Manifold learning on single-molecule trajectories of receptor proteins in the plasma membranes of living cells

Accepted 20th November, 2020

ABSTRACT

Retrieving valuable information from a large trajectory dataset is an important issue in single-molecule imaging and tracking. To help retrieve such information, it was developed two data analysis techniques that yield a better view of the diffusive motion of receptor proteins in live cells. It was shown that a contour plot on the plane of normalized variance and local mean-squared displacements can reveal the type of force and the distribution of hidden states encountered by receptor proteins in an inhomogeneous cellular environment. A spectral embedding analysis (SEA) provides further information about the diffusive eigenmodes of the molecules of interest. The combination of these two methods (a contour plot and SEA) enabled us to examine the detailed motion of liganded epidermal growth factor receptors (EGFRs) in live cells. It was found that two neighboring activated EGFRs moved correlatively with interaction that was regulated by membrane cholesterol. The SEA result of the present study showed that those correlated EGFRs may diffuse in a unidirectional force field, presumably stemming from the combined influence of the dynamic self-organizing actomyosin network and flows. The link of signal-transducing proteins, such as EGFRs, with active processes of actomyosin may offer additional insights into the regulation of cellular signaling processes.

Key words: Single-molecule imaging, trajectory, Langevin equation, mean-squared displacement, spectral embedding, epidermal growth factor receptor, correlated motion, cellular signaling.

INTRODUCTION

One of the core issues in cellular biology is to establish the relationship between cellular organization and functions (Hartwell et al., 1999). Tracking single molecules of interest in a living cell can yield valuable insights into this relationship that are inaccessible by ensemble-average methods (Douglass and Vale, 2005; Teramura et al., 2006; Pinaud et al., 2010; Yasui et al., 2018). The single-molecule tracking (SMT) technique is now sufficiently developed to routinely achieve a precision of a few tens of nanometers. The enhanced spatial resolution and non-ensemble-average nature make this technique attractive for investigating cellular dynamical processes (Rosch et al., 2018). SMT has been successfully applied to study various subcellular events, such as genomic dynamics (Wollman et al., 2020), viral infection (Brandenburg and Zhuang, 2007), cellular endocytosis (Jin et al., 2008), membrane protein trafficking (Cui et al., 2018), and cargo transport (Salman et al., 2005).

The mobility of cellular proteins in a living cell is affected by several factors including the viscosity of the cytosol, protein clustering, and binding to subcellular components (Tabaka et al., 2014). The analysis of single-molecule trajectories is commonly based on fits to the mean-squared displacement (MSD) versus time plot. Different modes of motion, such as Brownian diffusion, hopping, confined motion, or directed diffusion, have been discovered (Ritchie et al., 2005; Robson et al., 2012a). Brownian diffusion is characterized by a linear fit of the MSD over
time, yielding a slope proportional to the diffusion coefficient. However, deviations from linearity in the MSD versus time plot are ubiquitous in membrane-associated proteins (Robson et al., 2012b; Sungkaworn et al., 2017). Theoretical and experimental studies have disclosed that such deviations can be caused by cellular flow (Bressloff and Newby 2013; Trong et al. 2012), the presence of obstacles (Li et al., 2020), membrane compartmentalization, or changes in membrane lipid organization (Ramadurai et al., 2010; Grosjean et al., 2018).

Individual particle trajectories often show the evidence of heterogeneity that is not easily resolved (Bouzigues and Dahan, 2007; Kerketta et al., 2016). Therefore, new analytical methods are needed to extract information from single-molecule trajectories. To retrieve information about the local physical properties of molecules in their environment and interpret subtle changes in their diffusive behavior, hidden Markov models (HMMs) have been developed to resolve hidden state changes within single trajectories and deduce the diffusion coefficients of the individual states (Das et al., 2009; Linden and Elf, 2018). HMMs expand upon the standard MSD analysis and provide previously inaccessible information about heterogeneous diffusion. This technique has been successfully applied to analyze actomyosin and kinesin-microtubule movement data (Smith et al., 2001), DNA looping kinematics (Beausang et al., 2007), and the integrin receptor lymphocyte function-associated antigen-1 (Das et al., 2009) in single-molecule microscopy experiments.

However, in the extraction of hidden information from complex trajectories, advance knowledge of the number of hidden states is prerequisite. Counting hidden states in SMT data is a model selection problem that has thus far been solved with simplified noise models (Persson et al., 2013). The current study presents a data-driven method with solid theoretical foundation to provide the information needed for HMM analysis. Our method also yields a clear portrait of the underlying diffusive processes and the types of force fields involved (Lin et al., 2014). To illustrate the unique capabilities of our method, we used it to analyze the trajectories of paired epidermal growth factor receptors (EGFRs), diffusing in proximity in the plasma membranes of a living cell (Lin et al., 2015). We found that the paired liganded EGFRs moved cooperatively with each other, probably due to the correlated fluctuations in their lipid environments (Huang and Lin, 2015). We demonstrated that our method can capture dynamic receptor interactions at the single-molecule level, providing details that are often obscured in other methods.

**APPARATUS AND EXPERIMENTAL PROCEDURES**

**Optical setup**

The output from a blue (473 nm) solid-state laser was used to excite fluorophore-labeled biomolecules in live cells. The fluorescence was collected with a high-numerical aperture (NA) oil immersion objective lens (APON 60X OTIRFM, NA 1.45, Olympus Optical Co., Ltd., Tokyo, Japan) mounted on an inverted optical microscope (IX-71, Olympus Optical Co.) and filtered using a 473-nm notch filter. We detected the fluorescence signals with an electron-multiplying charge-coupled device (EMCCD, Cascade II 512, Photometrics Inc., Huntington Beach, CA, U.S.A.).

**Cell Culture and reagents**

Two cell lines were selected for this study. The HeLa cell line, which is derived from cervical cancer cells, is a standard cell line in cellular biology research worldwide. Typically, HeLa cells can express EGFRs at a level of 20,000 per cell (Zhang et al., 2015) and approximately contain 17 µg cholesterol per mg protein (Robinet et al., 2006). The second cell line, A431, derived from a human epidermal carcinoma, is a model system for the study of cancer-associated cellular signaling pathways. A431 cells can express EGF at an extremely high level (500,000 per cell, Zhang et al., 2015) and contain 32 µg cholesterol per mg protein (Westover et al., 2003).

HeLa and A431 cells were cultured in Dulbecco’s Modified Eagle’s medium with 10 % (v/v) fetal bovine serum without phenol red. Before performing single-molecule live-cell imaging, the cells were cultured in an 8-well slide. After reaching a 70–80 % confluence, HeLa and A431 cells were deprived of serum for 24 h.

A dye-labeled antibody to EGFR (Cy3-anti-EGFR) was prepared by conjugating amine-reactive Cy3-NHS ester (TOCRIS Bioscience, Minneapolis, MN, U.S.A.) to the primary amines of anti-EGFR (Thermo Scientific, Waltham, MA, U.S.A.) following the manufacturer’s protocol. Briefly, the antibody (Ab) was subjected to solvent exchange and concentrated 1 mg/ml. Cy3-NHS ester was dissolved in dried dimethyl sulfoxide and added in excess to the Ab solution at an ester:amines molar ratio of 8:1.

The free fluorophores were removed from the conjugated product using a Zeba spin desalting column (Thermo Scientific). To label the EGFRs, the cells were incubated with 10 nM Cy3-anti-EGFR for 15 min and washed three times with phosphate-buffered saline (PBS).

To synthesize fluorescent epidermal growth factor (EGF), Biotin-EGF (Invitrogen™, Thermo Scientific) was conjugated to Alexa-488-streptavidin in PBS. The resulting product was denoted as Alexa-488-EGF. Excess fluorophores were removed using an immobilized biotin resin column. The cells were stimulated by adding 100 ng/ml Alexa-488-EGF to the cell culture, and were imaged after 10 min. As shown in Figure 1, the diffusive dynamics of the dual-receptor complex was then explored by selecting a pair of liganded (Alexa-488-EGF-EGFR) and unliganded (Cy3-Ab-EGFR) EGFRs or a pair of liganded...
 EGFR complexes (Alexa-488-EGF-EGFR) and following their motions.

**Linking coordinates of localized molecules to generate trajectories**

Single-molecule optical images were recorded at a frame rate of 40 Hz. Position coordinates of the receptor proteins were then extracted from a stack of fluorescent images using a multiple-target tracking algorithm (Serge et al., 2008). The localization accuracy of our apparatus was approximately 40 nm. The resulting position coordinates were linked to form trajectories. To achieve that, we first converted this linking problem into a linear assignment problem (LAP) for the sets of coordinates at time \( t \) and \( t+1 \) using the Hungarian algorithm (Date and Nagi, 2016). We set up a cost function using Euclidean distances as the metric and then minimized the cost function to generate a linking diagram between the sets of position coordinates. However, non-negligible linking errors can occur between two sets of coordinates. For example, in Figure 2(a), a mislink may occur in the light blue and light green circles on the left side of the figure. In the third diffusion step, the third light green circle will be attached to the light blue trajectory as it is the closest circle to its light blue neighbor. To minimize such mislinking, we implemented a Kalman filter with a Brownian diffusion kernel (Wu et al., 2010) to predict the position coordinates (shown as the orange circles on the right side of Figure 2(a)) at \( t+1 \) using the values at \( t \). LAP will then connect the nearest neighbors of the predicted coordinates to the associated trajectories.
this way, we can not only accelerate the searching speed of LAP but also greatly reduce the linking errors (Huang, 2018).

More than 2000 trajectories were generated from stacks of fluorescent images. Typical single-molecule tracks of unliganded Cy3-Ab-EGFR and activated Alexa-488-EGF-EGFR receptors on live HeLa cells are shown in Figure 2(b). Both the EGFR species exhibited confined diffusion, interspaced by fast movement. Because confined diffusion is more relevant to the cellular signaling of EGFR, in the following analysis, we extracted these events of confined diffusion from single-molecule trajectories using the confinement quantification procedure (Meilhac et al., 2006).

RESULTS AND DISCUSSION

Simulation results

Stochastic analysis of paired molecular trajectories

Protein–lipid and lipid–lipid interactions in the plasma membranes of a living cell can drive proteins partitioning into self-associating clusters of protein-rich regions and result in membrane heterogeneity (Kusumi et al., 2005). As shown in Figure 1, the diffusion of cellular molecules can be confined by the cell cytoskeleton close to the membrane (Li et al., 2020) and the proteins anchored to the cytoskeleton, which form dynamic corollas of diffusive motion. Occasional hops between adjacent domains effectively slow down the diffusion of membrane proteins. The resulting trajectory data contain crucial information about the deterministic and random forces that drive the molecules (Lin et al., 2014).

Brownian motion in a force field is usually described using the Langevin equation (Huang and Lin, 2015; Holcman et al., 2015). However, experimentally acquired single-molecule trajectories are in fact spatially and temporally coarse-grained by the data-retrieving apparatus, which may cause some fine features of the single-molecule motion to be washed out in the process of data acquisition. We can coarse-grain the Langevin equation into an effective stochastic equation, which includes a position-dependent diffusion coefficient and a drift in a force field (Holcman et al., 2015). Using the diffusion model, it is possible to generate numerical simulations of trajectories and deduce relevant statistics.

As a receptor moves in the plasma membrane, it will induce an ordering of its surrounding lipid molecules via protein–lipid interaction (Nicolau et al., 2006; Turkcan et al., 2013). We had developed an energy-based model based on the generalized Langevin equation and Cahn-Hilliard equation to simulate the dynamic interactions of receptors in lipid domains (Lin et al., 2014). The receptor and the induced lipid ordering can be viewed as a lipid-dressed protein. When two nearby proteins move in the plasma membrane, they may interact with each other through the ordered lipid molecules (Lin et al., 2014, 2015). For simplicity, the interaction was simulated with a Lennard-Jones (LJ)-type potential of:

\[
U(r = |\vec{x}_1 - \vec{x}_2|) = \varepsilon\left[\left(\frac{r_m}{r}\right)^{12} - 2\left(\frac{r_m}{r}\right)^6\right]
\]

(1)

Here, \(\varepsilon\) is the well depth of the LJ potential and \(r_m\) denotes the distance where \(U(r_m) = -\varepsilon\). For simplicity, the diffusive behavior of two lipid-dressed proteins in proximity is described using the coupled Langevin equations in a single domain (Huang and Lin, 2015):

\[
\begin{align*}
\partial_t \vec{x}_1 &= -\frac{1}{m\gamma} \nabla U_{xx} \vec{x}_1 + \sqrt{2D}dW_{x_1} \\
\partial_t \vec{x}_2 &= -\frac{1}{m\gamma} \nabla U_{xx} \vec{x}_2 + \sqrt{2D}dW_{x_2}
\end{align*}
\]

(2)

where \(\vec{x}_k\) denotes the position of the \(k\)th protein with an effective mass \(m\). The frictional parameter \(\gamma\) in a lipid environment is relevant to the diffusion coefficient \(D\) of the protein by \(\gamma = k_BT/mD\) at temperature \(T\) and Boltzmann constant \(k_B\). We assumed that the fluctuating forces experienced by the proteins followed the Wiener process \(dW_{x_k}\) and the forces fluctuated with a degree of correlation \(\rho\) at the locations of the proteins. By invoking the stochastic chain rule (Øksendal, 2003), we can deduce \(dW_{x_k} = \rho dW_{x_{k-1}} + \sqrt{1-\rho^2} dW_{x_k}\). As the two receptors are far separated with \(\rho = 0\), the correlated Wiener process is reduced to \(dW_{x_k} = dW_{x_k}\).

We solved Equation 2 and generated a set of trajectory pairs. The simulations were performed with an initial condition that one protein was positioned at \(\vec{x}_1 = (-0.25, 0)\) and the other at \((0.25, 0)\). The following parameters were chosen to imitate a receptor in the plasma membrane: \(m\gamma = 4\times10^{-7} \text{ kg/s} (\text{that is, } D = 0.05 \mu \text{m}^2/\text{s})\), \(\varepsilon = 4.81\times10^{-22} \text{ J} (30 \text{ meV})\), \(r_m = 0.4 \mu \text{m}\), and \(\rho = 0.8\) when the separation of two proteins is less than 0.4 \(\mu\text{m}\); otherwise \(\rho = 0\). One hundred pairs of trajectories were generated, each containing 600 time steps with 20 ms per step. Figure 3(a) shows a pair of trajectories with time coordinates color-coded in hue-saturation-value (HSV), indicating that the particles diffuse together with a certain degree of correlation as time increases. Unfortunately, it is difficult to directly compare these simulated trajectories to empirical trajectories. What can be compared are the various statistics of the trajectories. Thus, we developed a two-dimensional contour plot to reveal these statistics of single-molecule trajectories.
Figure 3: (Color online) $V(R^2) - R^2$ contour plots of single-molecule trajectories. (a) A pair of particles (1 and 2) diffuse and interact mutually with a LJ potential of $\epsilon = 30$ meV. (b) $V(R^2) - R^2$ contour plots of 100 trajectories of particle 1 with different LJ potentials (red: $\epsilon = 10$ meV, green: 30 meV, and blue: 90 meV). Inset shows the trajectories used for the analysis. (c) $V(R^2) - R^2$ contour plots of the correlated segments.

**Two-dimensional analysis with normalized variances and mean-squared displacements**

To deduce the statistical properties of the recorded trajectories, we rewrote Equation 2 as

$$dx_t = (f_t/\gamma)dt + \sqrt{2D}dW_t$$

and invoked the stochastic chain rule to rewrite the squared displacement as (Lin et al., 2014):

$$(dx_t)^2 = 2f_t/\gamma \cdot x_t dt + 2D dt + 2\sqrt{2Dx_t}dW_t$$

(3)

The local MSD at time $t$ becomes:

$$\overline{R_t^2} = \overline{(dx_t)^2} = (2/\gamma)f_t x_t dt + 2D dt$$

(4)

Similarly, the variance of $(dx_t)^2$ can also be derived as:

$$\text{Var}(dx_t^2) = 8D\overline{x_t^2}dt - (4/\gamma)\text{Var}(f_t x_t)(dt)^2$$

(5)

where $\text{Var}(A) = \overline{A^2} - \overline{A}^2$ is the variance operation on a random variable $A$. We define a normalized variance $V(R_t^2) = \overline{\text{Var}(R_t^2)} / \langle R_t^2 \rangle^2$ of the local MSD, which yields (Lin et al., 2015):

$$V(R_t^2) = \frac{(2\gamma / dt)\overline{x_t^2} - \text{Var}(f_t x_t)}{\gamma D + f_t x_t}$$

(6)

For a freely diffusing particle with force $f_t = 0$, the
normalized variance \( V(R^2) \) will approach the limit of constant 2. As the particle diffuses in a corral with a short-range force field, it can be stalled briefly, which will cause large variances in the local MSD and thereby result in a high-valued \( V(R^2) \). In contrast, when a particle moves in an environment where it can polarize the surrounding medium, \( V(R^2) \) may fall below the free diffusion limit of \( V(R^2) = 2 \). The \( V(R^2) - R^2 \) contour plot can be prepared to show the statistics of trajectories while preserving single-molecule sensitivity. Here, the local MSDs can quantify the diffusive properties of the molecule in its cellular environment, and the normalized variance can reveal the nature of interactions. An attractive feature of the plot is that when a molecule repeatedly visits or stays in a spatial region, the characteristic \( V(R^2) \) and \( R^2 \) of the domain will be imposed on the trajectories, resulting in the formation of a peak at the corresponding position on the plot.

Figure 3(b) shows the \( V(R^2) - R^2 \) plots of particle 1 in 100 pairs of trajectories with LJ potentials \( \varepsilon = 10 \) (red), 30 (green), and 90 meV (blue). For \( \varepsilon = 10 \) meV, the central peak of \( V(R^2) - R^2 \) falls at \( V(R^2) = 0.2 \) and \( R^2 = 0.2 \) \( \mu m^2 \). Note that \( R^2 \) is equal to the time derivative of the sum of squared displacements, which yields an effective diffusion constant of \( \frac{R^2}{4} = 0.05 \mu m^2/s \). The value is identical to the diffusion constant used for this simulation. As \( \varepsilon \) increases to 30 meV, \( V(R^2) \) decreases slightly to 0.1, reflecting that the two particles are somewhat coupled and diffuse together, which yields a broader distribution profile than that in a weaker coupling. With a further increase in \( \varepsilon \) to 90 meV (three times the thermal energy), the two particles are effectively tied together and diffuse like a single particle, causing \( V(R^2) \) to shift toward the limit of free diffusion. From the result, it can be concluded that the \( V(R^2) - R^2 \) plot is more sensitive than MSD alone in reflecting the diffusive dynamics of molecules in a fluctuating environment.

**Correlative motions of paired particles**

Receptor dimerization plays an important role in initializing signaling cascades (Kerketta et al., 2016). As a receptor passes a nearby receptor, it may encounter an interaction force, which causes a correlated motion between the two proteins. To quantify the mutual correlation, we express the position vectors as a phasor form \( \mathbf{x}(t) = A_1(t)e^{i\phi_1(t)} \) and define a degree of correlation as follows:

\[
C(\tau) = \text{Re} \left[ \frac{\sum \mathbf{x}(t + \tau) \cdot \mathbf{x}(t)}{\sqrt{\sum |\mathbf{x}(t + \tau)|^2 \sum |\mathbf{x}(t)|^2}} \right] = \frac{\sum A_1(t)A_1(t + \tau + \tau) \cos(\phi_1(t + \tau) - \phi_1(t))}{\sqrt{\sum A_1^2 \sum A_1^2}}
\]

The summations were taken over a time mesh along single-molecule tracks. Segments of particle 1’s trajectory with a separation distance \( r < 0.5 \mu m \) from its partner and \( C(\tau) > 0.8 \) were selected for the \( V(R^2) - R^2 \) analysis. The result is shown in Figure 3(c). For \( \varepsilon = 10 \) meV (red), the peak position of \( R^2 \) moves from 0.2 to 0.003 \( \mu m^2 \), indicating much smaller squared displacements along these correlated segments. As \( \varepsilon \) increases, the differences in \( V(R^2) \) become smaller than those for complete trajectories. The three-peaks feature is visible even in this simulation with a single domain, suggesting that three hidden states may be embedded in these dynamically coupled segments.

**Spectral embedding of single-molecule trajectories**

Crucial information about inter-molecular interactions and the energy landscape of the cellular environment can be recorded from protein trajectories (Lin et al., 2017). It is valuable to retrieve a low-dimensional manifold from the dataset because the slow collective variables that form the manifold are relevant to the system functionality (Chiavazzo et al., 2017). However, retrieving such a manifold is typically a nonlinear procedure and is difficult to execute. The first issue encountered is how to identify a set of appropriate variables. In recent years, a number of machine learning approaches have been developed to infer the relevant slow variables (Sidky et al., 2003).

Spectral embedding analysis (SEA) enables an efficient construction of good collective variables for exposing the low-dimensional manifolds underlying the high-dimensional trajectory data (Luo et al., 2003). Here, we implemented a graph-based version of the SEA with a discretized approximation of the manifold to provide an efficient eigen-decomposition of the datasets. As a result, the first eigenvector was trivial because its eigenvalue gave only the data density in a cluster. Therefore, we focused on the next two eigenvectors, \( \phi_2 \) and \( \phi_3 \), to show the critical information about the interactions between diffusing molecules and their environments (Lin et al., 2017).

Figure 4(a)–(c) shows an overview of the spectral embedding (SE) of the correlated segments selected from the same trajectory dataset as that used in Figure 3(c). The relevant time coordinates are color-coded in parula color map. The SE exhibits somewhat circular distribution in the \( \phi_2 - \phi_3 \) plane with earlier coordinates lying at the center and then spreading out with time. For comparison, the correlated segments were displayed on the \( \phi_2 - \phi_3 \) plane with the same coloring scheme. In Figure 4(d), the normalized standard deviation of azimuthal distribution from a circle is presented as a function of the LJ potential strength. The error bars at each \( \varepsilon \) are estimated from the results of five simulations. From the plot, we concluded that inter-particle interaction exerts a negligible influence on SEA.

**Analysis of paired trajectories in a unidirectional force field**

Mounting evidence suggests that the cellular environments
of activated receptors play an important role in regulating signal transduction. To investigate the regulation of signal transduction, we view activated receptor proteins as nanoscale sensors, which record the influence of the surrounding hierarchical structure of the plasma membranes on their trajectories. From this viewpoint, it is interesting to know how a force field from the hierarchical membrane structure affects the movements of paired receptors.

We simulated the diffusive motion of paired receptors under a $y$-directed force field with Equation 2 using $D = 0.05 \, \mu m^2/s$ and $\varepsilon = 30$ meV. An ensemble of 100 pairs of trajectories was prepared, each trajectory containing 600 diffusion steps with a time step of $\Delta t = 0.02$ s. Figure 5(a) shows the $V(R^2) - \overline{R^2}$ plots of paired receptors under three different force fields of $f_y \Delta t / m\gamma = 0.003 \, \mu m$ (red), $0.01 \, \mu m$ (green) and $0.03 \, \mu m$ (blue). The real-space trajectories of $0.01 \, \mu m$ are shown in the inset of the figure. The $V(R^2) - \overline{R^2}$ profile slightly expands along the $\overline{R^2}$ direction as the field strength increases, presumably due to an effect of field-assisted diffusion. Figure 5(b) shows the $V(R^2) - \overline{R^2}$ plot of the correlated segments selected from the trajectories of particle 1. The peak position shifts from $\overline{R^2} = 0.2$ to $0.004 \, \mu m^2$, indicating a direct link between slow diffusion and correlated motion. The force field has also resulted in a significant decrease in the variance of the local MSD as $f_y \Delta t / m\gamma$ increases to $0.03 \, \mu m$ (blue).

Figure 6(a)--(c) shows the SEA results of the correlated segments, which are the same as that used in Figure 5(b). In the $\phi_2 - \phi_3$ plots, the time coordinates were color-coded in parula color map. For comparison, a real-space trajectory projected on the $\phi_2 - \phi_3$ plane is reproduced in the plots. The SE in a weak force field ($f_y \Delta t / m\gamma = 0.003 \, \mu m$) exhibits a nearly circular distribution with trajectories starting at the left side and evolving to the right of the distribution. A $\Lambda$ shape develops gradually with the increase in the force field.
**Figure 5:** (Color online) $V(R^3) - \overline{R^2}$ contour plots of single-molecule trajectories in a force field. (a) A pair of particles interact mutually with a LJ potential of 30 meV and diffuse under a force field of $f_y \Delta t / m \gamma = 0.003 \mu m$ (red), $0.01 \mu m$ (green), and $0.03 \mu m$ (blue). Inset shows the real-space trajectories at $f_y \Delta t / m \gamma = 0.01 \mu m$. (b) $V(R^3) - \overline{R^2}$ plots of the correlated segments from particle 1 under the three force fields.

**Figure 6:** (Color online) SEA of correlated trajectory segments of paired particles interacting mutually with a LJ potential of 30 meV and diffusing in a y-directed force field. A total of 100 simulated trajectories with a field strength (a) $f_y \Delta t / m \gamma = 0.003 \mu m$, (b) $0.01 \mu m$, and (c) $0.03 \mu m$ are shown on a low-dimensional spectral embedding manifold with the time coordinates color-coded in parula colormap. For each case, the correlated segments of a real-space trajectory are also displayed with the same coloring scheme. (d) Normalized azimuthal distribution from a circle is plotted as a function of the LJ potential strength.

In Figure 6(d), the normalized standard deviation from a circle is plotted as a function of field strength. The error bars are estimated from five simulations. The SEA strength, suggesting V- or Λ-shaped features to be a useful indicator of directed movement in a force field (Lin et al., 2017).
distribution changes significantly with force field as the field strength exceeds \( f \gamma \Delta f / m \gamma = 0.01 \mu m \), which corresponds to 1.29 eV/\( \mu m \).

**Experimental results**

**Trajectory analysis of single-molecule EGFRs in plasma membranes**

A widely accepted notion in cellular signaling is that ligand binding to a receptor causes dimerization of receptors and initiates a downstream signaling cascade. Recent identification of cholesterol-dependent nanoassemblies suggests that a cholesterol-mediated interaction plays a role in stabilizing the dimerization process of membrane receptor proteins and regulating the resulting function (Robson et al., 2012b; Orr et al., 2005; Lin et al., 2016). Thus, to determine the impact of the cholesterol-mediated interactions on the diffusion of EGF-bound EGFRs, we were interested in analyzing the trajectory segments of dual Alexa-488-EGF-EGFRs. Only segments belonging to the confined diffusion mode were selected for the analysis because they may be more relevant to cellular signaling. The resulting \( \langle R^2 \rangle - \langle R \rangle \) plots in the two cell lines are shown in Figure 7.

In Figure 7(a), the \( \langle R^2 \rangle - \langle R \rangle \) plots extracted from 2000 trajectories of independent Alexa-488-EGF-EGFRs in HeLa (red) and A431 (green) cells are shown. Multiple streaks are also observed, indicating that the Alexa-488-EGF-EGFRs encounter multiple domains in the plasma membranes of live cells. The plots also exhibit notable overlapping regions, which is surprising considering that the expression level of EGFR in the two cell lines differs significantly. This result indicates that the trajectory segments of independent Alexa-488-EGF-EGFRs are not sensitive to the expression level of EGFR.

The \( \langle R^2 \rangle - \langle R \rangle \) plots of correlated segments of Alexa-488-EGF-EGFRs in HeLa and A431 cells are presented in Figure 7(b). Slower diffusion and more scattered \( \langle R^2 \rangle \) than those of independent Alexa-488-EGF-EGFRs are visible, suggesting that the statistics of correlated segments are sensitive to receptor interaction. The \( \langle R^2 \rangle \) value of the correlated Alexa-488-EGF-EGFRs is considerably lower in A431 cells, attributable to highly effective receptor-ligand and receptor-receptor interactions in A431 cells that presumably stem from a higher level of membrane cholesterol in the cells. This conclusion is also supported by the observation that the \( \langle R^2 \rangle - \langle R \rangle \) of the correlated Alexa-488-EGF-EGFRs in A431 cells increased by two orders of magnitude from \( 10^{-2} \) for native cells (Figure 7(b)) to 1 for MβCD-treated cells (Figure 7(c)), reflecting a reduced cholesterol-mediated interaction after membrane cholesterol was depleted by MβCD. These observations identified the vital role of membrane cholesterol in mediating the interaction between liganded receptors in the cell lines studied herein.

In our previous study, we found that unliganded EGFRs may reside outside the cholesterol-rich lipid domains of plasma membranes and can move into lipid raft domains when they are bound to the ligand (Lin et al., 2015). Prasanna et al. (2014) recently performed a coarse-grained molecular dynamics simulation on \( \beta_2 \)-adrenergic receptor to investigate the interaction of the receptor with membrane cholesterol. They found that cholesterol can bind to the transmembrane helix IV of \( \beta_2 \)-adrenergic receptor and thereby regulate the dimer formation. Several experimental studies have also shown that membrane cholesterol can regulate ligand-induced activation of receptors (Robson et al., 2012b; Westover et al., 2003; Orr et al., 2005). Thus, for EGFRs, we may speculate that ligand binding causes a conformational change that exposes some residues with high cholesterol affinity and provokes nearby raft lipids to form a cholesterol-rich domain. When two
nearby lipid-dressed receptor proteins move in the plasma membrane, they interact with each other through the ordered lipid domains and are thus regulated by the cholesterol content. Our $V(R^2) - \overline{R^2}$ plots of single-molecule EGFR trajectories provide experimental evidence of the regulatory effect of membrane cholesterol on signaling receptors.

**Spectral embedding analysis (SEA) of paired EGFRs in the plasma membranes of living cells**

In the following section, we focus on manifold embedding in the diffusive motion of Alexa-488-EGF-EGFRs. SEA was used to extract the low-dimensional manifold from experimental trajectories. However, due to the limitation of our computer memory, we divided those trajectories into four groups. In Figure 8(a), the SEA outputs are superposed in the $\phi_2 - \phi_3$ plane with all of the trajectory coordinates being displayed in gray. For comparison, in the figure, we also present one of the trajectories with time coordinates coded using the HSV color scheme. The independent liganded EGFRs appear to wander outward without a preferred direction. Thus, many trajectory coordinates sum up to form a nearly circular distribution as shown in Figure 8(a).

When the correlated segments of Alexa-488-EGF-EGFRs are subjected to SEA, a different result emerges. Figure 8(b) presents an overview of those correlated segments from paired Alexa-488-EGF-EGFRs in HeLa cells. For comparison, a correlated trajectory is also displayed on the $\phi_2 - \phi_3$ plane with times of its appearance color-coded in HSV. Unlike the diffusive behaviors shown by independent Alexa-488-EGF-EGFRs, the SEA of correlated receptors displays a characteristic V-shaped distribution, suggesting that these receptors diffuse in a unidirectional force field.

Note that actin polymers form a meshwork that is juxtaposed with the plasma membrane as shown in Figure 1. In HeLa cells, the thickness of the cortical actin layer is approximately 250 nm (Chugh et al., 2017). Myosin motors acting on actin filaments can cause local contractions of initially isotropic actomyosin cortices, which was proposed to initiate coordinated directional motion of actomyosin clusters (Vogel et al., 2020). These directed actomyosin flows have been observed both in vitro (Vogel et al., 2020; Koster et al., 2016) and in living systems (Goehring et al., 2011). Advection in cytosol can drive transient nanoclustering of many actin-associated membrane components (Koster et al., 2016; Beach et al., 2014). A fine balance between the dynamic self-organizing actomyosin network and flows may yield a unidirectional force field. The resulting force field can exert its influence through either acting directly on the pairs of activated Alexa-488-EGF-EGFRs or through the promotion of the correlated motions of activated EGFRs by the unidirectional force field.

**CONCLUSION**

In summary, we developed two data analysis techniques to extract valuable information from single-molecule trajectories of receptor proteins in the plasma membranes of living cells. Our $V(R^2) - \overline{R^2}$ contour plot can expose the distribution profile of hidden states embedded in an inhomogeneous cellular environment with single-molecule sensitivity, whereas the SEA can provide further information about the diffusive eigenmodes of the molecules of interest. By combining these two methods, we
could investigate the motion of liganded EGFRs in living cells. We found that two nearby activated receptors moved correlatively with interaction that could be regulated by membrane cholesterol. Our SEA result showed that those correlated EGFRs may diffuse in a unidirectional force field, presumably stemming from a fine balance between the dynamic self-organizing actomyosin network and flows. The link of signal-transducing proteins, such EGFRs, with active processes arising from actomyosin may provide additional insight into the regulation of signal-transducing proteins in living cells.

ACKNOWLEDGMENT

This research is funded by the Ministry of Science and Technology (MOST) of the Republic of China (grant number MOST 108-2112-M-009-019-MY3).

REFERENCES

cholesterol delivery to the endoplasmic reticulum: Role in cholesterol homeostasis. Traffic. 7(7): 811–823.


Cite this article as:

Submit your manuscript at
http://www.academiapublishing.org/ajsr