

Receptor tyrosine kinase MET ligand-interaction classified via machine learning from single-particle tracking data

Sebastian Malkusch^{a,*}, Johanna V. Rahm^b, Marina S. Dietz^b, Mike Heilemann^b, Jean-Baptiste Sibarita^c, and Jörn Lötsch^{a,d}

^aInstitute of Clinical Pharmacology, Goethe University Frankfurt, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany; ^bInstitute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt am Main, Germany; ^cUniversity Bordeaux, CNRS, Interdisciplinary Institute for Neuroscience, IINS, UMR 5297, F-33000 Bordeaux, France; ^dFraunhofer Institute for Translational Medicine and Pharmacology ITMP, Theodor-Stern-Kai 7, 60596 Frankfurt am Main, Germany

ABSTRACT Internalin B–mediated activation of the membrane-bound receptor tyrosine kinase MET is accompanied by a change in receptor mobility. Conversely, it should be possible to infer from receptor mobility whether a cell has been treated with internalin B. Here, we propose a method based on hidden Markov modeling and explainable artificial intelligence that machine-learns the key differences in MET mobility between internalin B–treated and –untreated cells from single-particle tracking data. Our method assigns receptor mobility to three diffusion modes (immobile, slow, and fast). It discriminates between internalin B–treated and –untreated cells with a balanced accuracy of >99% and identifies three parameters that are most affected by internalin B treatment: a decrease in the mobility of slow molecules (1) and a depopulation of the fast mode (2) caused by an increased transition of fast molecules to the slow mode (3). Our approach is based entirely on free software and is readily applicable to the analysis of other membrane receptors.

Monitoring Editor

Alex Mogilner
New York University

Received: Oct 18, 2021

Revised: Feb 9, 2022

Accepted: Feb 11, 2022

INTRODUCTION

Receptor tyrosine kinases are a family of membrane-bound receptors that transmit information from the extracellular space across the membrane after ligand binding. Ligand-dependent signal transduc-

tion is initiated by the formation of ligand–receptor complexes and receptor dimerization (Dietz *et al.*, 2013). The activated signaling complexes are then rapidly internalized, mainly by clathrin-mediated endocytosis (Lemmon and Schlessinger, 2010; Trenker and Jura, 2020). These reactions are accompanied by a ligand-induced change in receptor mobility (Harwardt *et al.*, 2017). Conversely, this means that ligand-treated cells differ from untreated cells in their receptor mobility, allowing the hypothesis that ligand-labeled cells can be identified by their receptor mobility.

To obtain information about the mobility and activation of membrane receptors in living cells, fluorescence methods are particularly suitable, such as restoration of fluorescence after photobleaching (FRAP; Axelrod *et al.*, 1976), fluorescence correlation spectroscopy (FCS; Magde *et al.*, 1972; Haustein and Schwillie, 2007), and single-particle tracking (SPT; Saxton and Jacobson, 1997). However, because FRAP and FCS provide ensemble-average information on receptor mobility, function-dependent heterogeneity is lost. This shortcoming is overcome by using SPT (Shen *et al.*, 2017), in which the motion of a single molecule is tracked by creating a trajectory of the molecule's localization at different time points. Using single-molecule fluorescence microscopy, it is possible to record the trajectories of many molecules located on the cell membrane simultaneously

This article was published online ahead of print in MBoC in Press (<http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E21-10-0496>) on February 16, 2022.

Author contributions: S.M. conceived and designed the experiments; J.V.R., M.H., and M.S.D. acquired the data; S.M. wrote the software; J.V.R. and S.M. analyzed the data; J.B.S., J.L., J.V.R., M.H., M.S.D., and S.M. interpreted the data; J.B.S., J.L., J.V.R., M.H., M.S.D., and S.M. drafted the article.

Conflicts of interest: The authors declare no conflicts of interest.

*Address correspondence to: Sebastian Malkusch (malkusch@med.uni-frankfurt.de).

Abbreviations used: ANN, artificial neural network; BIC, Bayesian information criterion; CART, classification and regression tree; CME, chemical master equation; DBSCAN, density-based spatial clustering of applications with noise; HMM, hidden Markov model; InlB, internalin B; LOD, limit of detection; MSD, mean squared displacement; NeNA, nearest neighbor analysis; PCA, principal component analysis; PN, Petri net; RF, random forest; ROC-AUC, receiver operation characteristic—area under the curve; SNAP, Shapely additive explanations; SPT, single-particle tracking; XG-Boost, extreme gradient boosting tree.

© 2022 Malkusch *et al.* This article is distributed by The American Society for Cell Biology under license from the author(s). Two months after publication it is available to the public under an Attribution–Noncommercial–Share Alike 4.0 International Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/4.0>).

“ASCB®,” “The American Society for Cell Biology®,” and “Molecular Biology of the Cell®” are registered trademarks of The American Society for Cell Biology.

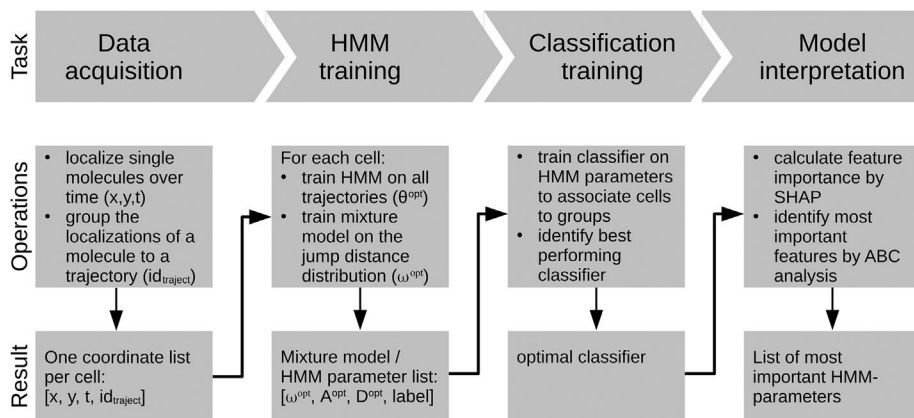


FIGURE 1: Schematic flow chart of the data analysis performed. The four main steps of data analysis are represented as columns. The rows represent the main task of the analysis step, the main analysis operations of the respective step, and the result of the respective step, as well as a possible forwarding of the results to subsequent.

with nanometer-scale lateral resolution and millisecond-scale temporal resolution (for a review, see Shen *et al.*, 2017). The resulting data sets contain the trajectories of hundreds of receptors per cell, providing a representative snapshot of cellular receptor diffusion. SPT data are usually analyzed by modeling the particles' mean squared displacement (MSD) for different time periods using linear models of Brownian motion (Saxton and Jacobson, 1997). Unfortunately, the time averaging during MSD calculation limits this technique to determining exactly one diffusion coefficient per trajectory. Using FCS- and MSD-based analysis of SPT data, a previous study demonstrated that the mobility of membrane-bound MET receptors decreased when cells were previously treated with the bacterial ligand internalin B (InIB; Harwardt *et al.*, 2017). Furthermore, analyzing the SPT data with nonlinear diffusion models showed that the receptors can adopt multiple mobility modes (Harwardt *et al.*, 2017). However, the analysis of possible transitions of an MET receptor molecule between mobility modes has so far been impossible due to the aforementioned limitations of the data analysis techniques used.

An alternative to MSD-based linear models for the analysis of SPT data are hidden Markov models (HMM; Eddy, 2004). Based on the work of Lawrence Rabiner, these are described as follows (Rabiner, 1989): HMMs are based on the hypothesis that the data-generating mechanism relies on a finite number of unobservable modes (N), each characterized by a molecular mobility (B) and each associated with a distinct diffusion coefficient (D). Thus, it is possible to infer the state of an observed molecule if its diffusion coefficient is known. Furthermore, an HMM assumes that mobility mode transitions only take place at the measurement times. The probability of such a change is described in an $N \times N$ matrix (A) of transition probabilities between the modes (Das *et al.*, 2009; Chung *et al.*, 2010; Low-Nam *et al.*, 2011; Ott *et al.*, 2013; Slator *et al.*, 2015; Sungkaworn *et al.*, 2017; Zhao *et al.*, 2019). Finally, the initial mode distribution (π) reports the initial probability of a molecule existing in one of the N modes at the beginning of the observation. HMMs are parameterized by unsupervised machine learning (Baum *et al.*, 1970). During parameterization the models learn the molecular mobilities (B) that characterize the hidden modes, the initial occupation of the modes (π), and the matrix of probabilities that a receptor changes its mobility mode between two measurements (A) directly from the SPT data. Thus, a HMM with N hidden modes is described by a parameter set $\Theta = (\pi, A, B)$ that comprises $2N + N^2$ parameters. A

trained HMM can assign each jump of a trajectory to a particular mobility mode (Viterbi, 1967).

Here, we apply a data-based analysis to a SPT data set of MET receptors on HeLa cells either treated with InIB or untreated to gain deeper insights into the InIB-mediated changes in MET mobility. We propose a data analysis pipeline that uses machine learning paired with explainable artificial intelligence (XAI) (Lundberg and Lee, 2017b) in combination with computed item categorization (Juran, 1975) to learn and interpret the main differences in receptor tyrosine kinase mobility between ligand-treated and untreated cells from the resulting HMM parameters.

The proposed data analysis pipeline is divided into four steps (Figure 1): first, the acquisition of SPT data from differently treated cells; second, the cellwise training of HMMs on the data; third, the training of classification algorithms on the HMM parameters to assign cells to one of the two groups; fourth, the identification of the main group differences by interpretation of the classifier decisions.

As a result, our method assigns MET diffusion to three mobility modes (immobile, slow, and fast) and identifies the three HMM parameters that are most affected by InIB treatment: the diffusion coefficient of the slow state decreases significantly (1). The fast diffusion mode is depopulated (2). The latter is due to an increase in the transition probability from the fast to the slow mode (3). Based on these three parameters, the two cell populations can be classified with a balanced accuracy of >99%.

RESULTS

Single-particle tracking data of MET

The data used within this study, together with experimental details, were published previously (Harwardt *et al.*, 2017). They include SPT data on the mobility of the membrane-bound MET receptor in a total of 117 HeLa cells. Of these, 57 were treated with InIB and 60 cells were untreated. The raw data are single-molecule movies with a length of 1000 frames recorded with an exposure time of 20 ms per frame. For InIB-treated cells, an average of 9168 ± 5435 trajectories per cell (mean \pm SD) were recorded with a trajectory length range of [5, 9, 19] (first, second, third quartile) frames and 4143 ± 2807 trajectories (mean \pm SD) with a trajectory length range of [5, 9, 20] (first, second, third quartile) frames were recorded for untreated cells. The higher degree of receptor labeling with InIB may be due to the high affinity of InIB for the MET receptor (Dietz *et al.*, 2014), which likely exceeds the binding affinity of the Fab fragment. (For more information on cellular statistics of trajectory lengths, see the dataset "met_diffusion_hmm_parameters.csv" included in the Supplemental Material of the paper.)

Training hidden Markov models on single-particle tracking data

HMMs were trained on MET receptor SPT data with the aim of reducing the information to the optimized HMM parameter set $\Theta = (\pi, A, B)$. The optimal model with N mobility modes was determined by assessing the HMM quality using the Bayesian information criterion (BIC) and likelihood ratio tests. Subsequently, the parameter space was expanded to include mode occupancy (ω) and the HMM mobility

parameter (B) was replaced by its associated diffusion coefficient (D) to increase the physiological interpretability of the parameter data. Cells with abnormal parameters were identified using isolation forests and density-based spatial clustering of applications with noise (DBSCAN) analyses and removed from the parameter data.

MET mobility is best described by hidden Markov models with three mobility modes

For each cell, HMMs with $N = 1-5$ mobility modes were trained using the respective single-molecule trajectories. This resulted in five HMMs per cell, each with a parameter set $\Theta = (\pi^{\text{opt}}, A^{\text{opt}}, B^{\text{opt}})$ that was optimized to describe the data and that is characterized by $N^2 + N + 1$ degrees of freedom. Based on the distribution of BIC values of each model, it appeared that the diffusion of membrane-bound MET receptors in both groups was best described by a three-mode model (Supplemental Figure 1, A–D). This result was reproduced by performing likelihood ratio tests between cellular HMMs of increasing complexity and selecting the first HMM whose performance did not improve significantly when complexity was increased by adding an additional mobility mode (Supplemental Figure 1, E and F).

Mobility mode population estimation

The initial mode distribution (π^{opt}) describes the probability of a molecule populating a distinct mobility mode at the beginning of the observation. This definition is rather theoretical and difficult to interpret physiologically. Equating the initial mode distribution with the overall mode population is not justifiable, as a diffusion mixture model built from the two HMM parameters ($P(r|\pi^{\text{opt}}, B^{\text{opt}})$) differs from the corresponding jump distance distribution (Supplemental Figure 2, blue model). Therefore, an additional parameter (ω^{opt}) is introduced that describes the overall mode population considering the mobilities learned from the HMM ($P(r|\omega^{\text{opt}}, B^{\text{opt}})$), Supplemental Figure 2, black model).

Associating mobility modes with diffusion coefficients

The mobility learned from the HMM is actually the expected MSD that molecules undergo between two consecutive images, or within the time interval defined by the camera integration time. The measured molecular displacement is caused by the superposition of two phenomena. The first phenomenon is the molecular diffusion. The second phenomenon is the erroneous determination of the molecule position. Because only diffusion is a molecular property and the positional error is a characteristic of the experimental setup, the latter must be factored out of the mobility. For this purpose, the theoretically estimated minimum detectable MSDs for InIB-treated and untreated cells were determined from the average detected photon distribution of the fluorescence probes. They will be referred to as the limits of detection (LODs; $\text{LOD}_{\text{Fab}} = 3043.57 \pm 431.57 \text{ nm}^2$, $\text{LOD}_{\text{InIB}} = 2430.6 \pm 392.73 \text{ nm}^2$). The expected MSDs of the slowest diffusing modes as determined by the HMM analysis for untreated cells ($\min\langle r^2 \rangle_{\text{Fab}} = 1175.35 \pm 217.03 \text{ nm}^2$) and InIB-treated cells ($\min\langle r^2 \rangle_{\text{InIB}} = 738 \pm 39 \text{ nm}^2$) were both below the respective LODs (Supplemental Figure 3). Therefore, their measured expected MSDs are dominated by the positional error and no information about molecular diffusion can be learned from the data. These modes were defined as immobile ($D := 0 \frac{\mu\text{m}^2}{\text{s}}$). Their measured expected MSDs are considered to be formed solely by the static error. Therefore, the static error for each model is calculated from the slowest mobility modes using the nearest neighbor analysis (NeNA) distribution ($\epsilon^{\text{Fab}} = 17.08 \pm 2.63 \text{ nm}$, $\epsilon^{\text{InIB}} = 14.06 \pm 2.24 \text{ nm}$; Endesfelder

et al., 2014). Finally, the diffusion coefficients of the remaining two states for both groups were corrected for static errors and dynamic measurement errors. As a result, the three HMM mobility modes will be referred to as immobile ($D_1^{\text{Fab}} = 0.00 \frac{\mu\text{m}^2}{\text{s}}$, $D_1^{\text{InIB}} = 0.00 \frac{\mu\text{m}^2}{\text{s}}$), slow ($D_2^{\text{Fab}} = 0.07 \pm 0.01 \frac{\mu\text{m}^2}{\text{s}}$, $D_2^{\text{InIB}} = 0.04 \pm < 0.01 \frac{\mu\text{m}^2}{\text{s}}$), and fast ($D_3^{\text{Fab}} = 0.25 \pm 0.02 \frac{\mu\text{m}^2}{\text{s}}$, $D_3^{\text{InIB}} = 0.22 \pm 0.01 \frac{\mu\text{m}^2}{\text{s}}$) based on their assigned diffusion coefficients in ascending order throughout the manuscript for reasons of physiological interpretability. (Values are given as mean \pm SD.)

Anomaly detection

Nine cells had an anomalous set of parameters ($\pi^{\text{opt}}, \omega^{\text{opt}}, A^{\text{opt}}, D^{\text{opt}}$) identified by isolation forests and DBSCAN-based analysis (Supplemental Figure 4). These abnormalities, of which six were untreated and three were InIB-treated cells, were excluded from further analysis, resulting in a balanced data set with 54 cells from each group. The resulting dataset is available as a text file in the Supplemental Material of the article ("met_diffusion_hmm_parameters.csv").

Statistical parameter exploration

The results of the HMM analysis is exemplarily visualized for two cells in Figure 2. The physiologically interpretable parameters ($\omega^{\text{opt}}, A^{\text{opt}}, D^{\text{opt}}$) of InIB-treated and untreated cells were compared by parameterwise statistical hypothesis tests. In a second explorative analysis, the parameter interdependence was demonstrated. The parameters originate from a HMM and therefore several parameter interdependencies are defined by the nature of the model (Rabiner, 1989). These result in the redundancy of one mode population probability as $\sum_{i=1}^N \omega_i^{\text{opt}} = 1$ and one entry of the transition probability matrix per mobility mode as $\sum_{i=1}^N \alpha_{i,j}^{\text{opt}} = 1$. Furthermore, the parameters are screened for pairwise interdependence by correlation analysis. Finally, the interdependence between $\omega^{\text{opt}}, A^{\text{opt}}$, and D^{opt} is demonstrated by simulations.

Internalin B-treated and untreated cells differ in nearly all HMM parameters

Comparison of HMM parameters between the two groups was performed using the Mann–Whitney U test, because several parameters were not normally distributed and pairs of parameters had unequal variances. Comparison showed that the groups differ significantly in 12 of 15 parameters (Figure 3 and Supplemental Table 1). The parameters that did not change significantly upon InIB treatment were the diffusion coefficient of the immobile mode ($D_1^{\text{Fab}} = 0.00 \frac{\mu\text{m}^2}{\text{s}}$, $D_1^{\text{InIB}} = 0.00 \frac{\mu\text{m}^2}{\text{s}}$), the transition probability from the immobile to the fast mode ($P^{\text{Fab}}(3|1) = 9.29 \times 10^{-5} \pm 3.48 \times 10^{-4}$, $P^{\text{InIB}}(3|1) = 1.64 \times 10^{-4} \pm 3.82 \times 10^{-4}$) and the transition probability from the fast to the immobile mode ($P^{\text{Fab}}(1|3) = 7.74 \times 10^{-5} \pm 1.78 \times 10^{-4}$, $P^{\text{InIB}}(1|3) = 1.46 \times 10^{-4} \pm 2.49 \times 10^{-4}$). Their values either were close to zero or were set to zero during error correction. Here, a problem with the interpretability of results from HMM analyses of SPT data becomes apparent: when two groups differ significantly on almost all parameters, it is difficult to narrow the differences down to variations in key parameters.

Bivariate correlation analysis reveals pairwise parameter interdependencies

A bivariate correlation of the parameters ($\omega^{\text{opt}}, A^{\text{opt}}, D^{\text{opt}}$) was performed using Spearman's rank correlation (Supplemental Figure 5).

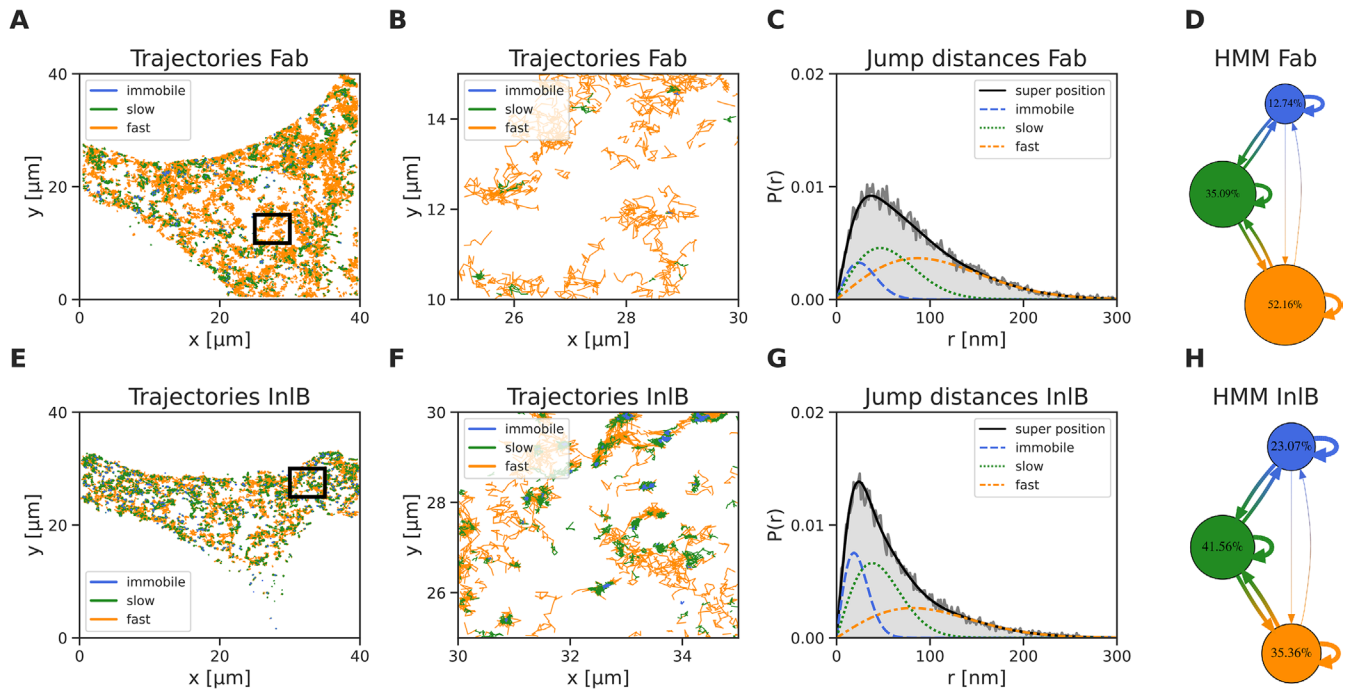


FIGURE 2: Hidden Markov model (HMM)-based analysis of two representative cells, unlabeled (Fab) and labeled with InIB. The trajectory data sets of the individual cells were each analyzed with a three-state HMM. Each state is characterized by a specific diffusion coefficient. Based on these diffusion coefficients, the receptor movement is either characterized as immobile (blue), slow (green), or fast (orange). Graphical representation of the HMM-based analysis of an untreated cell: (A, B) diffusion map of the whole cell membrane, A and zoom-in on a boxed region, B. (C) Overlay of the jump-distance probability density function and a three-state diffusion mixture model of the state occupancy $P(r|\omega^{\text{opt}}, B^{\text{opt}})$. (D) The resulting HMM of the untreated cell is depicted as a transition-state diagram with states color coded by their diffusion (immobile [blue], slow [green], or fast [orange]). The state occupancy (ω^{opt}) is encoded in the node size and the transition probability between states (A^{opt}) is encoded in the arrow widths. Graphical representation of the HMM-based analysis of an InIB-treated cell: (E, F) diffusion map of the whole cell membrane, E, and zoom-in on a boxed region, F. (G) Overlay of the jump-distance probability density function and a three-state diffusion mixture model of the state occupancy $P(r|\omega^{\text{opt}}, B^{\text{opt}})$. (H) The resulting HMM of the InIB-treated cell is depicted as a transition-state diagram with states color coded by their diffusion (immobile [blue], slow [green], or fast [orange]). The state occupancy (ω^{opt}) is encoded in the node size and the transition probability between states (A^{opt}) is encoded in the arrow widths.

A high correlation was defined by a correlation coefficient of $|r_S| \geq [0.7; 0.89]$ and a very high correlation was defined by a correlation coefficient of $|r_S| \geq [0.9; 0.1.0]$ (Hinkle *et al.*, 2003; Mukaka, 2012). However, thresholds of correlation coefficients are arbitrary and should be used judiciously (Schober *et al.*, 2018). The parameter set of the Fab-labeled cells included four pairs with a very high correlation while the parameter set of the InIB-labeled cells included three pairs with a very high correlation and 18 pairs with a high correlation. This indicates that there is a high pairwise interdependence among the parameters.

Stochastic simulations reveal multivariate parameter interdependencies

Assuming that the population of mobility modes described by the trained HMMs describes an equilibrium, a stochastic simulation was designed based on the HMM transition probabilities with the goal of reproducing the equilibrium. The experiment was performed for InIB-treated and untreated cells and the effect of InIB treatment on the equilibrium was analyzed (Figure 4, A and B). The analysis of the HMM-mode population parameters proposed different equilibrium distributions for untreated cells and InIB-treated cells (ω_{1-3} in Figure 3 and Figure 4C). Both InIB-treated and untreated cells showed more mobile MET receptors (slow and fast) than immobile. More MET receptors diffused fast than slowly

in untreated cells (Figure 4C blue distributions, Table 1, and Supplemental Table 2). This situation changed with InIB treatment. After this, most MET receptors diffused slowly. However, still more receptor molecules were assigned to a mobile mode than to the immobile one (Figure 4C orange distributions and Table 1). The populations of mobile modes for untreated and InIB-treated cells each gave the same ranking when the simulated results were compared with the analytical results (Figure 4, Table 1, and Supplemental Table 2). As a result, the InIB-induced differences found for the HMM parameters in the mobility mode population were reproduced by the simulation equilibrium, although these were based only on the transition probabilities. These simulations demonstrate that the information about the mode population is incorporated into the transition probabilities. Furthermore, it is known that certain reaction events are collision-driven and are therefore directly associated with the molecular diffusion (Gillespie, 2007). This means that differences in individual transition probabilities that are caused by changes in the diffusion shifted the response equilibrium, and thus parameters are interdependent. As a result, the demonstrated interdependencies rectify dimension reduction of the parameter space without loss of information. The complexity of the parameter interdependencies makes it difficult to narrow down the differences between the two groups to a few key parameters, since the specific change in a transition probability due to

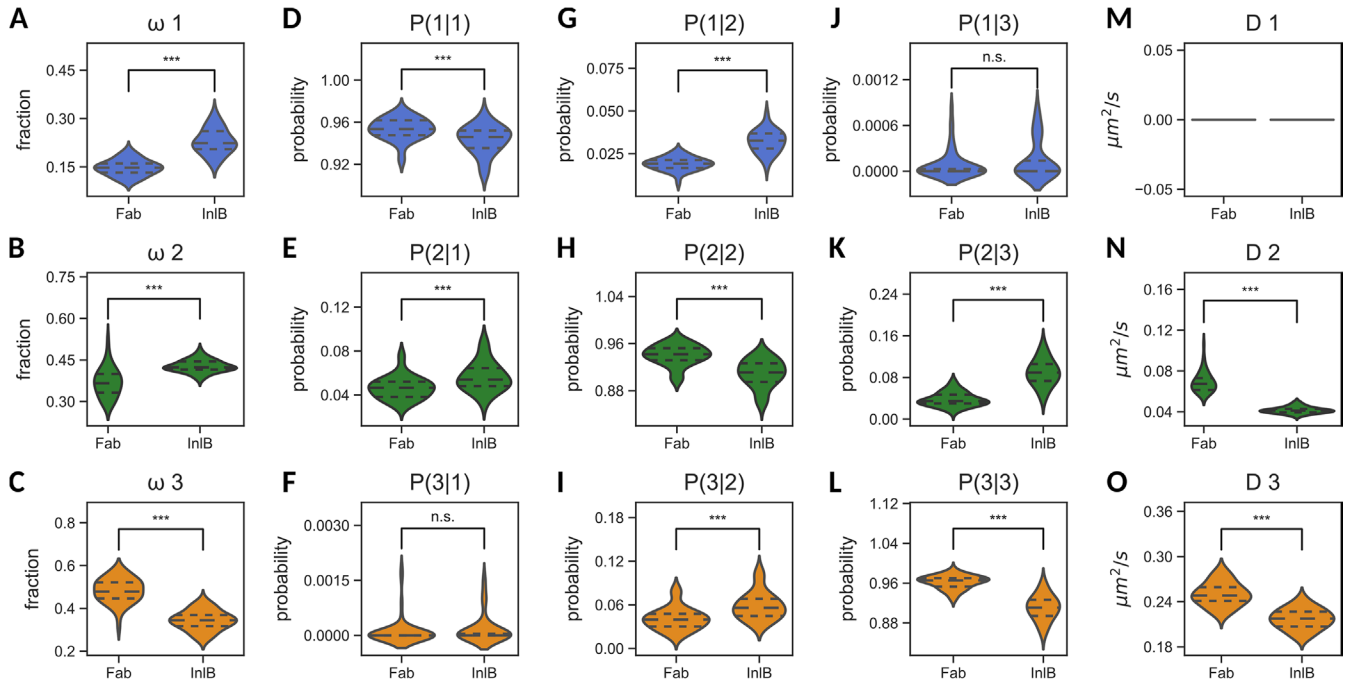


FIGURE 3: Distribution of parameters associated with membrane-bound MET mobility. The parameters were machine learned by 108 three-state HMMs and diffusion mixture models. For simplicity, the resulting list of optimized model parameters $\theta^{\text{opt}} = (\omega^{\text{opt}}, A^{\text{opt}}, D^{\text{opt}})$ are termed optimized HMM parameters. The distributions of the optimized HMM parameters are shown as violin plots. The quartiles of the distributions are highlighted as dashed lines. Each state is characterized by a specific diffusion coefficient. Based on these diffusion coefficients the states are either characterized as immobile (blue, A, D, G, J, M), slow (green, B, E, H, K, N) or fast (orange, C, F, I, L, O). The data set is split into 54 untreated cells (Fab) and 54 InIB-treated cells (InIB). The parameter distributions of the two groups are pairwise compared using a two-sided Mann–Whitney U test. The p -values are corrected for multiple testing using the method of Bonferroni. The number of tests performed is 14. The significance thresholds were set to $\alpha = 0.05$ (*), $\alpha = 0.01$ (**), and $\alpha = 0.01$ (***). Corrected p -values above a significance threshold of $\alpha = 0.05$ are termed not significant (n.s.).

the introduction of an external stimulus such as a ligand is accompanied by changes in several dependent parameters.

Training of classification algorithms on hidden Markov parameters

The multivariate changes in MET mobility induced by InIB treatment were determined using supervised machine learning. For this purpose, a data set was composed from the physiologically interpretable parameters that were derived from the former trained N -modal HMMs ($\omega^{\text{opt}}, A^{\text{opt}}, D^{\text{opt}}$). Thus, each cell formed one instance. This resulted in a balanced data set of cells that were labeled by their treatment (InIB-treated or not). Subsequently, different supervised classification algorithms (classification and regression tree [CART], random forest [RF], extreme gradient boosting tree [XG-Boost], and artificial neural network [ANN]) were trained on 2/3 of the data (training data) to identify InIB-treated cells. The performance of the trained algorithms was quantified on their ability to correctly identify cells from the remaining one-third of the data (validation data) that were so far not known to the algorithms. For performance quantification the receiver operator characteristic-area under the curve (ROC-AUC), balanced accuracy (Brodersen *et al.*, 2010), F1-score, precision, and recall metrics were used (Branco *et al.*, 2015; Luque *et al.*, 2019).

Defining training data from physiologically relevant parameters

Parameter exploration revealed that the parameter set comprises redundant parameters due to interdependencies. Unfortunately, it did not identify the key parameters that yield the desired informa-

tion about the physiological differences between InIB treated and untreated cells. For example, the nature of the model makes it possible to drop one entry of the mode occupation probability vector. However, it does not provide any guidance on how to choose the right parameter. Therefore, all physiologically interpretable parameters are first used for training the classifier. In summary, these were the mode occupation probability (ω^{opt}), the transition probabilities of a molecule to switch between mobility modes ($P(i|j)$ with $i \neq j$), and the diffusion coefficients that characterize the mobility modes (D^{opt}). Furthermore, the diffusion coefficients of the immobile modes of both groups, which were previously defined as $D_1^{\text{Fab}} := 0 \frac{\mu\text{m}^2}{\text{s}}$ and $D_1^{\text{InIB}} := 0 \frac{\mu\text{m}^2}{\text{s}}$, were removed from the dataset because they would not contribute to classification. The result was two balanced datasets with $n = 11$ features ($n = N + (N^2 - N) + (N - 1)$). The training data consisted of 72 instances, of which 36 represented untreated and 36 represented InIB-treated cells. The validation data consisted of 36 instances, of which 18 represented untreated and 18 represented InIB-treated cells.

Internalin B-treated and untreated cells can be classified based on MET mobility with a balance accuracy >90%

All classifiers trained with this dataset were able to distinguish InIB-treated cells from untreated cells with a balanced accuracy of >93% (Table 2). Nevertheless, the random forest classifier did solve the binary classification task best in all of the measured metrics (ROC-AUC, balanced accuracy, F1 score, precision score, and recall score) and was therefore chosen for further investigation of the differences

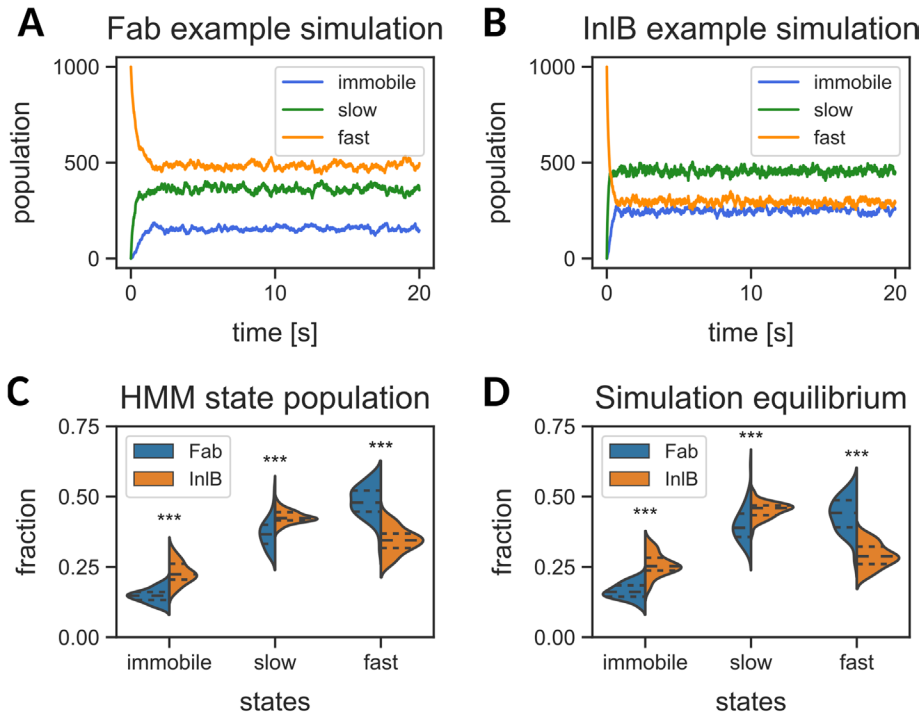


FIGURE 4: Comparison of the state occupancy machine-learned by HMMs with the equilibrium state occupancy generated by the simulation of stochastic PNs parameterized with the transition probabilities of the HMMs. The time evolution of small ensembles of 1000 molecules, initially all fast-diffusing, was simulated with the stochastic PNs until equilibrium was reached. For each simulation, the average equilibrium population ratio was calculated from the last 1000 time steps. (A, B) Simulated time evolution of a representative untreated cell, A, and a representative internalin B (InIB)-treated cell, B. The time traces represent the populations of the three diffusive states immobile (blue), slow (green), and fast (orange). (C) The effect of InIB treatment on the distribution of state occupation as learned by the 108 HMMs is shown in the form of a statewise comparison of the two groups. (D) The effect of InIB treatment on the equilibrium mobility state distribution of MET as generated by the simulation of 108 stochastic PNs. Here, the 54 untreated cells are depicted as blue violin plots (Fab) and the 54 InIB-treated cells are depicted as orange violin plots (InIB). The distribution quartiles are highlighted as dashed lines. Differences in state occupation are tested using a two-sided Mann–Whitney *U* test. For each experiment (HMM, stochastic PN), three tests were performed and the *p*-values were corrected for multiple testing using the method of Bonferroni. The significance thresholds were set to $\alpha = 0.05$ (*), $\alpha = 0.01$ (**), and $\alpha = 0.001$ (***). Corrected *p*-values above a significance threshold of $\alpha = 0.05$ are termed not significant (n.s.).

between the two groups. If all performance measurement metrics were ranked equally, then the XGBoost performed second best, followed by the ANN and finally the CART (Table 2). The random forest model achieved a ROC-AUC of 1.0 (Figure 5A) and a balanced accuracy of >99.9% on the validation data. As a result, the model misclassified 0.055% of the validation instances, each time classifying an activated cell as resting (Figure 5B). To check whether the classifiers learned the differences between the two groups based on the measured receptor mobility, a control experiment was conducted in which the classifiers were trained on permuted training data. As expected, the performance of the classifiers trained on permuted parameters was not better than guessing (Supplemental Table 3).

Model interpretation and identification of key parameters for classification

Shapely additive explanations (SHAP; Lundberg and Lee, 2017a) were calculated for each cell of the validation data using the assignment accuracy achieved by the random forest classifier (Figure 5C). The mean value of the absolute SHAP values for each feature was

calculated and interpreted as feature importance for the classification task (Figure 5D). Finally the most relevant parameters for the classification task were identified via computed ABC analysis of feature importance, which is an item categorization technique that divides each set of positive numeric items into three nonoverlapping subsets named “A,” “B,” and “C” (Juran, 1975). The parameters in ABC subset “A” were considered “the important few” (Juran, 1975) and therefore retained as the key parameters that contained the important information regarding the differences in MET mobility between the groups of InIB-treated and untreated cells (Figure 5D, Table 3, and Supplemental Figure 6).

The internalin B-induced changes in MET mobility are best described by three key parameters

The set size of “A” comprised the diffusion coefficient of the slowly diffusing population (D_2) with a mean absolute SHAP value of 0.19 ± 0.01 , the population of the fast-diffusing state (ω_3) with a mean absolute SHAP value of 0.09 ± 0.01 , and the transition probability from the fast-diffusing state to the slowly diffusing state ($P(2|3)$), with a mean absolute SHAP value of 0.07 ± 0.01 . The SHAP value distributions of all three variables were characterized by a bimodal distribution with a decision boundary formed at a SHAP value of 0 that also separated high and low feature values (Figure 5C). The transition probability from the fast state to the immobile state and the transition probability in the reverse direction had the least effect on the decision of the classifier. To prove the feature selection mediated by the ABC analysis, the classification task was performed with a reduced feature set that included only the three features in category “A.” As a result, the performance

of the random forest and ANN classifiers increased to the point where the ANN classifier now shows the second best performance after the random forest classifier, followed by the XGBoost classifier and the CART classifier (Supplemental Table 4). Repeating the experiment 100 times each with three randomly chosen parameters resulted in less accurate models (Supplemental Figure 7). Therefore, the increase in performance was due to the key parameters identified by ABC analysis and not to the dimension reduction. Thus, it is reasonable to assume that the InIB-mediated changes in features D_2 , $P(2|3)$, and ω_3 represent the main differences between InIB-treated and untreated cells. In conclusion, it was shown that the diffusion coefficient of the slow state (D_2) decreased significantly due to InIB treatment. Further, InIB treatment of MET receptors led to a depopulation of the fast-diffusing state (ω_3), which was caused by an increase of the transition probability from the fast to the slow state ($P(2|3)$).

DISCUSSION

To compress the information encoded in SPT data sets of membrane receptors, hidden Markov models are frequently used

State	rank(ω_{Fab}^{opt})	rank(ω_{InIB}^{opt})	rank(ω_{Fab}^{Sim})	rank(ω_{InIB}^{Sim})
Immobile	1	1	1	1
Slow	2	3	2	3
Fast	3	2	3	2

TABLE 1: Diffusion state occupancy distributions ranked by their distribution medians. The significance of the rank distributions is tested by pairwise-performed two-sided and one-sided Mann–Whitney U -tests of independent samples (Supplemental Table 2). A comparison of the rank vectors shows that for both the mode populations learned by the hidden Markov model (ω^{opt}) and the simulated equilibrium distributions (ω^{sim}), the rank order changes equally upon treatment with internalin B (InIB): The ascending rank order for untreated cells is [$\omega_1, \omega_2, \omega_3$] and for InIB-treated cells the rank order is [$\omega_1, \omega_3, \omega_2$].

(Das et al., 2009; Chung et al., 2010; Low-Nam et al., 2011; Ott et al., 2013; Slator et al., 2015; Sungkaworn et al., 2017; Zhao et al., 2019). These HMMs describe the membrane diffusion of single molecules by the use of three types of parameters (π : the initial modal weights; A : the transition probabilities; B : the molecular mobilities). A system with N mobility modes is described by an HMM with $2N + N^2$ parameters. These were extended by two additional parameters due to physiological interpretability (ω : the modal occupancy; D : the molecular diffusion coefficient). Most of these parameters are interdependent (Rabiner, 1989). The transition probability matrix, for example, describes the transition probabilities of single molecules between the individual mobility modes. The matrix can be transformed into the chemical master equation (Gillespie, 1977), which describes the system as a closed reaction network that is defined by first-order reaction equations. Because many reactions are based on collisions between two reactants, they are diffusion-limited. Their molecular diffusion is directly coupled to their reaction rate constants (Gillespie, 2007). Furthermore, the reaction rate constants of a reaction network also define the equilibrium population. The latter could be demonstrated experimentally in this report by reproducing the equilibrium mode distribution with stochastic Petri nets (PNs) that were parameterized with the transition probability matrix learned by HMMs. This interdependence of HMM parameters makes it difficult to narrow down the main differences in membrane diffusion between groups of differently treated cells. To identify the HMM parameters, which contain important information regarding the differences between the groups, a previously published approach based on machine-learned feature importance in

combination with computed item categorization was used in a slightly modified form (Lotsch and Malkusch, 2021):

The problem is first transformed into a classification problem where classifiers are trained to identify members of certain groups based on the HMM parameters. The decisions of the trained classifiers are then interpreted using SHAP values to determine the importance of each parameter for the classifier decision (Lundberg and Lee, 2017b). Subsequently the HMM parameters that are most important for the classifier decision are identified using ABC analysis as an item categorization method (Juran, 1975).

We demonstrated the validity of our approach on a single-particle trajectory data set that comprises diffusion information about MET receptor mobility on the membranes of InIB-treated and untreated cells. Using our approach, we could show that the lateral membrane diffusion of individual MET receptors is best described by a system of three mobility modes, with each mode characterized by an individual diffusion coefficient. All states exhibiting a diffusion coefficient below the spatial resolution of the experiment were assigned to an immobile state. The other two diffusion states, the slow diffusion state ($D_{2,Fab} = 0.07 \pm 0.01 \mu\text{m}^2/\text{s}$, $D_{2,InIB} = 0.04 \pm 0.003 \mu\text{m}^2/\text{s}$) and the fast diffusion state ($D_{2,Fab} = 0.25 \pm 0.01 \mu\text{m}^2/\text{s}$, $D_{3,InIB} = 0.21 \pm 0.01 \mu\text{m}^2/\text{s}$) can probably be assigned to more confined movement and freely diffusing particles, respectively (Harwardt et al., 2020). Confinement can occur due to receptor localization inside nanodomains on the cell membrane (Seveau et al., 2004) and it is also likely that receptors are confined before immobilization. MET receptors become immobilized due to interactions with the actin cytoskeleton and before endocytosis (Shen et al., 2000; Li et al., 2005; Orian-Rousseau et al., 2007). The diffusion coefficients of the mobile InIB-treated MET receptors are always lower than those of the untreated receptor, which indicates that the diffusion of ligand-bound receptors is in general slower. This could be due to ligand-induced dimerization of the receptors (Dietz et al., 2013) and the formation of signaling hubs by recruitment of adapter proteins and signaling molecules to the receptors (Furge et al., 2000; Trusolino et al., 2010; Niemann, 2013).

InIB treatment shifts the population of the states toward the less mobile and the immobile state. Increased immobilization upon activation with InIB was previously observed (Harwardt et al., 2017; Baldering et al., 2021). The diffusion state with the highest diffusion coefficient becomes less stable in the case of InIB and transitions to the less mobile population are more likely. These observations are reflected in the fact that the diffusion coefficient features D_2 , the probability $P(2|3)$ of switching from the highly mobile state to the less mobile state, and the occurrence ω_3 of the fast diffusion state are the most important features to distinguish between untreated

Classifier	Rank	ROC-AUC	Accuracy	F1-score	Precision	Recall
ANN	2	1.000 \pm 0.001	0.968 \pm 0.022	0.967 \pm 0.022	1.000 \pm 0.000	0.937 \pm 0.043
CART	1	0.969 \pm 0.030	0.931 \pm 0.054	0.926 \pm 0.058	0.966 \pm 0.052	0.894 \pm 0.088
RF	4	1.000 \pm 0.007	0.999 \pm 0.004	0.999 \pm 0.004	1.000 \pm 0.000	0.999 \pm 0.008
XGBoost	3	0.999 \pm 0.007	0.993 \pm 0.022	0.992 \pm 0.024	0.997 \pm 0.013	0.988 \pm 0.039

TABLE 2: Classifier performance measure based on validation metrics. The performance of different classifiers on classifying single cells as either untreated or internalin B (InIB)-treated was measured. The analysis was performed in three steps: First the hidden Markov model (HMM) parameter list (ω^{opt} , A^{opt} , D^{opt}) was filtered for highly correlated parameters and split into a training and a validation data set (2:1). Second, an ANN, a CART, an RF, and an XGBoost were trained on the training data set to predict whether cells are untreated or InIB-treated from the variables learned from the HMM. To rule out a random result, the experiment was repeated 100 times with different classifier initializations. Third, the classifier performance was validated on the validation data set by different metrics: The area under the receiver operating characteristic curve (ROC-AUC), the balanced accuracy, the f1-score, the precision, and the recall are shown. The values are given as mean \pm SD. The rank is calculated as the summed rank of all five validation metrics.

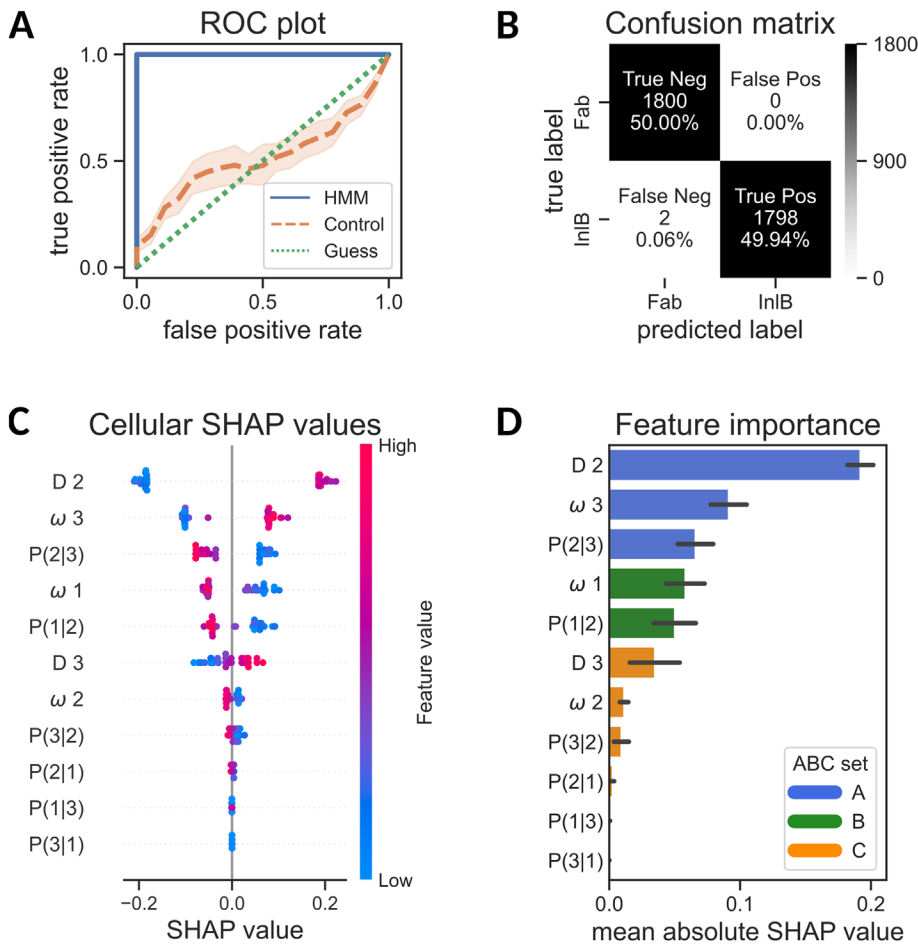


FIGURE 5: Machine learning of internalin B (InIB)-induced changes in MET mobility. (A) Receiver operating characteristic (ROC) curve calculated from the predictions for the validation data set (consisting of 36 cells, of which 18 were InIB-treated and 18 were untreated) made by a random forest classifier that was trained on the training data set (HMM, blue). To rule out random effects, the experiment was repeated 100 times with different seed-based classifier initializations. For the control experiment, the training data were permuted before training (Control, orange). The control experiment was repeated 100 times. The envelope represents the 95-confidence interval of the control experiment. The green line indicates an absolutely random guess. (B) Confusion matrix for the ability of the trained model to correctly classify Fab-labeled cells from the validation data set. Shapely additive explanations (SHAP) for the impact of the parameters on the classification task: (C) SHAP values for individual cells. The SHAP values indicate the relationship between a variable and a possible classification result for the individual cell. Here, positive SHAP values are indicative of untreated cells, while negative SHAP values are indicative of InIB-treated cells. The color code represents the relative variable value of the individual cell. Blue indicates a low value and red indicates a high value. As an example, cells with a high value in D 2 are more likely to be classified as untreated. (D) The model parameters are ranked by their importance for the classification task, which is calculated by the mean absolute SHAP values. Here, a higher absolute mean SHAP value indicates a higher parameter importance. The parameters are categorized by their importance based on an ABC analysis (set A: blue, set B: green, set C: orange).

and InIB-treated cells. These observations can be explained by InIB-induced MET dimerization, signaling transduction, and receptor internalization as described above.

It should be mentioned that a HMM will only reliably detect transitions from data if the transition rate is low enough on one hand so that assuming transitions within observation times is a reasonable approximation, but high enough on the other hand so that one will see some transitions in the data. If this is not the case, the experimental design must be adjusted. In the case of high transition rates,

the integration time of the camera can be reduced. In the case of low transition rates, the observation time can be increased either by introducing stable fluorescent probes such quantum dots instead of organic fluorophores (You *et al.*, 2014; Fricke *et al.*, 2015) or by reducing the probability of photobleaching events by using photostabilizing buffer systems (Vogelsang *et al.*, 2008; Wilmes *et al.*, 2015). The biocompatibility of both approaches should be verified to exclude an influence of the measuring system on the reaction to be investigated (Abraham *et al.*, 2017). If it is not possible to learn the transition probabilities reliably from the experimental data, it is also possible to apply the proposed method using only the parameters from the diffusion mixture model. For the MET SPT data set analyzed in this study, it was shown that changes in mobility modes in single-molecule trajectories can be observed by a combined application of segmentation and mean squared displacement analysis (Rahm *et al.*, 2021).

Post hoc modification of molecular mobility learned from the HMM should also be done with caution. The determination of the localization error via photon statistics is only an estimation (Savin and Doyle, 2005; Mortensen *et al.*, 2010). In the present study, the expected MSD caused by the localization error estimated by the photon distribution exceeds the measured expected MSD, which means that the method overestimates the true localization error, which in turn would result in a negative diffusion coefficient after error correction. This is why the localization error was estimated by the NeNA probability density function instead (Endesfelder *et al.*, 2014). Furthermore, the interpretation of a slow state as immobile is only permissible if the expected MSD caused by the localization error is greater than or equal to the measured expected MSD. In the case where the expected MSDs of all modes exceed the expected MSD caused by the localization error, the method is not applicable.

The procedure is highly modular. It is therefore possible to introduce alternative methods for each step depicted in Figure 1. Alternative methods for the acquisition of single-molecule positions and alternative tracking methods are available (Chenouard *et al.*, 2014; Sage *et al.*, 2015). Regarding the modeling step, there are several alternative approaches to the data-driven estimation of mobility modes (Linden and Elf, 2018; Falcao and Coombs, 2020; Karlsake *et al.*, 2021), as well as alternative HMMs that can handle diffusion types apart from Brownian motion (Chen *et al.*, 2021) and post hoc modifications of diffusivities associated with mobility modes (Michalet and Berglund, 2012). There are also alternative methods for the algorithm-independent interpretation of classifier decisions (Ribeiro *et al.*, 2016).

HMM parameter	Mean absolute SHAP value	ABC category
D_2	0.192 ± 0.010	A
ω_3	0.091 ± 0.014	A
$P(2 3)$	0.066 ± 0.014	A
ω_1	0.058 ± 0.015	B
$P(1 2)$	0.050 ± 0.016	B
D_3	0.035 ± 0.019	C
ω_2	0.011 ± 0.003	C
$P(3 2)$	0.009 ± 0.006	C
$P(2 1)$	0.003 ± 0.001	C
$P(1 3)$	<0.001	C
$P(3 1)$	<0.001	C

TABLE 3: ABC analysis–based identification of the most important parameters for the classification task. The importance of individual HMM parameters to the classification task of identifying untreated and internalin B–treated cells was determined by calculating the absolute mean SHAP value for each parameter. The values are given as mean \pm SD. The parameters were ordered by descending SHAP values and item categorization was performed using ABC analysis (see Supplemental Figure 6). The items rated A were considered very important for the classification task.

Given the abundance of possible algorithms, their optimal composition for a wide range of experimental conditions is proving to be a challenge in this research field.

In conclusion, the proposed method provides a modular reproducible interpretation of the results of HMM-based analyses of SPT data by attributing ligand treatment to a change in individual model parameters. On the biological side, the procedure is not limited to receptor tyrosine kinases and thus will provide new insights into membrane bound receptor–ligand interaction in general. However, the classification performances that will be achieved for different receptor–ligand pairs may vary. The cell-specific SHAP value distributions of feature fraction “A” (Figure 5C) showed a small overlap between activated and nonactivated cells. For feature subsets with a higher overlap between the SHAP value distributions, the classification performance dropped (Supplemental Figure 7). Because the proposed method determines and interprets differences from measurements of single-molecule mobilities, it in turn assumes that there are differences. Therefore, as long as receptor activation causes a measurable change in mobility, the classification procedure will probably be able to learn the differences.

METHODS

[Request a protocol](#) through *Bio-protocol*.

Experimental setup

The single-molecule localization microscopy data used within this study were previously published by our group (Harwardt *et al.*, 2017). A detailed description of the experiments can be found in the original publication. Briefly, HeLa cells were cultivated on glass coverslips with a thickness of 0.17 mm that were functionalized with poly-L-lysine-grafted polyethylene glycol modified with a CGRGDS peptide (VandeVondele *et al.*, 2003). The membrane-bound MET receptors were either stained by a 3H3-Fab antibody fragment that is able to bind but not activate MET or by the bacterial ligand InIB₃₂₁ that activates MET and initiates signaling pathways and endocytosis,

similarly to the physiological ligand HGF/SF. Both ligands were chemically coupled to ATTO 647N (ATTO-TEC, Siegen, Germany).

Data acquisition was performed using an N-STORM microscope (Nikon, Duesseldorf, Germany) equipped with a 647-nm laser, an oil immersion objective with a numerical aperture of $NA = 1.49$ (100 \times Apo TIRF oil), and an electron-multiplying charge-coupled device camera (DU-897U-CS0-BV; Andor Technology, Belfast, United Kingdom). Receptors labeled with ATTO 647N were excited by illuminating the sample with an evanescent field at an intensity of 0.1 kW/cm² in total internal reflection mode. The fluorescence signal was collected by the same objective, cleaned by appropriate optical filters, and detected by the camera. The image size was set to 256 \times 256 pixels with a pixel size of 0.158 μ m. The camera’s electron-multiplying gain was set to 200 and the integration time per frame was 20 ms. For each cell, a movie of 1000 frames was recorded without temporal delay between frames. A total of 117 cells was measured. Of these, 60 cells were labeled with 3H3-Fab and 57 with InIB.

Computational setup

All experiments were performed on an Intel Xeon E5-2620 v3 (12) at 3.200 GHz equipped with 128 GB of memory and a GeForce GTX 750 Ti GPU running Linux (openSUSE Leap 15.2). Machine learning on the GPU was enabled using the CUDA library (v. 10.2). Data-science pipelines are implemented in the python language (python 3.7.7) using the libraries matplotlib (Hunter, 2007) (v. 3.3.1), numpy (Harris *et al.*, 2020) (v. 1.19.1), pandas (McKinney *et al.*, 2010) (v. 1.1.1), seaborn (Waskom, 2021) (v. 0.11.1), scikit-learn (Pedregosa *et al.*, 2011) (v. 0.23.2), SciPy (Virtanen *et al.*, 2020) (v. 1.5.2), tensorflow (Abadi *et al.*, 2016) (v. 2.2.0), and xgboost (Chen and Guestrin, 2016) (v. 1.3.3). Experimental development of data science pipelines was performed using the Spyder-IDE (v. 5.0.0). Hidden Markov models were trained and stochastic PNs were simulated in parallel on multiple cores using GNU parallel (Tange, 2018) (v. 20180422).

Analysis of single-molecule localization microscopy data

Single-molecule localization. The positions of isolated MET ligand complexes were determined by modeling the photon distribution emitted from the ligand bound fluorescent probe with a normal distribution using the ThunderSTORM (Ovesny *et al.*, 2014) plugin of the free image processing software suite ImageJ (Schindelin *et al.*, 2015). Model parameter optimization was performed using maximum likelihood estimation. The option “Multi-emitter fitting analysis” was enabled with the parameter “maximum numbers of molecules per fitting region” set to 3. The parameter “limit intensity range” was set to the 2σ interval of the photon distribution in log-space of localizations found with “multi-emitter fitting analysis” disabled. The “remove duplicates filter” was applied. The quality of the experimental data was determined by means of the average localization precision.

Single-particle tracking. The localizations of individual MET–ligand complexes in consecutive frames were connected to trajectories using the swift tracking software (Turkowsky *et al.*, 2020; Version 0.4.2, used in this manuscript, and all subsequent versions of the swift software, as well as documentation and test data sets, can be obtained on the swift beta-testing repository <http://bit.ly/swiftracking> and upon request to the authors). Parameters for swift were estimated using the SPTAnalyser software (<https://github.com/JohannaRahm/SPTAnalyser>). The following parameters were set globally for all cells: “diffraction_limit” = 14 nm, “exp_displacement” = 85 nm (Fab) / 75 nm (InIB), “p_bleach” = 0.01 (Fab) / 0.014 (InIB), and “p_switch” = 0.01. The parameters “exp_noise_rate” and “precision” were determined individually per cell.

Hidden Markov modeling of single-particle tracking data

Hidden Markov modeling was performed using the *ermine* (estimate reaction rates by Markov-based investigation of nanoscopy experiments) package for python (<https://github.com/SMLMS/py-Ermine>). HMMs were used to map the diffusion behavior of membrane-bound MET receptors measured at discrete time points equally separated by Δt onto a hidden transition network. This transition network is defined by a set of N states $S = \{s_1, \dots, s_N\}$. The starting probability for each state is given by $\pi = \{\pi_1, \dots, \pi_N\}$ with $\sum_{i=1}^N \pi_i = 1$. The probability of interstate transitions between consecutive discrete time steps is given by an $N \times N$ -state transition probability matrix $A = \{a_{i,j}\}$. Given a discrete time sequence $H = (h_{\Delta t}, \dots, h_{M\Delta t})$ of M hidden states drawn from S , a state at time t is denoted by $h_t = s_i$. The probability of a state transition between two consecutive entries of H is given by $P(h_{t+\Delta t} = s_j | h_t = s_i)$ with $\sum_{j=1}^N P(h_{t+\Delta t} = s_j | h_t = s_i) = 1$. Therefore, the state of h_t solely depends on the state of $h_{t-\Delta t}$. The observed sequence of the HMM $O = (o_{\Delta t}, \dots, o_{M\Delta t})$ is a sequence of M Euclidean point-to-point distances of the molecules' displacements between two consecutive discrete time points and will be referred to as the jump distance sequence. A jump distance at time t is denoted by $o_t = r$. Each hidden state in S is associated with an individual diffusion model $B = \{\langle r^2 \rangle_1, \dots, \langle r^2 \rangle_N\}$ that is characterized by its expected MSD $\langle r^2 \rangle$. The probability of observing $o_t = r$ is given by $P(o_t = r, \Delta t | h_t = s_i)$. It follows a Markov property (Rabiner, 1989) and is therefore solely dependent on h_t (see equation 6). The parameter space of a HMM is given by $\Theta = (\pi, A, B)$ and has $\text{dof} = N - 1 + N(N - 1) + N \text{degrees of freedom}$.

Data preprocessing. Before modeling, the single-particle localization trajectories obtained from the *swift*-based tracking analysis were transformed into jump distance trajectories that are readable by the *ermine* package. For each particle's localization trajectory, the Euclidean point-to-point distances between all directly successive localizations in time were calculated. These distance trajectories are further referred to as jump distance trajectories. The calculation of jump distance trajectories from localization-based trajectories was performed using the "preprocess_swift_data" function of the *ermine* package. Jump distance trajectories with less than four jumps were filtered from the data by setting the "min_track_length" parameter to 4, as short trajectories seldom exhibit state transitions (Rahm et al., 2021).

Initial model parameter determination. Initial diffusion models B associated with the hidden states S were determined by optimizing the probability of observing the distribution of jump distances given a diffusion mixture model

$$\omega^{\text{init}}, B^{\text{init}} = \arg \text{Max} \left(P(r | \omega, B) = \sum_{i=1}^N \omega_i P(r, \Delta t | \langle r^2 \rangle_i) \right)$$

(Supplemental Material: Appendix A, Equations 1–7). The optimization process was performed using the expectation-maximization algorithm (Dempster et al., 1977) implemented in the "JumpDistanceMixtureModel" class of the *ermine* package with B and ω as free fit parameters. The initial starting probability of the HMM π^{init} was set to ω^{init} . Further, the states were assumed to be relatively stable. Therefore, the entries of the diagonal of the state transition matrix A^{init} were initially set to $P(h_{t+\Delta t} = s_i | h_t = s_{j=i}) = 0.9$ while the remaining entries of the state transition matrix were filled with a uniform distribution $P(h_{t+\Delta t} = s_j | h_t = s_{j \neq i}) = (1 - 0.9) / (N - 1)$.

Model parameterization. Starting with the initial HMM parameters $\Theta^{\text{init}} = (\pi^{\text{init}}, A^{\text{init}}, B^{\text{init}})$, the likelihood for observing the measured

jump distance trajectories O was maximized using the Baum–Welch algorithm $\arg \text{Max} P(O | \Theta)$ to obtain the model parameters with multiple observations $\Theta^{\text{opt}} = (\pi^{\text{opt}}, A^{\text{opt}}, B^{\text{opt}})$ (Rabiner, 1989; Li et al., 2000). The optimization process was performed using the "ErmineHMM" class of the "ermine" package, which uses a diffusion mixture HMM that is designed using the `_BaseHMM` class of the `hmmlearn` package (<https://github.com/hmmlearn/hmmlearn>) with a diffusion mixture model as custom emission probability.

Reaction state sequence estimation. For the optimal HMM, the most likely sequence of hidden states H^{opt} underlying the observation sequence of jump distances O was postulated using the Viterbi algorithm (Viterbi, 1967). Reaction state sequence estimation was performed using the "predict" function of the "ErmineHMM" class of the *ermine* package.

Model selection

The capability of different HMMs to describe a given set of jump distance trajectories was measured using the Bayesian information criterion (Schwarz, 1978) $BIC = k \ln(n) - 2 \ln(\hat{L}(O | \Theta^{\text{opt}}))$. Here, $k = N - 1 + N(N - 1) + N$ is the number of estimated parameters, n is the number of observed jump distances, and $\hat{L}(O | \Theta^{\text{opt}})$ is the maximized likelihood of observing O with HMM Θ^{opt} . The favoured model is the HMM that minimizes the BIC.

The number of estimated parameters falls below the number of HMM parameters due to the stochastic constraints of the optimized parameters as defined by Rabiner (1989). Based on these constraints, the initial mode population probability for state N can be calculated from the remaining initial mode probabilities ($\pi_N^{\text{opt}} = 1 - \sum_{i=1}^{N-1} \pi_i^{\text{opt}}$) and the probabilities that no transition is observed in two consecutive observation frames can be calculated from the probabilities that a transition occurs ($a_{i,j}^{\text{opt}} = 1 - \sum_{i=1}^N a_{i,j}^{\text{opt}}$ for $i \neq j$).

An alternative approach to model comparison is the likelihood ratio test (Neyman and Pearson, 2020). It legitimizes the use of a more complex model if increasing the model complexity is accompanied by a significant increase in the likelihood ratio ($\alpha = 0.05$). Here, the models are compared with each other along ascending complexity. As soon as a model does not satisfy the test parameter, the search is stopped and the most complex model that meets the search criterion is favored.

Using these two tests, the minimum number of states was determined by a majority vote, based on which there is an HMM for each cell that describes the diffusion of the MET receptor sufficiently well.

Post hoc hidden Markov model parameter analysis

Mobility mode population estimation. The overall populations of the mobility modes during the measurement ω^{opt} were determined by optimizing the likelihood of a diffusion mixture model $\omega^{\text{opt}} = \arg \text{Max} P(r | \omega, B^{\text{opt}})$ while keeping B^{opt} fixed.

Limit of detection estimation. The LOD was defined as the expected apparent MSD of an immobile particle ($\langle r, m^2 \rangle$). The LOD was determined by calculating the average localization error (σ) from the fluorescent probes photon distributions (Mortensen et al., 2010; LOD = $4\sigma^2$; see Supplemental Material: Appendix B). All HMM modes with an expected apparent MSD below the LOD were assumed to be immobile ($D : = 0 \frac{\mu\text{m}^2}{\text{s}}$).

Diffusion coefficient estimation. The mobility modes of the HMM (B) are directly characterized by a specific MSD ($\langle r^2 \rangle$), which is

defined by the state's diffusion coefficient and a global static error (ϵ) shared by all modes (see Supplemental Material: Appendix B). To assign a specific diffusion coefficient to each state, ϵ needs to be known. For this purpose the diffusion coefficient of the immobile state was defined as $D_{im} = 0 \frac{\mu\text{m}^2}{\text{s}}$. D_{im} and $\langle r_{im}^2 \rangle$ were inserted into $\langle r^2 \rangle = 4D \left(t - \frac{v}{3} \right) + 4\epsilon^2$ (Savin and Doyle, 2005) to obtain the static error ϵ . With ϵ at hand, the diffusion coefficients of the remaining modes were calculated similarly, resulting in the diffusion coefficient vector D^{opt} .

Anomaly detection. Anomalies within the two groups of the defined data set were identified by multivariate analyses using either an isolation forest (Liu et al., 2008) or the DBSCAN algorithm (Ester et al., 1996) trained on the features ω^{opt} , π^{opt} , A^{opt} , D^{opt} . Prior to algorithm training, the features were scaled into a range of [0, 1] by

$$\text{Min-Max scaling } X_s = \frac{X - \min(X)}{\max(X) - \min(X)}$$

analyses were visualized by centering the feature values to 0 and reducing the data set's dimensionality to two using principal component analysis (PCA) (Pearson, 1901) and plotting the first two principal components (PCs). Instances characterized as outliers by the isolation forest or as noise by DBSCAN analysis were discarded from the data set.

Parameter exploration

Parameterwise determination of differences between groups using hypothesis tests. For groupwise comparison of the parameter distributions, the choice of an appropriate hypothesis test was based on the distribution of the two test variables, as well as their variances. As most variables were not normally distributed (tested by the D'Agostino test; d'Agostino, 1971; D'Agostino and Pearson, 1973) and differed in their variances (tested by Levene's test for equal variances; Brown and Forsythe, 1974), the Mann-Whitney U test was used (Mann and Whitney, 1947). The significance level was set to $\alpha = 0.01$. The p -values were corrected for multiple testing using the Bonferroni method ($m p \leq \alpha$; $m = \text{number of tests}$; Bonferroni, 1936).

Demonstration of pairwise parameter interdependence. Spearman's rank correlation coefficient (Beyer, 1991) was calculated pairwise between all features of the data set. The analysis was performed individually for Fab- and InIB-labeled instances.

A generative approach to demonstrate the interdependence of the hidden Markov model parameters

Defining the chemical master equation from transition probabilities. The machine-learned HMM parameter A^{opt} defines the probability of a single molecule switching its diffusive behavior within the time span of two consecutive measurements $P(j|i)$. In a HMM, these transition probabilities are temporally stable over the measurement period and independent of the instantaneous weighting of the states S . Therefore, the transition matrix A^{opt} can be interpreted as a closed system of first-order chemical reactions.

The reaction rate constant $k_{i \rightarrow j} [s^{-1}]$ was calculated from Δt and $P(j|i)$ (Supplemental Material: Appendix C). With the reaction rate constants ($k_{(i \rightarrow j)} [s^{-1}]$ with $i \neq j$) at hand, the system is then fully described by the chemical master equation (CME)

$$\frac{dx_i}{dt} = \sum_{j \neq i} (j \rightarrow i) [x_j] - k(i \rightarrow j) [x_i]$$

Stochastic simulation of the chemical master equation. The CME was transformed into a stochastic PN (Bause and Kritzing, 1999) and Monte Carlo simulations of the state trajectories were simulated using the Stochastic Simulation Algorithm (SSA; Gillespie, 1977) as implemented in the *fossa* (free objective-c stochastic simulation algorithm) software (<https://github.com/SMLMS/fossa>). For first-order chemical reactions the reaction rate constant equals the stochastic rate constant ($k_{(i \rightarrow j)} [s^{-1}] = c_{(i \rightarrow j)} [s^{-1}]$ with $i \neq j$; Gillespie, 2007). For each cell a stochastic PN was set up and parameterized by calculating k^{sim} from HMM learned transition matrix A^{opt} . Simulations were performed on ensembles comprising 1000 molecules, all of which were defined as free at the initiation of the experiment ($t = 0$ s). The chosen simulation period guaranteed that an equilibrium was reached toward the end.

Validation of the machine-learned chemical master equation by verification of the effect of receptor labeling on simulated diffusion state populations.

To prove whether the stochastic PNs that were parameterized using the HMM learned transition probabilities could reproduce the InIB-caused repopulation of HMM states, the equilibrium state population of the stochastically simulated CMEs was calculated as the mean population of the last 1000 simulation time steps measured in the steady state (ω^{sim}). The repopulation effect was qualified by determining the differences in state populations using hypothesis tests (see *Parameterwise Determination of Differences between Groups Using Hypothesis Tests*). The simulated effect of InIB-mediated cell activation was compared with the InIB-mediated effect learned from the data (ω^{opt}).

Machine learning associated model interpretation

Data set definition. A labeled data set was created from the extracted physiologically interpretable parameters obtained by HMM and diffusion mixture model training with optimal mobility mode number. From these optimized parameters (ω^{opt} , A^{opt} , D^{opt}), the mode occupation probability (ω^{opt}), the transition probabilities of a molecule switching between mobility modes ($P(j|i)$ with $i \neq j$), and the diffusion coefficients that characterize the mobility modes ($D_{i,i}^{\text{opt}}$, $i \neq \text{immobile}$) were defined as features and the measured cells as instances. Therefore, the resulting data set of an N -modal HMM comprises $n = N + (N^2 - N) + (N - 1)$ extracted physiologically interpretable parameters.

Model training. The data set was split in a two to one ratio into a training set and a validation set. Care was taken to maintain the ratio of activated to nonactivated instances in training and validation data. The hyperparameters of four different classifiers—CART (Breiman et al., 2017), RF (Ho, 1995), XG-Boost (Chen and Guestrin, 2016), and ANN in the form of a multilayer perceptron (Bradley, 1995)—were optimized by performing a randomized grid search through the hyperparameter space, during which the model performance was measured by using the balanced accuracy as a metric. Hyperparameter optimization was performed using a stratified 10-fold cross validation on the training set to guarantee that the model never saw instances of the validation set during training.

Model validation. The performance of the four classifiers with optimized hyperparameters was measured by training the models on the complete training set and measuring the model's capability to predict the classes of the validation set correctly by calculating the balanced accuracy (an accuracy score that is not affected by class imbalances within the data; Brodersen et al., 2010), ROC-AUC score, F1 score, precision score, and recall score (Branco et al., 2015);

Luque *et al.*, 2019; Supplemental Material: Appendix D). To reduce random effects caused by model initialization values, the procedure was repeated 100 times with randomly chosen model initialization.

To verify the learning capability, the models were trained on the training set with permuted feature values. This procedure was repeated 100 times with different permutations. It is expected that the models cannot learn anything from the permuted training data.

The model with the optimal performance on the validation set is chosen for further analyses.

Model interpretation and feature importance. Model interpretation was performed post hoc using the SHAP framework (Lundberg and Lee, 2017a), which calculates cell-individual Shapley additive explanation values for each feature. The feature importance was given by the mean absolute SHAP value of a feature. The higher the mean absolute SHAP value of a feature, the more the feature contributes to the decision of the binary classifier. Based on the SHAP value distribution, the feature set yielding the optimal amount of information about the data was identified by ABC analysis, which identifies the features that achieve the maximum effect with the lowest possible effort. This analysis is a generalization of the Pareto 80/20 rule, which claims that 80% of the results are achieved with 20% of the total effort.

To verify the model interpretation, the classification procedure was repeated with a reduced data set that only contained the HMM parameters rated as “A” by the ABC analysis. The classification performance was evaluated using the balanced accuracy, ROC-AUC score, F1 score, precision score, and recall score. To exclude effects due to general dimension reduction, the validation procedure was repeated 100 times with randomly chosen HMM parameter subsets.

Package implementation and data availability

Ermine. The software for estimation of reaction rates by Markov-based investigation of nanoscopy experiments (*ermine*) is implemented as a python library and thus requires the python software package (<https://pypi.org/project/pyErmine/>). The methods included in the *ermine* library can be used in python scripts to produce individual problem-specific workflows or in interactive jupyter notebooks that follow a specific preprocessing workflow. The *ermine* package is free and open source software. The source code is freely available at GitHub (<https://github.com/SMLMS/pyErmine>) and can be redistributed and/or modified under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 (GPLv3) of the License, or any later version. The programming work was performed in the python language and is based upon the library *hmmlearn* (v. 0.2.4). Tutorials on the usage of the *ermine* package can be obtained from our github site (<https://github.com/SMLMS/ermine-tutorial>).

fossa. The free open-source stochastic simulation algorithm (*fossa*) was written in objective-c for the purpose of simulation execution speed (Blankenbecler, 1990). It uses the open source standard library framework “gnustep” (v. 1.25.1; Botto *et al.*, 2001) and was compiled using *gnustep-make* (v. 2.7.0) and *gcc* (v. 7.5.0). The “*fossa*” package is free and open-source software. The source code is freely available at GitHub (<https://github.com/SMLMS/fossa>) and can be redistributed and/or modified under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 (GPLv3) of the License, or any later version.

SPT data. The SPT data of InIB-treated and -untreated cells are available from <https://www.ebi.ac.uk/biostudies/studies/S-BSST712> (Heilemann *et al.*, 2021).

MET diffusion hidden Markov model parameter data. The data set of hidden Markov model parameters of MET diffusion is available as a text document within the Supplemental Material of the article (“*met_diffusion_hmm_parameters.csv*”).

ACKNOWLEDGMENTS

We thank Ulrike Endesfelder for providing the *swift* software. We thank Jakob T. Bullerjahn for fruitful discussions about single particle-tracking theory. We thank Jörg Ackermann for the kind introduction in stochastic PNs. S.M. acknowledges funding by the Federal Ministry of Education and Research (BMBF) in the form of a “Forschungspraktikum für Systembiologen” (Grant e:Bio). M.H., J.V.R., and M.S.D. acknowledge funding by the German Science Foundation (Grant iMol, RTG 2566). J.B.S. acknowledges funding by the National Infrastructure France-Bioluminescence supported by the French National Research Agency (ANR-10-INBS-04). J.L. acknowledges funding by the German Science Foundation (Grant DFG LO 612/16-1).

REFERENCES

- Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, *et al.* (2016). Tensorflow: a system for large-scale machine learning. 12th USENIX symposium on operating systems design and implementation (OSDI 16), Savannah, USA, 265–283.
- Abraham L, Lu HY, Falcao RC, Scurll J, Jou T, Irwin B, Tafteh R, Gold MR, Coombs D (2017). Limitations of Qdot labelling compared to directly-conjugated probes for single particle tracking of B cell receptor mobility. *Sci Rep* 7, 11379.
- Axelrod D, Koppel DE, Schlessinger J, Elson E, Webb WW (1976). Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys J* 16, 1055–1069.
- Baldering TN, Karathanasis C, Harwardt MIE, Freund P, Meurer M, Rahm JV, Knop M, Dietz MS, Heilemann M (2021). CRISPR/Cas12a-mediated labeling of MET receptor enables quantitative single-molecule imaging of endogenous protein organization and dynamics. *iScience* 24, 101895.
- Baum LE, Petrie T, Soules G, Weiss N (1970). A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics* 41, 164–171.
- Bause F, Kritzinger PS (1999). Stochastic Petri nets: An introduction to the theory. Vieweg+Teubner Verlag.
- Beyer WH (1991). Standard probability and statistics: tables and formulae. CRC Press.
- Blankenbecler R (1990). Object oriented programming for simulation problems in physics.
- Bonferroni C (1936). Teoria statistica delle classi e calcolo delle probabilità. Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze 8, 3–62.
- Botto F, Frith-Macdonald R, Pero N, Robert A (2001). Objective-C language and GNUstep base library programming manual. Free Software Foundation 2004.
- Bradley JB (1995). Book review: Neural networks: a comprehensive foundation: S. Haykin. *Inform Process Manag* 31, 786–786.
- Branco P, Torgo L, Ribeiro R (2015). A survey of predictive modelling under imbalanced distributions. *arXiv preprint arXiv:1505.01658*.
- Breiman L, Friedman JH, Olshen RA, Stone CJ (2017). Classification and regression trees. Routledge.
- Brodersen KH, Ong CS, Stephan KE, Buhmann JM (2010). The balanced accuracy and its posterior distribution. 2010 20th International Conference on Pattern Recognition, 3121–3124.
- Brown MB, Forsythe AB (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association* 69, 364–367.
- Chen T, Guestrin C (2016). Xgboost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.
- Chen Z, Geffroy L, Biteen JS (2021). NOBIAS: analyzing anomalous diffusion in single-molecule tracks with nonparametric Bayesian inference. *Front Bioinform* 1.

- Chenouard N, Smal I, de Chaumont F, Maska M, Sbalzarini IF, Gong Y, Cardinale J, Carthel C, Coraluppi S, Winter M, et al. (2014). Objective comparison of particle tracking methods. *Nat Methods* 11, 281–289.
- Chung I, Akita R, Vandlen R, Toomre D, Schlessinger J, Mellman I (2010). Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* 464, 783–787.
- D'Agostino R, Pearson ES (1973). Tests for departure from normality. empirical results for the distributions of b_2 and $\sqrt{b_1}$. *Biometrika* 60, 613–622.
- d'Agostino RB (1971). An omnibus test of normality for moderate and large size samples. *Biometrika* 58, 341–348.
- Das R, Cairo CW, Coombs D (2009). A hidden Markov model for single particle tracks quantifies dynamic interactions between LFA-1 and the actin cytoskeleton. *PLoS Comput Biol* 5, e1000556.
- Dempster AP, Laird NM, Rubin DB (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39, 1–22.
- Dietz MS, Fricke F, Kruger CL, Niemann HH, Heilemann M (2014). Receptor–ligand interactions: binding affinities studied by single-molecule and super-resolution microscopy on intact cells. *Chemphyschem* 15, 671–676.
- Dietz MS, Hasse D, Ferraris DM, Gohler A, Niemann HH, Heilemann M (2013). Single-molecule photobleaching reveals increased MET receptor dimerization upon ligand binding in intact cells. *BMC Biophys* 6, 6.
- Eddy SR (2004). What is a hidden Markov model? *Nat Biotechnol* 22, 1315–1316.
- Endesfelder U, Malkusch S, Fricke F, Heilemann M (2014). A simple method to estimate the average localization precision of a single-molecule localization microscopy experiment. *Histochem Cell Biol* 141, 629–638.
- Ester M, Kriegel H-P, Sander J, Xu X (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *KDD*, 96, 226–231.
- Falcao RC, Coombs D. (2020). Diffusion analysis of single particle trajectories in a Bayesian nonparametrics framework. *Phys Biol* 17, 025001.
- Fricke F, Dietz MS, Heilemann M (2015). Single-molecule methods to study membrane receptor oligomerization. *Chemphyschem* 16, 713–721.
- Furge KA, Zhang YW, Vande Woude GF (2000). Met receptor tyrosine kinase: enhanced signaling through adapter proteins. *Oncogene* 19, 5582–5589.
- Gillespie DT (1977). Exact stochastic simulation of coupled chemical-rReactions. *J Phys Chem-US* 81, 2340–2361.
- Gillespie DT (2007). Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* 58, 35–55.
- Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, et al. (2020). Array programming with NumPy. *Nature* 585, 357–362.
- Harwardt MLIE, Schroder MS, Li YQ, Malkusch S, Freund P, Gupta S, Janjic N, Strauss S, Jungmann R, Dietz MS, Heilemann M (2020). Single-molecule super-resolution microscopy reveals heteromeric complexes of MET and EGFR upon ligand activation. *Int J Mol Sci* 21, 2803.
- Harwardt MLIE, Young P, Bleytmuller WM, Meyer T, Karathanasis C, Niemann HH, Heilemann M, Dietz MS (2017). Membrane dynamics of resting and internalin B-bound MET receptor tyrosine kinase studied by single-molecule tracking. *Febs Open Bio* 7, 1422–1440.
- Hauastein E, Schwillie P (2007). Fluorescence correlation spectroscopy: novel variations of an established technique. *Annu Rev Bioph Biom* 36, 151–169.
- Heilemann M, Dietz MS, Rahm JV, Harwardt M-LIE (2021). Single-particle tracking (SPT) data of MET with a non-activating Fab ligand and the activating ligand InlB321 recorded in HeLa cells.
- Hinkle DE, Wiersma W, Jurs SG (2003). *Applied statistics for the behavioral sciences*. Houghton Mifflin Hi Marketing (distributor): Boston, MA. London.
- Ho TK (1995). Random decision forests. *Proceedings of 3rd International Conference on Document Analysis and Recognition*, 1, 278–282.
- Hunter JD (2007). Matplotlib: a 2D graphics environment. *Comput Sci Eng* 9, 90–95.
- Juran JM (1975). The non-Pareto principle: mea culpa. *Quality Progress* 8, 8–9.
- Karslake JD, Donarski ED, Shelby SA, Demey LM, DiRita VJ, Veatch SL, Biteen JS (2021). SMAUG: analyzing single-molecule tracks with non-parametric Bayesian statistics. *Methods* 193, 16–26.
- Lemmon MA, Schlessinger J (2010). Cell signaling by receptor tyrosine kinases. *Cell* 141, 1117–1134.
- Li N, Xiang GS, Dokainish H, Ireton K, Elferink LA (2005). The Listeria protein internalin B mimics hepatocyte growth factor-induced receptor trafficking. *Traffic* 6, 459–473.
- Li XL, Parizeau M, Plamondon R (2000). Training hidden Markov models with multiple observations—a combinatorial method. *IEEE T Pattern Anal* 22, 371–377.
- Linden M, Elf J (2018). Variational algorithms for analyzing noisy multistate diffusion trajectories. *Biophys J* 115, 276–282.
- Liu FT, Ting KM, Zhou Z-H (2008). Isolation forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422.
- Lotsch J, Malkusch S (2021). Interpretation of cluster structures in pain-related phenotype data using explainable artificial intelligence (XAI). *Eur J Pain* 25, 442–465.
- Low-Nam ST, Lidke KA, Cutler PJ, Roovers RC, Henegouwen PMPVE, Wilson BS, Lidke DS (2011). ErbB1 dimerization is promoted by domain co-confinement and stabilized by ligand binding. *Nat Struct Mol Biol* 18, 1244–U1288.
- Lundberg SM, Lee S-I (2017a). A unified approach to interpreting model predictions. *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 4768–4777.
- Lundberg SM, Lee SI (2017b). A unified approach to interpreting model predictions. *Adv Neur In* 30.
- Luque A, Carrasco A, Martín A, de las Heras A (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition* 91, 216–231.
- Magde D, Webb WW, Elson E (1972). Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy. *Phys Rev Lett* 29, 705–8.
- Mann HB, Whitney DR (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 50–60.
- McKinney W (2010). Data structures for statistical computing in python. *Proceedings of the 9th Python in Science Conference*, 445, 51–56.
- Michalet X, Berglund AJ (2012). Optimal diffusion coefficient estimation in single-particle tracking. *Phys Rev E Stat Nonlin Soft Matter Phys* 85, 061916.
- Mortensen KI, Churchman LS, Spudich JA, Flyvbjerg H (2010). Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nat Methods* 7, 377–381.
- Mukaka MM (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 24, 69–71.
- Neyman J, Pearson ES (2020). On the use and interpretation of certain test criteria for purposes of statistical inference. Part I. University of California Press.
- Niemann HH (2013). Structural basis of MET receptor dimerization by the bacterial invasion protein InlB and the HGF/SF splice variant NK1. *Biochim Biophys Acta* 1834, 2195–2204.
- Orian-Rousseau W, Morrison H, Matzke A, Kastilan T, Pace G, Herrlich P, Ponta H (2007). Hepatocyte growth factor-induced Ras activation requires ERM proteins linked to both CD44v6 and F-actin. *Mol Biol Cell* 18, 76–83.
- Ott M, Shai Y, Haran G (2013). Single-particle tracking reveals switching of the HIV fusion peptide between two diffusive modes in membranes. *J Phys Chem B* 117, 13308–13321.
- Ovesny M, Krizek P, Borkovec J, Svindrych ZK, Hagen GM (2014). ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics* 30, 2389–2390.
- Pearson K (1901). On lines and planes of closest fit to systems of points in space. *Philos Mag* 2, 559–572.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. (2011). Scikit-learn: machine learning in Python. *J Mach Learn Res* 12, 2825–2830.
- Rabiner LR (1989). A tutorial on hidden Markov-models and selected applications in speech recognition. *P IEEE* 77, 257–286.
- Rahm JV, Malkusch S, Endesfelder U, Dietz MS, Heilemann M (2021). Diffusion state transitions in single-particle trajectories of MET receptor tyrosine kinase measured in live cells. *Front Comput Sci* 3.
- Ribeiro MT, Singh S, Guestrin C (2016). “Why should I trust you?” Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144.
- Sage D, Kirshner H, Pengo T, Stuurman N, Min J, Manley S, Unser M (2015). Quantitative evaluation of software packages for single-molecule localization microscopy. *Nat Methods* 12, 717–724.
- Savin T, Doyle PS (2005). Static and dynamic errors in particle tracking microrheology. *Biophysical Journal* 88, 623–638.
- Saxton MJ, Jacobson K (1997). Single-particle tracking: applications to membrane dynamics. *Annu Rev Bioph Biom* 26, 373–399.

- Schindelin J, Rueden CT, Hiner MC, Eliceiri KW (2015). The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol Reprod Dev* 82, 518–529.
- Schober P, Boer C, Schwarte LA (2018). Correlation coefficients: appropriate use and interpretation. *Anesth Analg* 126, 1763–1768.
- Schwarz G (1978). Estimating the dimension of a model. *Annals of statistics*, 461–464.
- Seveau S, Bierne H, Giroux S, Prevost MC, Cossart P (2004). Role of lipid rafts in E-cadherin- and HGF-R/Met-mediated entry of *Listeria monocytogenes* into host cells. *J Cell Biol* 166, 743–753.
- Shen H, Tautzin LJ, Baiyasi R, Wang W, Moringo N, Shuang B, Landes CF (2017). Single particle tracking: from theory to biophysical applications. *Chem Rev* 117, 7331–7376.
- Shen Y, Naujokas K, Park M, Ireton K (2000). InlB-dependent internalization of *Listeria* is mediated by the Met receptor tyrosine kinase. *Cell* 103, 501–510.
- Slator PJ, Cairo CW, Burroughs NJ (2015). Detection of diffusion heterogeneity in single particle tracking trajectories using a hHidden Markov model with measurement noise propagation. *PLoS One* 10, e0140759.
- Sungkaworn T, Jobin ML, Burnecki K, Weron A, Lohse MJ, Calebiro D (2017). Single-molecule imaging reveals receptor-G protein interactions at cell surface hot spots. *Nature* 550, 543–.
- Tange O (2018). GNU Parallel 2018. Lulu. com.
- Trenker R, Jura N (2020). Receptor tyrosine kinase activation: from the ligand perspective. *Curr Opin Cell Biol* 63, 174–185.
- Trusolino L, Bertotti A, Comoglio PM (2010). MET signalling: principles and functions in development, organ regeneration and cancer. *Nat Rev Mol Cell Bio* 11, 834–848.
- Turkowsky B, Schreiber S, Wortz J, Segal ES, Mevarech M, Duggin IG, Marchfelder A, Endesfelder U (2020). Establishing live-cell single-molecule localization microscopy imaging and single-particle tracking in the archaeon *Haloferax volcanii*. *Front Microbiol* 11, 583010.
- VandeVondele S, Voros J, Hubbell JA (2003). RGD-grafted poly-L-lysine-graft-(polyethylene glycol) copolymers block non-specific protein adsorption while promoting cell adhesion. *Biotechnol Bioeng* 82, 784–790.
- Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python (vol 33, pg 219, 2020). *Nature Methods* 17, 352–352.
- Viterbi AJ (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *Ieee T Inform Theory* 13, 260–269.
- Vogelsang J, Kasper R, Steinhauer C, Person B, Heilemann M, Sauer M, Tinnefeld P (2008). A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. *Angew Chem Int Ed Engl* 47, 5465–5469.
- Waskom ML (2021). Seaborn: statistical data visualization. *Journal of Open Source Software* 6, 3021.
- Wilmes S, Beutel O, Li Z, Francois-Newton V, Richter CP, Janning D, Kroll C, Hanhart P, Hotte K, You C, et al. (2015). Receptor dimerization dynamics as a regulatory valve for plasticity of type I interferon signaling. *J Cell Biol* 209, 579–593.
- You C, Richter CP, Lochte S, Wilmes S, Piehler J (2014). Dynamic submicroscopic signaling zones revealed by pair correlation tracking and localization microscopy. *Anal Chem* 86, 8593–8602.
- Zhao R, Yuan JH, Li N, Sun YH, Xia T, Fang XH (2019). Analysis of the diffusivity change from single-molecule trajectories on living cells. *Anal Chem* 91, 13390–13397.