Inferring diffusion in single live cells at the single-molecule level

Alex Robson1, Kevin Burrage2,3 and Mark C. Leake1,4

1 Clarendon Laboratory, Department of Physics, Oxford University, Parks Road, Oxford OX1 3PU, UK
2 Department of Computer Science, Oxford University, Wolfson Building, Parks Road, Oxford OX1 3QD, UK
3 Mathematics Department, QUT, Brisbane, Queensland 4001, Australia
4 Department of Biochemistry, Oxford University, South Parks Road, Oxford OX1 3QU, UK

The movement of molecules inside living cells is a fundamental feature of biological processes. The ability to both observe and analyse the details of molecular diffusion in vivo at the single-molecule and single-cell level can add significant insight into understanding molecular architectures of diffusing molecules and the nanoscale environment in which the molecules diffuse. The tool of choice for monitoring dynamic molecular localization in live cells is fluorescence microscopy, especially so combining total internal reflection fluorescence with the use of fluorescent protein (FP) reporters in offering exceptional imaging contrast for dynamic processes in the cell membrane under relatively physiological conditions compared with competing single-molecule techniques. There exist several different complex modes of diffusion, and discriminating these from each other is challenging at the molecular level owing to underlying stochastic behaviour. Analysis is traditionally performed using mean square displacements of tracked particles; however, this generally requires more data points than is typical for single FP tracks owing to photophysical instability. Presented here is a novel approach allowing robust Bayesian ranking of diffusion processes to discriminate multiple complex modes probabilistically. It is a computational approach that biologists can use to understand single-molecule features in live cells.

1. Introduction

Biological processes in the cell membrane are hard to replicate in artificial biomimetic membranes in vitro as the native protein–lipid architectures and dynamics in the membrane environment are far from well understood, even in the simplest prokaryotic organisms such as bacteria, let alone in more complex eukaryotic cells. An emerging paradigm for membrane substructure has changed from that of a freely mixed system embodied by the classic Singer–Nicholson model [1] to the concept of a compartmentalized fluid [2–4]. It is the interactions between diffusing proteins and the underlying membrane substructure that maintains the observed heterogeneity. Several observations have led to this hypothesis; on a macroscopic length scale of several hundred nanometres, the diffusion coefficients of proteins are one to two orders of magnitude lower than those observed in artificial membranes [4–9], also the observation that membrane proteins have dramatic drops in diffusion rates upon oligomerization or aggregation [4,10,11], incommensurate with Saffman–Delbrück modelling [12,13] which represents the standard analytical method for characterizing the frictional drag of protein molecules in lipid bilayers. Non-specific interactions are also attributed to membrane heterogeneity; for example, in simple lipid bilayers protein–lipid and lipid–lipid interactions can cause proteins to partition into self-associating clusters [14], creating protein-rich or poor regions in cells. Also, there is some evidence for regions of lipid micro- and nanoscale structure identified in some eukaryotic membranes commonly referred to as lipid ‘rafts’, which often appear to be consistent with mobile regions of phase-separated membrane that exist in an ordered, dense liquid phase surrounded by a more fluidic phase [15,16]. These may be of functional
advantage to signalling systems as well as being implicated in protein partitioning.

What is apparent is that there exists significant heterogeneity in local membrane architecture for a range of important biological functions. A key method for investigating the complex environment of the cell membrane is to monitor the fine details of diffusion of single molecules and complexes in native membranes. A tool of choice is fluorescence microscopy. This offers relatively minimal perturbation to native physiology while presenting an exceptional imaging contrast at single-molecule sensitivity levels that can allow the movement of individual fluorophore-tagged molecules, such as proteins and lipids, to be tracked with nanoscale precision [17-19].

Single-particle tracking (SPT) approaches in general are powerful for interrogating dynamic membrane processes. Earlier studies involved colloidal gold for tracking [4,20]. This has a clear advantage of an exceptionally high signal-to-noise ratio for particle detection with no danger of probe photobleaching, which permits longer tracks to be obtained with very short sampling time intervals at the submillisecond level. However, a significant disadvantage is the size of the probe at typically tens to hundreds of nanometres—this is often larger than underlying substructures of the membrane. SPT of fluorescently labelled particles in the membrane offers significant advantages in using a much smaller probe on the nanometre scale. This was first applied using organic dye labelling [21,22], but the recent use of genomically encoded fluorescent protein (FP) reporters, such as green fluorescent protein (GFP) and its different coloured variants, has enabled many SPT studies to be performed on living cells with exceptional tagging specificity for the protein under investigation [23].

The most robust fluorescence imaging method for probing molecular level localization in the cell membrane is total internal reflection fluorescence (TIRF) microscopy (see [24] for a discussion). This uses typically laser excitation at a highly oblique angle of incidence to generate an evanescent excitation field in the water-based environment of the sample—this can be thought of as an ‘optical slice’ of approximately 100 nm thickness on the surface of the glass microscope slide/cover slip on which a cell sample is mounted. This results in significant excitation of fluorescently labelled molecules in the cell membrane in the vicinity of the slide/cover slip surface. There is minimal excitation of components beyond this, either in the cell or from background fluorescence in the physiological buffer; therefore, the signal-to-noise ratio for imaging membrane components is increased substantially. The result is a very high detection contrast for fluorescently labelled molecules and complexes in the cell membrane.

Four principal different diffusive modes are illustrated schematically in figure 1a,b, with figure 1c plotting idealized mean square displacement (MSD) versus time interval $t$ [25]. Brownian motion represents ‘normal’ diffusion and is the simplest mode of diffusion characterized by a linear relation between MSD and $t$. However, a tracked protein trajectory for which the MSD reaches an asymptote at high $t$ is indicative of confined diffusion, suggesting that the tracked protein is being trapped by its local environment—such corrals have been hypothesized as being important to forming nanoscale reaction chambers, thereby greatly enhancing chemical efficiency [4,9,26-28]. Directed motion has an upwardly parabolic MSD function versus $t$ and is seen for example during active diffusive processes such as those that occur when molecular motors walk on microtubules [29,30]. The final class of diffusion has many examples in biology and is defined by the so-called anomalous or subdiffusive behaviour [31]. This motion is usually modelled as an MSD being proportional to $t^a$, where $a$ is a coefficient between 0 and 1; this mode may, for example, represent the percolation of a protein through the disordered media of the membrane, hopping across corrals or interactions with specialized domains [32-35].

What is apparent from the in vivo SPT studies that have emerged over the past decade is that diffusion of molecules and complexes in living cells is in general not simple when viewed over the broad time scale of milliseconds to seconds relevant to many essential biological processes, and the reason for this may be fundamentally linked to critical substructural features of cells that are characterized over a length scale of a few to several hundred nanometres. There is therefore a compelling biological need to try to understand these complex diffusive processes.

Several analytical approaches have been attempted for characterizing diffusion in the cell membrane in addition to standard Saffman–Delbrück modelling, often involving heuristic approaches [11,36]. A common approach has been to measure the ratio of the MSD to that expected from simple Brownian motion, embodied by a ‘relative deviation’ parameter [26]. For directed motion, this parameter increases with $t$, whereas for confined motion, it decreases. However, these types of MSD analysis approaches are weak on several levels when applied to tracks generated from FP-tagged molecular complexes in vivo. First, trajectories are determined from a low-signal-to-noise ratio environment in which only tracks of short duration are able to be measured owing primarily to the poor photophysics of the fluorophore resulting in scant, imprecise MSD information at high $t$ [37]. Second, trajectories are generated from a stochastic process, implying significant deviations from the idealized graphs of figure 1c. In addition, this method is highly reliant upon an accurate measurement of the diffusion coefficient, which in the noisy, heterogeneous environment of the cell membrane may prove very challenging.

An appeal of MSD analysis lies in its relative simplicity to address qualitative questions concerning the membrane environment, for example effects of molecular crowding or confinement [38,39]. However, as figure 1d,e illustrates, there is an expected statistical spread of the MSD traces for pure Brownian diffusion simply on the basis of diffusion being a stochastical process, that could be erroneously interpreted as a different diffusive behaviour judged by any individual MSD curve from a single particle trajectory. Population averaging can smooth out such variation to obtain average behaviour, but with the unfortunate result that we lose informative data concerning the biological heterogeneity of the ensemble of diffusing molecular complexes. Recent improvements to MSD analyses have involved applications of some diffusion propagators directly. The propagators define the probability distributions that a diffusing particle will be at a given distance from its origin after a given time. These methods have been used in estimating cumulative probability distributions to substantiate the presence of different non-Brownian diffusive modes [33,40,41].

Our new approach here is to present an inference scheme that can separate the distinct types of diffusive modes of
individual trajectories without population averaging, and do so in a probabilistic fashion, given conditions imposed on real experimental data. This can then be combined with photophysical information to quantify molecular stoichiometry of diffusing complexes, thus allowing probing of non-trivial relations between the size of a molecular complex and how fast it moves \textit{in vivo}. The inference of these diffusive modes is done using a Bayesian approach, incorporating \textit{a priori} knowledge, based on both simulation and experiment. We denote this as \textit{Bayesian ranking of diffusion} (BARD).

Our study outlines the principles of the inference in light of the theory of diffusive processes. We describe the details of simulation using the different diffusive modes, and the inference algorithm used in separating out diffusive modes in a quantitative, probabilistic manner. We validate the inference using realistic simulated data, and apply to two different cell strains expressing FP fusion constructs to different membrane proteins. We obtain these live-cell data using TIRF microscopy. One cell strain expresses a single transmembrane helix probe in the cell membrane with a GFP fusion protein. The second strain is a yellow fluorescent protein (YFP) fusion to single twin-arginine translocation (Tat) protein complexes expressed in the cytoplasmic membranes of living bacteria, which exhibit significant real heterogeneity in terms of molecular stoichiometry, architecture and mobility.

The key concepts of these analyses are described in §2, with the specifics in the electronic supplementary material.

2. Material and methods

\textbf{(a) Bacterial cell strain and preparation}

Two different \textit{Escherichia coli} strains were used in our \textit{in vivo} microscopy investigations. One was cell strain AyBC, as studied previously [11], using identical cell preparation conditions. This represents a heterogeneous, oligomeric membrane protein system. The cell strain contained a construct specifying a C-terminal enhanced YFP tag (Clontech Laboratories Inc., Mountain View, CA) to the native \textit{E. coli} protein TatA on the cytoplasmic side of the membrane. The Tat system of bacteria translocates natively folded protein substrates across the cytoplasmic membrane through a nanopore whose walls are...
composed of subunits of the TatA protein (figure 2a). In addition to TatA, there are two other essential proteins in the Tat system, TatB and TatC, implicated both in substrate recruitment and gating of the TatA nanopore (figure 2b).

A second fusion construct was also investigated, denoted Helix1021–GFP (figure 2b). This represents a far less complex membrane protein system, which consisted of just a simple model membrane protein of a single membrane-spanning alpha-helix fused to a GFP tag on the cytoplasmic side of the membrane [42]. The fusion gene coding for this model membrane protein used the open reading frame sll1021 in the cyanobacterium Synechocystis sp. PCC6803 as a start point, but expressing this as a membrane protein in E. coli for which there were no identified orthologues. The protein has an undetermined function but has been identified in the plasma membrane of Synechocystis [43], with the predicted gene product consisting of 673 amino acids with a single predicted transmembrane alpha-helix close to the N-terminus. A portion of the sll1021 sequence coding for 38 amino acids including the predicted transmembrane alpha-helix was fused in-frame to the gene coding for GFPmut3* [44] with a linker of five asparagine residues. This construct was expressed in E. coli cells from the arabinose-inducible pBAD24 vector [45], with a predicted topology for the TatA TatC complex of (a) TatA TatB TatC.

Cells of both strains were grown in Luria–Bertani medium [46] aerobically with shaking overnight at 37°C, and supplemented with 50 μg ml⁻¹ ampicillin for correct antibiotic-resistant colony selection. Cells were diluted by 1:100 from the saturated culture into M63 minimal media for subculturing, and were grown to mid-exponential phase typically for 3.5 h at 30°C. For the Helix1021–GFP strain, t-arabinose was added to the culture at a final concentration of 2 mM. Cells were injected into a 5–10 μl flow-cell with poly-l-lysine-coated glass coverslips as the lower surface, allowed to settle for 10 min, washed with excess M63 and incubated with a 0.1 per cent suspension of 202 nm diameter latex microspheres (Invitrogen Ltd., Paisley, UK) for 2 min to mark the coverslip surface, and washed with excess M63.

(b) TIRF microscopy and single-particle tracking

A home-built inverted TIRF microscope was used with either a 473 nm laser for GFP excitation, or a 532 nm excitation wavelength for YFP excitation, with excitation intensity in the range 250–500 W cm⁻² and measured depth of evanescent field penetration 110 ± 10 nm, with specifications as described previously [7,11,28,47–50], using either 473 or 332 nm laser diode mirrors and notch-rejection filters (Semrock) as appropriate. The focal plane was set at 100 nm from the coverslip surface to image the cell membrane conjugated to the glass coverslip. Fluorescence emission was imaged at approximately 40 nm per pixel in frame-transfer mode at 25 Hz by a 128 × 128 pixel, cooled, back-thinned electron-multiplying charge-coupled device camera (EMCCD; iXon+ DV860-BI, Andor Technology). Images were sampled for typically approximately 8 s. Fluorescent particle positions on each time-stamped image frame were detected and fitted using automated custom-written image-analysis software that fitted a two-dimensional radial Gaussian function plus planar local background to the image intensity data for each candidate particle.

This generated the fluorescence intensity for each distinct diffusing ‘spot’ of fluorescence in the cell of typical width 300–400 nm, either owing to a TatA–YFP complex or to an assemblage of Helix1021–GFP molecules, plus the local background intensity per pixel due to any autofluorescence and/or diffuse fluorescence components, and outputted the intensity centroid to a subpixel precision of approximately 40 nm for single FP molecules, and down to 5–10 nm for molecular complexes/assemblages containing more typically approximately tens of FP molecules.

Tracks were generated from each particle provided tolerance criteria in subsequent image frames were satisfied on the basis of size, intensity and position of detected particles in subsequent image frames, for at least five consecutive image frames. The MSD versus time-interval relation was then calculated for each particle trajectory, as described previously [11]. Using a Fourier spectral approach, we were able to estimate the stoichiometry of these complexes through step-wise photobleaching of the relevant FP molecule [7].

(c) Implementing and validating the BARD algorithm

(i) Generation of synthetic tracks for validation

Two-dimensional simulated tracks for use in validation were generated in a standard way by a stochastic random walk process in MATLAB (The MathWorks, Natick, MA) to approximate real diffusion for the fluorescently labelled proteins in cytoplasmic membranes of E. coli cells, sampling at the same 40 ms video-rate time interval as for experimental imaging, with track durations of typically 0.8 s (see the electronic supplementary material, figure S1).
(ii) Bayesian formulation

The general principle of Bayesian inference is to quantify the present state of knowledge and refine this on the basis of new data, underpinned by Bayes’ theorem, emerging from the definition of conditional probabilities (further details, see the electronic supplementary material). In words, this is simply:

\[
\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}}.
\]

There are two stages in our statistical inference; parameter inference and model selection. Both use an application of Bayes’ theorem. The first stage infers the posterior distributions about each model parameter, which is defined as

\[
P(w|M) = \frac{P(w|M)P(M)}{P(d|M)}.
\]

Here, \(M\) is a specific diffusion model, \(w\) is a model parameter and \(d\) represents SPT data, and a phrase \(P(A|B)\) means ‘the probability of \(A\) occurring given that \(B\) has occurred’. This stage is independent of other models, but is conditioned on one single model, \(M\). Both the posterior and likelihood are conditioned upon the data \(d\). We can now explain the three names of the terms above:

- The likelihood, \(P(d|w,M)\): the probability distribution of the data for a given parameter.
- The prior, \(P(w|M)\): the initial distribution prior to any conditioning by the data. Priors embody our initial estimate of the system, such as distribution of the parameters or the expected order of magnitude.
- The posterior, \(P(w|d,M)\): the distribution of the parameter following the conditioning by the data.

The second stage in our statistical inference is model selection. This invokes another application of Bayes’ theorem:

\[
P(M|d) = \frac{P(d|M)P(M)}{P(d)}.
\]

\(P(M|d)\) is a number which is the model posterior, or probability. \(P(M)\) is a number which is the model prior, \(P(d|M)\) is a number which is the model likelihood and \(P(d)\) is a number which is a normalizing factor which accounts for all possible models. This now generates the posterior (i.e. probability) for a specific model.

Linking the two stages in our statistical inference is the term \(P(d|M)\), the model likelihood. This is also the normalization term in the first stage. As model priors are usually flat (i.e. all models are expected equally), \(P(d|M)\) is often referred to as the ‘evidence’, a portable unitless quantity. In the general case, comparing the \(P(d|M)\) values for each independent model allows us to rank and select models (see the electronic supplementary material).

(iii) Diffusion models

As proof of principle, we used four standard diffusion models that are typical of observed molecular scale motion in living cells (see the electronic supplementary material, tables S1 and S2). These were Brownian, anomalous, confined and directed diffusion, and we used the underlying propagators associated with each diffusion model directly (full details in the electronic supplementary material).

(iv) Inference in BARD

The inference scheme was split into two forms. One uses the likelihoods based on the MSD distribution of each track, which we call the MSD method. The second uses the probability distribution functions directly on the individual frame-by-frame spatial displacements measured for each track, which we call the PDF method. These form the likelihoods, \(P(w|d,M)\) (see the electronic supplementary material).

As discussed in §3, the PDF method performed more accurately in many applications for comparing just two different non-confined diffusion models, such as anomalous diffusion with Brownian, but could not be applied to cases of confined diffusion, in which circumstance the MSD method was applied. Both approaches result in an estimate for the preliminary likelihood associated with each given single particle track.

The prior distribution for the diffusion coefficients \(D\) for Brownian diffusion, and the equivalent transport coefficient \(K_a\) for anomalous diffusion, were modelled as Gamma distributions (see the electronic supplementary material, and see MacKay [51] for a discussion of using a Gamma distribution). The prior distributions for the effective characteristic confinement radius \(R\) for the confined diffusion model, and for the mean drift speed \(v\) for the directed diffusion model, were both approximated as exponential distributions with expected sizes in the range of values that had been measured from several earlier studies in other biological systems (see the electronic supplementary material, table S3). The \(\alpha\) factor in the anomalous diffusion model was assumed to be uniform (i.e. flat) in the range 0.5–1.0 without further modelling. We have no a priori expectations to indicate how this factor would be distributed. The literature at present suggests multiple models of subdiffusion, so the sensible consensus prior in light of this would be flat. However, an extension would be to discriminate between these different families. Either way, our uniform assumption can account for the experimental observations of anomalous diffusion with an anomalous coefficient of approximately 0.7–0.8.

(v) BARD implementation

To implement our BARD algorithm, the following steps were taken:

1. Quantify all of the microscopic diffusion coefficients, \(D_m\), from the SPT data (shown here for simulated Brownian diffusion tracks in figure 3e). Here, \(D_m\) gives a measure of the short time-scale rate of diffusion and is estimated from a linear fit of the MSD data of each individual track using the first four data points (full details in the electronic supplementary material).

2. Fit a Gamma distribution to the distribution of all \(D_m\) (figure 3f) and use this fit to generate the two characteristic shape parameters of this function. Then, use these shape parameters to generate the diffusion coefficient prior (see equation S9 for the MSD method, see electronic supplementary material).

3. Calculate the other parameter priors for \(a\), \(R\) and \(v\) for the anomalous, confined and directed diffusion models (electronic supplementary material, table S3).

4. For each separate single particle track, we then calculated the likelihood (either using equation S8 for the PDF method, or equation S9 for the MSD method, see electronic supplementary material).

5. We then estimated the unnormalized posterior for each single particle track against each diffusion model, taken for the pure Brownian diffusion model as:

\[
\text{Posterior} = \text{likelihood (Brownian propagator)} \times a \times \text{prior(}\alpha\text{)}.\]

For the anomalous diffusion model as:

\[
\text{Posterior} = \text{likelihood (anomalous propagator)} \times a \times \text{prior(}\alpha\text{)}.\]

For the confined diffusion model as:

\[
\text{Posterior} = \text{likelihood (confined propagator)} \times a \times \text{prior(}\alpha\text{)}.\]

And for the directed diffusion model as:

\[
\text{Posterior} = \text{likelihood (directed propagator)} \times a \times \text{prior(}\alpha\text{)}.\]
Implementing BARD. (a) Simulated Brownian tracks (grey) all with diffusion coefficient \( D = 0.01 \mu \text{m}^2 \text{s}^{-1} \), with one of these tracks highlighted (black) for BARD analysis. (b) Distribution of measured microscopic diffusion coefficient \( D_m \) values from tracks in (a) with Gamma fit indicated (dashed line). (c) Constructing probability distributions used in BARD for the highlighted track of (a) tested against a Brownian diffusion model showing the unnormalized (i) prior, (ii) likelihood and (iii) posterior. Testing against the three other diffusion model generates two-dimensional unnormalized posterior distributions for (d) anomalous, (e) confined and (f) directed diffusion models. For the highlighted track shown in (a), the highest relative ranking probability was measured at 65\% for the Brownian mode, thus correctly identified with a probability that was more than twice as much as the next ranked mode of anomalous diffusion.

An example of the unnormalized Brownian model posterior distribution for a typical simulated track is shown in figure 3c(iii). The posteriors for the other diffusion models are shown for the same example track in figure 3d–f.

(6) Normalize the parameter posterior distributions from stage 4 (details in the electronic supplementary material, equation S4), calculating the evidence term. This is the final step in the parameter inference section, which bridges to the second inference stage (i.e. model selection).

(7) Model selection: calculate the model posterior. Rank the models on the basis of the size of the model posterior (a numeric probability). This final step then yields a probability estimate for a given model, relative to all the other models investigated: \( P(M|d) \). For example, for the sample track shown in figure 3a, which was simulated using a pure Brownian diffusion propagator function, the inference ranking probabilities that were generated from the four candidate diffusion models of anomalous, Brownian, directed and confined are 33.1 per cent, 65.6 per cent, 1.1 per cent and 0.2 per cent, respectively, and so in this instance Brownian diffusion is the favoured model. This is not to say that the absolute probability that the Brownian diffusion model is the correct one is approximately 66 per cent, but rather that it has the highest probability of being true from the set of candidate models investigated.

(8) For the top-ranked diffusion model for each single particle track, we then automatically locate the centroid of the posterior, to indicate the specific value of the transport parameter for that particular diffusion model. This is done using a Gaussian fit about the posterior peak.

(9) Repeat this process for all single particle tracks in the dataset.

(vi) Modelling mobility changes due to switches in diffusion coefficient

In order to demonstrate that the framework presented here can be extended to even more complicated cases of heterogeneous diffusion environments, we simulated a change in lateral diffusion coefficient as might be experienced by a single molecular complex undergoing transitions to multiple kinetic states. This may occur in signalling systems with transitions between ligand-bound and unbound states, or be due to a change in lateral mobility attributable to interactions with the underlying membrane such as local changes in viscosity [16] or interactions with the membrane cytoskeleton.

3. Results

(a) TIRF microscopy on live bacterial cells

Bespoke video-rate TIRF microscopy at 40 ms per frame (figure 4r) was performed on GFP-labelled Helix1021 and YFP-labelled TatA membrane protein complexes, resulting in the appearance of multiple distinct diffusing fluorescent
‘spots’ in each cell that could be tracked automatically from frame to frame. These spots were typically approximately 300–400 nm in width. This was larger than we measured for the point spread function width from single FP molecules immobilized to the surface of the coverslip by approximately 100 nm [7]. The measured point spread function width of single FP molecules of approximately 200–300 nm is equivalent to the optical resolution limit of our microscope and is an inevitable feature due to diffraction of emitted fluorescence when the detector, in our case an EMCCD camera, is physically more than a few wavelengths distance away.

The TatA system had been characterized previously using epifluorescence microscopy that indicated multiple spots per cell (mean of approx. 15) with a range of fluorescence intensities, diffusing over the cytoplasmic membrane surface [11]. Our aim in the present study was to use TIRF illumination to improve the imaging contrast sufficiently to generate single particle trajectories in the TIRF evanescent field in the specimen focal plane, corresponding to localization of either the Helix1021 or TatA in the cytoplasmic membrane. This would then permit analysis of the transport properties of these proteins at the single molecule/single molecular complex level for a relatively simple membrane protein probe at one extreme and for a complex heterogeneous membrane protein molecular complex at the other, both in functional, living cells.

Using automated SPT [11], we were able to track individual fluorescent spots to a super-resolution precision of approximately 40 nm or less. Experimental single particle tracks were collated and MSD values estimated (full details in the electronic supplementary materials). The longest duration tracks lasted typically approximately 1 s, but in most cases the tracks were shorter, with approximately 10 data points per track being more typical.

For the TatA–YFP data, cells contained typically approximately two to three fluorescent spots in TIRF images (figure 4b), suggesting approximately 12–18 spots per cell since the TIRF evanescent field of our microscope we estimate encapsulates roughly one-sixth of the E. coli cell membrane. Most MSD traces indicated putative evidence for Brownian diffusion, with typical values of $D \approx 0.01 \mu m^2 s^{-1}$. This was consistent with the earlier investigation, but with qualitative evidence from some putatively asymptotic MSD traces for a smaller subpopulation of relatively immobile spots, as had been reported in the previous study but not robustly quantified [11]. Prior, likelihood and posterior distributions were estimated for each single particle track.

The Helix1021–GFP cells contained typically approximately four to six fluorescent spots per TIRF image (figure 4c), suggesting more like approximately 30 spots in total in the whole cell membrane. MSD data again indicated putative evidence for two populations in terms of diffusive modes, one of Brownian diffusion with typical values of diffusion coefficient higher by a factor of approximately 5 to 10 than the TatA–YFP data, and the other mode again qualitatively suggesting confined diffusion.

(b) Model ranking and parameter estimation

To validate our approach, we tested the inference method using realistic simulated two-dimensional SPT input data using mobility parameters with characteristic values comparable to those estimated qualitatively for the experimental Helix1021–GFP and TatA–YFP data from the MSD plots. We then analysed both the parameter estimates and model rankings outputs. The correctness of the model ranking was assessed by classification matrices. A classification matrix represents different ‘input’ simulated diffusion models down the rows, $i$, while the different ‘output’ diffusion models from the ranking inference are represented across the columns, $j$, and then each location in the matrix is given an associated number for the percentage of tracks that are included in that particular ($i,j$) class combination.

In figure 5, we show the results of two example classification matrices: one corresponding to likelihood estimation using the MSD method in figure 5a, the other to the PDF method in figure 5b. Previous experiments on other biological systems that involve examples of directed diffusion, for example with putative protein treadmilling studies in vivo [29], suggested different values for characteristic diffusion coefficients, and so directed diffusion model ranking was done separately (see the electronic supplementary material). In this case, we investigated a range of different drift speeds from 1 to 20 nm s$^{-1}$. This indicated that at typical drift speeds used, the directed diffusion model can be correctly identified against a Brownian model with a relative probability of approximately 60–70%.

Using the MSD method, confined and Brownian diffusion could clearly be identified correctly with greater than 50 per cent relative accuracy, though the true identification of anomalous behaviour was poorer (approx. 30%), probably because subtle subdiffusive behaviour is not apparent for such typically short track lengths of only 10–20 data points as used here. The inference output for confined diffusion in particular was unsurprisingly found to be a function of track length, with the 50 per cent threshold of correct inference for tracks being composed of at least approximately 16 data points (see the electronic supplementary material, figure S2), though the change in correct relative inference probability for confined diffusion was found to be only a few per cent.
when the confinement radius was varied across a relatively large range of 50–200 nm in estimating the posterior distribution. This is not to say that the choice in prior function has little effect on the final outcome; if we use a naïve ‘flat’ prior function for the confined diffusion model (in effect, taking an infinitely large value for the confinement radius) then we estimate that the correct relative inference probability is over 20 per cent lower compared against the non-flat priors used. In other words, using physically sensible prior functions makes a substantial difference to correctly inferring the underlying type of diffusion (see the electronic supplementary material).

The PDF method is an approach that uses information from the relative displacements of a tracked molecule or complex from frame to frame, and so cannot be applied to a confined diffusion model without a priori knowledge of the absolute position of the diffusing particle relative to the boundaries of the putative confinement zone, which in general is not the case. Therefore, for the PDF method we display in figure 5b the relevant classification matrix between just anomalous and Brownian diffusion models. Here, anomalous diffusion was correctly discriminated with an accuracy of at least 62 per cent, performing better than the MSD method for corresponding diffusion models (e.g. Brownian diffusion was correctly identified with a relative accuracy of 95% using the PDF method compared with 52% for the MSD method).

(c) Identifying switches in molecular mobility

Tracks were simulated to mimic a sudden change in effective lateral diffusion coefficient, which have been observed previously in biological systems where multiple diffusion states exist [52]. Such multiple diffusion states might, for example, be due to either a dramatic change in lipid viscosity for the micro- or nanoscale environment in which a protein molecule or complex is diffusing, or conversely through a rapid oligomerization or molecular assembly process of the diffusing complex. In this simple generalization, we assumed that the time scale of the transitional step between different lateral diffusion coefficients was much less than the sampling period. For simplicity, we assumed that diffusing particles make this mobility switch at the halfway point of their full simulated trajectory. At this point, particles were assumed to switch to a higher diffusion coefficient (from 0.01 μm² s⁻¹ to either 0.05 or 0.10 μm² s⁻¹), assuming true Brownian diffusion in each case and a video-rate sampling time interval of 40 ms for which the number of data points in each half of a trajectory is n = 10.

A switching inference model was formulated by separating the displacement data at each time point and allowing for two separate mobility measurements to be inferred either side of this. Figure 6 illustrates the typical simulated individual and time-averaged MSD outputs with model ranking predictions. This relatively simple switching modification can correctly predict switching behaviour characterized by two separate microscopic diffusion coefficients over a simple Brownian diffusion mode characterized by just a single microscopic diffusion coefficient, with a relative ranking probability in the range 65–85%, depending upon the size of the switch in diffusion coefficient. Using the PDF approach under the same conditions generated a slight improvement to correct identification, and in doing so we found that the correct switching model was identified in preference to simple Brownian motion (i.e. a ranking probability in excess of 50%) down to as small a change as approximately threefold in the microscopic diffusion coefficient.

(d) Application of BARD to live-cell experimental data

Preliminary inspection of the MSD traces generated from automated SPT from both the Helix1021–GFP and the TatA–YFP E. coli cell strains suggested a predominantly mobile population with roughly linear MSD versus time-interval traces, in addition to a relatively immobile population characterized by putatively asymptotic MSD versus time-interval traces, which could be indicative of two possible populations corresponding predominantly to Brownian diffusion and confined diffusion. In the first instance, we ran a BARD analysis using all four standard diffusion models of anomalous, Brownian, confined and directed diffusion, which clearly indicated for both cell strains that Brownian and confined were the two most inferred diffusion models. We
then pooled the combined inferred results from anomalous, Brownian and directed diffusion as constituting ‘mobile’ tracks, and compared this with the inferred confined track data on MSD versus time-interval plots.

Simulated realistic track data using our standard set of mobility parameters (see the electronic supplementary material, table S3) indicated that a mixture of such mobile and confined tracks could be successfully discriminated, with both the imposed values for microscopic diffusion coefficient and confinement radius agreeing with those inferred from the BARD analysis to within the measurement error (figure 7a).

Applying BARD analysis to the Helix1021–GFP track data indicated that 50–60% of all tracks exhibited confined diffusion with an estimated confinement radius of 110 ± 50 nm (± s.d.), with the mobile population characterized by a microscopic diffusion coefficient typically in the range 0.01–0.05 μm² s⁻¹ (figure 7b). BARD analysis applied to the TatA–YFP track data indicated a smaller but still significant proportion of 30–40% of all tracks exhibiting confined diffusion with a mean confinement radius of 60 ± 40 nm, and the mobile population characterized by a smaller typical microscopic diffusion coefficient in the range 0.002–0.01 μm² s⁻¹ (figure 7c).

### Figure 6. Inferring a sudden change in diffusion coefficient. (a) MSD relation for a single track simulated assuming Brownian diffusion (black) and the mean average of 20 such tracks (blue) for which the diffusion coefficient $D$ switches from 0.01 to 0.10 μm² s⁻¹, with (b) associated ranking inference probabilities for a single $D$ Brownian (B) and two $D$ switching (S) model. (c) MSD relation for a single track (black) and average of several tracks (blue) for which $D$ switches from 0.01 to 0.05 μm² s⁻¹, with (d) associated ranking inference predictions.

### Figure 7. Comparing mobile and immobile tracks. (a–c) MSD versus time interval traces showing mean values assuming at least $n = 3$ data points at each time-interval value, for all ‘mobile’ tracks (blue, solid line) and confined tracks (red, solid line), s.e.m. error bounds shown (dotted lines), individual tracks shown in grey, for (a) simulated tracks ($n = 50$ mobile tracks here simulated using a Brownian diffusion propagator function, $n = 50$ confined tracks), (b) Helix1021–GFP ($n = 20$ mobile tracks, $n = 27$ confined tracks) and (c) TatA–YFP ($n = 258$ mobile tracks, $n = 164$ confined tracks). (d–f) Corresponding unbiased kernel density estimations for the distribution of predicted confinement radius from the inferred confined tracks.
For both the Helix1021–GFP data (A. Robson, A. Nenninger, C. Mullineaux, M. C. Leake, unpublished data) and the TatA–YFP data (figure 8), we were able to estimate the molecular stoichiometry of the diffusing fluorescent spots using a Fourier spectral technique that used the step-wise photobleaching of FPs [7]. This indicated a difference between the mobile and confined track populations suggesting that confined tracks were associated with a greater typical number of FP subunits. For example, we estimated the median stoichiometry from the mobile TatA–YFP spot population as being in the range approximately 20 TatA–YFP molecules per spot, whereas that of the confined population was higher by approximately 50 per cent (figure 8a). We saw no obvious differences in microscopic diffusion coefficient between the confined and mobile populations (figure 8b) nor of any clear correlation between molecular stoichiometry in each fluorescent spot and the inferred size of the confinement radius (figure 8c).

4. Discussion

The ability to monitor single molecules or complexes diffusing in living cells is an excellent example of the ‘next-generation’ single-molecule cellular biophysics approaches that have emerged over the past decade. What some researchers are now trying to do with such exceptionally precise molecular-level data is to use them to increase our understanding of the functional architecture of both the diffusing molecules themselves and of their local cellular environment. However, to do so requires a development of novel computational methods that can accurately measure the underlying modes of diffusion from the typically noisy and limited data from these tracked molecules in vivo.

In this study, we describe a novel analytical method to discriminate different modes of diffusion, applicable to data obtained from SPT of fluorescently labelled proteins in the cell membrane. Although the biocomputational algorithm in itself is complex it should find broad application for researchers in the cell biology field. We report our approach based on both modelling and stochastic simulation of multiple biologically relevant diffusive modes experienced by proteins in different underlying micro- and nanoscale environments. Priors are formulated from both simulation and experimental work. We demonstrate how the use of the correct propagator functions can permit discrimination between Brownian, directed, confined and anomalous diffusion, even for relatively sparse data tracks. When comparing two diffusive modes in a pair-wise fashion our results indicate that model ranking predicts the correct diffusive mode for a single video-rate sampled track as short as approximately 0.4 s in duration. Furthermore, the model can be extended to permit discrimination for a diffusion model involving sudden switching of the diffusion coefficient during a particle’s trajectory.

Two approaches were investigated, one using the MSD and the other using the PDF method. In each case, a prior formulation was used to describe the expected distribution of the parameters. Although neither approach could effectively discriminate between anomalous and confined modes of motion, which from the MSD curves have qualitatively similar shapes for noisy short tracks, we find that the PDF and MSD methods in tandem have different resolving power. The MSD method effectively identifies confined from simple Brownian motion, whereas the PDF method effectively identifies anomalous diffusion from Brownian diffusion. In addition, the PDF approach has a strong resolving power in that it can identify dynamics within a single track, as observed in the simulations of diffusion coefficient switching.

The PDF method does not take into account the full effect of experimental noise, as distinct from random fluctuations due to the stochastic nature of the diffusion processes. Levels of experimental noise are likely to vary between different experimental equipment and need to be properly characterized for each individual case. However, this was qualitatively incorporated into the MSD approach, where Gaussian errors are assumed. Experimental noise arising in tracking would be included in the error, and would add by quadrature to the expected fluctuations due to experimental noise, as distinct from random fluctuations due to stochastic noise (the time-interval zero point in our case assumes an MSD error of around 40 nm).

Our in vivo video-rate particle trajectories contain approximately 10–30 times fewer data points than those used in previous studies using tracking of gold particles [4,26,27,53], organic dye labelling of clusters containing hundreds of molecules [39] or quantum-dot tracking [41], and are of comparable duration to those obtained previously using single-molecule fluorescence microscopy either in artificial lipid layers or in vivo [11,33,40]. These have implemented a variety of different methodologies to analyse single particle trajectories involving either regression fitting of the MSD versus time-interval relation, application of a relative deviation parameter or constructed probability distributions representative of the modes of interest.

The propagator functions in effect model the likelihood of an observed track. What our study includes is how the
distribution of the parameters that formulate these models can be used to aid in discrimination of the diffusion processes. In fluorescence microscopy, with typically very short tracks observed, it is generally infeasible to analyse such trajectories without some form of population averaging using conventional techniques. Exceptions are made of course to the occasional long track which is observed, or tracks which appear representative, but a majority of the body of data captured is noisy, and unrepresentative if multiple modes of behaviour are under investigation.

Our study was aimed at being able to discriminate, without population averaging, such molecular-level tracks. Once individual particle trajectories are categorized into different modes of diffusive behaviour, models can be built on how they behave collectively, potentially allowing greater physiological interpretation of the protein mobility characteristics in functional, living cells, and hence to have a greater understanding on their underlying membrane micro- and nanoscale structure in a biologically relevant context. We have performed a validation across the approximately biological relevant parameters for the datasets presented. However, extrapolating these to any real system should come with the caveat that the classifications can only really be used as a guide for the particular set of algorithm parameters and system parameters used.

Ultimately, because the inference scheme is probabilistic, there will inevitably be some trajectories that are falsely categorized with the wrong behaviour, most often into simple Brownian motion, as shown in the classification matrices. We included details on how model ranking varies with respect to the number of data points to demonstrate that there will often be a crossover between mis-categorization and the correct identification. A caveat then, for interpreting any model selection on the experimental data would be that there is no evidence of heterogeneity under the given experimental conditions. If this crossover is unreachable in the experimental framework it will at least inform the experimentalist on the typical minimum duration of track that needs to be detected to permit reliable discrimination (perhaps thereby directing them to change the characteristics of the optical setup and/or the related biological and physico-chemical conditions, such as the type of fluorophore used and whether the application of anti-bleaching reagents is required).

In an earlier SPT study on the Tat system using non-TIRF illumination, the presence of an ‘immobile’ population of TatA protein complexes was reported, but not investigated further [11]. In our study, BARD analysis reveals that a significant proportion of TatA–YFP complex tracks have a confinement radius of 60 ± 40 nm. The measured localization precision on our microscope for tracking a single YFP molecule is approximately 40 nm. However, TatA complexes were observed to have a broad range of stoichiometry, with a median value equivalent to approximately 20–30 TatA–YFP subunits, consistent with that reported previously [11]. These complexes are therefore brighter than a single YFP molecule by a factor of approximately 20–30, with the localization precision following iterative Gaussian fitting of the intensity profile of these fluorescent spots scaling approximately by the square-root of this factor, or approximately 5 [54], so the localization tracking precision for most TatA–YFP complexes is more like 5–10 nm. Therefore, the estimated confinement radius here is substantially higher than the localization precision for diffusing complexes, which strongly suggests that the majority of the ‘immobile’ TatA complexes previously reported were in fact exhibiting true confined diffusion.

Similarly, we observed a significant subpopulation of tracks for the Helix1021–GFP strain which exhibited confined diffusion, here with a mean confinement radius of 110 ± 50 nm, within experimental error of that measured for the TatA–YFP strain. The fact that the transmembrane helix probe has no known specific interaction with molecular systems in E. coli suggests that the confinement domains in both cell strains may represent an intrinsic feature of the cell membrane itself. Similar size putative confinement domains were observed previously in SPT studies of an unrelated bacterial oxidative phosphorylation (OXPHOS) membrane protein [28]. This behaviour had been previously attributed to a possible ‘respirazone’ effect [49] in which different OXPHOS enzymes were pooled together into the same confinement domains to improve electron transfer efficiency throughout the OXPHOS system. However, our work here may point to a more generic confining feature of the cell membrane.

The diffusion models illustrated here are not exclusive as such—there is a risk that none of the models is actually the physically ‘correct’ one. BARD analysis will provide probabilistic rankings of these models, but these probabilities can strictly only be interpreted in the context of the other models considered, and do not represent an absolute probability [55–57]. Model selection is open-ended; the models presented here do not take into account the full degree of potential heterogeneity that may exist in the cell membrane, and other models can be considered. For example, there are several theorized models of anomalous subdiffusion, each with a unique PDF. There may also be complex dynamic behaviour that has not been taken into account, such as hopping diffusion, reaction kinetics and molecular assembly effects. A natural extension of this BARD approach as we present it here is to incorporate more complex behaviour that may better capture the real, physiological behaviour of diffusion in living cells.

Separating different mobility characteristics into different categories will clearly facilitate insight into several important biological questions: for example, how proteins partition dynamically in the cell membrane, whether signalling events are linked to membrane architecture, the precise manner in which motor proteins shuttle in or near to cell membranes, and the extent to which interacting proteins rely upon random collisions or are putative confined ‘solid-state’ reaction zones. Such new diffusion analysis tools that we report here might indeed also be further extended to larger length-scale investigations beyond that of the single molecule and single cell, such as rheological or cell migration studies at the level of cellular populations in normal tissue development and tumour formation in cancer.

The authors thank Philip Maini and Marcus Tyndall for preliminary discussions concerning diffusion simulation; Nick Greene and Ben Berk for the donation of bacterial cell strain AyBC; and Anja Nenninger and Conrad Mullineaux for the donation of bacterial cell strain Helix1021-GFP. This work was supported via a research grant to M.C.L. (EP/G061009). M.C.L. was supported by a Royal Society University Research Fellowship. A.R. was supported by the Research Councils UK.
References


