

# PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

# Imaging and quantification of trans-membrane protein diffusion in living bacteria

## Perspective

Felix Oswald<sup>1,2,3</sup>, Ernst Bank<sup>3</sup>, Yves J.M. Bollen<sup>1,3</sup>, and Erwin J.G. Peterman<sup>1,2,\*</sup>

<sup>1</sup>LaserLaB Amsterdam, VU University, Amsterdam

<sup>2</sup>Department of Physics and Astronomy, VU University, Amsterdam

<sup>3</sup>Department of Molecular Cell Biology, VU University, Amsterdam

\* Corresponding author. De Boelelaan 1081, 1081 HV, Amsterdam. Email: e.j.g.peterman@vu.nl

## Abstract

The cytoplasmic membrane forms the barrier between any cell's interior and the outside world. It contains many proteins that enable essential processes such as the transmission of signals, the uptake of nutrients, and cell division. In case of prokaryotes, which do not contain intracellular membranes, the cytoplasmic membrane also contains proteins for respiration and protein folding. Mutual interactions and specific localization of these proteins depend on two-dimensional diffusion driven by thermal fluctuations. The experimental investigation of membrane-protein diffusion in bacteria is challenging due to their small size, only a few times larger than the resolution of an optical microscope. Here, we review fluorescence microscopy-based methods to study diffusion of membrane proteins in living bacteria. The main focus is on data-analysis tools to extract diffusion coefficients from single-particle tracking data obtained with single-molecule fluorescence microscopy. We introduce a novel approach, IPODD (inverse projection of displacement distributions), to obtain diffusion coefficients from the usually

obtained 2-D projected diffusion trajectories in the highly 3-D curved bacterial membrane. This method provides, in contrast to traditional mean-squared-displacement methods, correct diffusion coefficients and allows unravelling heterogeneously diffusing populations.

## Introduction

One of the universal features of life is the use of membranes to separate the cell's interior from its environment<sup>1,2</sup>. Membranes allow for high concentrations of macromolecules, building blocks and nutrients inside the cell and keep unwanted molecules, such as waste products and toxic molecules, outside.

Compartmentalization is not the only function of biological membranes: essential processes, such as respiration, signalling and uptake of nutrients, take place in membranes, at the interface of two compartments. Such processes are facilitated or catalysed by proteins that are embedded in the membrane.

Eukaryotic cells contain many specialized membranes that enclose small intracellular compartments with specific function, such as mitochondria where respiration takes place, lysosomes where macromolecules are degraded, the nucleus that contains the genome and the transcription machinery, and the endoplasmic reticulum where extra-cytoplasmic or membrane-embedded proteins are folded. In contrast, prokaryotic cells do not contain intracellular membranes and the protein machineries that catalyse those processes are all embedded in the cytoplasmic membrane (Figure 1). As a consequence, prokaryotic membranes are among the most crowded membranes known, with up to two-third of their mass consisting of proteins<sup>3</sup>. Membranes are fluidic; driven by thermal fluctuations the lipids and proteins forming the membrane move around in the plane of the membrane, in random fashion, resulting in diffusive motion. Diffusive motion of proteins in crowded bacterial cytoplasmic membranes is the focus of this perspective. We will start with a brief discussion of the biological importance of bacterial membrane protein diffusion. Next we will address the experimental approaches necessary to accurately measure and quantify diffusive motion of membrane proteins in bacteria. We will in particular focus on single-particle tracking (SPT) using single-molecule fluorescence microscopy (SMF). The major focus of this perspective will be on data-analysis tools that can be used to extract diffusion constants from experimental data. We will discuss approaches that have been applied to date and present a novel method that is particularly useful to quantify diffusion in the highly curved

membrane of bacteria and allows for direct determination of heterogeneity in diffusion behaviour.

### The biological relevance of membrane-protein diffusion

Although some bacterial membrane proteins perform their function on their own, many processes require the formation of complexes of multiple membrane-bound proteins. For example, the insertion of proteins in the membrane<sup>4</sup>, the translocation of proteins across the membrane<sup>5</sup> or signal transduction by two-component systems<sup>6</sup> requires transient formation of membrane-bound complexes involving multiple proteins. Recently, it was shown that the proteins responsible for oxidative phosphorylation in bacteria form a number of distinct mobile complexes, around 50 nm in size. Ubiquinone diffuses rapidly between these complexes and shuttles electrons from one complex to another<sup>7</sup>. This is markedly different than in mitochondria, where all components of the system are assembled in large supercomplexes. Complex formation between proteins in the membrane critically depends on the temperature-driven diffusive motion of the proteins in the plane of the membrane, bringing partner proteins together.

For some proteins, complex formation is not sufficient to fulfil their function; they also need to be present at a specific location in the cell. A number of different mechanisms has been described<sup>8</sup>. Some proteins localize to the poles of rod-shaped bacteria by binding to remnants of the division machinery. Also many over-expressed membrane proteins accumulate in the poles, via mechanisms that are not fully understood. Interestingly, however, there is increasing evidence that many proteins are recruited to the correct location by a diffusion-and-capture mechanism<sup>9, 10</sup>, in which the proteins diffuse in the plane of the membrane, until they reach their target location, where they are captured by other membrane proteins that have already been targeted to this site. This mechanism requires, however, accurate localization of certain proteins, which provides position cues to other proteins. One example of the hierarchical assembly of large protein complexes using diffusion-and-capture is the cell-division machinery that needs to localize at the centre of the bacterial cell to form the so-called z-ring. After formation, the z-

ring contracts, dividing the mother cell into two daughters. How does the first component of the z-ring, the tubulin-like protein FtsZ, localize correctly? A central role in how most gram-negative bacteria, like *Escherichia coli*, solve this problem is played by the protein MinC. MinC prevents the polymerization of FtsZ, required for the formation of the z-ring. Consequently, the z-ring can only be formed in the absence of MinC. The localization of MinC on the membrane is controlled by the proteins MinD and MinE, which shuttle, driven by diffusive motion, between the two poles in an oscillating manner. In this way they leave the membrane in the centre of the bacterial cell free of MinC, only allowing FtsZ to polymerize there and form a z-ring<sup>11</sup>.

Another cue that can be used by proteins to find a certain location in the cell is the geometry of the cell. Some bacterial proteins specifically accumulate in regions of the membrane with a particular curvature. For example, *Bacillus subtilis* DivIVA has a strong preference for concave, or negatively curved, membranes, and hence localizes to the poles of the cells, where it acts as a beacon for other proteins that need to localize there<sup>12</sup>. In contrast, SpoVM (also *Bacillus subtilis*) prefers convex, or positively curved membranes and localizes to the inward-facing side of a nascent spore, where it attracts other components of the sporulation machinery<sup>13</sup>.

These are only a couple of examples of processes taking place in bacteria, demonstrating the importance of membrane-protein diffusion driving complex formation or protein localization. The diffusion of membrane proteins in bacteria is, however, poorly understood. For example, it is not known whether bacterial membranes are homogeneous or contain micro-domains with distinct physical properties. It is also unclear how the very high concentration of proteins in the bacterial membrane exactly affects diffusion. It has, however, been reported that diffusion constants of membrane proteins are in general significantly smaller in the bacterial cytoplasmic membrane than in reconstituted lipid bilayers with a far lower amount of proteins embedded<sup>14-16</sup>. A key reason for this lack of understanding is the small size of bacteria (about 1  $\mu\text{m}$ ), only a few times larger than the resolution of optical microscopes (250 nm), which makes *in vivo* measurements of diffusive processes in bacteria substantially more complicated than in much larger eukaryotic

cells. In the next paragraph, we will discuss the microscopy methods that can be applied to unravel membrane-protein diffusion in bacteria.

### Microscopy methods for diffusion characterization

Fluorescence microscopy has become the most popular method for the characterization of membrane-protein diffusion in living cells<sup>2</sup>. Key advantages of fluorescence microscopy are its relatively non-invasiveness and compatibility with living cells. Crucial for fluorescence microscopy is that the protein of interest is fluorescent. Since most membrane proteins are not fluorescent by themselves, labels need to be added. The most widely used approach, in particular for studies of living bacteria, is to make genetic fusions of the gene coding for the protein of interest with a gene coding for a fluorescent protein, such as EGFP, Venus or YFP<sup>17, 18</sup>. This can be done on a plasmid, in the presence or absence of the wild-type membrane protein, using a different promoter (that can for example be regulated more precisely to obtain the very low expression levels required for single-molecule imaging). Another approach is to modify the chromosomal gene, which allows using the natural promoter and flanking sequences. This latter approach has the advantage that (close to) natural expression levels are obtained. In cases of medium to high copy numbers of fluorescently tagged proteins, however, the resulting protein density can be too high for single-molecule detection due to overlapping point-spread functions. In addition, this approach can be critical for essential proteins, since their function might be altered by the fluorescent label (fluorescent proteins (typically 27 kD)<sup>18</sup> are often larger than a typical bacterial membrane protein). Care has to be taken to check whether aspects like localization, oligomerisation and biological function are unaltered by fluorescent labelling; this cannot be taken for granted. We typically check localization of bacterial proteins using membrane fractioning and western blotting and confocal fluorescence microscopy. To test proper biological function, assays appropriate for each protein of interest need to be applied. An additional issue with the use of fluorescent proteins is that they take a while (tens of minutes) to mature, i.e. to properly fold and form a fluorescent chromophore. For many experiments this is not a critical

issue, but for some it is. Special variants of fluorescent proteins (such as the yellow-fluorescent variant Venus) have been developed that are optimized for fast maturation, but this still amounts to many minutes<sup>17, 18</sup>.

Once bacteria are available with the membrane protein of interest fluorescently tagged, they can be subjected to fluorescence microscopy. Three distinct methods have been developed in particular to characterize the diffusion of membrane proteins in living bacteria<sup>2</sup>: fluorescence recovery after photobleaching (FRAP), fluorescence correlation spectroscopy (FCS) and wide-field single-molecule fluorescence microscopy (SMF).

In FRAP, a fluorescence microscope, confocal or wide field, is equipped with a unit that allows instantaneous photobleaching of part of the imaged sample, with an intense burst of excitation light<sup>16, 19, 20</sup>. After photobleaching, the sample is imaged and the recovery of the fluorescence intensity inside the photobleached area is measured as a function of time. In case the fluorescently labelled proteins in the sample are static, the fluorescence signal will not recover. In case of diffusing proteins, the photobleached fluorophores will be replaced by intact ones, resulting in recovery of the fluorescence signal. From the fluorescence intensity time trace, generally two parameters are obtained: (i) the recovery time constant, which depends on the size of the photobleached area and the diffusion coefficient of the labelled proteins and (ii) the relative amplitude of the recovery, which provides insight in the fraction of proteins that are immobile. This approach has been successfully applied to determine diffusion coefficients of membrane proteins in eukaryotic and prokaryotic cells<sup>19</sup>. Key advantage is that no single-molecule sensitivity is required, which puts less stringent demands on the microscope equipment and number of fluorophores in the sample. A disadvantage is that it is a bulk technique, which might average out the heterogeneous behaviour of (sub)classes of proteins. Another disadvantage is that FRAP is quite difficult to apply to small samples, but FRAP has been applied successfully to bacteria<sup>21</sup>. A smart solution to this issue has been devised by Mullineaux and co-workers: they applied cephalaxin in order to inhibit cell division, resulting in elongated bacteria (many  $\mu\text{m}$  long)<sup>22</sup>.

In FCS<sup>23, 24</sup>, a confocal microscope is used. The confocal volume is positioned on top of the sample of interest and the fluorescence intensity is measured as a function of time, with (sub)microsecond time resolution. In case the labelled molecules in the sample are free to diffuse, they will move in and out of the confocal volume, resulting in a fluctuating number of molecules and hence fluctuating fluorescence intensity. The time scale of the fluctuations can be obtained by autocorrelating the fluorescence intensity. From a fit to the decay of the autocorrelation the diffusion coefficient of the moving particles can be measured accurately, if the exact dimensions of the confocal volume are well known. Applying FCS to study diffusion in bacterial membranes is possible, but not straightforward due to the small size of bacteria. Crucial is to align the confocal spot carefully on the bacterium. In particular, care needs to be taken in the reproducible axial focusing of the confocal spot on the bacterial membrane, since insufficient control results in variations in (axial) size of the intersection of confocal volume and membrane, resulting in ill determination of diffusion coefficients.

Over the last decade, following the pioneering work of the Xie and Moerner groups<sup>10, 25, 26</sup>, the approach of choice has become to track single, fluorescently labelled proteins, using SMF, while they are moving through living cells. In this approach, use is made of hypersensitive fluorescence microscopes, in most case wide-field instruments that allow the recording of time sequences of images with single-fluorophore sensitivity<sup>17</sup>. These approaches have become feasible thanks to the availability of sensitive and fast cameras, based on back-illuminated, electron multiplying CCD or scientific CMOS detectors. In addition, the advent of relatively inexpensive continuous-wave, visible laser sources, for ordinary or total internal reflection wide-field SMF has been crucial. In ordinary epi-illuminated wide-field fluorescence microscopy a nearly parallel excitation beam is created in the sample, exciting a substantial area of the specimen, over the full thickness of the sample<sup>27</sup>. Advantage of this approach is its simplicity and its ability to provide a constant intensity all over the sample. A key disadvantage is that, although only a single plane in the object is imaged in focus, fluorescence signals arising from above or below of this plane will contribute to an out-of-focus back ground, which can overwhelm the

signal from the focal plane. A solution to this problem is to apply total internal reflection fluorescence microscopy (TIRF)<sup>28</sup>. In TIRF, excitation is provided by the evanescent wave generated by a laser beam that is totally internally reflected at the glass cover slip sample interface. Total internal reflection is obtained by using a highly inclined beam, hitting the surface at an angle larger than the critical angle. This highly inclined excitation geometry can be readily obtained by entering the laser beam off-axis in a very high numerical aperture microscope objective (1.45 or higher). TIRF solves the aforementioned background issues, but it is important to realize that it can be only applied close to the coverslip (typically up to a depth of  $\sim 100$  nm in the sample) and that the excitation intensity decreases exponentially with penetration depth in the sample. It should be noted that bacteria are in general thin enough (less than  $1 \mu\text{m}$ ), to not cause severe background issues, in contrast to mammalian cells. Still, autofluorescence of the cells can be a significant source of background, in particular when excitation wavelengths below about 500 nm are used<sup>29</sup>. Also in this case, TIRF can be used to increase the signal-to-noise ratio. TIRF, however, can cause complications in determining the stoichiometry of multimeric assemblies. One complication is that the excitation intensity is depth dependent, another that it requires coherent laser illumination, which can cause speckle and fringes leading to position dependent intensity variations. These latter problems are less prominent in conventional, laser, epi-illuminated wide-field fluorescence microscopy. In our and other's experience this approach can yield images of sufficient quality in case diffusing proteins, present at low copy numbers, are imaged in bacteria<sup>10, 30</sup>.

With SMF, image sequences are obtained that show spots due to individual molecules. For further processing it is necessary to first obtain, on- or offline, the trajectories of the individual molecules (i.e. spatial coordinates as a function of time)<sup>17</sup>. Several methods exist to robustly detect fluorescence spots and to determine the location of their centre with sub-diffraction accuracy. A widely used approach is to fit a 2D-Gaussian to the intensity profile of spots<sup>31</sup>. The Gaussian fit not only yields x- and y-coordinates of the intensity maximum, but also amplitude and width of the underlying intensity distribution. The accuracy with which single

fluorophores can be localized depends on the signal-to-noise ratio<sup>31</sup>. For these low-intensity signals, the most prominent source of noise is shot noise, the relative contribution of which decreases with the square root of the number of photons detected. Other sources of error include camera dark and read noise, pixilation noise, and background light in the sample<sup>32</sup>. Depending on integration time and excitation intensity, localization accuracies in the order of 30 nm can be readily achieved using fluorescent proteins as fluorophore<sup>17</sup>. After the locations of the maxima of all spots in all frames are determined, particle positions in subsequent image frames need to be linked to trajectories, typically using a nearest-neighbour algorithm. In case of high labelling densities, robust identification of trajectories is not trivial because of particle overlap and crossing. Linking algorithms can therefore only obtain the most probable trajectories, making use of the history of individual particles, including intensity<sup>33, 34</sup>.

It is important to note that obtaining optimal trajectories puts conflicting demands on the experiments. Best is to obtain the coordinates of individual proteins, very accurately, with very short time intervals, for a long duration of time. Accurate localization requires a high signal-to-noise ratio, which can be obtained by applying high excitation intensities to obtain a large number of photons. But high excitation intensities result in enhanced photobleaching, shortening the trajectories. An optimal combination of single image frame integration time and excitation intensity needs to be applied for optimal tracking of the individual proteins. Using stroboscopic illumination, the time scale probed can be made independent of the excitation time required for a single localization, providing access to longer time scales<sup>26</sup>.

#### Analysis of trajectories due to pure Brownian diffusion

In case of Brownian motion in 2D of a single homogeneous population of biomolecules, the motion is purely random, driven by thermally activated particle-particle collisions, the probability density of a biomolecule to move over distance  $\vec{r}$  in 2D within time  $t$  is a Gaussian function<sup>35</sup>:

$$P(\vec{r}, t) = \frac{1}{4\pi Dt} e^{-\frac{\vec{r}^2}{4Dt}} \quad (1).$$

The first moment (the average displacement) of this so-called Gaussian propagator vanishes, reflecting that a Brownian particle moves in any direction with equal probability, resulting in zero net displacement. The second moment (the variance or mean squared displacement (MSD)), however, is not zero and is determined by the diffusion constant as a measure of the position fluctuations of the particle:

$$\langle \vec{r}(t)^2 \rangle = 4Dt \quad (2).$$

The linear increase of MSD over time is traditionally and still widely used to estimate the diffusion coefficient (Figure 2)<sup>35</sup>. In later paragraphs we will discuss deviations from this ideal behaviour due to confinement, heterogeneity and other complications. In this section we will further focus on how MSD analysis can be applied soundly and what its limitations are.

Two important factors complicate MSD analysis<sup>36</sup>. First, the relative statistical weight of MSD values obtained at different time lags is not well defined. This is caused by correlations in the data, since in most cases the same trajectories are used to calculate MSD's at different time lags<sup>35</sup>. Discussing the effect of different sampling strategies using averages over independent or averages over all pairs of squared displacements, Saxton<sup>37</sup> found that in either case the correlations cause the fitted diffusion constant to depend in a non-trivial way on the number of time lags considered. In conclusion, Saxton demonstrated that the best estimate of the diffusion coefficient is obtained by only considering a small number of time lags, but taking into account and averaging all displacements, while trying to obtain the longest trajectories possible.

The second factor complicating MSD analysis is the experimental uncertainty in determining particle localizations. As discussed above the localisation error  $\sigma$  depends on signal intensity, pixel size and various noise sources. This static localisation error leads to an offset in the MSD as function of time<sup>38</sup>:

$$\langle \vec{r}(t)^2 \rangle = 4Dt + 4\sigma^2 \quad (3).$$

In case of fast diffusion compared to the camera exposure time, the localization error is increased by an additional factor due to motion blurring of the particles moving during the finite camera exposure time ( $T_{exp}$ ). This motion blur leads to an additional uncertainty<sup>39</sup>:

$$\langle \vec{r}(t)^2 \rangle = 4Dt + 4\sigma^2 - \frac{4}{3}DT_{exp} \quad (4).$$

A statistically thorough analysis of the optimal number of time lags to be considered in classical MSD-analysis to obtain the best estimate of the diffusion coefficients has been provided by Michalet<sup>40</sup>. He demonstrated that this is crucial: a non-optimal number of time lags were shown to result in substantial deviations of the diffusion coefficients obtained. Another approach, based on a maximum likelihood estimator (MLE) was developed by Berglund<sup>36</sup>. This approach relies on maximizing the probability distribution of correlated displacements by directly optimizing the parameters in the covariance matrix. The advantage of this approach is that, in contrast to Michalet's approach, no initial estimates of  $D$  and  $\sigma$  are required. Recently, an extension of the MLE approach was developed taking into account blinking of the fluorophores<sup>41</sup>. This demonstrates the general potential of the MLE method to be extended to more complex forms of diffusion.

#### Analysis of anomalous diffusion

In biological systems researchers frequently observe diffusion processes that substantially deviate from the pure Brownian motion described above (Figure 2). These forms of non-Brownian, anomalous diffusion are typically evident from non-linear scaling of MSD with time<sup>35, 42</sup>. Different behaviours can be discerned. In confined or corralled diffusion, proteins can diffuse freely inside a constrained space, formed by for example cytoskeletal structures, or lipid domains. In this form of constrained diffusion, the MSD starts linear as a function of time, but plateaus,

converging asymptotically to an upper limit (set by the size of the corrals/confined regions) at long time lags<sup>35</sup>. In case of super- or subdiffusive behaviour the MSD scales in a non-linear fashion with time, with exponent  $\alpha$ :

$$\langle \vec{r}(t)^2 \rangle = 4Dt^\alpha \quad (5).$$

Processes characterized by an exponent  $\alpha$  larger than 1 are superdiffusive, involving in many cases active cellular transport. Processes characterized by an exponent smaller than 1 are subdiffusive. Subdiffusion can be caused by transient immobilization or geometrical inhomogeneities<sup>35, 43</sup>. Models have been proposed that describe the non-classical behaviour for these two underlying mechanisms. Diffusion in the presence of passive traps or immobilized binding partners can be modelled with a continuous-time random walk (CTRW) exhibiting a broad distribution of waiting times between random steps<sup>44</sup>. Fractional Brownian motion (FBM)<sup>45</sup> and diffusion on percolation clusters<sup>46</sup> are models that describe the effect of molecular crowding or barrier-like structures. Even though there is general agreement that molecular crowding reduces lateral protein mobility<sup>42, 47, 48</sup>, there is substantial controversy whether crowding intrinsically causes subdiffusive behaviour<sup>48, 49</sup>. In contrast to experimental findings that suggest a general correlation between molecular crowding and subdiffusion<sup>49-52</sup>, there are several reports on simple Brownian motion in crowded environments that support the notion that anomalous diffusion is caused by other reasons than crowding<sup>48, 53, 54</sup>. In order to obtain insight in the underlying molecular mechanism of subdiffusive behaviour it is desirable to identify what model is most probable. This is, however, a difficult task, since it involves detailed analysis of the propagator and the ergodicity of the diffusion process<sup>43, 49, 55</sup>.

In many biological systems multiple modes of diffusion coexist, demanding classification of individual trajectories. Most of the analytical methods developed over the past decades to quantify such behaviour look for deviations between the experimental single-trajectory MSDs from classical behaviour<sup>42</sup>. The short

trajectories obtained with single-particle tracking of fluorescence proteins, however, yield MSDs that are not accurately defined and thus show considerable spread. Such methods thus require strict significance criteria, making them rather insensitive to subtle deviations from simple Brownian behaviour. Recently, Leake and co-workers introduced a novel method, Bayesian ranking of diffusion (BARD), which allows probability analysis of single-molecule trajectories in conjunction with robust ranking of diffusion models underlying the data<sup>56</sup>. In this approach, the model parameters are first estimated in a probabilistic manner, inferring single-molecule trajectories from an *a priori* assumed diffusion model. In a subsequent step the most probable diffusion model is then identified by Bayesian ranking using a prior calculated model likelihood. BARD has been shown to be able to identify mobility switches in individual trajectories, allowing characterization of heterogeneous behaviour in experimental diffusion data in a robust, probabilistic way<sup>56</sup>. Limitations of the approach are that it relies on *a priori* assumptions and its sensitivity is limited in case of short trajectories, like the ones typically obtained from single molecules.

#### Analysis of heterogeneous diffusion

As discussed in the introduction, many biological processes involve the transient formation of membrane-protein complexes, in which the proteins of interest undergo transitions between different states of mobility (for example switching from static to dynamic or from slower to faster diffusion). Another cause of heterogeneity could be inhomogeneities in the environment, for example with respect to membrane viscosity. In traditional MSD analysis, MSD values are averaged over all time lags and all particles, which can only yield an average diffusion coefficient. A better approach to unravel heterogeneity is to calculate diffusion coefficients for individual trajectories and consider the distribution of single-trajectory diffusion coefficients. This approach requires long trajectories of single-molecule motion<sup>35</sup>. Single fluorophores, however, typically yield short trajectories, resulting in broad distributions of diffusion constants, in which the width of the distribution is a combination of the heterogeneity and the error of

determining individual diffusion coefficients from limited data<sup>57</sup>. Therefore, global-analysis strategies using the whole displacement distribution of the entire population of trajectories can improve sensitivity and robustness<sup>58</sup>.

One approach is the determination of cumulative probability distributions (CPDs). Integrating the propagator for Brownian motion, Eq. (1), over the angular space results in a Rayleigh-distributed probability distribution function (PDF) of squared displacements (Figure 3)<sup>59</sup>:

$$PDF(r, t) = \frac{r}{2Dt} e^{-\frac{r^2}{4Dt}} \quad (6).$$

In most cases, not the PDF, but the related CPD is calculated, which is defined as the probability that a particle remains within a circle with radius  $r$  after time lag  $t$ <sup>60</sup>.

The CPD of a single species diffusing with diffusion coefficient  $D$ , plotted as a function of  $r^2$  is characterized by an exponential decay, which can be readily fitted (Figure 3):

$$CPD(r^2, t) = 1 - e^{-\frac{r^2}{4Dt}} \quad (7).$$

The CPD of a measurement of two diffusing populations with relative occurrence  $\gamma$  and diffusion coefficients  $D_1$  and  $D_2$  Eq. (6) is the sum of two exponentials<sup>60</sup>:

$$CPD(r^2, t) = 1 - \left[ \gamma \cdot e^{-\frac{r^2}{4D_1t}} + (1 - \gamma) \cdot e^{-\frac{r^2}{4D_2t}} \right] \quad (8).$$

By applying multi-exponential fitting, CPD analysis can robustly identify and quantify multiple diffusing components, as has been demonstrated *in vitro* as well as *in vivo*<sup>30, 60</sup>. Extracting transition rates between the distinct states, however, requires more elaborate analysis. Matsuoka and co-workers determined the autocorrelation function of MSD's for Brownian motion characterized by transition between two states with distinct diffusion coefficients<sup>61</sup>. Fitting this autocorrelation

function to experimental data allowed for discrimination between distinct models underlying the data and quantification of the transition kinetics between the two diffusive states.

An alternative approach to characterize heterogeneous diffusion makes use of hidden-Markov modelling (HMM). This approach has been used to retrieve diffusion coefficients and transition rates for a 2-state Brownian diffusion model<sup>57</sup>. It was demonstrated that the sensitivity of this approach is better than that of the classical MSD approach and that it can be applied successfully to experimental data. A comparison or incorporation with a CPD-based approach might have potential but has, however, not yet been performed. Recently, a variant of HMM, variational Bayesian treatment (vbSPT) has been introduced<sup>62</sup>. This approach has the advantage that no presumption of the number of states is required and that the approach is capable of selecting the model with the most probable number of states. The probably most general approach in unravelling complex diffusive behaviour is to compare experimental data with computer-generated data. This has the advantage that any diffusion model can be tested, even models for which no analytical solution is available. This approach in combination with quantitative statistical comparison of experimental and simulated data has been successfully applied to determine diffusion coefficients and transition rates in a bimodal Brownian diffusion system<sup>58</sup>.

#### Analysis of diffusion along the curved surface of the bacterial membrane

So far we have ignored the shape of the membrane of cell or subcellular compartments where diffusion takes place, which is, in general not flat. In most SPT experiments, the location of the biomolecule of interest is only tracked in 2 dimensions, effectively measuring 2-dimensional projections (on the focal plane) of the 3-dimensional diffusion trajectories. Mammalian cells are in general so large that part of their membrane is effectively flat and horizontal (on the scale of the field of view of the microscope). In case of bacteria, which are much smaller, tracking particles in highly curved stretches of membranes cannot be avoided and corrections need to be applied. Simulations of diffusion in the inner membrane of *Caulobacter crescentus*<sup>10</sup> and

*E. coli*<sup>30, 63</sup> demonstrated that MSD analysis of projected trajectories yields a 30% lower diffusion coefficient compared to the 3D trajectories. In addition, the distribution of displacements is substantially distorted: the CPD of a single, homogeneously diffusing species cannot be represented by a single exponential<sup>30</sup>. Several 3D-tracking methods have been developed to avoid this issue<sup>64-67</sup>. The 3<sup>rd</sup> dimension comes, however, at a substantial cost of time resolution and sensitivity, which, in our hands, hinders application to single, fast diffusing fluorescent-protein markers. In case the probed cell geometry is well known, an alternative is to obtain diffusion parameters by comparing experimental with computer-simulated data<sup>30</sup>. This approach, however, requires building an extensive simulation library representing the complete parameter (and diffusion model) space that can be expected in the experiments.

#### **A new method to analyse diffusion on curved surfaces: inverse projection of displacement distributions (IPODD).**

Here, we present a new method, IPODD, that makes use of the direct conversion of a 2D-projected distribution of diffusional displacements back into the most likely 3D-displacement distribution. Crucial for our approach is that the surface geometry is well determined and homogeneous over all cells considered (which is typically the case for healthy *E. coli*). Our method consists of 4 consecutive steps. In the first step a model surface is selected and triangulated. For our application we describe the *E. coli* shape as a cylinder with diameter 0.5  $\mu\text{m}$  and length 1  $\mu\text{m}$ , capped on both end with a half sphere with diameter 0.5  $\mu\text{m}$  (Supplemental Information), the size we typically observe in our experiments. Triangulation was favoured over a curvilinear model<sup>30, 68</sup> because of its general applicability, in particular in cases of complex shapes that cannot be parameterised<sup>69, 70</sup>. Triangulation is effectively a finite-element approach that, to be a faithful representation of the real surface, has to be performed with enough and small enough elements (triangles). We tested the validity of our triangulated spatial model by performing pure Brownian diffusion simulations over the model surface (Supplemental Information) and by comparing the diffusion coefficient put in the simulation, with the one obtained from the 3D trajectories (using a fit to a Rayleigh distribution).

In the second step, we sampled displacement vectors of a given length, distributed randomly over the model surface, projected them on a 2D image plane and determined the projected length. In this way, a 2D-projected displacement histogram is obtained for a given 3D displacement, the "projected displacement distribution" (PDD). For the spatial model we used, PDDs are zero for length larger than the 3D displacement, peak exactly at the length of the 3D displacement, and show a complex pattern towards shorter displacements (Figure 4). The normalized PDD is the probability distribution for finding a 2D-projected length given a 3D displacement with particular length on the surface of the spatial model.

In the third step PDDs are generated for many different 3D displacements (from 0 to 1000 nm, in 5 nm increments), yielding a transformation matrix (Figure 4) that can convert any distribution of 3D displacements (with 5 nm bin size) over the surface of the spatial model in a 2D-projected displacement distribution (with 5 nm bin size) by simple multiplication of input distribution vector with the transformation matrix. To obtain a representative, low-noise transformation matrix, sufficiently smooth PDDs have to be generated. To this end, for each PDD, it is crucial to sample enough 3D displacements in order to probe the complete model surface with randomly oriented displacement vectors in a statistically favourable way (Supplemental Information) To test the applicability of the matrix-multiplication approach and the representativeness of our transformation matrix, we simulated pure Brownian diffusion over the triangulated model surface. The distribution of 3D-displacement lengths (a Rayleigh distribution) was transformed using our transformation matrix into a distribution of 2D-projected displacements (Figure 5), yielding an indiscernible histogram as obtained from direct projection of the 3D displacements. The distribution of displacements is substantially distorted by the projection and can no longer be described well by a Rayleigh distribution. It can still be fitted, yielding a diffusion coefficient of  $103 \pm 1 \text{ nm}^2/\text{ms}$ , about one third lower than the value used for the simulation ( $156 \text{ nm}^2/\text{ms}$ ), in agreement with previous observations<sup>10, 30, 63</sup>. This distortion is purely due to the projection and should not be mistaken for heterogeneous or anomalous diffusion.

In the fourth, final step, the projection matrix is inverted using Gaussian elimination. The inverted projection matrix can be used to convert a measured 2D-projected displacement distribution into a most probable 3D displacement distribution over the model surface, given that the input distribution is random and unbiased. Multiplication of the inverted projection matrix with the 2D-projected displacement distribution used in the previous step, results in a 3D-displacement distribution that is indiscernible from the original Rayleigh-like distribution (Figure 5). A fit yields a diffusion coefficient of  $155 \pm 2 \text{ nm}^2/\text{ms}$ , within the error margin identical to the input diffusion coefficient of the simulations ( $156 \text{ nm}^2/\text{ms}$ ).

To test whether the proposed method works for experimental data, we measured diffusion of an artificial monomeric transmembrane peptide WALP-KcsA consisting of two trans-membrane helices (for experimental details see supplementary information). Using automated single-particle tracking we recorded about two thousand trajectories with an average length of six time lags of 32 ms. The experimentally obtained displacement distribution (which is a 2D projection) is not Rayleigh distributed and the corresponding CPD does not decay exponentially (Figure 5). A Rayleigh fit to the displacement distribution yields a diffusion coefficient of  $164 \pm 3 \text{ nm}^2/\text{ms}$ . Multiplying the distribution with the inverted projection matrix yields a 3D displacement distribution that is Rayleigh distributed (Figure 5) and a CPD that decays single exponentially, yielding a diffusion coefficient of  $246 \pm 9 \text{ nm}^2/\text{ms}$ , substantially larger than the value obtained from the measured, projected displacements. This indicates that normal Brownian motion is the underlying model for the diffusion of WALP-KcsA in the bacterial membrane.

### Conclusions and outlook

In this perspective we have reviewed methods to study the diffusion of membrane proteins in living bacteria. We have in particular focused on analysis methods to obtain quantitative insight in diffusion parameters and underlying mechanisms from trajectories obtained using single-molecule fluorescence microscopy. We hope to have convinced the readers that careful, statistically sound analysis is required to obtain reliable results. Traditional MSD analysis has, when performed correctly, its

merits, mainly thanks to its apparent simplicity and wide acceptance. It has, however, its limitations, since it does not consider the shape of the distributions underlying the displacements, which can contain key supplemental information. Over the last years analysis of diffusion data has regained attention and many advanced statistical approaches have been proposed, based, among others, on maximum likelihood estimators<sup>36</sup>, Bayesian analysis<sup>56</sup> and Hidden-Markov Modelling<sup>57</sup>. In this perspective we have proposed another approach, IPODD, in particular useful for application to experimental data that is distorted due to 2D projection of the 3D trajectories. Crucial is that the shape of the membrane imaged is well defined and well known, which is often the case for bacteria and other single-cellular organisms. We have shown the validity of this approach by applying it to simulated and experimental data. In principle, our approach can unravel distinct populations of particles represented by different diffusion coefficients. In addition, it allows determination of deviations from Brownian behaviour by considering displacement distributions at different time lags. Besides a substantial reduction in computation time, the particular strength of IPODD compared to explicit computer simulations<sup>10, 30, 70, 71</sup> is that it reconstructs 3D-displacement distributions which can be readily analysed with the standard analysis methods discussed above (MSD analysis, CPD analysis or direct propagator fitting). Moreover, when the experimental data set is of insufficient size, IPODD can help to screen the potential parameter space by directly converting 3D-displacement distributions to 2D data sets using the projection matrix. In that way, 2D-projected, computer-generated data sets can be readily compared with experimental data without explicitly simulating diffusion over the model-surface in 3D. Using triangulation of the spatial model surface makes IPODD directly applicable to diffusion in more complex surface structures such as dendritic spines<sup>71</sup> or mitochondrial cristae<sup>70</sup>.

We foresee that in the years to come additional, improved methods will be presented. The approaches presented here and to come will help researchers to better understand the fascinating and biologically important process of membrane-protein diffusion in the membranes of bacteria and other organisms.

### Acknowledgements

We thank the members of our research groups for stimulating discussion and acknowledge financial support from LaserLaB Amsterdam, the Netherlands Organisation for Scientific Research (NWO) with a Vici and an NWO-Groot grant (E.J.G.P.), and with a grant within the Dutch Technology Foundation (STW) research program "Nanoscopy" (E.J.G.P.).

## References

1. B. Alberts, J. H. Wilson and T. Hunt, *Molecular biology of the cell*, 5th edn., Garland Science, New York, 2008.
2. D. Marguet, P.-F. Lenne, H. Rigneault and H.-T. He, *Embo Journal*, 2006, **25**, 3446-3457.
3. K. Mitra, I. Ubarretxena-Belandia, T. Taguchi, G. Warren and D. M. Engelman, *Proc Natl Acad Sci U S A*, 2004, **101**, 4083-4088.
4. J. Luirink, Z. Yu, S. Wagner and J. W. de Gier, *Biochimica et biophysica acta*, 2012, **1817**, 965-976.
5. K. E. Chatzi, M. F. Sardis, S. Karamanou and A. Economou, *The Biochemical journal*, 2013, **449**, 25-37.
6. K. Jung, L. Fried, S. Behr and R. Heermann, *Current opinion in microbiology*, 2012, **15**, 118-124.
7. I. Llorente-Garcia, T. Lenn, H. Erhardt, O. L. Harriman, L.-N. Liu, A. Robson, S.-W. Chiu, S. Matthews, N. J. Willis, C. D. Bray, S.-H. Lee, J. Y. Shin, C. Bustamante, J. Liphardt, T. Friedrich, C. W. Mullineaux and M. C. Leake, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 2014.
8. D. Z. Rudner and R. Losick, *Cold Spring Harbor Perspectives in Biology*, 2010, **2**.
9. D. Z. Rudner, Q. Pan and R. M. Losick, *Proc Natl Acad Sci U S A*, 2002, **99**, 8701-8706.
10. J. Deich, E. M. Judd, H. H. McAdams and W. E. Moerner, *Proceedings of the National Academy of Sciences of the United States of America*, 2004, **101**, 15921-15926.
11. J. Lutkenhaus, *Trends in microbiology*, 2012, **20**, 411-418.
12. H. Strahl and L. W. Hamoen, *Current opinion in microbiology*, 2012, **15**, 731-736.
13. K. S. Ramamurthi, *Current opinion in microbiology*, 2010, **13**, 753-757.
14. S. Ramadurai, A. Holt, V. Krasnikov, G. van den Bogaart, J. A. Killian and B. Poolman, *Journal of the American Chemical Society*, 2009, **131**, 12650-12656.
15. M. Kumar, M. S. Mommer and V. Sourjik, *Biophysical journal*, 2010, **98**, 552-559.
16. J. T. Mika and B. Poolman, *Current Opinion in Biotechnology*, 2011, **22**, 117-126.
17. X. S. Xie, P. J. Choi, G.-W. Li, N. K. Lee and G. Lia, in *Annual Review of Biophysics*, 2008, vol. 37, pp. 417-444.
18. G.-J. Kremers, S. G. Gilbert, P. J. Cranfill, M. W. Davidson and D. W. Piston, *Journal of Cell Science*, 2011, **124**, 157-160.
19. G. Rayan, J.-E. Guet, N. Taulier, F. Pincet and W. Urbach, *Sensors*, 2010, **10**, 5927-5948.
20. E. A. J. Reits and J. J. Neefjes, *Nature Cell Biology*, 2001, **3**, E145-E147.
21. M. B. Elowitz, M. G. Surette, P. E. Wolf, J. B. Stock and S. Leibler, *Journal of Bacteriology*, 1999, **181**, 197-203.
22. C. W. Mullineaux, A. Nenninger, N. Ray and C. Robinson, *Journal of Bacteriology*, 2006, **188**, 3442-3448.
23. S. Chiantia, J. Ries and P. Schwille, *Biochimica Et Biophysica Acta-Biomembranes*, 2009, **1788**, 225-233.

24. E. Haustein and P. Schwille, in *Annual Review of Biophysics and Biomolecular Structure*, 2007, vol. 36, pp. 151-169.
25. J. Yu, J. Xiao, X. J. Ren, K. Q. Lao and X. S. Xie, *Science*, 2006, **311**, 1600-1603.
26. J. Elf, G.-W. Li and X. S. Xie, *Science*, 2007, **316**, 1191-1194.
27. W. E. Moerner and D. P. Fromm, *Review of Scientific Instruments*, 2003, **74**, 3597-3619.
28. D. Axelrod, in *Biophysical Tools for Biologists, Vol 2: In Vivo Techniques*, eds. J. J. Correia and H. W. Detrich, 2008, vol. 89, pp. 169-221.
29. M. C. Leake, J. H. Chandler, G. H. Wadhams, F. Bai, R. M. Berry and J. P. Armitage, *Nature*, 2006, **443**, 355-358.
30. S. M. van den Wildenberg, Y. J. Bollen and E. J. Peterman, *Biopolymers*, 2011, **95**, 312-321.
31. R. E. Thompson, D. R. Larson and W. W. Webb, *Biophysical Journal*, 2002, **82**, 2775-2783.
32. R. E. Thompson, D. R. Larson and W. W. Webb, *Biophysical journal*, 2002, **82**, 2775-2783.
33. K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid and G. Danuser, *Nature Methods*, 2008, **5**, 695-702.
34. A. Serge, N. Bertaux, H. Rigneault and D. Marguet, *Nature Methods*, 2008, **5**, 687-694.
35. H. Qian, M. P. Sheetz and E. L. Elson, *Biophysical journal*, 1991, **60**, 910-921.
36. A. J. Berglund, *Physical review. E, Statistical, nonlinear, and soft matter physics*, 2010, **82**, 011917.
37. M. J. Saxton, *Biophysical journal*, 1997, **72**, 1744-1753.
38. D. S. Martin, M. B. Forstner and J. A. Kas, *Biophysical journal*, 2002, **83**, 2109-2117.
39. T. Savin and P. S. Doyle, *Biophysical journal*, 2005, **88**, 623-638.
40. X. Michalet, *Physical review. E, Statistical, nonlinear, and soft matter physics*, 2010, **82**, 041914.
41. B. Shuang, C. P. Byers, L. Kisley, L. Y. Wang, J. L. Zhao, H. Morimura, S. Link and C. F. Landes, *Langmuir*, 2013, **29**, 228-234.
42. M. J. Saxton and K. Jacobson, *Annual review of biophysics and biomolecular structure*, 1997, **26**, 373-399.
43. A. V. Weigel, B. Simon, M. M. Tamkun and D. Krapf, *Proceedings of the National Academy of Sciences*, 2011.
44. J. Klafter, A. Blumen and M. F. Shlesinger, *Physical review. A*, 1987, **35**, 3081-3085.
45. B. B. Mandelbrot and J. W. V. Ness, *SIAM Review*, 1968, **10**, 422-437.
46. D. Ben-Avraham and S. Havlin, *Diffusion and reactions in fractals and disordered systems*, Cambridge University Press, Cambridge ; New York, 2000.
47. A. Kusumi, C. Nakada, K. Ritchie, K. Murase, K. Suzuki, H. Murakoshi, R. S. Kasai, J. Kondo and T. Fujiwara, *Annual review of biophysics and biomolecular structure*, 2005, **34**, 351-378.
48. J. A. Dix and A. S. Verkman, *Annual Review of Biophysics*, 2008, **37**, 247-263.

49. M. R. Horton, F. Hofling, J. O. Radler and T. Franosch, *Soft Matter*, 2010, **6**, 2648-2656.
50. S. B. Zimmerman and A. P. Minton, *Annual review of biophysics and biomolecular structure*, 1993, **22**, 27-65.
51. P. Schwille, J. Korlach and W. W. Webb, *Cytometry*, 1999, **36**, 176-182.
52. M. Weiss, H. Hashimoto and T. Nilsson, *Biophysical journal*, 2003, **84**, 4043-4052.
53. M. Vrljic, S. Y. Nishimura, S. Brasselet, W. E. Moerner and H. M. McConnell, *Biophysical journal*, 2002, **83**, 2681-2692.
54. J. M. Crane and A. S. Verkman, *Biophysical journal*, 2008, **94**, 702-713.
55. M. Magdziarz and J. Klafter, *Physical review. E, Statistical, nonlinear, and soft matter physics*, 2010, **82**, 011129.
56. A. Robson, K. Burrage and M. C. Leake, *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 2013, **368**, 20120029.
57. R. Das, C. W. Cairo and D. Coombs, *PLoS computational biology*, 2009, **5**, e1000556.
58. V. Ruprecht, M. Axmann, S. Wieser and G. J. Schutz, *Current protein & peptide science*, 2011, **12**, 714-724.
59. C. M. Anderson, G. N. Georgiou, I. E. Morrison, G. V. Stevenson and R. J. Cherry, *Journal of cell science*, 1992, **101 ( Pt 2)**, 415-425.
60. M. A. Deverall, E. Gindl, E. K. Sinner, H. Besir, J. Ruehe, M. J. Saxton and C. A. Naumann, *Biophysical journal*, 2005, **88**, 1875-1886.
61. S. Matsuoka, T. Shibata and M. Ueda, *Biophysical journal*, 2009, **97**, 1115-1124.
62. F. Persson, M. Linden, C. Unoson and J. Elf, *Nature methods*, 2013, **10**, 265-269.
63. M. C. Leake, N. P. Greene, R. M. Godun, T. Granjon, G. Buchanan, S. Chen, R. M. Berry, T. Palmer and B. C. Berks, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 15376-15381.
64. B. Huang, W. Wang, M. Bates and X. Zhuang, *Science*, 2008, **319**, 810-813.
65. M. F. Juette, T. J. Gould, M. D. Lessard, M. J. Mlodzianoski, B. S. Nagpure, B. T. Bennett, S. T. Hess and J. Bewersdorf, *Nature Methods*, 2008, **5**, 527-529.
66. S. R. P. Pavani, M. A. Thompson, J. S. Biteen, S. J. Lord, N. Liu, R. J. Twieg, R. Piestun and W. E. Moerner, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 2995-2999.
67. G. Shtengel, J. A. Galbraith, C. G. Galbraith, J. Lippincott-Schwartz, J. M. Gillette, S. Manley, R. Sougrat, C. M. Waterman, P. Kanchanawong, M. W. Davidson, R. D. Fetter and H. F. Hess, *Proceedings of the National Academy of Sciences of the United States of America*, 2009, **106**, 3125-3130.
68. N. J. Delalez, G. H. Wadhams, G. Rosser, Q. Xue, M. T. Brown, I. M. Dobbie, R. M. Berry, M. C. Leake and J. P. Armitage, *Proceedings of the National Academy of Sciences*, 2010, **107**, 11347-11351.
69. C. Simon, I. Hepburn, W. Chen and E. De Schutter, *J Comput Neurosci*, 2013, 1-15.
70. V. M. Sukhorukov and J. Bereiter-Hahn, *PLoS ONE*, 2009, **4**, e4604.
71. R. Kusters, Lukas C. Kapitein, Casper C. Hoogenraad and C. Storm, *Biophysical journal*, 2013, **105**, 2743-2750.

### Figure captions

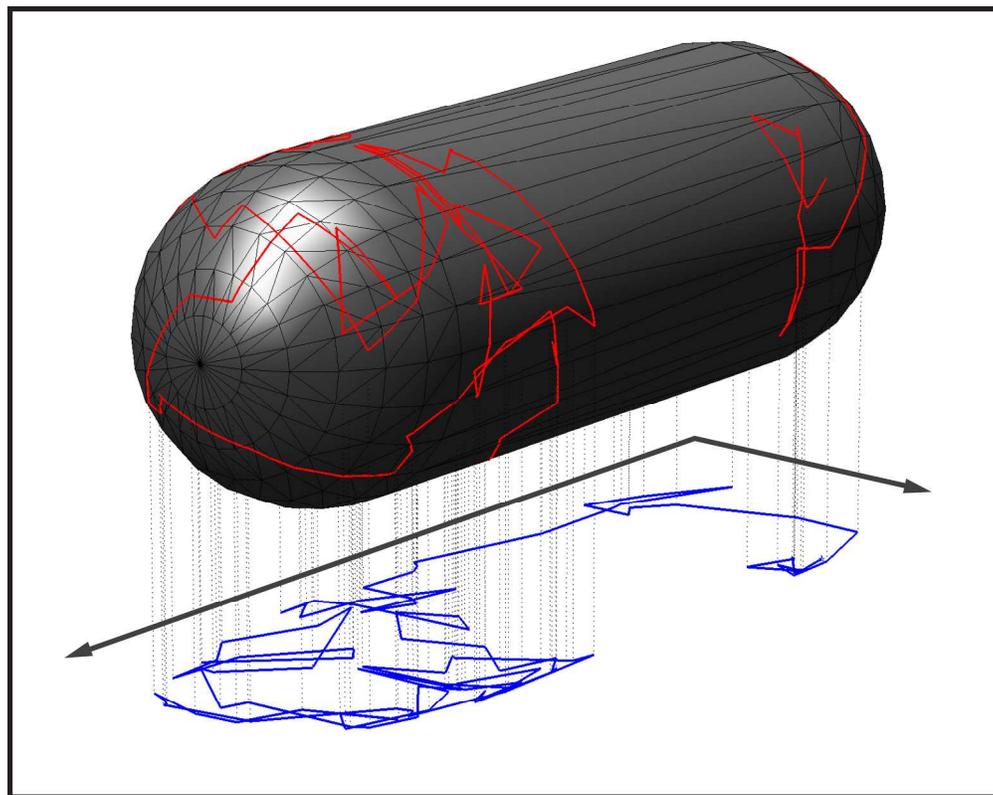
**Figure 1.** Left: schematic representation of a biological membrane, consisting of phospholipids and proteins. Right: membrane structure of the bacterium *E. coli*. The cytoplasm of this bacterium is enclosed by two membranes: the cytoplasmic membrane and the outer membrane. The space between these membranes is the periplasmic space.

**Figure 2.** Plots of mean squared displacements (MSD) as a function of time for distinct diffusion modes. **Black:** normal, Brownian diffusion, characterized by a linear increase of MSD with time:  $\langle r^2 \rangle = 4D\Delta t$ . **Green:** confined diffusion results in levelling off of the MSD curve to an asymptotic value for longer time intervals. **Red:** anomalous superdiffusion results in an upward curvature:  $\langle r^2 \rangle = 4D\Delta t^\alpha$ ,  $\alpha > 1$ . **Blue:** anomalous subdiffusion results in a downward curvature:  $\langle r^2 \rangle = 4D\Delta t^\alpha$ ,  $\alpha < 1$ .

**Figure 3.** Brownian diffusion results in specifically shaped probability distributions of displacements within a fixed time interval. Left: the PDF (probability distribution function) of finding a displacement  $r$  within time interval  $t$  can be represented by a Rayleigh distribution (equation 6). The red curve represents particles with a diffusion coefficient twice that of the black curve. Right: the CPD (cumulative probability distribution) of finding a squared displacement  $r^2$  within time interval  $t$  for the same particles is exponential (note the logarithmic vertical axis), equation 7.

**Figure 4. A:** visual representation of the projection matrix to transform a distribution of 3D displacements along a model surface (C) into the corresponding distributions of 2D-projected displacements. **B:** example of a projected displacement distribution (PDD, at 750 nm 3D displacement, corresponding to the vertical line at 750 nm in A). **C:** representation of the surface used to model an *E. coli* bacterium.

**Figure 5.** Diffusion analysis using IPODD. Shown are PDFs (probability distribution functions) with corresponding CPDs (cumulative probability distributions) as inset for simulated data (diffusion coefficient  $156 \text{ nm}^2/\text{ms}$ , **A-D**) and experimental data (WALP-KcsA-eGFP in *E. coli*; **E, F**). **A:** histogram of the length of  $10^5$  3D displacements, showing a Rayleigh distribution (fit:  $155 \pm 1 \text{ nm}^2/\text{ms}$ ). **B and D:** histograms of the length of the 2D-projected displacements from A, obtained by direct projection of individual 3D displacements (B) and multiplication of the distribution in A with the projection matrix (D). Both projection methods yield identical displacement distributions that deviate substantially from a Rayleigh distribution. Fits yield  $103 \pm 1 \text{ nm}^2/\text{ms}$  (B) and  $102 \pm 2 \text{ nm}^2/\text{ms}$  (D). The CPDs are in both cases not well described by an exponential. **C:** Most probable distribution of 3D displacements calculated from the 2D-projected data (B) using IPODD, by multiplying with the inverted projection matrix. The resulting displacement distribution reproduces the original, 3D displacement distribution, can be fitted well by a Rayleigh distribution yielding a diffusion coefficient of  $155 \pm 2 \text{ nm}^2/\text{ms}$ , within the error margin equal to the diffusion coefficient used for the simulation. **E:** The experimentally obtained displacement distribution of WALP-KcsA-eGFP also displays substantial deviations from a Rayleigh distribution (a fit yields a diffusion coefficient of  $164 \pm 3 \text{ nm}^2/\text{ms}$ ) and the CPD does not decay mono-exponentially, both similar to the simulated data (B). Note that the measured displacements are intrinsically 2D projected due to experimental limitations. **F:** Application of IPODD to the WALP-KcsA-eGFP experimental data of (E) yields a most probable displacement distribution that is well described by a Rayleigh distribution (a fit yields a diffusion coefficient of  $246 \pm 9 \text{ nm}^2/\text{ms}$ ). The corresponding CPD shows a single-exponential decay.



Analysis of single-membrane-protein trajectories obtained using single-molecule fluorescence microscopy in living bacteria provides profound insight in protein diffusive motion.  
220x175mm (300 x 300 DPI)