  \hspace{1cm} (A6)

Then,

\[ P(t) = 1 - e^{-\lambda t}. \]  \hspace{1cm} (A7)

It follows that for a complex with \( n \) fluorophore, the probability that all \( n \) are bleached by time \( t \) is

\[ P(t)^n = (1 - e^{-\lambda t})^n. \]  \hspace{1cm} (A8)

In turn, the lifetime distribution in this case is

\[ \frac{d}{dt} (P(t)^n) = n \lambda e^{-\lambda t} (1 - e^{-\lambda t})^{n-1} \approx n \lambda^n t^{n-1}, \]  \hspace{1cm} (A9)

consistent with Eq. A5.

References

1. R.W. Albers, Biochemical aspects of active transport. Ann. Rev. Biochem. 36(1), 727–756 (1967). https://doi.org/10.1146/annurev.bi.36.070167.003455

2. R.L. Post, C. Hegvary, S. Kume, Activation by adenine triphosphate in the phosphorylation kinetics of sodium and potassium ion transport adenosine triphosphatase. J. Biol. Chem. 247(20), 6530–6540 (1972). https://doi.org/10.1016/s0021-9258(19)44725-x

3. E. Goormaghtigh, C. Chadwick, G.A. Scarborough, Monomers of the neurospora plasma membrane H^+-ATPase catalyze efficient proton translocation. J. Biol. Chem. 261, 7466–7471 (1986)

4. P. Morosmaa, C.W. Slayman, A. Goffeau, Mutagenic study of the structure, function and biogenesis of the yeast plasma membrane H^+-atpase. Biochim. Biophys. Acta Rev. Biomembr. 1469(3), 133–157 (2000). https://doi.org/10.1016/s0304-4157(00)00015-0

5. K. Malinská, J. Malinský, M. Opekarová, W. Tanner, Visualization of protein compartmentation within the plasma membrane of living yeast cells. Mol. Biol. Cell 14(11), 4427–4436 (2003). https://doi.org/10.1091/mbc.e03-04-0221

6. W. Kühlbrandt, Biology, structure and mechanism of p-type ATPases. Nat. Rev. Mol. Cell Biol. 5(4), 282–295 (2004). https://doi.org/10.1038/nrm1354

7. F. Spira, N.S. Mueller, G. Beck, P. von Olshausen, J. Beig, R. Widlich-Soeldner, Patchwork organization of the yeast plasma membrane into numerous coexisting domains. Nat. Cell Biol 14, 640–648 (2012)

8. A. Mason, C. Slayman, Encyclopedia of Biological Chemistry (Elsevier, Berlin, 2013), pp.688–692. https://doi.org/10.1016/b978-0-12-378630-2.00194-8

9. K.A. Henderson, A.L. Hughes, D.E. Gottschling, Mother–daughter asymmetry of pH underlyes aging and rejuvenation in yeast. eLife (2014). https://doi.org/10.7554/eLife.05504

10. A. Athanasopoulos, B. Andre, V. Sophianopoulou, C. Gournas, Fungal plasma membrane domains. FEMS Microbiol. Rev. 43, 642–673 (2019)

11. S. Heit, M.M.G. Geurts, B.J. Murphy, R.A. Corey, D.J. Mills, W. Kühlbrandt, M. Bublitz, Structure of the hexameric fungal plasma membrane proton pump in its autoinhibited state. Sci. Adv. 7(46), eabj5255 (2021). https://doi.org/10.1126/sciadv.abj5255

12. P. Zhao, C. Zhao, D. Chen, C. Yun, L. Bai, Structure and activation mechanism of the hexameric plasma membrane H^+-ATPase. Nat. Commun. 12, 6439 (2021)

13. S.E. Pratt, E.B. Speltz, S.G.J. Mochrie, L. Regan, Designed proteins as novel imaging reagents in living Escherichia coli. ChemBioChem 17(17), 1652–1657 (2016). https://doi.org/10.1002/cbic.201600252

14. C. Oi, Z. Gidden, L. Holyoake, O. Kantelberg, S. Mochrie, M.H. Horrocks, L. Regan, A step-by-step protocol for performing LIVE-PAINT super-resolution imaging of proteins in live cells using reversible peptide–protein interactions (2020). https://doi.org/10.21203/rs.3.pex-1043/v1

15. M. Hinrichsen, M. Lenz, J.M. Edwards, O.K. Miller, S.G.J. Mochrie, P.S. Swain, U. Schwarz-Linek, L. Regan, A new method for post-translationally labeling proteins in live cells for fluorescence imaging and tracking. Protein Eng. Des. Sel. 30(12), 771–780 (2017). https://doi.org/10.1093/protein/gzx059

16. C. Oi, Z. Gidden, L. Holyoake, O. Kantelberg, S. Mochrie, M.H. Horrocks, L. Regan, LIVE-PAINT allows super-resolution microscopy inside living cells using reversible peptide–protein interactions. Commun. Biol. (2020). https://doi.org/10.1038/s42003-020-01188-6

17. C. Oi, S.G.J. Mochrie, M.H. Horrocks, L. Regan, PAINT using proteins: a new brush for super-resolution artists. Protein Sci. 29(11), 2142–2149 (2020). https://doi.org/10.1002/pro.3953

18. W.K. Huh, J.V. Falvo, L.C. Gerke, A.S. Carroll, R.W. Howson, J.S. Weissman, E.K. O’Shea, Global analysis of protein localization in budding yeast. Nature 425(6959), 686–691 (2003). https://doi.org/10.1038/nature02026

19. D.L. Coutu, T. Schroeder, Probing cellular processes by long-term live imaging—historic problems and current solutions. J. Cell Sci. (2013). https://doi.org/10.1242/jcs.118349

20. C. Li, A. Tebo, A. Gautier, Fluorogenic labeling strategies for biological imaging. Int. J. Mol. Sci. 18(7), 1473 (2017). https://doi.org/10.3390/ijms18071473

21. C.P. Toseland, Fluorescent labeling and modification of proteins. J. Chem. Biol. 6(3), 85–95 (2013). https://doi.org/10.1016/s12154-013-0094-5

22. E.C. Jensen, Use of fluorescent probes: their effect on cell biology and limitations. Anat. Rec. Adv. Integr. Anat. Evolut. Biol. 295(12), 2031–2036 (2012). https://doi.org/10.1002/ar.22602

23. F. Montecinos-Franjola, B.L. Bauer, J.A. Mears, R. Ramachandran, GFP fluorescence tagging alters dynamin-related protein 1 oligomerization dynamics and creates disassembly-refractory puncta to mediate mitochondrial fission. Sci. Rep. (2020). https://doi.org/10.1038/s41598-020-71655-x

24. C.S. Lisenbee, S.K. Kaurik, R.N. Trelease, Overexpression and mislocalization of a tail-anchored GFP refines the identity of peroxisomal ER. Traffic 4(7), 491–501 (2003). https://doi.org/10.1046/j.1600-0854.2003.00107.x

25. M.L.P. Bailey, H. Yan, I. Surovtsev, J.F. Williams, M.C. King, S.G.J. Mochrie, Covariance distributions in single
particle tracking. Phys. Rev. E (2021). https://doi.org/10.1103/physreve.103.032405

26. P.K. Koo, M. Weitzman, C.R. Sabanaygam, K.L. van Golen, S.G. Mochrie, Extracting diffusive states of Rho GTPase in live cells: towards in vivo biochemistry. PLoS Comput. Biol. 11(10), e1004297 (2015)

27. P.K. Koo, S.G. Mochrie, Systems-level approach to uncovering diffusive states and their transitions from single-particle trajectories. Phys. Rev. E 94(5), 052412 (2016)

28. I. Rey-Suarez, B.A. Wheatley, P. Koo, A. Bhanja, Z. Shu, S. Mochrie, W. Song, H. Shroff, A. Upadhyaya, WASP family proteins regulate the mobility of the B cell receptor during signaling activation. Nat. Commun. (2020). https://doi.org/10.1038/s41467-020-14335-8

29. C.L. Vestergaard, P.C. Blainey, H. Flyvbjerg, Optimal estimation of diffusion coefficients from single-particle trajectories. Phys. Rev. E 89(2), 022726 (2014)

30. B. Justesen, R. Hansen, H. Martens, L. Theorin, M. Palmgren, K. Martinez, T. Pomorski, A. Fuglsang, Active plasma membrane p-type H\(^+\)-ATPase reconstituted into nanodiscs is a monomer. J. Biol. Chem. 288, 26419–26429 (2013)

31. M. Saxton, Single-particle tracking: the distribution of diffusion coefficients. Biophys. J. 72(4), 1744–1753 (1997). https://doi.org/10.1016/s0006-3495(97)78820-9

32. T. Savin, P.S. Doyle, Statistical and sampling issues when using multiple particle tracking. Phys. Rev. E (2007). https://doi.org/10.1103/physreve.76.021501

33. N. Monnier, Z. Barry, H.Y. Park, K.C. Su, Z. Katz, B.P. English, A. Dey, K. Pan, I.M. Cheeseman, R.H. Singer, M. Bathe, Inferring transient particle transport dynamics in live cells. Nat. Methods 12(9), 838–840 (2015). https://doi.org/10.1038/nmeth.3483

34. R. Das, C.W. Cairo, D. Coombs, A hidden Markov model for single particle tracks quantifies dynamic interactions between LFA-1 and the actin cytoskeleton. PLoS Comput. Biol. 5(11), e1000556 (2009). https://doi.org/10.1371/journal.pcbi.1000556

35. P.J. Slator, C.W. Cairo, N.J. Burroughs, Detection of diffusion heterogeneity in single particle tracking trajectories using a hidden Markov model with measurement noise propagation. PLoS ONE 10(10), e0140759 (2015). https://doi.org/10.1371/journal.pone.0140759

36. R. Zhao, J. Yuan, N. Li, Y. Sun, T. Xia, X. Fang, Analysis of the diffusivity change from single-molecule trajectories on living cells. Anal. Chem. 91(21), 13390–13397 (2019). https://doi.org/10.1021/acs.analchem.9b01005

37. H.D. Pinholt, S.S.R. Boehr, J.F. Iversen, W. Boomsma, N.S. Hatzakis, Single-particle diffusional fingerprinting: a machine-learning framework for quantitative analysis of heterogeneous diffusion. Proc. Natl. Acad. Sci. 118(31), 2014624118 (2021). https://doi.org/10.1073/pnas.2104624118

38. M. Hinrichsen, Applying protein–peptide interactions to fluorescently label proteins in S. cerevisiae. Ph.D. Thesis, Yale University (2018)

39. Y. Zhang, M. Lara-Tejero, J. Bewersdorf, J.E. Galán, Visualization and characterization of individual type III protein secretion machines in live bacteria. Proc. Natl. Acad. Sci. 114(23), 6098–6103 (2017). https://doi.org/10.1073/pnas.1705823114

40. J.C. Crocker, D.G. Grier, Methods of digital video microscopy for colloidal studies. J. Colloid Interface Sci. 179(1), 298–310 (1996)

41. C. Oi, Reversible peptide–protein interactions inside cells: enabling a new approach for achieving super-resolution imaging in live cells. Ph.D. Thesis, Yale University (2021)