Full title:

FISIK: Framework for the Inference of in Situ Interaction Kinetics from single-molecule imaging data

Short title:

Interaction kinetics from single-molecule imaging data

Luciana R. de Oliveira[1] and Khuloud Jaqaman[1,2]

[1]Department of Biophysics, [2]Lyda Hill Department of Bioinformatics, UT Southwestern Medical Center, Dallas, TX

Abstract

Recent experimental and computational developments have been pushing the limits of live-cell single-molecule imaging, enabling the monitoring of inter-molecular interactions in their native environment with high spatiotemporal resolution. However, interactions are captured only for the labeled subset of molecules, which tends to be a small fraction. As a result, it has remained a challenge to calculate molecular interaction kinetics, in particular association rates, from single-molecule tracking data. To overcome this challenge, we developed a mathematical modeling-based Framework for the Inference of in Situ Interaction Kinetics from single-molecule imaging data (termed “FISIK”). FISIK consists of (1) devising a mathematical model of molecular movement and interactions, mimicking the biological system and data-acquisition setup, and (2) estimating the unknown model parameters, including molecular association and dissociation rates, by fitting the model to experimental single-molecule data. Due to the stochastic nature of the model and data, we adapted the method of indirect inference for model calibration. We validated FISIK using a series of tests, where we simulated trajectories of diffusing molecules that interact with each other, considering a wide range of model parameters, and including resolution limitations and tracking errors. We found that FISIK has the sensitivity to determine association and dissociation rates, where its accuracy depends on the labeled fraction of molecules and the extent of molecule tracking errors. For cases where the labeled fraction is relatively low (e.g. to afford accurate tracking), combining dynamic but sparse single-molecule imaging data with almost whole-population oligomer distribution data improves FISIK’s performance. All in all, FISIK is a promising approach

for the derivation of molecular interaction kinetics in their native environment from single-molecule imaging data.

Introduction

Light microscopy is providing information on molecular activities in cells with ever-increasing sensitivity and spatial and temporal resolution. Modalities such as single-molecule imaging provide the highest sensitivity and resolution, reporting on the dynamic activities of individual molecules. Recent experimental and computational developments have been pushing the limits of single-molecule imaging to capture not only molecular movement – the traditional strength of single-molecule imaging [1-3] – but also inter-molecular interactions [4-9]. Capturing molecular interactions in their native environment provides critical spatial and temporal context for the observed interactions. For example, receptor interactions in their native plasma membrane environment can have different kinetics from the interactions observed via traditional 3D in vitro methods [5, 10, 11]. These kinetics might also depend on the cellular environment and subcellular context [12]. Additionally, microscopy approaches allow the study of interactions between proteins, such as transmembrane proteins, that are difficult to purify for traditional biophysical studies.

However, most single-molecule approaches, from direct one-color and multi-color imaging [4, 6-8, 13, 14] to single-molecule Förster resonance energy transfer (smFRET) [5, 15], suffer from the inherent limitation that only a small fraction of molecules can be visualized at a time. Such sub-stoichiometric labeling is necessary because of diffraction, which results in each fluorescent molecule appearing as a Point Spread Function (PSF) in the image. Therefore, to visualize individual molecules, the labeled molecules must be separated by ~0.5 μm and ~1 μm from each other in the lateral and axial directions,

respectively, putting an upper limit on the labeled molecule density of, at best, 4 molecules/μm^3 in 3D imaging and similarly 4 molecules/μm^2 in 2D imaging. Additional limitations are imposed by the need to also track the imaged molecules over time, thus generally requiring even sparser labeling [16].

As a consequence of sub-stoichiometric labeling, it is challenging to calculate molecular interaction rates from single-molecule experiments, in spite of the rich information they might provide. This is because only a small subset of interaction events can be captured in single-molecule imaging experiments. Importantly, this subset is biased toward oligomers/complexes/clusters with a smaller number of molecules. For example, if 10% of molecules are labeled, then only 1% of dimers would be visualized, 0.1% of trimers (if they exist), etc. The few studies that have thus far estimated molecular interaction rates from single-molecule data were primarily focused on the special cases of dimerization [8, 17] and bimolecular interactions [14]. Additionally, to derive association rates, they employed single-molecule experiments with a high-labeled fraction, greater than 75% [14, 17]. Such high degree of labeling is generally not compatible with single-molecule imaging for most molecules at their endogenous expression levels.

To fill this technological void, here we present a generic, mathematical modeling-based approach to calculate interaction kinetics from single-molecule data, termed FISIK (Framework for the Inference of in Situ Interaction Kinetics). FISIK consists of two components: (I) devising a mathematical model that mimics the dynamics, interactions and sub-stoichiometric labeling of the molecular system, and (II) calibrating the model with single-molecule data to estimate the unknown model parameters, including the

molecular association and dissociation rates. To test the feasibility, strengths and limitations of FISIK, here we applied it to the case of movements and interactions in 2D, applicable to e.g. cell-surface receptors, a system often studied with single-molecule imaging [4-9]. To the best of our knowledge, this is the first attempt to develop a generic framework to derive molecular interaction rates from single-molecule data, not limited to any specific oligomerization model (e.g. only dimerization). Testing FISIK has also helped shed light on the limits of what quantitative information can be learned about molecular interactions from single-molecule data.

Methods

1. Overall workflow of FISIK

FISIK consists of two components (Fig 1):

Component I. Build a stochastic mathematical model of molecular movement and interactions, mimicking the biological system and data acquisition setup, including sparse labeling. Many parameters in this model are derived explicitly from the experimental data and directly input into the model, such as molecular movement properties and most data acquisition parameters (e.g. sampling rate). This constrains the modeling and leaves only a few unknown model parameters, including the rates of molecular association and dissociation, to be estimated by fitting the model to experimental single-molecule data. The specific model employed to test the feasibility of FISIK is described in subsection 2 below.

Component II. Calibrate the model with experimental single-molecule data to estimate molecular association and dissociation rates. The stochasticity of molecular movement and interactions, compounded by the labeling of a random subset of molecules, makes it meaningless to compare model-generated and experimental molecular trajectories and interaction instances point-by-point for model calibration, as would be done for a deterministic system. Therefore, we have devised a stochastic model calibration algorithm by adapting the method of Indirect Inference [18-20]. In this approach, the model-generated (termed "probe") and experimental (termed "target") data are first analyzed to extract from them descriptors, called intermediate statistics, which characterize them. The specific intermediate statistics used for calibrating our test model, which are closely related to the unknown model parameters, are discussed in subsection 4 below. The difference between probe and target is then taken as the difference between their respective intermediate statistics, calculated as the Mahalanobis distance between them (Eq. 1). Specifically, if θ_{prb} and θ_{tar} are, respectively, the intermediate statistics of the probe and target data, with associated variance-covariance matrices V_{prb} and V_{tar}, then the Mahalanobis distance S is calculated as:

$$S = \left(\theta_{prb} - \theta_{tar}\right)'\left(V_{prb} + V_{tar}\right)^{-1}\left(\theta_{prb} - \theta_{tar}\right). \quad (1)$$

Casting the comparison in terms of the Mahalanobis distance has two advantages: (i) The difference between intermediate statistics gets weighted by their variance-covariance matrices, thus accounting for the statistics' uncertainties when calculating the difference between target and probe. (ii) Under the null hypothesis that $\theta_{prb} = \theta_{tar}$, and given that the intermediate statistics are normally distributed (as verified using a Lilliefors

test, which gave p-values in the range 0.1153-0.5), S follows a χ^2-distribution with number of degrees of freedom equal to number of intermediate statistics. This allows us to calculate a comparison p-value for any given *S*, allowing us to interpret the value of *S* [21, 22]. Specifically, probes with comparison p-value < 0.05 are considered statistically different from the target, and are thus outside the solution range. Probes with p-value ≥ 0.05 are statistically indistinguishable from the target, and are thus possible solutions.

Note that, at the conceptual level, FISIK is not limited to using Indirect Inference for stochastic model calibration. Other methods, for example Approximate Bayesian Computation [23, 24], are expected to be equally applicable.

2. Stochastic model of molecular movement, interactions and sub-stoichiometric labeling

To test the feasibility of FISIK, we employed a simple stochastic model of molecular diffusion and interactions in 2D, as well as sub-stoichiometric labeling to mimic single-molecule sampling. A 2D model is less computationally intensive to simulate than a 3D model, while at the same time providing equal power to test the feasibility of FISIK. Additionally, a 2D model is a reasonable approximation of e.g. receptor movement on the cell surface, a system that is particularly accessible and heavily studied by single-molecule imaging approaches [4-9]. Membrane proteins are also challenging to purify for traditional in vitro biophysical approaches, making cell-based imaging approaches particularly useful to study them. The model contains only one molecular species (e.g. a particular receptor type), which can oligomerize into dimers, trimers, etc. Note that in the

following and subsequent descriptions, oligomer(n) indicates an oligomer of size n, where n = 1 means monomer, n = 2 means dimer, n = 3 means trimer, etc. The terms monomer and oligomer(1) will be used interchangeably, as needed for compactness of description. In this model, molecules have a 2D density ρ and undergo 2D free diffusion with diffusion coefficient D. When a monomer and an oligomer($n \geq 1$) encounter each other, i.e. their pairwise distance $\leq d_{\text{encounter}}$ (taken as 10 nm), they associate with probability $p_a(n+1)$ to form an oligomer(n+1). The case of n = 1 is the special case of two monomers associating to form a dimer. For the sake of simplicity, association and dissociation are limited to one molecule at a time. Molecules in an oligomer diffuse together, with the same diffusion coefficient D (for the sake of keeping the model simple). For an oligomer($n \geq 2$), a molecule can dissociate from it with rate $k_{off}(n)$, to produce a monomer and an oligomer(n-1). To mimic single-molecule imaging, a fraction f of the molecules is labeled, with individual fluorophore intensity $I_{ind} \sim N(\mu_{ind},\sigma_{ind})$ (i.e. the intensity of an individual fluorophore fluctuates over time, mimicking realistic fluorophore intensity fluctuations). The remaining fraction of molecules (1-f) is invisible. Therefore, while all molecules diffuse and interact, only the trajectories and interactions of the labeled subset are visible.

The unknown parameters in this model are the motion parameter D, the interaction parameters $p_a(n)$ and $k_{off}(n)$ (n = 2, 3, 4, …), the population parameters ρ and f, and the fluorophore intensity parameters μ_{ind} and σ_{ind}. Among these parameters, D can be estimated directly from the single-molecule data, as it is not affected by under-sampling. The fluorophore intensity parameters can be estimated experimentally as well, using e.g. monodispersion experiments [6]. The other parameters, however, must be estimated by

fitting the model to experimental data, as described in the general framework above (subsection 1), and using the intermediate statistics described in subsection 4 below.

3. Kinetic Monte Carlo simulations

To generate molecular trajectories and interactions from the above model, we used kinetic Monte Carlo simulations explicit in both space and time (effectively particle-based reaction-diffusion simulations, similar to [25, 26], although in 2D instead of 3D). In these simulations, space is treated as a continuum, while time is discretized into a series of time steps Δt, where Δt is small enough to minimize discretization artifacts (more on this below). To simulate molecules at density ρ within a simulation area A, $N_{mol} = \rho \times A$ molecules are initially placed randomly within the simulation area, all starting as monomers. Reflecting boundary conditions are used to keep the molecules within the simulation area. Each simulation starts with an initialization time T_{init} (usually 10 s) so that the system reaches steady state, followed by the desired simulation time T_{sim} (S1 Fig; S1 Table, Row 1). Only the post-initialization time (i.e. steady-state) trajectories and interactions are output for further processing and analysis.

At every time point t after the initial time point, molecules can dissociate, move and associate, in this order, as described in the following:

(1) Molecule dissociation: For each oligomer($n \geq 2$) at the previous time point ($t - \Delta t$), one molecule can dissociate from it at time point t with probability $p_{off}(n) = k_{off}(n) \times \Delta t$, producing a monomer and an oligomer(n-1). If a dissociation happens, the involved monomer and oligomer are not allowed to associate with each other or with any other

monomer or oligomer in the current time point (step 3 below), to avoid what will appear as simultaneous swapping of molecules between oligomers in one time point. This constraint is acceptable as long as the time step is small enough.

(2) Initial update of molecule positions: Under the model of free diffusion, each oligomer($n \geq 1$) takes a step from time point $t - \Delta t$ to time point t with x- and y-components, s_x and s_y, $\sim N(0, \sqrt{2D\Delta t})$, where D is the diffusion coefficient. As mentioned above, all molecules within an oligomer move together. Note that if a molecule has dissociated from an oligomer (step 1 above), it moves independently of the oligomer to which it used to belong. The resulting positions are considered "initial positions" because, as described next, these positions might be altered by association.

(3) Molecule association and final update of positions: Fundamentally, if the distance between two molecules $\leq d_{\text{encounter}}$ (= 10 nm) at time point t, they are considered to encounter each other and thus might associate. However, due to time discretization in the simulations, molecules might pass by each other as they move from time point $t - \Delta t$ to time point t, but the distance between them might be greater than $d_{\text{encounter}}$ at time point t. Therefore, a monomer and an oligomer($n \geq 1$) are considered to encounter each other if they satisfy any of the following three conditions:

(i) The monomer-oligomer($n \geq 1$) distance $\leq d_{\text{encounter}}$ at time point t.

(ii) The monomer and oligomer($n \geq 1$) paths cross, assuming a straight-line displacement for each object between its position at time point $t - \Delta t$ and its position at time point t.

(iii) There is overlap in the monomer's and oligomer($n \geq 1$)'s ranges of non-straight-line paths from time point $t - \Delta t$ to time point t. The range for each object is approximated by a circle centered midway between its positions at time points $t - \Delta t$ and t, with diameter = the displacement magnitude from time point $t - \Delta t$ to time point t. The probability of monomer-oligomer($n \geq 1$) encounter is then taken as the ratio of the area of overlap between their circles to the area of the larger of the two circles.

In checking each of these encounter conditions, if any object satisfies the condition in relation to more than one object, graph matching is used to resolve these conflicts globally [27] and impose that each object may interact with only one other object per time point. Also, objects satisfying earlier encounter conditions are taken out of the pool for later conditions (e.g., two objects that satisfy condition (i) are not considered for conditions (ii) and (iii)). With this, monomer-oligomer($n \geq 1$) pairs that encounter each other can associate with probability $p_a(n+1)$ to form an oligomer(n+1). When an association happens, the position of the resulting oligomer is the average position of the two associating objects. This yields the final positions of molecules at time t of the simulation.

Particle tracks output of kinetic Monte Carlo simulations.

The trajectories output by the simulation code to use for further analysis within FISIK are in the same form, and contain the same limited information, as tracks obtained via multiple particle tracking of one-color single-molecule data (using e.g. u-track [7]). Every object (monomer or oligomer) becomes a particle, and the only information stored about it is its position and intensity over time (as with experimental single-molecule data). When an

association happens, two particles merge into one. When a dissociation happens, one particle splits into two. The oligomeric state of particles is not stored explicitly, as this information is not a direct output of single-molecule imaging experiments. The time step for the output particle tracks is a user-specified parameter, generally chosen to match the experimental data time step.

4. Single-molecule data analysis to calculate intermediate statistics

Estimation of oligomeric state from particle intensities and merging and splitting history

The first step for calculating the intermediate statistics describing the probe or target data (whether simulated or experimental) is to estimate the oligomeric state of each particle over time using the available information, namely the particle intensities and their merging and splitting history. For this, particle tracks are divided into segments, where each segment starts via an appearance or a split, and ends via a disappearance or a merge. The oligomeric state of each segment is then determined using constrained least squares minimization. Specifically, suppose a group of N_{seg} segments undergo merging and splitting events with each other (in the example shown in S2A Fig, $N_{\text{seg}} = 5$). Given a mean intensity I_j for each segment j ($j = 1, 2, \ldots, N_{\text{seg}}$), the oligomeric states s_j ($j = 1, 2, \ldots, N_{\text{seg}}$) of all segments in the group are determined by minimizing the least squares function F:

$$F = \sum_{j=1}^{N_{\text{seg}}} \left(s_j \mu_{\text{ind}} - I_j \right)^2 , \tag{2}$$

where μ_{ind} is the (known) mean intensity of an individual fluorophore. The minimization is subject to constraints from the merging and splitting history of the segments, imposing conservation of number of molecules. We will use the example shown in S2A Fig to illustrate the constraints: In that example, segments 1 and 2 merge to form segment 3, and then segment 3 splits into segments 4 and 5. This imposes the constraints that:

$$s_1 + s_2 = s_3 \text{ and } s_3 = s_4 + s_5. \quad (3)$$

Thus, for the example shown in S2A Fig, F in Eq. 2 will be minimized subject to constraints in Eq. 3. Given the integer nature of the unknown oligomeric states, the constrained least squares minimization is solved as a Mixed Integer Quadratic Programming problem (using the Package YALMIP [28] (https://yalmip.github.io/)). This analysis yields the dynamic oligomerization history of the labeled molecules throughout the simulation/observation time.

Calculation of intermediate statistics

Within the framework of Indirect Inference, FISIK uses the labeled oligomer densities and the labeled molecule association and dissociation rates as intermediate statistics. These intermediate statistics were chosen because they are closely related to the unknown model parameters, namely the interaction parameters $p_a(n)$ and $k_{off}(n)$ (n = 2, 3, 4, …) and the population parameters ρ and f. Using the labeled molecules' dynamic oligomerization history (generated as described above), the intermediate statistics are calculated as follows for each simulation (or experimental time lapse in the case of experimental data):

(i) Labeled oligomer densities: Having the oligomeric state of every particle at every time point, the average density of monomers ($\rho_{oligo}^{lab}(1)$), dimers ($\rho_{oligo}^{lab}(2)$), etc. is calculated. The superscript "lab" is to emphasize that this is the density from the labeled subset of molecules.

(ii) Labeled molecule dissociation rates: These are derived by treating transitions between labeled oligomeric states as a Markov process. If there are labeled molecule oligomers up to oligomeric state O_{max}^{lab}, then the Markov process consists of O_{max}^{lab} states (S2B Fig). For each oligomeric state $n \geq 2$, we calculate from the dynamic oligomerization history its average lifetime $\tau(n)$ and its probability $p_{n\rightarrow n-1}$ to transition to state n-1 (i.e. a labeled molecule dissociates; S2B Fig). Inverting the Gillespie algorithm for the simulation of stochastic processes [29], the labeled molecule dissociation rate is then calculated as

$$k_{off}^{lab}(n) = \frac{p_{n\rightarrow n-1}}{\tau(n)}, \qquad n = 2, 3, \ldots, O_{max}^{lab}. \tag{4}$$

(iii) Labeled molecule association rates: These are calculated from the labeled molecule oligomer densities and labeled molecule dissociation rates, assuming that the system is at steady state:

$$k_{on}^{lab}(n) = \frac{k_{off}^{lab}(n)\rho_{oligo}^{lab}(n)}{\rho_{oligo}^{lab}(n-1)\rho_{oligo}^{lab}(1)}, \qquad \mathrm{n} = 2, 3, \ldots, O_{max}^{lab}. \tag{5}$$

With this, the vector of intermediate statistics describing each simulation/experimental time lapse is constructed as

$$\theta = \left[\rho_{oligo}^{lab}(1), \rho_{oligo}^{lab}(2), \ldots, \rho_{oligo}^{lab}\left(O_{\max}^{lab}\right), k_{off}^{lab}(2), \ldots, k_{off}^{lab}\left(O_{\max}^{lab}\right), k_{on}^{lab}(2), \ldots, k_{on}^{lab}\left(O_{\max}^{lab}\right)\right]' \tag{6}$$

For a probe/target represented by multiple simulations (or multiple experimental time lapses), its vector of intermediate statistics $\theta_{prb}/\theta_{tar}$ and associated variance-

covariance matrix V_{prb}/V_{tar} are calculated as the mean and covariance, respectively, of the intermediate statistics vectors from its multiple simulations. The calculated θ_{prb} and θ_{tar} and associated V_{prb} and V_{tar} are then used in Eq. 1 to calculate the Mahalanobis distance to compare probes and targets.

Results and Discussion

The intermediate statistics accurately characterize the modeled molecular interactions

To validate our choice of intermediate statistics, we employed simulations where all molecules were labeled, in which case there are straightforward relationships between the input model parameters and the calculated intermediate statistics. We systematically varied the model parameters, i.e. ρ, p_a, k_{off} and D and investigated how the intermediate statistics responded. Note that, for the sake of simplicity, in these and the following tests p_a refers to $p_a(2 \leq n \leq 5)$ (all equal) and k_{off} refers to $k_{off}(2 \leq n \leq 5)$ (also all equal). In all tests, $p_a(n > 5) = 0$, i.e. the maximum oligomer that could be formed was a pentamer. The employed simulation parameters are shown in S1 Table, Row 1.

As expected, the calculated total molecule density (ρ^{lab}) and dissociation rate (k_{off}^{lab}) were equal to their corresponding input values, and they only varied when their input values were changed (Fig 2, 1st and 2nd columns). The calculated association rate (k_{on}^{lab}) varied approximately linearly with the input p_a and D, and was largely independent of the input ρ and k_{off} (Fig 2, 3rd column). Finally, the distribution of oligomeric states shifted toward higher values as the input p_a or D increased (leading to an increase in k_{on}^{lab})

or the input k_{off} decreased, as expected. These results validate our simulation and intermediate statistics calculation strategy, and support our choice of intermediate statistics to capture the relevant information about the modeled system for model calibration. With this, we proceeded to test FISIK when the labeled fraction $f < 1$.

FISIK can estimate interaction parameters from single-molecule trajectories, provided a sufficient fraction/number of molecules is labeled

Most 2D single-molecule imaging experiments operate at a labeled molecule density roughly in the range 0.1-1 mol/μm^2, in order to achieve reasonably accurate particle tracking [7, 9, 16]. Therefore, as a first test of FISIK, we focused on simulations with relatively low molecule densities (ρ = 1-16 mol/μm^2), in order to afford low-medium labeled fractions (f = 0.1-0.6). Taking the example of cell-surface receptors, such densities are realistic for various receptor types, such as CD36 [6], EpoR [30], LDLR [31] and various GPCRs [14, 17]. Because the parameter inference problem is expected to get more difficult as the labeled fraction gets smaller – an expectation that is met in our tests as will be discussed shortly – these initial studies test FISIK under relatively favorable conditions.

For these studies we used the model parameters in S2 Table, Row 1 (see also Fig. 3) and simulation parameters in S1 Table, Row 2. The probe parameter landscape consisted of 3360 parameter combinations, each represented by 10 simulations. This probe landscape was tested against several targets, each also represented by 10 simulations. However, to mimic molecule density variability between cells, the target simulations differed from the probe simulations in that the density varied between the 10

simulations of each target, following a normal distribution with mean = the specified density and standard deviation = 0.1 × mean.The target parameters were chosen to span a wide range of values for the interaction parameters p_a and k_{off} and the population parameters ρ and f.

In the following we discuss our major findings from these tests, using representative targets for demonstration. Taking the target ρ = 4 ± 0.4 mol/μm^2, f = 0.2, p_a = 0.5 and k_{off} = 1/s as our reference point, we will first explain how the parameter inference results are displayed (Fig. 3A). As explained in Methods, Section 1, the difference between a probe and a target is calculated as the Mahalanobis distance between their intermediate statistics (Eq. 1), which is converted to a p-value. Thus the results are shown as p-value landscapes. However, because we are searching in a 4-dimensional parameter space, the p-value landscapes are shown in two steps. In the first step, the ρ vs. f landscape is shown on the left, where each rectangle is a "summary view" of the k_{off} vs. p_a landscape corresponding to each ρ-f pair (specifically, it shows the maximum p-value obtained for the underlying k_{off} vs. p_a landscape). White rectangles indicate ρ-f pairs containing no matches with the target (for any tested k_{off}-p_a combination). Non-white rectangles indicate ρ-f pairs containing some matches with the target (for some k_{off}-p_a pairs, as explained next). Darker rectangles indicate matches with higher p-value. In the second step, the non-white rectangles in the ρ vs. f landscape are "expanded" on the right, to show the full k_{off} vs. p_a landscapes corresponding to those ρ-f pairs. Again, white rectangles indicate k_{off}-p_a pairs (for each ρ-f pair) that do not match the target (p-value < 0.05), and non-white rectangles indicate k_{off}-p_a pairs (for each ρ-f

pair) that are considered to match the target (p-value ≥ 0.05). Darker rectangles indicate higher p-values.

For the reference target (Fig. 3A), FISIK was able to identify the correct parameters (indicated by white circles; p-value=0.99), within a range of solutions. This range of solutions was in part due to the expected coupling between the population parameters ρ and f, such that the estimated $\rho \times f \approx$ the average density of labeled molecules in the target (= 0.8 labeled mol/μm^2 for this target). Note, however, that this coupling did not extend indefinitely. For example, comparing the probe landscape to a target with ρ = 100 ± 10 mol/μm^2, f = 0.03, p_a = 0.5 and k_{off} = 1/s (described in detail in the next section) did not yield any matches between them, even though 100 × 0.03 = 3 labeled mol/μm^2, a labeled molecule density that was achievable by multiple ρ-f combinations in the probe landscape. This demonstrates that while FISIK yields a range of estimated ρ and f values, the range is finite and contained within one order of magnitude at most.

In spite of the observed ρ-f coupling, which reduced the determinability of ρ and f, the estimated k_{off} and p_a, i.e. the interaction parameters of interest, were relatively consistent for the different ρ-f combinations. k_{off} exhibited a narrow range, between 0.75/s and 1.5/s (target k_{off} = 1/s), with a mild inverse relationship with the estimated f (lower f was associated with higher k_{off}). This was most likely because, for probes with lower labeled fractions, less interaction events were visible, leading to estimated k_{off}^{lab} (i.e. probe intermediate statistics) < true probe k_{off} (S3 Fig), thus resulting in matching of target to probes with higher true k_{off}. p_a had a well-defined lower bound, 0.4, across all ρ-f combinations (target p_a = 0.5). Its upper bound was slightly less well-defined, increasing

from 0.6 to 0.7 to 0.8 as the estimated f decreased from ≥ 0.3 to 0.2 to 0.1. The inverse relationship between p_a and f was most likely due to the inverse relationship between k_{off} and f. Testing FISIK against targets of different population parameters f and ρ in the neighborhood of the reference parameters indicated that targets with higher f or ρ led to narrower solution ranges, while targets with lower f or ρ led to wider solution ranges, reducing the determinability of some parameters, especially p_a (Fig. 3B, C and S4 Fig).

Next we tested the sensitivity of FISIK to distinguish between targets with different interaction parameters. Varying k_{off} from 1/s to 0.5/s or to 1.5/s led to clear shifts in the solution range, from 0.75-1.5/s for target k_{off} = 1/s to 0.25-0.75/s for target k_{off} = 0.5/s and to 1-2/s for target k_{off} = 1.5/s (Fig. 3A, D, E). Varying p_a from 0.5 to 0.3 or to 0.7 shifted the solution range, for most probe ρ-f combinations, from 0.3-0.7 for target p_a = 0.5 to 0.2-0.5 for target p_a = 0.3 and to 0.5-0.8 for target p_a = 0.7 (Fig. 3A, F, G). The inverse relationship between the estimated p_a and the estimated f made the shifts in the p_a solution range a little less clear than those in the k_{off} solution range. However, when the population parameters ρ and f were estimated correctly, the p_a solution ranges were quite narrow and the shift in the estimated p_a was very clear.

All in all, these tests demonstrate that FISIK is able to estimate molecular interaction parameters from single-molecule data, as long as there is a sufficient fraction/number of labeled molecules. Combining FISIK with strong prior knowledge about the molecule density – which is experimentally achievable via methods such as flow cytometry – is expected to aid FISIK with this task.

The fraction of labeled molecules in single-molecule trajectory data is the most critical factor for FISIK's ability to estimate interaction parameters

While many transmembrane proteins exist at a surface density in the 1-10 mol/μm^2 range, others are much more highly expressed. Examples of receptors with high surface densities are CD8 [32] and receptors of the ErbB family [31, 33]. In such cases, the labeled fraction f will be below 0.1 to maintain single-molecule trackability. Therefore, as a next test of FISIK, we simulated targets and probes with $\rho \approx 100$ mol/μm^2 and f close to 0.1 and below. These tests not only investigated the performance of FISIK when f is very low, but also allowed us to determine more conclusively which property is more important: the labeled fraction f or the labeled molecule density = $\rho \times f$ (the tests above showed some interdependence between the minimum necessary f and ρ; see Fig 3A, B, C and S4 Fig).

For these studies we used the model parameters in S2 Table, Row 2 and the simulation parameters in S1 Table, Row 3. The probe parameter landscape consisted of 2352 parameter combinations (S2 Table), where each parameter combination was represented by 30 simulations (here the simulation area was smaller for the sake of simulation efficiency, and thus we compensated for the smaller area, i.e. smaller number of events, by combining 30 simulations instead of 10). The target parameters were ρ = 100 ± 10 mol/μm^2, $p_a = 0.5$, $k_{off} = 1$/s and a series of values for f to investigate its effect on interaction parameter determinability (also 30 simulations per target).

First we investigated the solution range when the target f = 0.03 (Fig. 4A), because a target with ρ = 100 ± 10 mol/μm^2 and f = 0.03 has the same labeled molecule density as a target with ρ = 10 ± 1 mol/μm^2 and f = 0.3, a target for which FISIK yielded an almost perfect solution (S4D Fig). However, in contrast to the ρ = 10 ± 1 mol/μm^2 and f = 0.3 case, the ρ = 100 ± 10 mol/μm^2 and f = 0.03 case had a wide range of solutions, with p_a practically undeterminable. This test clearly demonstrates that the labeled fraction f, and not the labeled density per se, is the most critical parameter influencing the performance of FISIK.

To determine the minimum labeled fraction needed to accurately estimate the interaction parameters, next we increased the target f from 0.03 to 0.06, 0.09 and 0.12 (Fig. 4B-D). As f increased, the solution range decreased. By f = 0.06, k_{off} was estimated accurately within a narrow range (Fig 4B). The range of p_a also gradually decreased as f increased, although it was still quite wide even at f = 0.12 (Fig 4D).

Putting together the results of the tests here and in the previous section, we conclude that the labeled fraction of molecules, more than the number/density of labeled molecules, is a critical parameter for FISIK's performance. When the labeled fraction is too small, fundamental information about multi-molecular processes in the system, such as association, is lost and seems to be irrecoverable. Indeed, determining k_{off} is possible with relatively low labeled fractions, down to ~0.06 (Fig 4B), as dissociation is a uni-molecular process. However, determining p_a, which characterizes a mutli-molecular process (association), requires labeled fractions of at least ~0.2 (Fig 3A).

Static oligomer distribution data can complement low-labeled-fraction single-molecule trajectory data and increase FISIK's power

Our results thus far pose a dilemma. The inference of interaction parameter values, especially p_a, improves as the labeled fraction of molecules increases. Yet increasing the labeled fraction often increases particle tracking errors [7, 16], thus reducing the data quality for model calibration and parameter inference (the effect of tracking errors will be investigated in the next section). One approach to address this challenge is to employ multi-color single-molecule imaging (e.g. 2- or 3-color, or even more with hyperspectral imaging [4]), where the labeled fraction in each channel affords good trackability, but the sum of all channels gives a high enough total labeled fraction. This approach is conceptually similar to the above tests, and thus will not be explored further here. Another approach to address this challenge is to combine the dynamic but sparse information from single-molecule tracking with dense but static information from approaches such as Number and Brightness (N&B) analysis [34, 35], Spatial Intensity Distribution Analysis (SpIDA) [36], or Single-Molecule Localization Microscopy (SMLM) [37, 38]. These approaches yield a dense snapshot of the oligomeric state distribution of the majority of molecules of interest, thus potentially complementing single-molecule tracking data and improving the ability of FISIK to infer interaction parameters.

As a first step, we investigated the performance of FISIK when using only static oligomer distribution data. For these tests we used the model parameters in S2 Table, Row 3 and the simulation parameters in S1 Table, Row 4. In terms of model parameters (S2 Table), we replaced the labeled fraction f with the observed fraction ϕ. The observed

fraction accounts for the fact that, even when all molecules are labeled, not all molecules are observed, due to incomplete fluorophore maturation [17], or e.g. photoconversion and localization in the case of SMLM [39, 40]. In the targets we used ϕ = 0.85, and made ϕ an unknown parameter to estimate; in the probes ϕ was varied in the range 0.75-0.95 [17, 39, 40]. The output of the static snapshot simulations was the particle intensities and positions at the end of the simulation, equivalent to the output of the full trajectory simulations, but for one time point (Methods, Section 3). As there is no merging and splitting history in the static snapshot data, the oligomeric state of each particle in this case was determined solely from the particle intensities, i.e. by minimizing the objective function in Eq. 2 but without any constraints (Methods, Section 4). Additionally, the only intermediate statistics to be used for target and probe comparison in the case of static distribution data were the observed oligomer densities: $\rho_{oligo}^{obs}(1)$ (monomers), $\rho_{oligo}^{obs}(2)$ (dimers), etc.

These tests demonstrated that implementing FISIK with static oligomer distribution data instead of dynamic single-molecule trajectory data yielded an accurate estimate of the ratio of k_{off} to p_a (Fig 5A). Note that p_a is linearly related to the association rate k_{on} (Fig 2C); as a result, $k_{off} / p_a \propto k_{off} / k_{on} = K_d$, the dissociation constant. In other words, applying FISIK with dense, static oligomer distribution data extracts from these data the maximum information possible, namely the dissociation constant(s) describing molecular interactions. The accuracy of the estimated K_d's is expected to depend on the accuracy of the experimental and analytical procedures employed to calculate the distribution of oligomeric states. For example, in the case of SMLM, fluorophore blinking and multiple

appearances must be accounted for properly to avoid over-counting that results in exaggerated apparent oligomerization [37].

Our tests thus far indicate that FISIK with low-labeled-fraction single-molecule trajectory data yields accurate estimates of k_{off}, while FISIK with static oligomer distribution data yields accurate estimates of k_{off}/p_a. Therefore, we surmised that combining the two data types should yield an accurate estimate of both k_{off} and p_a, along with estimating ρ, f and ϕ (i.e. 5 parameters in total). We implemented the combination of data types at the level of FISIK's solution space. Specifically, we defined the combined data comparison p-value as

$$\begin{aligned}&\text{combined p-value}\left(p_a, k_{off}, \rho, f, \phi \mid \text{dynamic trajectory data + static distribution data}\right) = \\ &\min\begin{pmatrix}\text{p-value}\left(p_a, k_{off}, \rho, f \mid \text{dynamic trajectory data}\right), \\ \text{p-value}\left(p_a, k_{off}, \rho, \phi \mid \text{static distribution data}\right)\end{pmatrix}\end{aligned} \tag{7}$$

where the p-values are calculated from the Mahalanobis distance as done previously (Eq. 1). With this, a combination of p_a, k_{off} and ρ values will only be accepted as a possible solution if it is a possible solution in both the dynamic comparison and the static comparison. This strategy indeed enabled the accurate estimation of p_a and k_{off}, as well as ρ, with a single-molecule labeled fraction as low as 0.06 (Fig. 5B-C). These tests demonstrate that combining these two types of complementary data is indeed a viable approach to enhance the accuracy of FISIK, without the need to increase the labeled fraction in single-molecule trajectory data.

FISIK can estimate molecular interaction parameters from single-molecule trajectories subject to resolution limitations and tracking errors

While single-molecule imaging can reveal colocalization of molecules, suggestive of interactions, on its own it cannot distinguish between molecules interacting with vs. passing by each other, largely because of the resolution limit. Note that for the purposes of this section we are using the term interactions in a general sense, thus including direct interactions, indirect interactions and clustering within a common nanodomain. Distinguishing between these different types of interactions depends on the known biology and further experimentation of the system under study. Here we are concerned with distinguishing between them and coincidental colocalization due to molecules passing by each other. One-color single-molecule imaging suffers the most from the limited resolution of light microscopy. Multi-color imaging suffers less due to the ability to use e.g. co-movement in addition to colocalization to assess specific interactions [8]. FRET-based experiments suffer the least from the resolution limit, due to the requirement of very close proximity for FRET. Nevertheless, fundamentally all light microscopy approaches face the challenge of distinguishing between proximity due to interactions vs. due to passing-by events.

Here we tested FISIK's performance when applied to single-molecule trajectories obtained from one-color imaging data, as the most challenging scenario. We and others have previously developed particle tracking algorithms to capture merging and splitting events between imaged molecules in one-color time-lapses [7, 41, 42], reflecting molecular interactions (in the general sense defined above) [6, 13]. For these tests, we

used relatively low density and low-medium labeled fraction simulations (S2 Table, Row 1), as FISIK had an overall good performance under these conditions in the absence of particle tracking errors (Fig. 3). We generated synthetic single-molecule image series that mimicked experimental time-lapses from simulated target trajectories (see S3 Table for image generation parameters). Then we detected and tracked the molecules in the synthetic image series using u-track [7], as we would detect and track actual experimental data (see S4 and S5 Tables for detection and tracking parameters). Finally, we used the tracking output, after a minor "clean up" (S1 Appendix), as the target data for model calibration within FISIK. Note that the probe trajectories were used directly as output by the Monte Carlo simulations.

With the above strategy, we first tested FISIK by generating images with an average signal-to-noise ratio (SNR) of 7 (S1 Video), which is achievable with many organic dyes imaged via total internal reflection fluorescence microscopy (TIRFM), a common combination for single-molecule imaging. We used targets with ρ = 4 ± 0.4 mol/μm^2, f = 0.2 or 0.1, p_a = 0.5 and k_{off} = 1/s (as in Fig 3A and C). A labeled molecule density of 0.8 mol/μm^2 (when f = 0.2) was at the upper limit of what is trackable [7, 16], and FISIK was unable to estimate the target parameters (not shown). A labeled-molecule density of 0.4 mol/μm^2 (when f = 0.1) afforded better trackability (S2 Video). In this case, FISIK's performance when applied to the u-track output was comparable to its performance when applied to the direct simulation targets (Fig 6A), although the solution range was slightly shifted downward toward smaller k_{off} values. We surmised that this was due to the tracking software constraint that merges and splits were only allowed when

there was no possibility of gap closing (S5 Table). This constraint helped reduce the assignment of merge-to-split events to molecules merely passing by each other, but at the same time it might eliminate short-lived specific interactions between molecules. To investigate this issue further, we tested FISIK against targets with lower and higher k_{off} values. Indeed, the downward shift was more pronounced for higher target k_{off} values (Fig 6B, C). Overall, however, the estimated k_{off} range contained the target k_{off}.

The above tests indicated that FISIK's performance was sensitive to tracking errors resulting from detection errors. Specifically, in order to get solutions even with f = 0.1, in the detection step it was necessary to employ frame-averaging with a rolling window to detect dimmer objects [7] (S4 Table). Note that while the average SNR was 7, with the utilized single molecule intensity mean and standard deviation (S3 Table), some objects could have an SNR < 1 in some frames (when their intensity = mean - 3×standard deviation). Therefore, we investigated the performance of FISIK under higher SNR conditions, to minimize detection issues. High SNR conditions are not very common in single-molecule data, but they are achievable with, for example, quantum dot labeling of the molecule(s) of interest [6, 8]. For this we used a background standard deviation of 50, resulting in average and "minimum" SNRs of 20 and 2, respectively. In this case, FISIK was successful in finding a solution without using a rolling window for detection, even for f = 0.2 (i.e. the more challenging tracking scenario) (S5 Fig; S3 and S4 Videos). As observed with direct simulations (Fig 3 and 4), the solution range was narrower with f = 0.2 compared to f = 0.1, although it remained quite wide for p_a.

These tests demonstrate that FISIK is able to extract molecular interaction parameters from single-molecule data, in spite of the resolution limit inherent to light microscopy. Accurate estimation of interaction parameters requires accurate molecule tracking, an important aspect of which is to capture merge and split events between the imaged molecules, and to distinguish between these events and molecules passing by each other. This task seems to critically depend on the accuracy of molecule detection, such that there are as few false negatives as possible, as undetected molecules add to the complexity of tracking by introducing track gaps that must be closed as well.

Conclusion

In conclusion, here we have developed a framework for the inference of in situ interaction kinetics from single-molecule tracking data (FISIK), overcoming their inherent limitation of sub-stoichiometric labeling. Our tests demonstrate that FISIK can indeed estimate molecular interactions rates, provided a sufficient fraction of molecules is labeled. Specifically, dissociation and association rates can be estimated with labeled fractions as low as ~0.06 and ~0.2, respectively. Accurate estimation of interaction parameters relies on accurate molecule tracking. In practice, the system properties (e.g. molecule density and movement) and imaging conditions (e.g. SNR and frame rate) impose an upper limit on the labeled fraction of molecules in order to achieve accurate tracking in one-color single-molecule time-lapses. If this labeled fraction is insufficient for interaction rate estimation, then combining tracking data from multiple channels, and/or tracking data with dense but static oligomer distribution data (as obtained from techniques such as N & B

analysis, SpIDA or SMLM), would provide sufficient information for FISIK to estimate accurately the unknown interaction rates. Overall, FISIK, combined with system-specific motion and interaction models, is expected to allow us to estimate the interaction rates between molecules in their native cellular environment, taking full advantage of the rich spatial and temporal information present in live-cell imaging data.

Acknowledgments

We thank R. Yirdaw for contributing toward initial implementations of the simulation code. This work was supported by funding from The Welch Foundation (I-1901), NIH/NIGMS (R35 GM119619) and CPRIT (R1216) to KJ.

Figure legends

Fig 1. FISIK workflow. FISIK consists of two components: (I) Building a mathematical model and generating simulated data (probe) that mimic the biological system and acquired experimental data (target). (II) Fitting the model to target data in order to determine the unknown model parameters, including molecular association and dissociation rates. This requires a stochastic model calibration framework, as both model and data are stochastic.

Fig 2. The intermediate statistics accurately reflect the modeled interactions. (**A-D**) Influence of varying molecule density ρ (**A**), dissociation rate k_{off} (**B**), association probability p_a (**C**), and diffusion coefficient D (**D**) on the intermediate statistics ρ_{oligo}^{lab}, $\mathrm{k}_{\mathrm{off}}^{\mathrm{lab}}$

and $\mathrm{k}_{\mathrm{on}}^{\mathrm{lab}}$, for labeled fraction f = 1. Shown are the mean and standard deviation from 10 simulations per parameter combination. Each panel explicitly states the varied parameter values, while the parameters not stated take on the following values: ρ = 4 mol/μm^2, p_a = 0.5, k_{off} = 1/s and D = 0.1 μm^2/s. In the first column, the right-most measurement (with gray background) is the total molecule density (ρ^{lab}) as calculated from the oligomer densities: $\rho^{lab} = \sum_{n\geq 1} n \times \rho_{oligo}^{lab}(n)$. Note that the left and right y-axes have different ranges.

Fig 3. FISIK can estimate interaction parameters from single-molecule trajectories, provided a sufficient fraction/number of molecules is labeled. (**A**) Four dimensional p-value landscape for comparing probe trajectories to the indicated target, as a function of the population parameters ρ and f and the interaction parameters p_a and k_{off} (total of 3360 probes). The p-value is shown as a heat map, going from white (p-value < 0.05) to black (p-value = 1). The ρ vs. f landscape is shown on the left, where each rectangle is a "summary view" of the k_{off} vs. p_a landscape corresponding to each ρ-f pair. Specifically, it shows the maximum p-value obtained for the underlying k_{off} vs. p_a landscape. The non-white rectangles in the ρ vs. f landscape are then "expanded" on the right, showing the corresponding full k_{off} vs. p_a landscapes. The specific ρ-f pairs being expanded are indicated above each k_{off} vs. p_a landscape, with the target ρ-f pair in bold. The target parameter values are also indicated by a white open circle in the ρ vs. f landscape on the left and a white open circle in the k_{off} vs. p_a landscape on the right. (**B-G**) Same as (A), but with target f increased (**B**) or decreased (**C**), target k_{off} decreased (**D**) or increased

(**E**), or target p_a decreased (**F**) or increased (**G**) to the indicated values. The correct probe parameters (= target parameters) had the following p-values: 0.9997 (rank 3/66 possible solutions) in (A), 0.9230 (rank 1/25 possible solutions) in (B), 0.9521 (rank 6/160 possible solutions) in (C), 0.9996 (rank 1/34 possible solutions) in (D), 0.9937 (rank 1/ 37 possible solutions) in (E), 0.9385 (rank 2/32 possible solutions) in (F) and 0.9998 (rank 1/44 possible solutions) in (G).

Fig 4. The fraction of labeled molecules in single-molecule trajectory data is the most critical factor for FISIK's ability to estimate interaction parameters. (**A**) Four dimensional p-value landscape for comparing probe trajectories to the indicated target, as a function of the population parameters ρ and f and the interaction parameters p_a and k_{off} (total of 2352 probes). Detailed description as in Fig 3A. (**B-D**) Same as (A), but with target f gradually increased to the indicated values. The correct probe parameters (= target parameters) had the following p-values: 0.9700 (rank 10/219 possible solutions) in (A), 0.9994 (rank 4/184 possible solutions) in (B), 0.9999 (rank 1/122 possible solutions) in (C) and 0.9999 (rank 1/89 possible solutions) in (D).

Fig 5. Dense but static oligomer distribution data increase FISIK's power in the case of sparse single-molecule trajectory data. (**A**) Four-dimensional p-value landscape for comparing probe oligomer distribution data to the indicated target oligomer distribution data, as a function of the population parameters ρ and ϕ and the interaction parameters p_a and k_{off} (total of 840 probes). Detailed description as in Fig 3A, except that

the labeled fraction f is replaced by the observed fraction ϕ (see text for details). (**B**) Five-dimensional p-value landscape for comparing combined probe data (trajectories + distribution data) to the indicated target (also trajectories + distribution data), as a function of the population parameters ρ, f and ϕ and the interaction parameters p_a and k_{off} (total of 11760 probes). Detailed description as in Fig 3A, except that the parameter space includes both f and ϕ. (**C-D**) Same as (B), but with target f gradually increased to the indicated values. The correct probe parameters (= the target parameters) had the following p-values: 0.9998 (rank 1/19 possible solutions) in (A), 0.9994 (rank 2/24 possible solutions) in (B), 0.9998 (rank 1/24 possible solutions) in (C) and 0.9998 (rank 2/18 possible solutions) in (D).

Fig 6. FISIK can estimate interaction parameters from single-molecule trajectories subject to resolution limitation and tracking errors. (**A**) Four dimensional p-value landscape for comparing probe trajectories to the indicated target, as a function of the population parameters ρ and f and the interaction parameters p_a and k_{off} (total of 3360 probes). Top row (titled “Direct simulations”) shows landscape for comparing to target trajectories directly as output by the Monte Carlo simulations (as done for all figures prior to this one). Bottom row (titled “u-track”) shows landscape for comparing to target trajectories as output by u-track (see text for image generation and tracking details). Detailed description as in Fig 3A. (**B**, **C**) Same as (A), but with k_{off} decreased (**B**) or increased (**C**) to the indicated values. In (A), the “Direct simulations” result is a repeat of the result shown in Fig 3C, shown here to facilitate direct comparison with the “u-track”

result. The correct probe parameters (= target parameters) had the following p-value: 0.9521 (rank 6/160 possible solutions) for direct simulations and 0.5812 (rank 13/93 possible solutions) for u-track in (A), 0.8308 (rank 14/50 possible solutions) for direct simulations and 0.5605 (rank 16/65 possible solutions) for u-track in (B), and 0.8308 (rank 3/110 possible solutions) for direct simulations and 0.4397 (rank 18/147 possible solutions) for u-track in (C).

Supplementary figure legends

S1 Fig. Simulation strategy considerations. (**A**, **B**) Labeled oligomer density as a function of simulation time for the indicated model parameters, shown as mean ± 1 standard deviation from 10 simulations. Yellow-green background indicates initialization time (T_{init}). White background indicates simulation time for output (T_{sim}). T_{init} = 10 s is sufficient to reach steady state even for the model simulated in (A), which has one of the slowest dynamics among the simulated parameter combinations in our study. (**C**) Labeled molecule association rate as obtained by analyzing (as in Methods, Section 4) the output of Monte Carlo simulations run with the indicated simulation time steps (Δt), without (left) or with (right) correcting for the finite simulation time step to capture molecular encounters (as explained in the description of association events in Methods, Section 3). Correcting for the finite Δt yields equivalent simulation output for Δt in the range 10^{-4}-10^{-2}. $\Delta t = 0.1$ is however too large, resulting in overcompensation and artifacts. Note that, for analysis, all

trajectories are downsampled to an output time step of 0.1, regardless of the simulation time step.

S2 Fig. Single-molecule data analysis and intermediate statistics calculation considerations. (**A**) Example of two particles merging and splitting. For assignment of particle oligomeric state, this "compound track" (which includes information for both particles) is divided into segments s_1 and s_2 before the merge, segment s_3 after the merge (but before the split), and then segments s_4 and s_5 after the split. This merge and split history provides constraints (Eq. 3) for the objective function minimization (Eq. 2) to determine the particle oligomeric states. (**B**) Markov process representation of transitions between oligomeric states due to molecule association and dissociation (also called oligomeric sizes in the text), for calculating the labeled molecule dissociation rate (Eq. 4).

S3 Fig. Dependence of estimated dissociation rate for the labeled molecules (k_{off}^{lab}) on labeled fraction *f*. Results shown are for simulations with ρ = 10 ± 1 mol/μm^2, p_a = 0.5, k_{off} = 1/s and the indicated labeled fractions. Shown are the mean and standard deviation from 10 simulations each.

S4 Fig. Further tests of FISIK by varying molecule density and labeled fraction under conditions of relatively low density and low-medium labeled fraction (as in Fig 3). Four dimensional p-value landscape for comparing probe trajectories to the indicated targets, as a function of the population parameters ρ and f and the interaction

parameters p_a and k_{off}. Detailed description as in Fig 3A. (**A**-**C**) Same as in Fig 3A, but with target *f* increased (**A**), or target ρ increased (**B**) or decreased (**C**) to the indicated values. (**D**) Example of a target with a very narrow solution range. The correct probe parameters (= target parameters) had the following p-values: 0.9817 (rank 1/8 possible solutions) in (A), 0.9885 (rank 1/26 possible solutions) in (B), 0.9984 (rank 1/126 possible solutions) in (C), and 0.9990 (rank 1/2 possible solutions) in (D).

S5 Fig. FISIK's performance when applied to tracking data from high SNR images. (**A**, **B**) Four dimensional p-value landscape for comparing probe trajectories to the indicated targets, as a function of the population parameters *ρ* and *f* and the interaction parameters p_a and k_{off} (total of 3360 probes). Target trajectories were obtained via u-track analysis of images with average SNR = 20 (see text for image generation and tracking details). Detailed description as in Fig 3A. The correct probe parameters (= target parameters) had the following p-values for, respectively, k_{off} = *1/s, 0.5/s and 1.5/s*: 0.0808 (rank 48/51 possible solutions), 0.2277 (rank 8/21 possible solutions), and 0.0509 (rank 48/49 possible solutions) in (A); and 0.9536 (rank 4/93 possible solutions), 0.9216 (rank 9/65 possible solutions), and 0.8515 (rank 22/147 possible solutions) in (B).

S6 Fig. Illustration of simultaneous merges and splits that get "cleaned up," as described in S1 Appendix. In this example, 2 merges (M(1) and M(2)) and 3 splits (S(1), S(2) and S(3)) occur with the base track at the same time point. After the clean-up, two

merge-split pairs (M(1)-S(1) and M(2)-S(2)) are separated from the base track, leaving only one excess split, S(3), connected to the base track.

Supplementary video legends

S1 Video. Synthetic time-lapse (left) and detected objects (right), with an average SNR = 7. Time-lapse corresponds to a target simulation with parameters ρ = 4 ± 0.4 mol/μm^2, f = 0.1, p_a = 0.5, k_{off} = 1/s and D = *0.1 μm^2/s*. Frame rate = 10 Hz. Image size = 27 μm × 27 μm. See text for image generation details.

S2 Video. Particle tracking in time-lapse with an average SNR = 7. Time-lapse same as in S1 Video. Trajectories are shown as dragtails of length 15 frames. Merges are shown as yellow diamonds, splits as green diamonds, and closed gaps as cyan circle.

S3 Video. Synthetic time-lapse (left) and detected objects (right), with an average SNR = 20. Time-lapse corresponds to a target simulation with parameters ρ = 4 ± 0.4 mol/μm^2, f = 0.2, p_a = 0.5, k_{off} = 1/s and D = 0.1 μm^2/s. Frame rate = 10 Hz. Image size = 27 μm × 27 μm. See text for image generation details.

S4 Video. Particle tracking in time-lapse with an average SNR = 20. Time-lapse same as in S3 Video. Trajectories are shown as dragtails of length 15 frames. Merges are shown as yellow diamonds, splits as green diamonds, and closed gaps as cyan circle.

Supplementary Tables

A (μm^2)	Δt (s)	T_{sim} (s)	T_{init} (s)	Δt_{out} (s)	N_{sim}
10 × 10	0.01	10	10	0.01	10
25 × 25	0.01	10	10	0.1	10
12 × 12	0.01	10	10	0.1	30
12 × 12	0.01	1 snapshot	20	N/A	30

S1 Table: Simulation parameters, for probe and target simulations. Shown are the values for simulation area (A), simulation time step (Δt), simulation time (T_{sim}), initialization time (T_{init}), output trajectory time step (Δt_{out}), and number of simulations (N_{sim}). Per the requirements of indirect inference [18], a fixed set of random number generator seeds was used in the simulations, in order to reduce the jaggedness of the objective function landscape when varying model parameters. Probe and target simulations employed different random number generator seeds.

D (μm^2/s)	ρ (mol/μm^2)	f (rows 1, 2) ϕ (row 3)	$p_a(2 \leq n \leq 5)$	$k_{off}(n \geq 2)$ (/s)	[μ_{ind}, σ_{ind}]	# of parameter comb.

0.1	1, 2, 3, 4, 6, 8, …, 16	0.1, 0.2, …, 0.6	0.2, 0.3, …, 0.8	0.25, 0.5, …, 2	[1, 0.3]	3360
0.1	80,100,120	0.01, 0.02, …, 0.14	0.2, 0.3, …, 0.8	0.25, 0.5, …, 2	[1, 0.3]	2352
0.1	80,100,120	0.75, 0.8, …,0.95	0.2, 0.3, …, 0.8	0.25, 0.5, …, 2	[1, 0.3]	840

S2 Table: Model parameters for generating probe landscapes. Shown are the values for diffusion coefficient (D), molecule density (ρ), labeled fraction (f) or observed fraction (ϕ), association probability (p_a), dissociation rate (k_{off}) and individual fluorophore intensity mean (μ_{ind}) and standard deviation (σ_{ind}). Note that $p_a(n > 5) = 0$. The number of probe parameter combinations is shown in the last column. Target simulations used a subset of these parameter combinations, spanning the whole range.

Image area (pixels^2)	Pixel size (nm)	PSF std (pixels)	Bit depth	Single-molecule intensity [mean, std] (a.u.)	Background mean (a.u.)	Background std (i.e. noise) (a.u.)	Mean SNR
302x302	90	1.2	16	[1000, 300]	5000	140	7
302x302	90	1.2	16	[1000, 300]	5000	50	20

S3 Table: Synthetic image generation parameters. Shown are the values for image area, pixel size, standard deviation (std) of Gaussian approximating the point spread function (PSF), image bit depth, single-molecule intensity above background, background mean, background std, and resulting signal-to-noise ratio (SNR).

Detection parameters		**Mean SNR 7**	**Mean SNR 20**
Gaussian standard deviation (pixels)		1.2	1.2
Local maxima detection	Rolling window for time averaging (frames)	1 (i.e. no time averaging), 3	1 (i.e. no time averaging)
	alpha-value for comparison with local background	0.1, 0.1	0.1
Gaussian fitting at local maxima	Do iterative Gaussian mixture-model fitting	Yes	Yes
	Alpha-value for residuals test	0.1	0.1

S4 Table: U-track detection parameters (Gaussian Mixture-Model Fitting). Shown are the parameters with non-default values (based on u-track Version 2.2.1).

Tracking parameters	**SNR 7 and 20**

Overall	Maximum gap to close (frames)	3
	Do segment merging	Yes
	Do segment splitting	Yes
Frame-to-frame linking cost function	Brownian search radius lower bound	7
	Brownian search radius upper bound	7
Gap closing, merging and splitting cost function	Brownian search radius lower bound	7
	Brownian search radius upper bound	7
	Merging and splitting: For an end/start, the possibility of merging/splitting is allowed only if there is no possibility of gap closing	Yes (see legend for more information)
	Birth and death cost	Derived from gap closing and merging/splitting cost formulae

S5 Table: U-track tracking parameters. Shown are the parameters with non-default values (based on u-track Version 2.2.1). As explained in the original publication [7], u-track uses particle movement and intensity information to decide on merging and splitting events. After the original publication, a new constraint was added to help further distinguish between specific interactions and particles passing by each other. Specifically,

merging and splitting events can be limited to ends and starts that do not have the possibility of gap closing, a constraint used in our work here.

S1 Appendix: Clean-up of the u-track output to supply target data to FISIK

A minor clean-up of the tracking output handled the following two conditions. These conditions were overall rather infrequent. Nevertheless, we corrected them for the sake of completeness.

(1) Simultaneous merges and splits with one track (~0.3% of tracks): Occasionally, a track got merged with and split from at the same time point. In this case, merge-split pairs canceled out each other. Specifically, suppose N_M merges and N_S splits occured with the same track (call it the "base track") at the same time (S6 Fig). Then, min(N_M, N_S) merges and splits were paired with each other and separated from the base track. Any excess merges (if $N_M > N_S$) or excess splits (if $N_S > N_M$) stayed connected to the base track. The merge-split pairing was formulated as a linear assignment problem (LAP) [43], with the cost for pairing merge m with split s calculated from their mean intensities (I_m and I_s, respectively) and mean step sizes (Δr_m and Δr_s, respectively):

$$pairing\ cost(m,s) = \frac{\max(I_m, I_s)}{\min(I_m, I_s)} \times \frac{\max(\Delta r_m, \Delta r_s)}{\min(\Delta r_m, \Delta r_s)}. \quad (8)$$

(2) Merges and splits often related to detection false positives (~1% of merges and splits): Occasionally a particle detected in only one frame got linked to the track of another particle as either a merge, a split, or a split followed by a merge in the next frame. One-frame detections are mostly false positives, as real molecules overwhelmingly last for more than one frame. Therefore, such merge, split, and spit-to-merge events were

eliminated from the tracks, such that the one-frame detection was separated from the other particle.

References

1. Fujiwara T, Ritchie K, Murakoshi H, Jacobson K, Kusumi A. Phospholipids undergo hop diffusion in compartmentalized cell membrane. J Cell Biol. 2002;157(6):1071-81. doi: 10.1083/jcb.200202050. PubMed PMID: WOS:000176188100017.

2. Gebhardt JCM, Suter DM, Roy R, Zhao ZQW, Chapman AR, Basu S, et al. Single-molecule imaging of transcription factor binding to DNA in live mammalian cells. Nat Methods. 2013;10(5):421-+. doi: 10.1038/nmeth.2411. PubMed PMID: WOS:000318220000022.

3. Yildiz A, Forkey JN, McKinney SA, Ha T, Goldman YE, Selvin PR. Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. Science. 2003;300(5628):2061-5. doi: 10.1126/science.1084398. PubMed PMID: WOS:000183774900037.

4. Cutler PJ, Malik MD, Liu S, Byars JM, Lidke DS, Lidke KA. Multi-Color Quantum Dot Tracking Using a High-Speed Hyperspectral Line-Scanning Microscope. PLoS One. 2013;8(5). doi: 10.1371/journal.pone.0064320. PubMed PMID: WOS:000320362700137.

5. Huppa JB, Axmann M, Mortelmaier MA, Lillemeier BF, Newell EW, Brameshuber M, et al. TCR-peptide-MHC interactions in situ show accelerated kinetics and increased affinity. Nature. 2010;463(7283):963-7. Epub 2010/02/19. doi: 10.1038/nature08746. PubMed PMID: 20164930; PubMed Central PMCID: PMCPMC3273423.

6. Jaqaman K, Kuwata H, Touret N, Collins R, Trimble WS, Danuser G, et al. Cytoskeletal control of CD36 diffusion promotes its receptor and signaling function. Cell. 2011;146(4):593-606. Epub 2011/08/23. doi: 10.1016/j.cell.2011.06.049. PubMed PMID: 21854984; PubMed Central PMCID: PMCPMC3160624.

854 7. Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL, et al. Robust single-particle
855 tracking in live-cell time-lapse sequences. Nat Methods. 2008;5(8):695-702. Epub 2008/07/22. doi:
856 10.1038/nmeth.1237. PubMed PMID: 18641657; PubMed Central PMCID: PMCPMC2747604.
857 8. Low-Nam ST, Lidke KA, Cutler PJ, Roovers RC, Henegouwen P, Wilson BS, et al. ErbB1 dimerization
858 is promoted by domain co-confinement and stabilized by ligand binding. Nat Struct Mol Biol.
859 2011;18(11):1244-U88. doi: 10.1038/nsmb.2135. PubMed PMID: WOS:000296734200011.
860 9. Calebiro D, Sungkaworn T. Single-Molecule Imaging of GPCR Interactions. Trends Pharmacol Sci.
861 2018;39(2):109-22. doi: 10.1016/j.tips.2017.10.010. PubMed PMID: WOS:000423456500004.
862 10. Axmann M, Huppa JB, Davis MM, Schutz GJ. Determination of interaction kinetics between the T
863 cell receptor and peptide-loaded MHC class II via single-molecule diffusion measurements. Biophys J.
864 2012;103(2):L17-9. Epub 2012/08/03. doi: 10.1016/j.bpj.2012.06.019. PubMed PMID: 22853916; PubMed
865 Central PMCID: PMCPMC3400783.
866 11. Huang J, Zarnitsyna VI, Liu B, Edwards LJ, Jiang N, Evavold BD, et al. The kinetics of two-
867 dimensional TCR and pMHC interactions determine T-cell responsiveness. Nature. 2010;464(7290):932-6.
868 Epub 2010/04/02. doi: 10.1038/nature08944. PubMed PMID: 20357766; PubMed Central PMCID:
869 PMCPMC2925443.
870 12. Liu BY, Chen W, Natarajan K, Li ZH, Margulies DH, Zhu C. The cellular environment regulates in situ
871 kinetics of T-cell receptor interaction with peptide major histocompatibility complex. Eur J Immunol.
872 2015;45(7):2099-110. doi: 10.1002/eji.201445358. PubMed PMID: WOS:000357687700023.
873 13. Calebiro D, Rieken F, Wagner J, Sungkaworn T, Zabel U, Borzi A, et al. Single-molecule analysis of
874 fluorescently labeled G-protein-coupled receptors reveals complexes with distinct dynamics and
875 organization. Proc Natl Acad Sci U S A. 2013;110(2):743-8. doi: 10.1073/pnas.1205798110. PubMed PMID:
876 WOS:000313906600072.

877 14. Sungkaworn T, Jobin ML, Burnecki K, Weron A, Lohse MJ, Calebiro D. Single-molecule imaging
878 reveals receptor-G protein interactions at cell surface hot spots. Nature. 2017;550(7677):543-+. doi:
879 10.1038/nature24264. PubMed PMID: WOS:000413697800048.
880 15. Suzuki KGN, Kasai RS, Hirosawa KM, Nemoto YL, Ishibashi M, Miwa Y, et al. Transient GPI-anchored
881 protein homodimers are units for raft organization and function. Nat Chem Biol. 2012;8(9):774-83. doi:
882 10.1038/nchembio.1028. PubMed PMID: WOS:000308077600009.
883 16. Chenouard N, Smal I, de Chaumont F, Maska M, Sbalzarini IF, Gong YH, et al. Objective comparison
884 of particle tracking methods. Nat Methods. 2014;11(3):281-U47. doi: 10.1038/nmeth.2808. PubMed
885 PMID: WOS:000332086100020.
886 17. Kasai RS, Suzuki KGN, Prossnitz ER, Koyama-Honda I, Nakada C, Fujiwara TK, et al. Full
887 characterization of GPCR monomer-dimer dynamic equilibrium by single molecule imaging. J Cell Biol.
888 2011;192(3):463-80. doi: 10.1083/jcb.201009128. PubMed PMID: WOS:000287778900010.
889 18. Gourieroux C, Monfort A, Renault E. Indirect inference. Journal of Applied Econometrics.
890 1993;8(S1):S85-S118. doi: doi:10.1002/jae.3950080507.
891 19. Jaqaman K, Danuser G. Linking data to models: data regression. Nat Rev Mol Cell Biol.
892 2006;7(11):813-9. Epub 2006/09/29. doi: 10.1038/nrm2030. PubMed PMID: 17006434.
893 20. Jiang WX, Turnbull B. The indirect method: Inference based on intermediate statistics - A synthesis
894 and examples. Stat Sci. 2004;19(2):239-63. doi: 10.1214/088342304000000152. PubMed PMID:
895 WOS:000227042800003.
896 21. Schrangl L, Gohring J, Schutz GJ. Kinetic analysis of single molecule FRET transitions without
897 trajectories. J Chem Phys. 2018;148(12). doi: 10.1063/1.5006038. PubMed PMID:
898 WOS:000428866500030.

899 22. Wieser S, Axmann M, Schutz GJ. Versatile Analysis of Single-Molecule Tracking Data by
900 Comprehensive Testing against Monte Carlo Simulations. Biophys J. 2008;95(12):5988-6001. doi:
901 10.1529/biophysj.108.141655. PubMed PMID: WOS:000261559300044.
902 23. Jagiella N, Rickert D, Theis FJ, Hasenauer J. Parallelization and High-Performance Computing
903 Enables Automated Statistical Inference of Multi-scale Models. Cell Syst. 2017;4(2):194-+. doi:
904 10.1016/j.cels.2016.12.002. PubMed PMID: WOS:000395786100011.
905 24. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate Bayesian computation scheme
906 for parameter inference and model selection in dynamical systems. J R Soc Interface. 2009;6(31):187-202.
907 doi: 10.1098/rsif.2008.0172. PubMed PMID: WOS:000262757200006.
908 25. Andrews SS, Bray D. Stochastic simulation of chemical reactions with spatial resolution and single
909 molecule detail. Phys Biol. 2004;1(3):137-51. doi: 10.1088/1478-3967/1/3/001. PubMed PMID:
910 WOS:000232136300001.
911 26. Frazier Z, Alber F. A Computational Approach to Increase Time Scales in Brownian Dynamics-Based
912 Reaction-Diffusion Modeling. J Comput Biol. 2012;19(6):606-18. doi: 10.1089/cmb.2012.0027. PubMed
913 PMID: WOS:000305335100004.
914 27. West DB. Introduction to Graph Theory, Chapter 3. 2nd ed: Prentice Hall; 1999.
915 28. Lofberg J. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. Proceedings of the
916 CACSD Conference. 2004.
917 29. Gillespie DT. EXACT STOCHASTIC SIMULATION OF COUPLED CHEMICAL-REACTIONS. J Phys Chem.
918 1977;81(25):2340-61. doi: 10.1021/j100540a008. PubMed PMID: WOS:A1977EE49800008.
919 30. Elliott S, Sinclair A, Collins H, Rice L, Jelkmann W. Progress in detecting cell-surface protein
920 receptors: the erythropoietin receptor example. Ann Hematol. 2014;93(2):181-92. doi: 10.1007/s00277-
921 013-1947-2. PubMed PMID: WOS:000330208100001.

922 31. Shankaran H, Resat H, Wiley HS. Cell surface receptors for signal transduction and ligand
923 transport: A design principles study. PLoS Comput Biol. 2007;3(6):986-99. doi:
924 10.1371/journal.pcbi.0030101. PubMed PMID: WOS:000249105500009.
925 32. Moskalensky AE, Chernyshev AV, Yurkin MA, Nekrasov VM, Polshchitsin AA, Parks DR, et al.
926 Dynamic quantification of antigen molecules with flow cytometry. J Immunol Methods. 2015;418:66-74.
927 doi: 10.1016/j.jim.2015.02.001. PubMed PMID: WOS:000352660400008.
928 33. Eladdadi A, Isaacson D. A mathematical model for the effects of HER2 overexpression on cell
929 proliferation in breast cancer. Bull Math Biol. 2008;70(6):1707-29. doi: 10.1007/s11538-008-9315-4.
930 PubMed PMID: WOS:000258085500007.
931 34. James NG, Digman MA, Gratton E, Barylko B, Ding XD, Albanesi JP, et al. Number and Brightness
932 Analysis of LRRK2 Oligomerization in Live Cells. Biophys J. 2012;102(11):A41-A3. doi:
933 10.1016/j.bpj.2012.04.046. PubMed PMID: WOS:000305003100001.
934 35. Plotegher N, Gratton E, Bubacco L. Number and Brightness analysis of alpha-synuclein
935 oligomerization and the associated mitochondrial morphology alterations in live cells. Biochim Biophys
936 Acta-Gen Subj. 2014;1840(6):2014-24. doi: 10.1016/j.bbagen.2014.02.013. PubMed PMID:
937 WOS:000336012700042.
938 36. Godin AG, Rappaz B, Potvin-Trottier L, Kennedy TE, De Koninck Y, Wisemant PW. Spatial Intensity
939 Distribution Analysis Reveals Abnormal Oligomerization of Proteins in Single Cells. Biophys J.
940 2015;109(4):710-21. doi: 10.1016/j.bpj.2015.06.068. PubMed PMID: WOS:000359882100007.
941 37. Fricke F, Beaudouin J, Eils R, Heilemann M. One, two or three? Probing the stoichiometry of
942 membrane proteins by single-molecule localization microscopy. Sci Rep. 2015;5. doi: 10.1038/srep14072.
943 PubMed PMID: WOS:000361039700004.

944 38. Kruger CL, Zeuner MT, Cottrell GS, Widera D, Heilemann M. Quantitative single-molecule imaging
945 of TLR4 reveals ligand-specific receptor dimerization. Sci Signal. 2017;10(503). doi:
946 10.1126/scisignal.aan1308. PubMed PMID: WOS:000414025800003.
947 39. Holden SJ, Uphoff S, Kapanidis AN. DAOSTORM: an algorithm for high-density super-resolution
948 microscopy. Nat Methods. 2011;8(4):279-80. doi: 10.1038/nmeth0411-279. PubMed PMID:
949 WOS:000288940300005.
950 40. Min JH, Vonesch C, Kirshner H, Carlini L, Olivier N, Holden S, et al. FALCON: fast and unbiased
951 reconstruction of high-density super-resolution microscopy data. Sci Rep. 2014;4. doi:
952 10.1038/srep04577. PubMed PMID: WOS:000333797300004.
953 41. Genovesio A, Olivo-Marin JC. Split and merge data association filter for dense multi-target
954 tracking. In: Kittler J, Petrou M, Nixon M, editors. Proceedings of the 17th International Conference on
955 Pattern Recognition, Vol 4. International Conference on Pattern Recognition. Los Alamitos: Ieee Computer
956 Soc; 2004. p. 677-80.
957 42. Park C, Woehl TJ, Evans JE, Browning ND. Minimum Cost Multi-Way Data Association for
958 Optimizing Multitarget Tracking of Interacting Objects. IEEE Trans Pattern Anal Mach Intell.
959 2015;37(3):611-24. doi: 10.1109/tpami.2014.2346202. PubMed PMID: WOS:000349626200010.
960 43. Jonker R, Volgenant A. A Shortest Augmenting Path Algorithm for Dense and Sparse Linear
961 Assignment Problems. Computing. 1987;38(4):325-40. PubMed PMID: ISI:A1987J648900004.

962

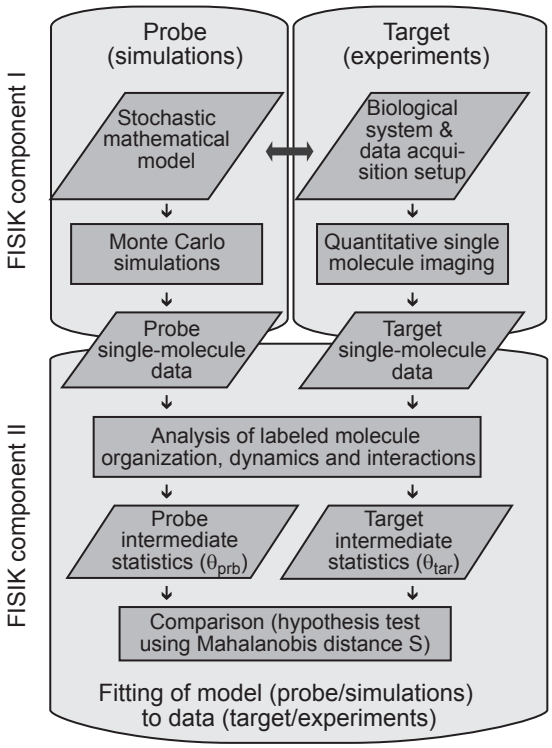

FISIK component I
Probe
(simulations)
Target
(experiments)
Stochastic mathematical model
Biological system & data acqui-sition setup
Monte Carlo simulations
Quantitative single molecule imaging
Probe single-molecule data
Target single-molecule data
FISIK component II
Analysis of labeled molecule organization, dynamics and interactions
Probe intermediate statistics (θ_{prb})
Target intermediate statistics (θ_{tar})
Comparison (hypothesis test using Mahalanobis distance S)
Fitting of model (probe/simulations) to data (target/experiments)

Figure 1

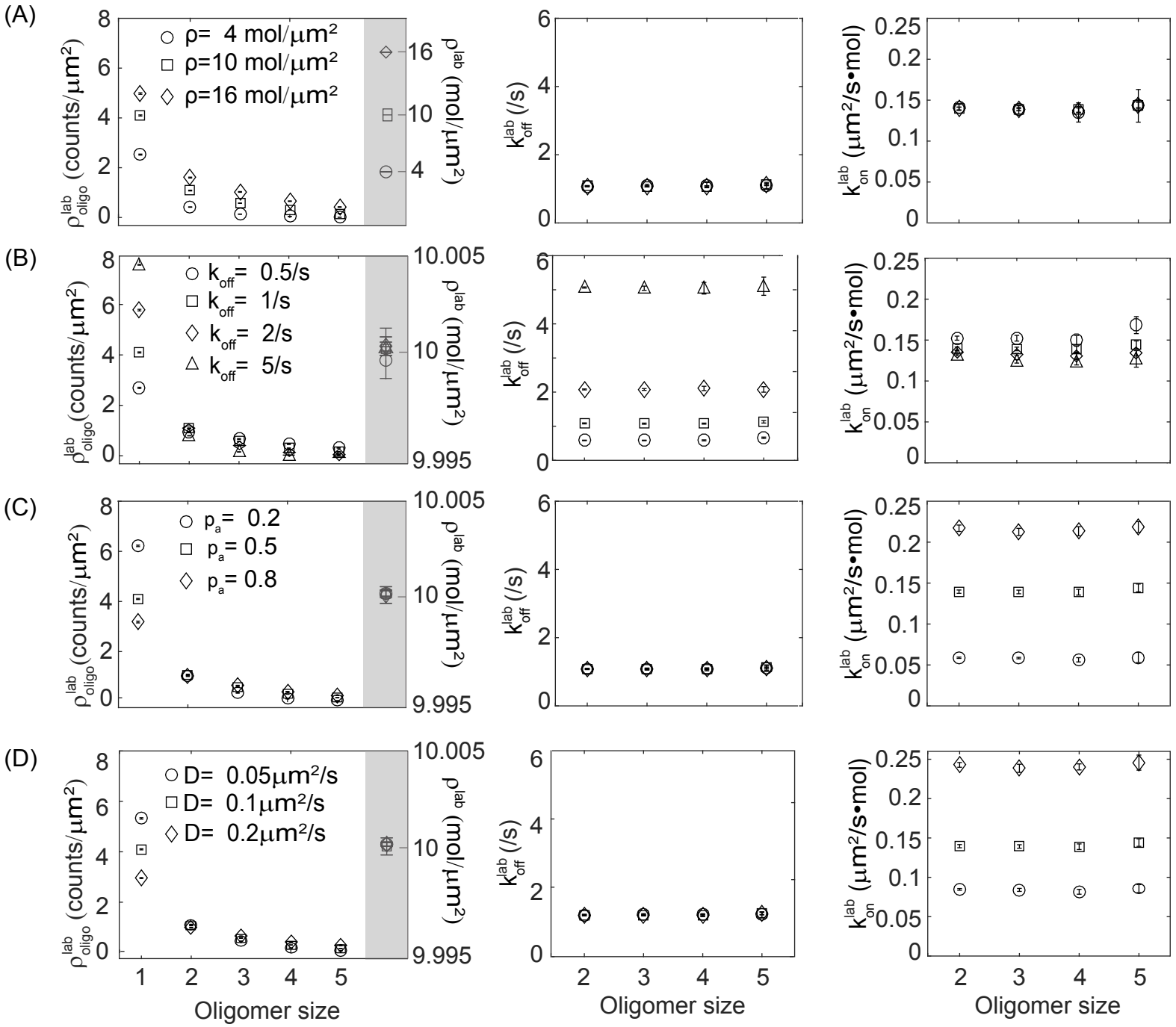

(A)
ρ= 4 mol/μm²
ρ=10 mol/μm²
ρ=16 mol/μm²
(B)
k_off= 0.5/s
k_off= 1/s
k_off= 2/s
k_off= 5/s
(C)
p_a= 0.2
p_a= 0.5
p_a= 0.8
(D)
D= 0.05μm²/s
D= 0.1μm²/s
D= 0.2μm²/s
ρ^lab_oligo (counts/μm²)
ρ^lab (mol/μm²)
k^lab_off (/s)
k^lab_on (μm²/s•mol)
Oligomer size

Figure 2

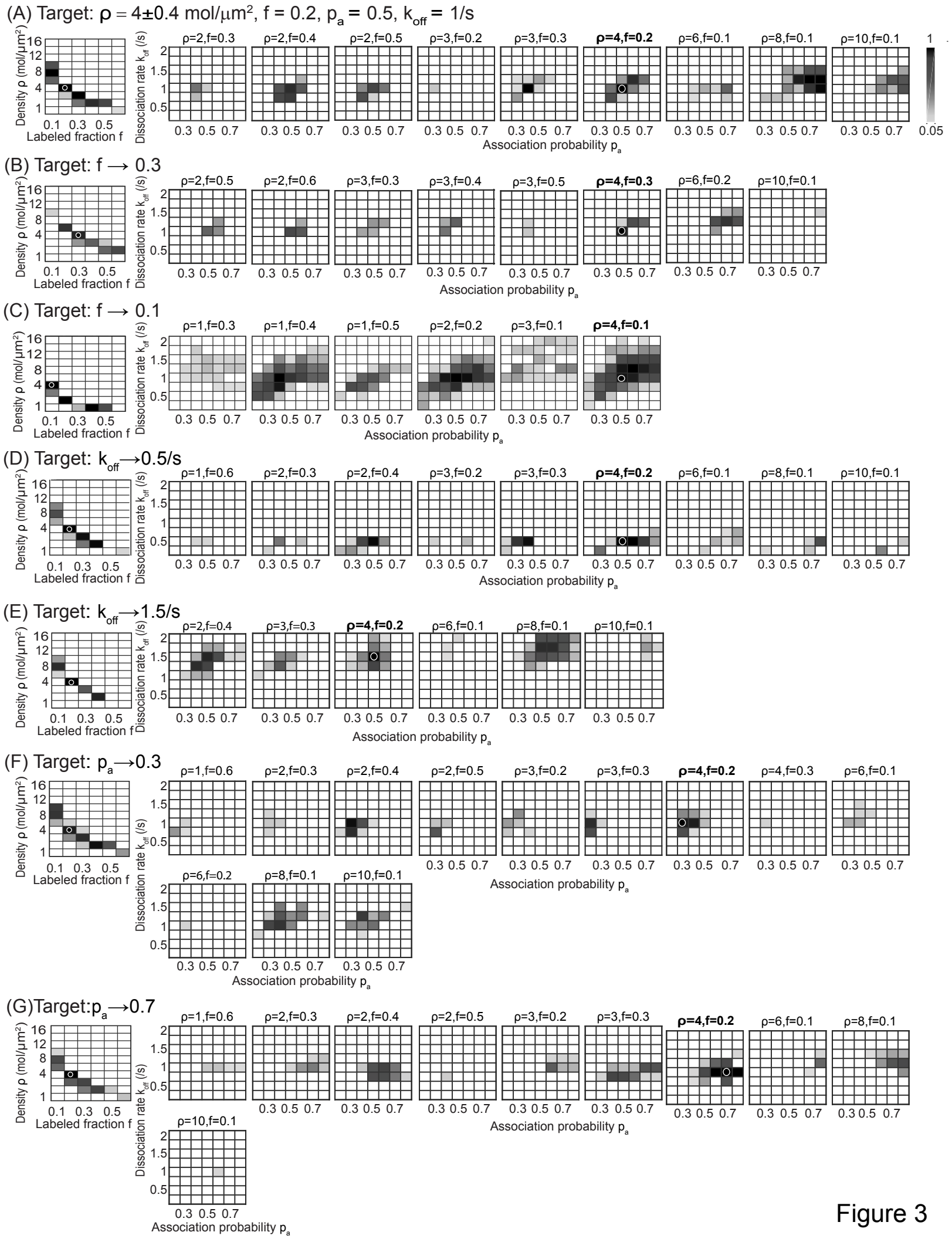

(A) Target: ρ = 4±0.4 mol/μm², f = 0.2, p_a = 0.5, k_off = 1/s
ρ=2,f=0.3
ρ=2,f=0.4
ρ=2,f=0.5
ρ=3,f=0.2
ρ=3,f=0.3
ρ=4,f=0.2
ρ=6,f=0.1
ρ=8,f=0.1
ρ=10,f=0.1
(B) Target: f → 0.3
ρ=2,f=0.5
ρ=2,f=0.6
ρ=3,f=0.3
ρ=3,f=0.4
ρ=3,f=0.5
ρ=4,f=0.3
ρ=6,f=0.2
ρ=10,f=0.1
(C) Target: f → 0.1
ρ=1,f=0.3
ρ=1,f=0.4
ρ=1,f=0.5
ρ=2,f=0.2
ρ=3,f=0.1
ρ=4,f=0.1
(D) Target: k_off → 0.5/s
ρ=1,f=0.6
ρ=2,f=0.3
ρ=2,f=0.4
ρ=3,f=0.2
ρ=3,f=0.3
ρ=4,f=0.2
ρ=6,f=0.1
ρ=8,f=0.1
ρ=10,f=0.1
(E) Target: k_off → 1.5/s
ρ=2,f=0.4
ρ=3,f=0.3
ρ=4,f=0.2
ρ=6,f=0.1
ρ=8,f=0.1
ρ=10,f=0.1
(F) Target: p_a → 0.3
ρ=1,f=0.6
ρ=2,f=0.3
ρ=2,f=0.4
ρ=2,f=0.5
ρ=3,f=0.2
ρ=3,f=0.3
ρ=4,f=0.2
ρ=4,f=0.3
ρ=6,f=0.1
ρ=6,f=0.2
ρ=8,f=0.1
ρ=10,f=0.1
(G) Target: p_a → 0.7
ρ=1,f=0.6
ρ=2,f=0.3
ρ=2,f=0.4
ρ=2,f=0.5
ρ=3,f=0.2
ρ=3,f=0.3
ρ=4,f=0.2
ρ=6,f=0.1
ρ=8,f=0.1
ρ=10,f=0.1
Density ρ (mol/μm²)
Labeled fraction f
Dissociation rate k_off (/s)
Association probability p_a
1
0.05

Figure 3

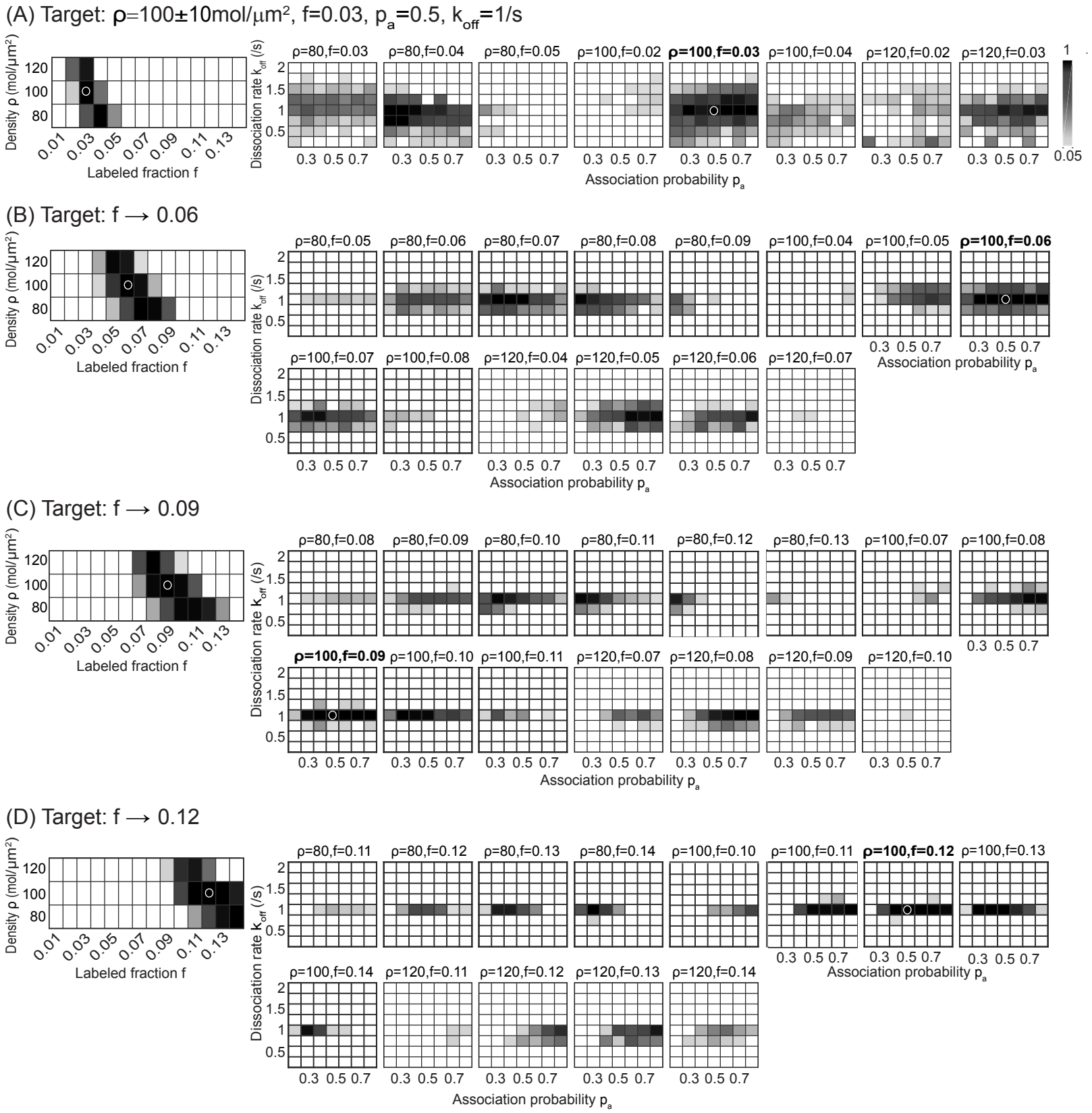
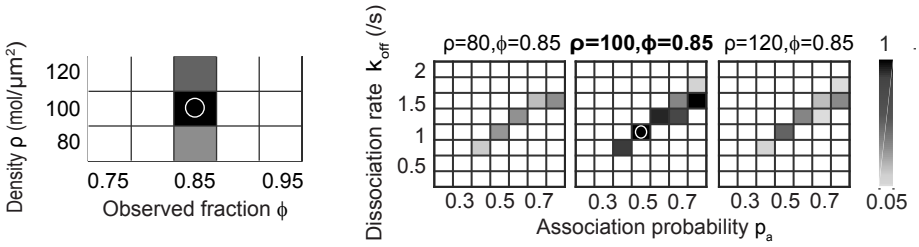
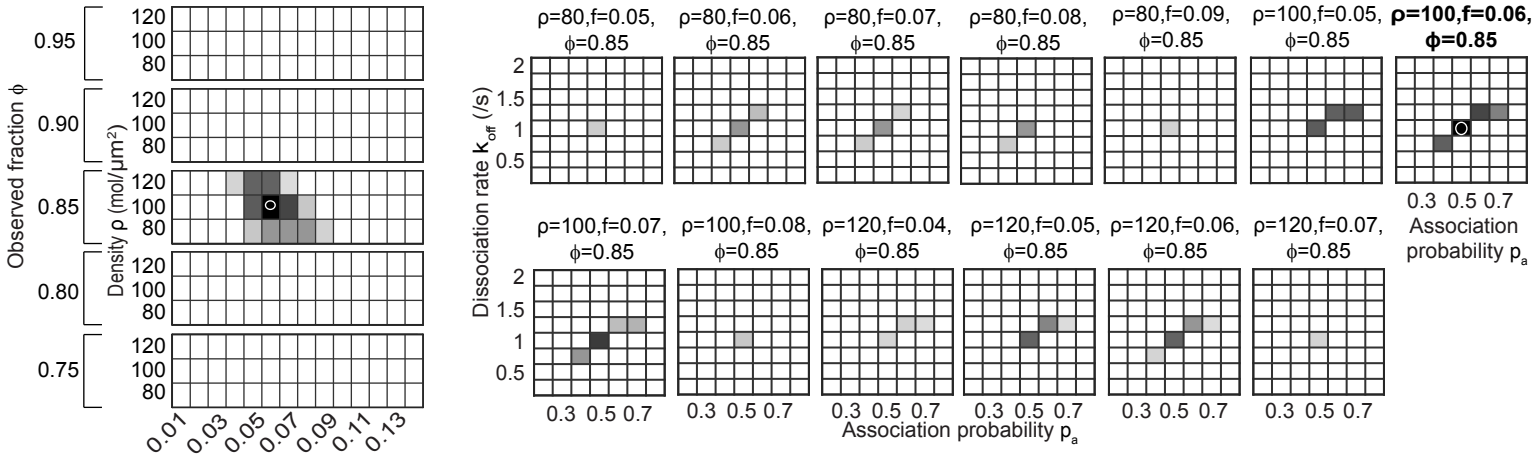

(A) Target: ρ=100±10mol/μm², f=0.03, pa=0.5, koff=1/s
Density ρ (mol/μm²)
Labeled fraction f
Dissociation rate koff (/s)
Association probability pa
ρ=80,f=0.03
ρ=80,f=0.04
ρ=80,f=0.05
ρ=100,f=0.02
ρ=100,f=0.03
ρ=100,f=0.04
ρ=120,f=0.02
ρ=120,f=0.03
(B) Target: f → 0.06
ρ=80,f=0.05
ρ=80,f=0.06
ρ=80,f=0.07
ρ=80,f=0.08
ρ=80,f=0.09
ρ=100,f=0.04
ρ=100,f=0.05
ρ=100,f=0.06
ρ=100,f=0.07
ρ=100,f=0.08
ρ=120,f=0.04
ρ=120,f=0.05
ρ=120,f=0.06
ρ=120,f=0.07
(C) Target: f → 0.09
ρ=80,f=0.08
ρ=80,f=0.09
ρ=80,f=0.10
ρ=80,f=0.11
ρ=80,f=0.12
ρ=80,f=0.13
ρ=100,f=0.07
ρ=100,f=0.08
ρ=100,f=0.09
ρ=100,f=0.10
ρ=100,f=0.11
ρ=120,f=0.07
ρ=120,f=0.08
ρ=120,f=0.09
ρ=120,f=0.10
(D) Target: f → 0.12
ρ=80,f=0.11
ρ=80,f=0.12
ρ=80,f=0.13
ρ=80,f=0.14
ρ=100,f=0.10
ρ=100,f=0.11
ρ=100,f=0.12
ρ=100,f=0.13
ρ=100,f=0.14
ρ=120,f=0.11
ρ=120,f=0.12
ρ=120,f=0.13
ρ=120,f=0.14

Figure 4

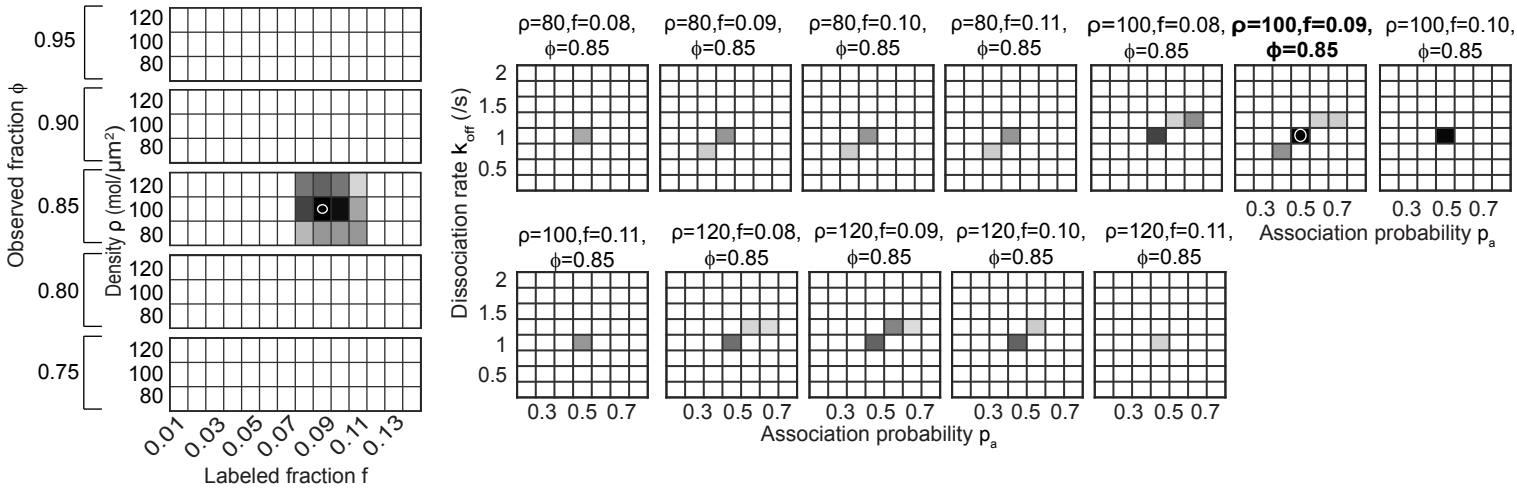
(A) Target: ρ=100±10mol/μm^2, ϕ=0.85, p_a=0.5, k_{off}=1/s



(B) Target: ρ=100±10mol/μm^2, f=0.06, ϕ=0.85, p_a=0.5, k_{off}=1/s



(C) Target: f→0.09



(D) Target: f→0.12

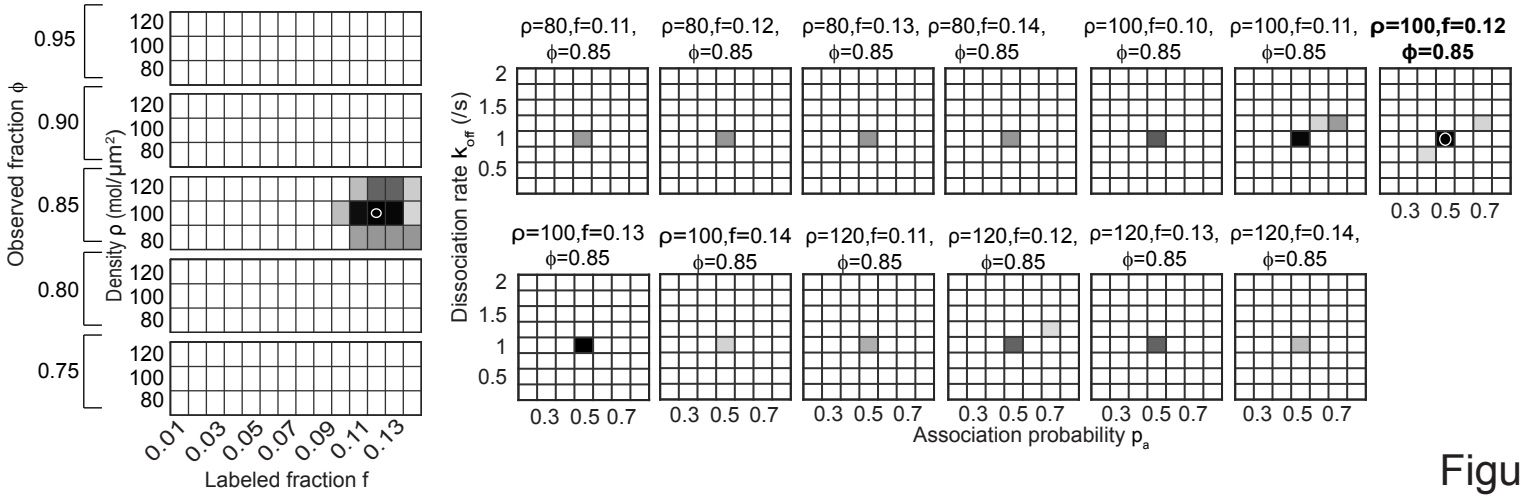
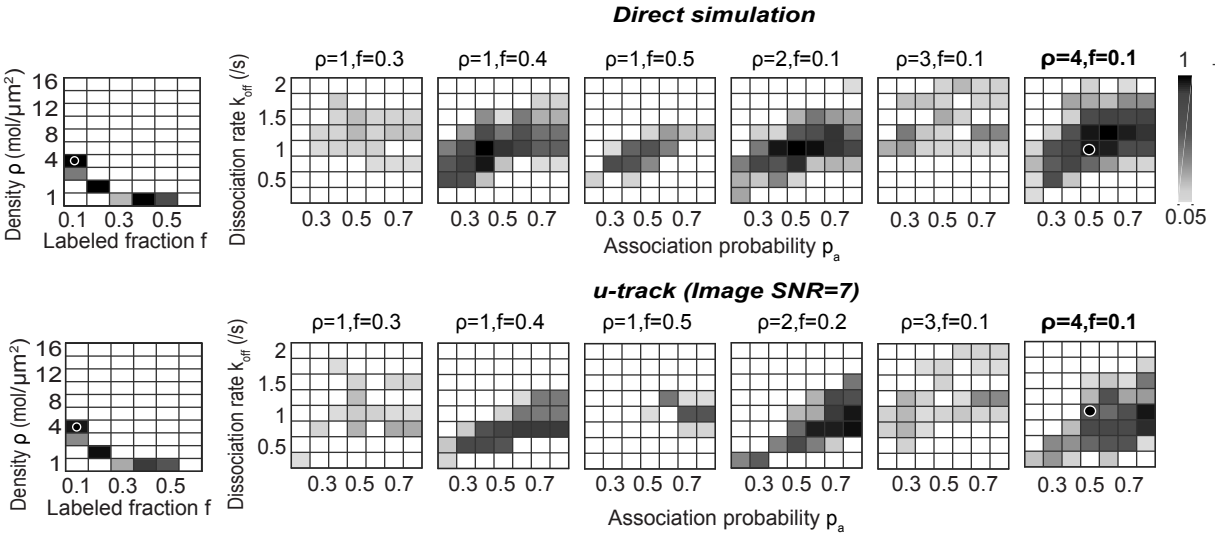
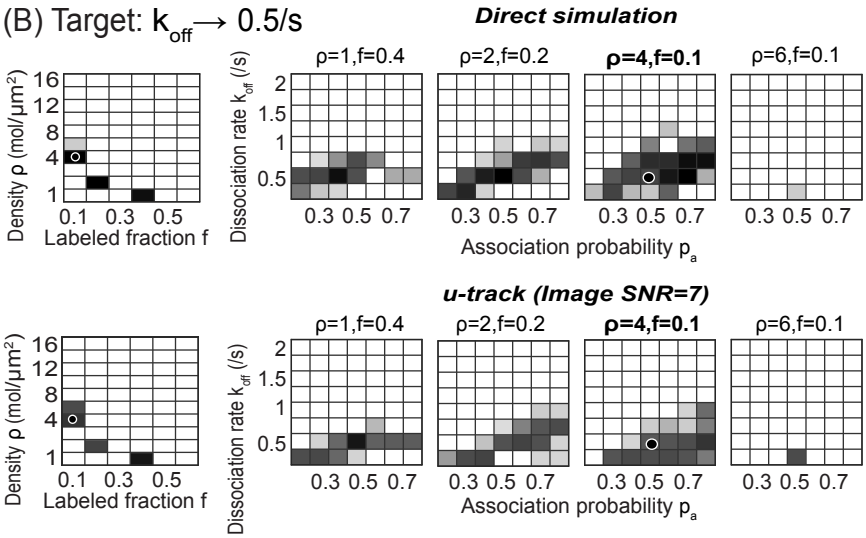


Figure 5

(A) Target: ρ=4±0.4 mol/μm^2, f=0.1, p_a=0.5, k_{off}=1/s



(B) Target: $k_{off} \rightarrow$ 0.5/s



(C) Target: $k_{off} \rightarrow$ 1.5/s

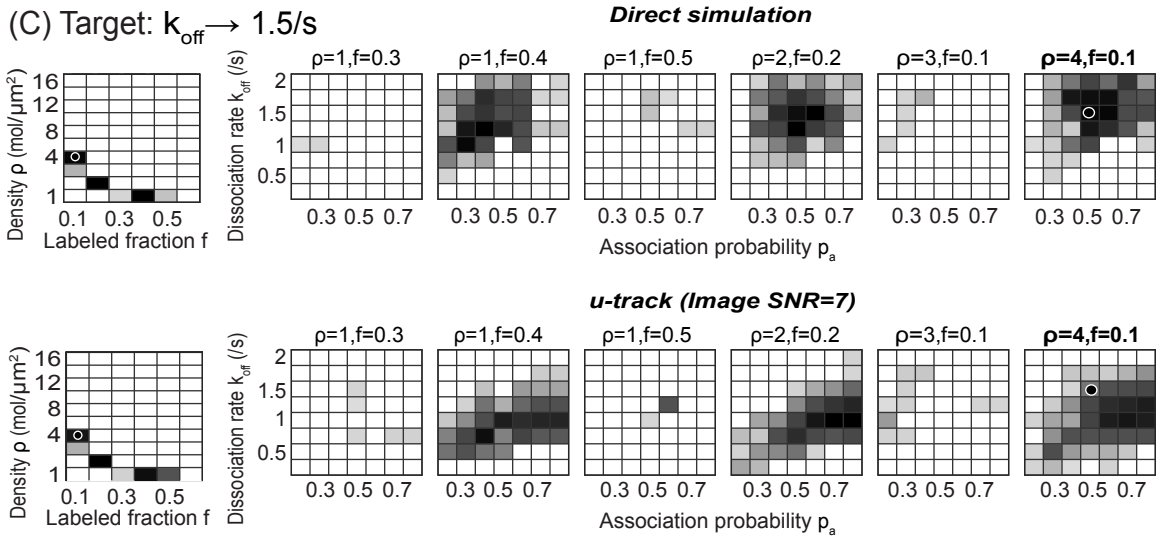
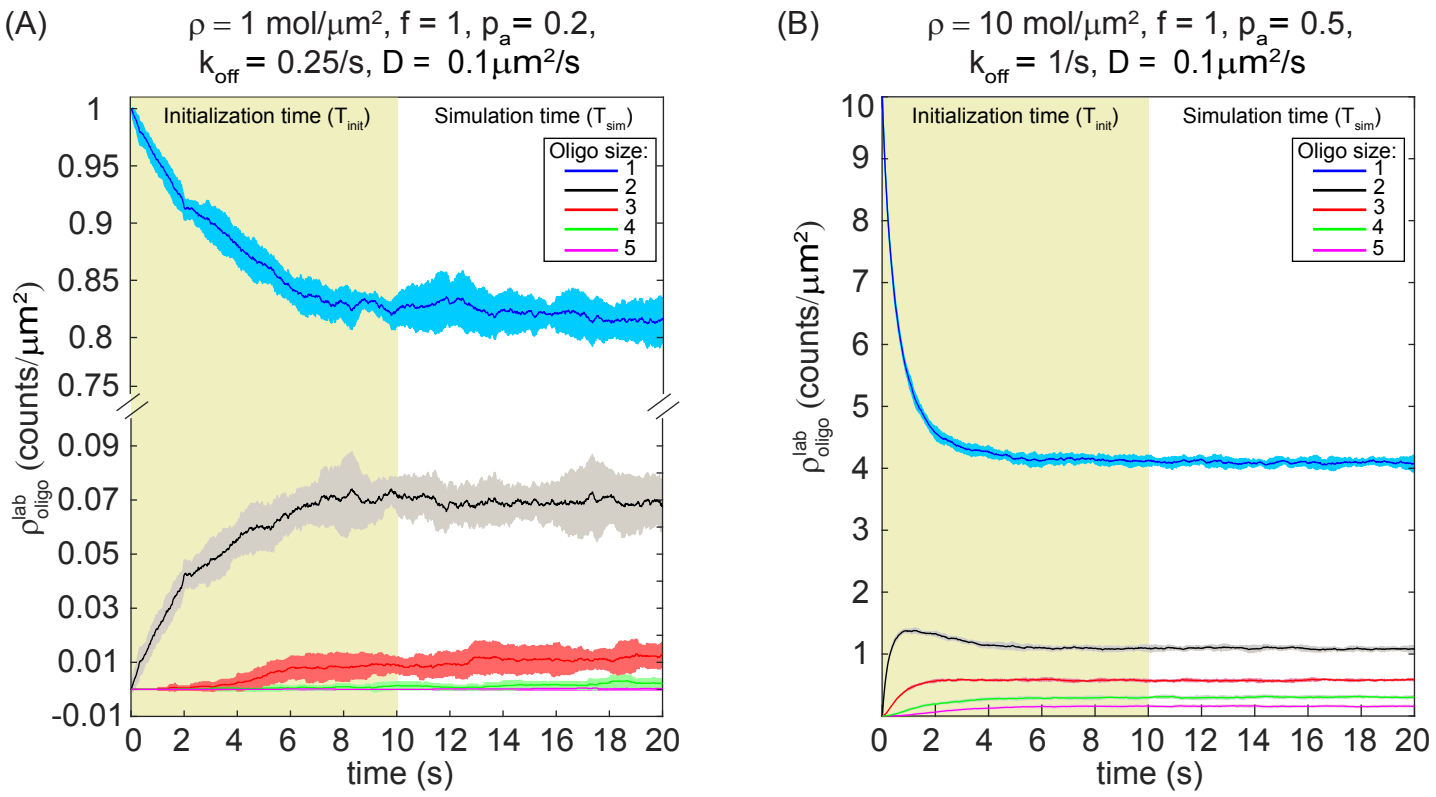
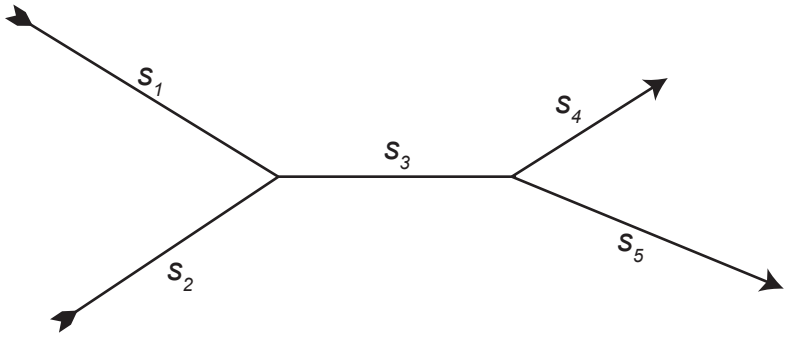


Figure 6

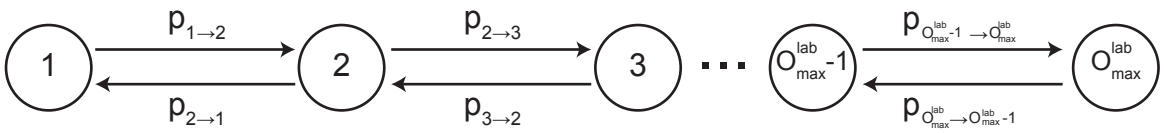


Supplementary figure 1

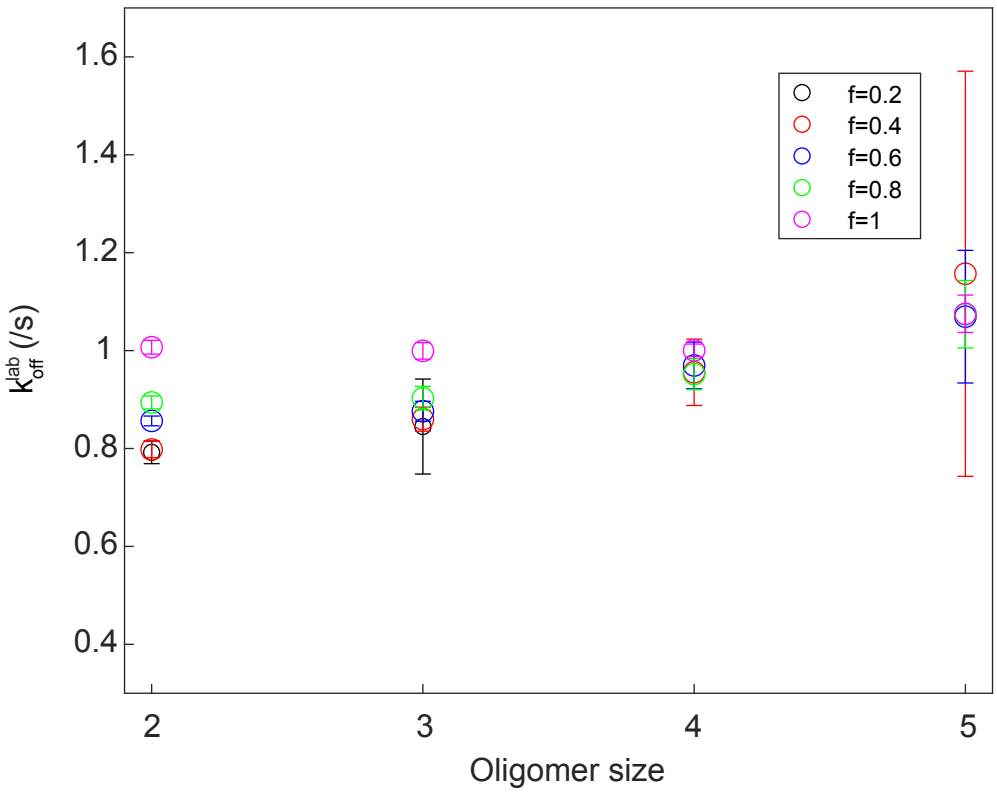
(A)



(B)

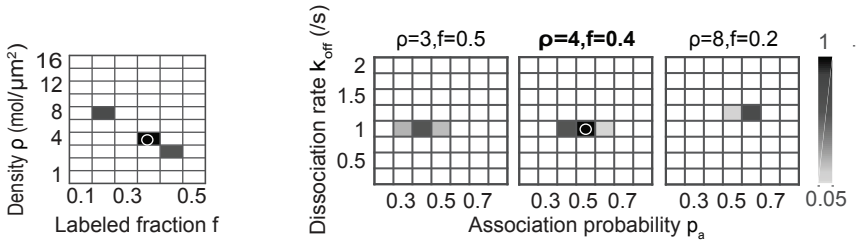


Supplementary figure 2

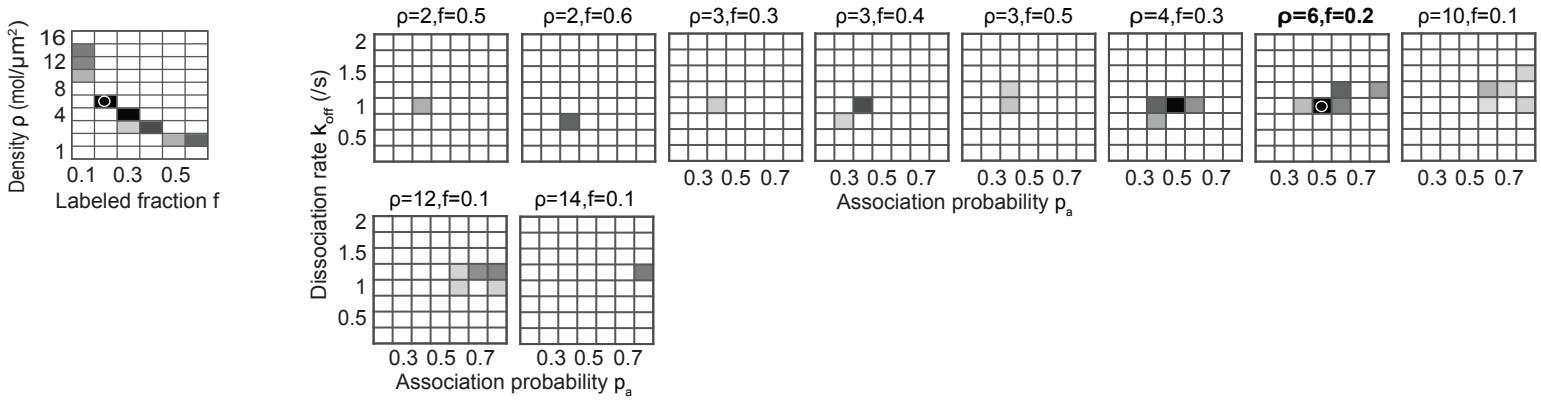


Supplementary figure 3

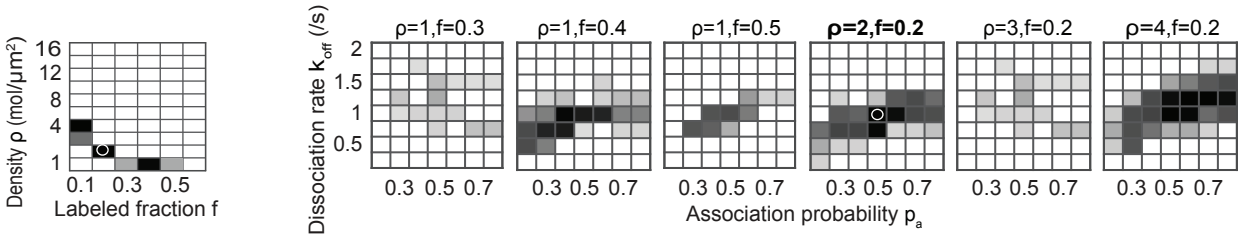
(A) Target: $\rho = 4 \pm 0.4$ mol/μm^2, f = 0.4, p_a = 0.5, k_{off} = 1/s



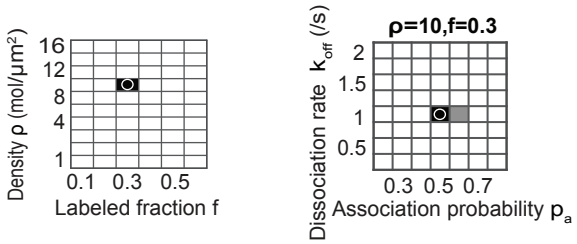
(B) Target: $\rho = 6 \pm 0.6$ mol/μm^2, f = 0.2, p_a = 0.5, k_{off} = 1/s



(C) Target: $\rho = 2 \pm 0.2$ mol/μm^2, f = 0.2, p_a = 0.5, k_{off} = 1/s

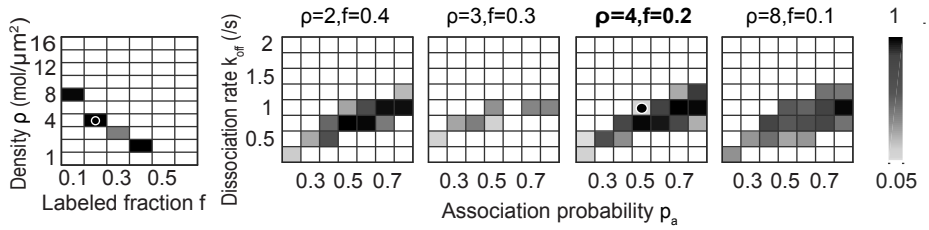


(D) Target: $\rho = 10 \pm 1$ mol/μm^2, f=0.3, p_a=0.5, k_{off}=1/s

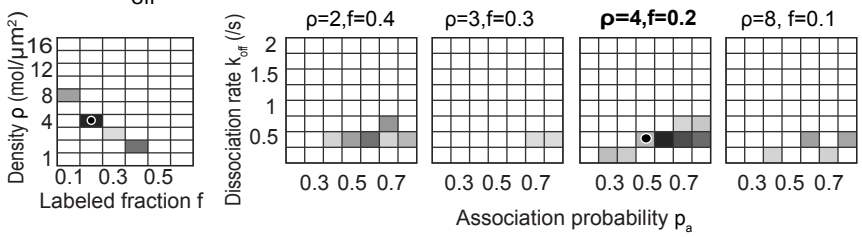


Supplementary figure 4

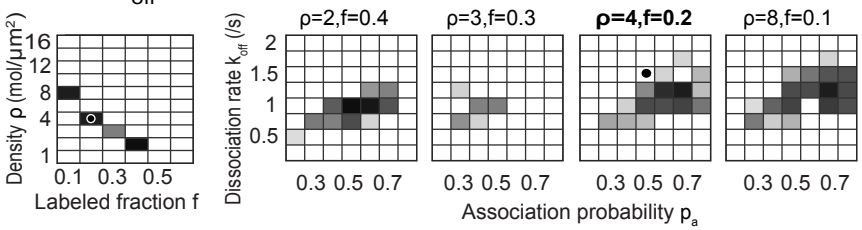
(A) Target: $\rho = 4 \pm 0.4$ mol/μm^2, f = 0.2, p_a = 0.5, k_{off} = 1/s, u-track (SNR=20)



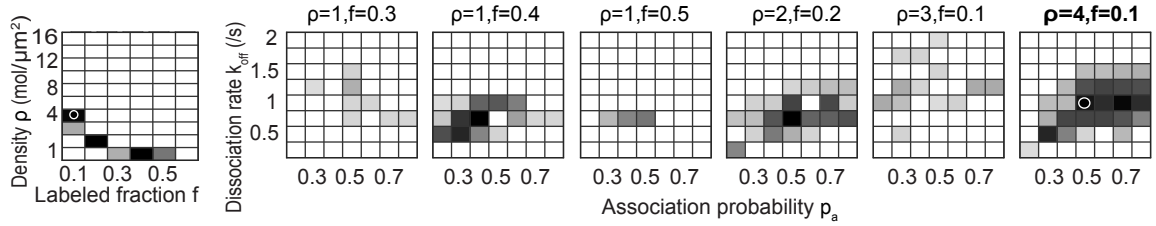
Target: $k_{off} \rightarrow$ 0.5/s



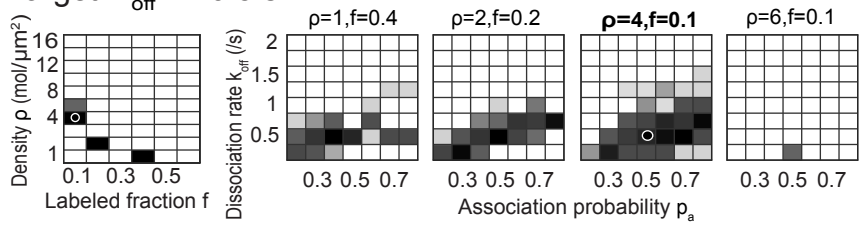
Target: $k_{off} \rightarrow$ 1.5/s



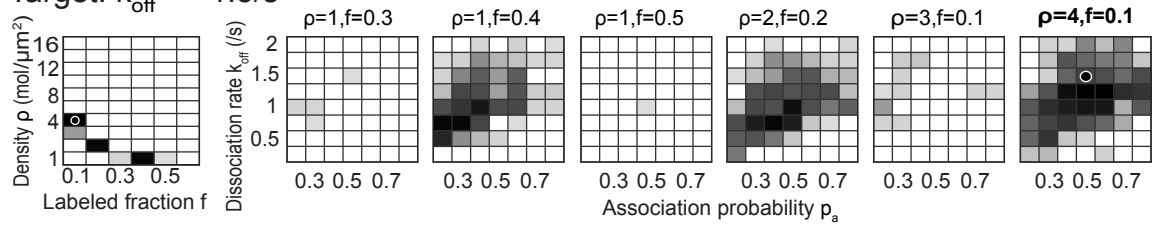
(B) Target: $\rho = 4 \pm 0.4$ mol/μm^2, f = 0.1, p_a = 0.5, k_{off} = 1/s, u-track (SNR=20)



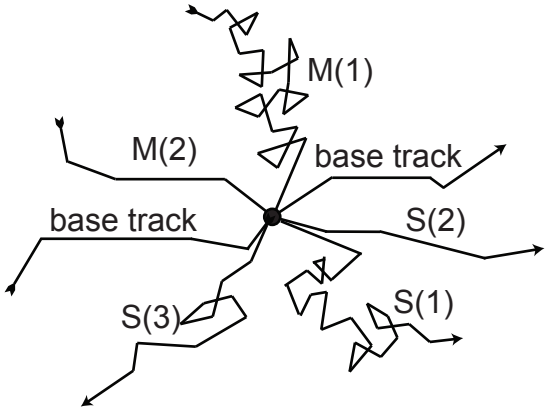
Target: $k_{off} \rightarrow$ 0.5/s



Target: $k_{off} \rightarrow$ 1.5/s



Supplementary figure 5



Supplementary figure 6