RESEARCH

Computational analysis of protein synthesis, diffusion, and binding in compartmental biochips

Stefanie Förste[1], Ohad Vonshak[2], Shirley S. Daube[2], Roy H. Bar-Ziv[2], Reinhard Lipowsky[1] and Sophia Rudorf[3*]

Abstract

Protein complex assembly facilitates the combination of individual protein subunits into functional entities, and thus plays a crucial role in biology and biotechnology. Recently, we developed quasi-twodimensional, silicon-based compartmental biochips that are designed to study and administer the synthesis and assembly of protein complexes. At these biochips, individual protein subunits are synthesized from locally confined high-density DNA brushes and are captured on the chip surface by molecular traps. Here, we investigate single-gene versions of our quasi-twodimensional synthesis systems and introduce the trap-binding efficiency to characterize their performance. We show by mathematical and computational modeling how a finite trap density determines the dynamics of protein-trap binding and identify three distinct regimes of the trap-binding efficiency. We systematically study how protein-trap binding is governed by the system's three key parameters, which are the synthesis rate, the diffusion constant and the trap-binding affinity of the expressed protein. In addition, we describe how spatially differential patterns of traps modulate the protein-trap binding dynamics. In this way, we extend the theoretical knowledge base for synthesis, diffusion, and binding in compartmental systems, which helps to achieve better control of directed molecular self-assembly required for the fabrication of nanomachines for synthetic biology applications or nanotechnological purposes.

Background

The plethora of tasks that biomolecules autonomously fulfill inside living cells inspired and triggered many studies on the use of peptides [1–3], nucleic acids [4–9], or polysaccharides [10–12] as building blocks for artificial nanodevices. Important properties of biomolecular nanodevices are bio-compatibility, stimuli-responsiveness,

and the capability to self-assemble [13]. These devices have the potential to serve for example as delivery systems for drugs or genetic material [14, 15], for molecular computation [16], or as nanosensors [17]. However, for many applications the precise and reliable fabrication of nanomachines *via* self-assembly still remains a major challenge [18, 19]. One particular problem is the inherent complexity of the self-assembly of any multi-component nanomachine [20, 21]. Freely diffusing components can interact in many different ways, which results in unwanted side-reactions and the formation of alternative, non-functional complexes. These off-pathway reactions cause a reduction of product quality and yield by diverting and sequestering resources [22]. There are different strategies to tackle these challenges. As an example, spatio-temporal control of the assembly process is achieved by compartmentalization of the reaction chamber [23].

*Correspondence:
Sophia Rudorf
rudorf@cell.uni-hannover.de
[1] Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany
[2] Department of Chemical and Biological Physics, The Weizmann Institute of Science, Rehovot 7610001, Israel
[3] Institute of Cell Biology and Biophysics, Leibniz University Hannover, 30419 Hannover, Germany

Förste *et al. Microbial Cell Factories*　(2023) 22:244

Page 2 of 13

Furthermore, fine-tuning the stoichiometry of the individual components [24] or using molecular assembly lines [25] can reduce the probability of off-pathway reactions.

Recently, we presented systems for the in-vitro synthesis and assembly of protein complexes in quasi-twodimensional silicon chips [26]. In these systems, we separate component synthesis from component assembly by applying a soft, membrane-less compartmentalization strategy: First, we attach DNA-brushes to confined regions on the surfaces of the chips. Second, the coding sequences that are comprised by the brushes are used as templates for the synthesis of protein subunits *via* cell-free transcription and translation. Third, assembly of the subunits into protein complexes is coupled to their synthesis and occurs either in solution or by scaffolding on the surfaces of the chips, which are coated with molecular traps.

In these systems, the dynamics of the assembly of nanomachines are determined by a specific set of parameters. This set comprises the rates of component synthesis, the diffusion constants of the components, and the rate constants for component-component binding and component-trap binding. An efficient fabrication of nanomachines is only possible when these system parameters are well-adjusted to facilitate correct assembly and to suppress off-pathway reactions. Therefore, understanding how each of these parameters influences the assembly of nanomachines is a fundamental prerequisite for the effective genetic and geometric control of this process. In practice, depending on the synthesis process it might not be feasible to adjust all of the parameters or the range of possible parameter values might be highly restricted: Protein complexes assemble from defined subunits, which means that the options for molecular modifications are limited by the requirement that the functionality and integrity of the complex need to be preserved [27]. However, synthesis rates, diffusion constants and binding affinities of subunits and surface traps are in principle adjustable, e.g. *via* changing temperature, DNA-brush densities and compositions, the viscosity of the reaction solution, and the particular choice of the molecular tags for trap binding. Furthermore, the biochip allows for adjustable positioning of the DNA brushes, introducing a distinct spatial component that is unique to the system.

Their adjustability turns synthesis rates, diffusion constants and binding affinities into system parameters that are particularly interesting to study. In multi-component systems, the impact of the synthesis rate, the diffusion constant and the trap-binding affinity may be obscured by the complexity of the reaction network. To better understand the fundamental process, we here focus on a simplified version of the biochip studied previously [26]. In this simplified system, only one type of protein subunit is actively expressed whereas subunits of a second type are tethered to the surface of the silicon chip and serve themselves as traps for the actively expressed subunits, see Fig. 1. Consequently, assembly of two-subunit complexes can only occur at the surface but not in solution and the performance of these systems in terms of two-subunit interaction can be characterized by their trap-binding efficiency.

Here, we model these simplified biochips for protein synthesis and binding as spatially inhomogeneous reaction-diffusion systems with finite trap densities and three adjustable system parameters: (i) synthesis rate, (ii) diffusion constant, and (iii) trap-binding affinity of the expressed protein subunit. We show how the trap-binding efficiency is determined by the aforementioned system parameters and study the time evolution of protein-trap
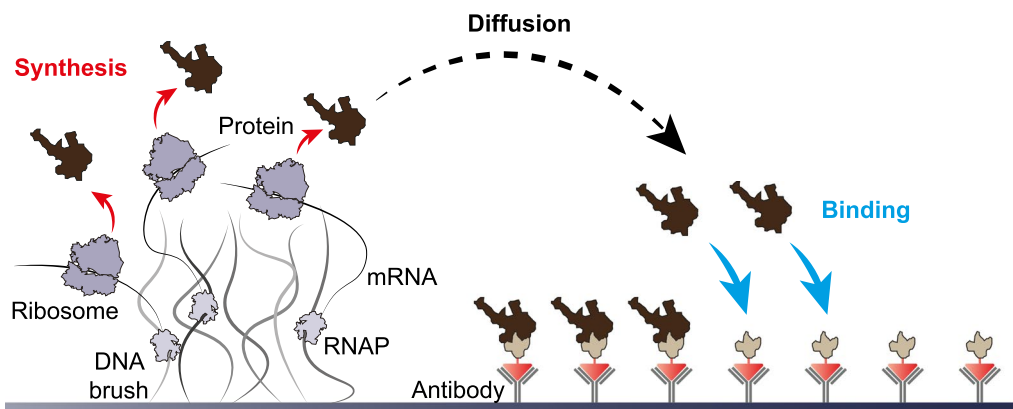

Fig. 1 Schematic representation of a compartmentalized biochip. A DNA brush is attached to the chip surface. The encoded sequences are transcribed by RNA polymerases (light gray) and translated by ribosomes (dark gray). The synthesized proteins (brown) spread by diffusion and bind to a second type of proteins (light brown) tethered to the chip surface by antibodies

binding. In particular, we identify three distinct regimes of the trap-binding efficiency that arise from the interplay of protein synthesis and binding. Furthermore, we study the spatial distribution of bound proteins and investigate the influence of trap patterning on the trap-binding efficiency. Finally, we test the predictive power of our computational modeling by comparison with experimental data.

Results and discussion

Trap-binding efficiency in one-box systems with finite trap density undergoes a transition from a binding-dominated to a synthesis-dominated regime

We consider a quasi-twodimensional system with height h and length l for which $h \ll l$, i.e., a very flat box as sketched in Fig. 2a). The box is filled with a solution of

proteins, where n new proteins are added to the box per unit of time $[t]$ at the constant synthesis rate $\alpha_{syn} \equiv n/[t]$. Protein synthesis is realized by in-vitro synthesis within the solution. The bottom of the box is coated with traps to which free proteins bind in a quasi-irreversible manner with probability p per unit of time $[t]$, i.e., with binding affinity $\kappa_{bin} \equiv p/[t]$. The trap density is finite and each trap can only bind one protein. Therefore, the number $N_b(t)$ of bound proteins is always less or equal than a maximal value N_b^{max}, which corresponds to the total number of traps provided the lateral size of the proteins is smaller than the trap separation. We assume that the solution is perfectly mixed at all times so that the time evolution of the numbers $N_u(t)$ and $N_b(t)$ of unbound and bound proteins is described by
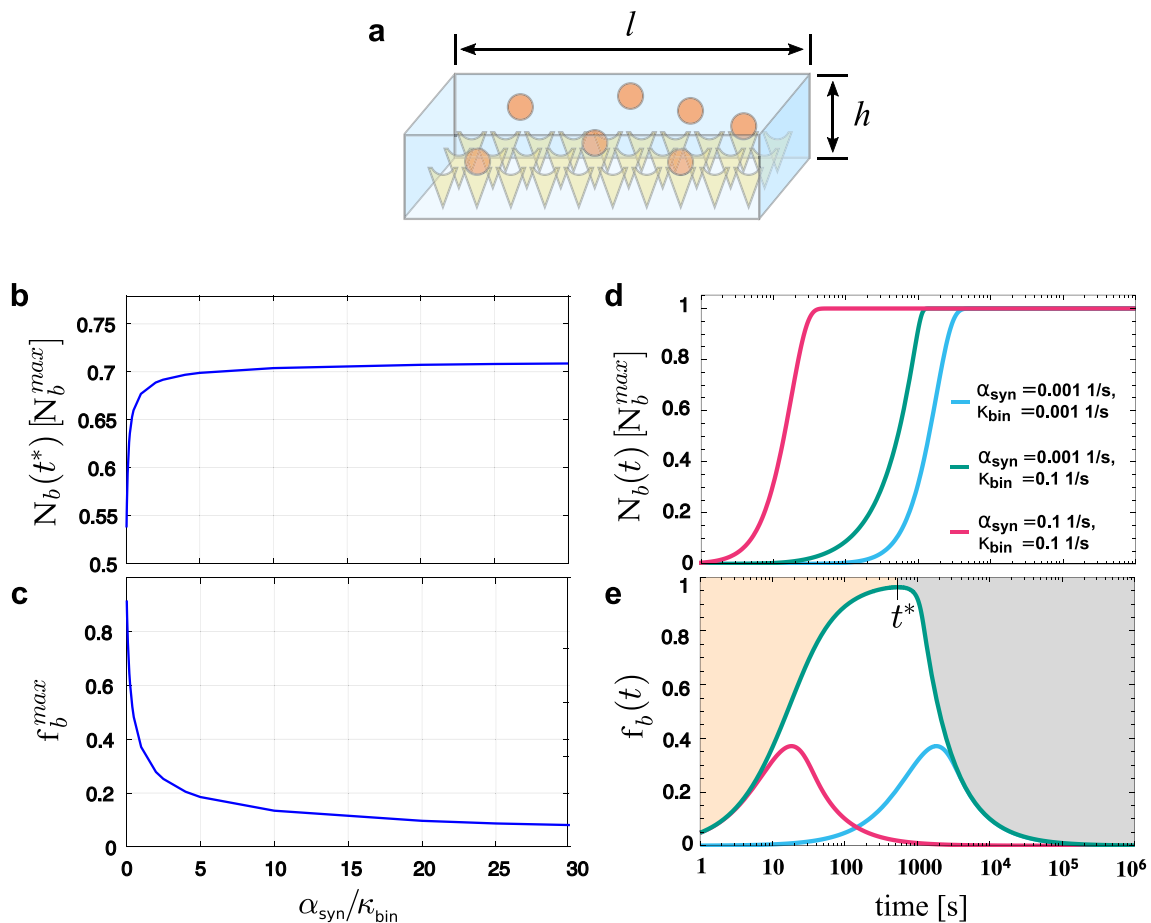

Fig. 2 Binding dynamics of the one-box system with limited trap density. **a** One-box system with finite trap density. The quasi-twodimensional synthesis chip of height h and length l with $h \ll l$ is filled with a solution of proteins (orange spheres) that bind to molecular traps (yellow cones) to form protein-trap complexes. **b** Number $N_b(t^*)$ of bound proteins at $t = t^*$ and **c** maximum of the trap-binding efficiency $f_b^{max} = f_b(t^*)$ for different ratios of synthesis rate α_{syn} and binding affinity κ_{bin}. **d** Number $N_b(t)$ of proteins bound to traps normalized to the maximal value N_b^{max} and **e** time-evolution of the trap-binding efficiency $f_b(t)$ for different choices of synthesis rate α_{syn} and binding affinity κ_{bin}. Note that the maximum of the trap-binding efficiency f_b^{max} is determined by the ratio of synthesis rate α_{syn} and binding affinity κ_{bin}, see also (**c**). For the light blue and magenta lines this ratio is one, thus f_b^{max} is identical but reached at different time points. Background color indicates binding-dominated (orange) and synthesis-dominated (gray) regime for $\alpha_{syn} = 0.001$ s^{-1} and $\kappa_{bin} = 0.1$ s^{-1}. Transition from the former to the latter occurs at $t = t^*$, i.e., when the maximum f_b^{max} of the trap-binding efficiency is reached

Förste *et al. Microbial Cell Factories*   (2023) 22:244

Page 4 of 13

$$\frac{\mathrm{d}}{\mathrm{d}t}N_u(t) = -\kappa_{bin}N_u(t)\left(N_b^{max} - N_b(t)\right) + \alpha_{syn}, \qquad (1)$$

$$\frac{\mathrm{d}}{\mathrm{d}t}N_b(t) = \kappa_{bin}N_u(t)\left(N_b^{max} - N_b(t)\right), \qquad (2)$$

and mass conservation

$$N(t) \equiv N_b(t) + N_u(t) = \alpha_{syn}t + N_b(0) + N_u(0) \quad (3)$$

with initial numbers $N_b(0)$ and $N_u(0)$ of bound and unbound proteins. A solution of this initial value problem is given by Eq. (15) in the Methods section. In the following, we investigate the time evolution of the fraction $f_b(t)$ of bound proteins in the box

$$f_b(t) \equiv N_b(t)/N(t) \qquad (4)$$

and refer to $f_b(t)$ as the *trap-binding efficiency* of the system. For more efficient systems, the maximum f_b^{max} of the trap-binding efficiency gets closer to the highest possible value 1, is reached at an earlier point in time $t = t^*$ after protein synthesis has started and stays at a high level for a longer time than for less efficient systems. Figure 2 shows examples for the number $N_b(t^*)$ of bound proteins at $t = t^*$ (Fig. 2b), the maximum $f_b^{max} = f_b(t^*)$ of the trap-binding efficiency (Fig. 2c), the time evolution of the number $N_b(t)$ of bound proteins (Fig. 2d), and the time evolution of the trap-binding efficiency $f_b(t)$ (Fig. 2e) for different synthesis rates α_{syn} and binding affinities κ_{bin}: As expected, the number $N_b(t)$ of bound proteins increases and approaches the plateau value N_b^{max}. The trap-binding efficiency $f_b(t)$ increases after the onset of protein synthesis at a rate that is determined by the binding affinity κ_{bin}. For longer timescales, the synthesis rate α_{syn} determines how fast $f_b(t)$ approaches zero as the long-term behavior of the trap-binding efficiency is governed by $f_b(t) \propto N_b^{max}\left(\alpha_{syn}t\right)^{-1}$. Thus, for each particular choice of κ_{bin} and α_{syn}, we can distinguish a binding-dominated and a synthesis-dominated regime for the time evolution of the trap-binding efficiency. The transition from the first to the latter regime occurs at time t^* when the trap-binding efficiency has reached its maximum f_b^{max} and starts to decrease, see Fig. 2e). At this time, many traps are already saturated, whereas the total number of proteins keeps to increase linearly. The time t^*, the trap-binding efficiency maximum f_b^{max} and the number $N_b(t^*)$ of bound proteins at $t = t^*$ depend on the binding affinity κ_{bin} and the synthesis rate α_{syn}, see Figs. 2 b), c), and e). They can be determined by solving the initial value problem defined in Eqs. (1) and (2) together with $\frac{\mathrm{d}}{\mathrm{d}t}f_b(t)|_{t=t^*} = 0$.

Trap-binding efficiency in two-box systems with inhomogeneous synthesis and slow diffusion exhibits three distinct regimes

We extend the one-box system introduced above by a second box that is identical and directly connected to the first one, see Fig. 3a). As in the one-box system, both boxes 1 and 2 of the two-box system are filled with a protein solution and their bottom is coated with traps that the proteins can bind to. Both boxes of the system are open in the sense that proteins can freely diffuse between them with *diffusion rate* $\widetilde{D}$. The diffusion rate $\widetilde{D}$ is defined as the diffusion constant D of the proteins scaled by the box length l, $\widetilde{D} \equiv D/l^2$, and is expressed by the inverse unit of time $[t]^{-1}$. In contrast to the one-box system, protein synthesis in the two-box system is not homogeneous. Instead, proteins are only synthesized in box 1 but not in box 2. Thus, the dynamics of the numbers $N_{i,u}(t)$ and $N_{i,b}(t)$ of unbound and bound proteins in boxes $i = 1$ and 2 are described by

$$\frac{\mathrm{d}}{\mathrm{d}t} = \widetilde{D}\left(N_{2,u}(t) - N_{1,u}(t)\right) - \kappa_{bin}N_{1,u}(t) \\ \left(N_{1,b}^{max} - N_{1,b}(t)\right) + \alpha_{syn}, \qquad (5)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} = \kappa_{bin}N_{1,u}(t)\left(N_{1,b}^{max} - N_{1,b}(t)\right), \qquad (6)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} = \widetilde{D}\left(N_{1,u}(t) - N_{2,u}(t)\right) - \kappa_{bin}N_{2,u}(t) \\ \left(N_{2,b}^{max} - N_{2,b}(t)\right), \quad \text{and} \qquad (7)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} = \kappa_{bin}N_{2,u}(t)\left(N_{2,b}^{max} - N_{2,b}(t)\right), \qquad (8)$$

with synthesis rate α_{syn} and binding affinity κ_{bin} defined in the context of Eq. (1). In addition, mass conservation leads to the relationship

$$N(t) = N_{1,u}(t) + N_{1,b}(t) + N_{2,u}(t) + N_{2,b}(t) \\ = \alpha_{syn}t + N_{1,u}(0) + N_{1,b}(0) + N_{2,u}(0) + N_{2,b}(0). \qquad (9)$$

The time evolution of the numbers of bound proteins in box 1 and 2 is shown in Fig. 3b): Due to the spatially inhomogeneous synthesis of proteins in the two-box system, the traps in box 2 are saturated at a later point in time than the traps in box 1. For a given synthesis rate α_{syn} and binding affinity κ_{bin}, the lag time depends on the diffusion rate $\widetilde{D}$ and is negligible for very fast diffusion (i.e., in well-mixed systems) as expected. A diffusion that is slow enough to cause a lag in the saturation of traps in box 2 has also an impact on the trap-binding efficiency
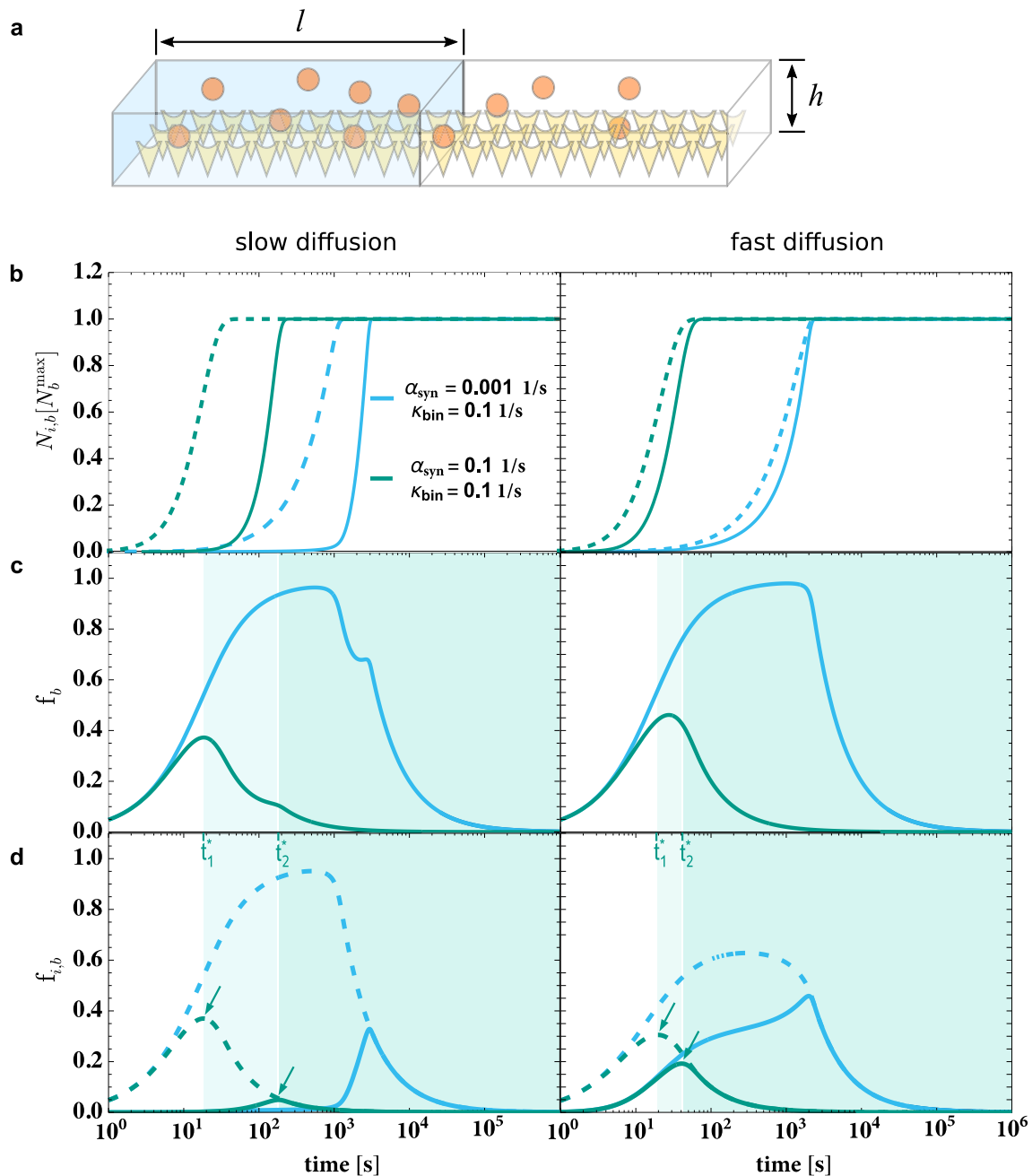
Fig. 3 Binding dynamics in a two-box system with diffusion rates $\widetilde{D} = 0.001 \ \mathrm{s^{-1}}$ (left) and $\widetilde{D} = 0.1 \ \mathrm{s^{-1}}$ (right), and synthesis rates α_{syn} and binding affinities κ_{bin} as indicated. **a** The two-box system with box height h, box length l, and $h << 2 \, l$. Proteins (orange spheres) can freely diffuse between both boxes and bind to traps (yellow cones). New proteins are synthesized only in the left box (blue). **b** Numbers $N_{i,b}(t)$ of bound proteins in box $i = 1$ (dashed lines) and $i = 2$ (solid lines). Because synthesis is restricted to box 1, saturation of traps in box 2 is lagging behind. This lag is not observable in well-mixed systems with high diffusion rates. **c** In two-box systems, the trap-binding efficiency $f_b(t)$ exhibits a pure binding-dominated ($0 \leq t < t_1^*$), a mixed ($t_1^* \leq t < t_2^*$), and a pure synthesis-dominated regime ($t \geq t_2^*$) as indicated by background shading. d) Individual trap binding efficiencies $f_{i,b}(t) \equiv N_{i,b}(t)/N_i(t)$ of box $i = 1$ (dashed lines) and $i = 2$ (solid lines). Transitions from the binding-dominated to the synthesis-dominated regimes in the boxes 1 and 2 occur at time points t_1^* and t_2^*, respectively. Arrows point to the maxima of $f_{1,b}(t)$ and $f_{2,b}(t)$ at $t = t_1^*$ and $t = t_2^*$, respectively

Förste *et al. Microbial Cell Factories*   (2023) 22:244

Page 6 of 13

$f_b(t) \equiv (N_{1,b}(t) + N_{2,b}(t))/N(t)$ of the two-box system. The time evolution of $f_b(t)$ can be now divided into three regimes, see Fig. 3c): At first, the system is in a *pure binding-dominated regime* (with boxes 1 and 2 being both in the binding-dominated regime). In this regime, the trap-binding efficiency $f_b(t)$ monotonically increases. At $t = t_1^*$ box 1 undergoes a transition from the binding-dominated to the synthesis-dominated regime, see Fig. 3d). At this point in time, the two-box system enters a *mixed regime*, because box 2 is still in the binding-dominated regime. The system remains in the mixed regime until $t = t_2^*$ when box 2 also undergoes a transition to the synthesis-dominated regime and the two-box system enters the *pure synthesis-dominated regime*. For fast diffusion, the mixed regime is short-lived because the lag time $t_2^* - t_1^*$ becomes small. Note that the time evolution of the trap-binding efficiency in the mixed regime strongly varies depending on the system parameters. In contrast, in the pure synthesis-dominated regime the trap-binding efficiency always monotonically decreases and asymptotically approaches $f_b(t) \propto \left(N_{1,b}^{max} + N_{2,b}^{max}\right)\left(\alpha_{syn}t\right)^{-1}$.

Binding and diffusion in quasi-realistic systems with finite trap densities and inhomogeneous synthesis

To study diffusion and binding for finite trap densities in a more realistic system, we investigate a system comprised of $M \gg 1$ boxes. The number $N_i(t)$ of proteins in box $i = 1, ..., M$ changes over time due to the synthesis of new proteins and their diffusion between neighboring boxes. The proteins freely diffuse with diffusion rate $\widetilde{D}$ until they bind to traps, which coat the bottom of the boxes, with binding affinity κ_{bin}. Binding can only occur as long as free traps are available, which implies that the number $N_{i,b}(t)$ of bound proteins in box i cannot exceed its maximum $N_{i,b}^{max}$. Like in the two-box system introduced above, the synthesis of new proteins can be restricted to a specific subset of boxes. We here focus on a system with $M = 100$ boxes and - corresponding to the experimental system - boxes $i = i_{syn} = 12, ..., 19$ serving as sources for new proteins of which each provides new proteins at a synthesis rate α_{syn}. As an example, the 100-box system could be a spatially discretized model of a quasi-twodimensional biochip of length $L = 1\,\text{mm}$, where each box represents a section with length $l = 10\,\mu\text{m}$ of the chip. Note that in this case the box index i indicates the discretized position relative to the biochip and $L = l \cdot M$.

Equivalently to the two-box system described above, the time-evolution of the numbers $N_{i,u}(t)$ and $N_{i,b}(t)$ of free and bound proteins in box i is given by

$$\frac{d}{dt} = \kappa_{bin}N_{i,u}(t)\left(N_{i,b}^{max} - N_{i,b}(t)\right) \quad \text{for i=1,...,M},$$

(10)

$$\frac{d}{dt} = \widetilde{D}\left(N_{i+1,u}(t) + N_{i-1,u}(t) - 2N_{i,u}(t)\right)$$
$$- \kappa_{bin}N_{i,u}(t)\left(N_{i,b}^{max} - N_{i,b}(t)\right) + \alpha_{syn}\left[i = i_{syn}\right]$$

(11)

for $i = 2, ..., M - 1$, and

$$\frac{d}{dt} = \widetilde{D}\left(N_{2,u}(t) - N_{1,u}(t)\right)$$
$$- \kappa_{bin}N_{1,u}(t)\left(N_{1,b}^{max} - N_{1,b}(t)\right)$$
$$+ \alpha_{syn}\left[1 = i_{syn}\right],$$

(12)

$$\frac{d}{dt} = \widetilde{D}\left(N_{M-1,u}(t) - N_{M,u}(t)\right)$$
$$- \kappa_{bin}N_{M,u}(t)\left(N_{M,b}^{max} - N_{M,b}(t)\right)$$
$$+ \alpha_{syn}\left[M = i_{syn}\right],$$

(13)

where the Iverson bracket $\left[i = i_{syn}\right] = 1$ if box i is a source for new proteins and 0 otherwise. In addition, mass conservation must hold, which implies

$$\sum_{i=1}^{M}\left(N_{i,u}(t) + N_{i,b}(t)\right) = \sum_{i=1}^{M}\left(N_{i,u}(0) + N_{i,b}(0) + \alpha_{syn}t\left[i = i_{syn}\right]\right).$$

(14)

We solved the system of equations (10) to (14) numerically by using Euler's method with a fixed discrete time step.

Trap-binding efficiency in quasi-realistic systems

At first, we study the time-dependent trap-binding efficiency $f_b(t)$ of the 100-box system, which is defined equivalently to Eq. (4) by the ratio of the total number $N_b(t) \equiv \sum_{i=1}^{100} N_{i,b}(t)$ of bound proteins and the total number $N(t) \equiv \sum_{i=1}^{100} N_i(t)$ of all proteins in the system. Like in the simple two-box system described above, the time evolution $f_b(t)$ of the trap-binding efficiency in the 100-box system can also be divided into three distinct regimes, see Fig. 4: (i) a pure binding-dominated regime, where all boxes are in the binding-dominated regime; (ii) a mixed regime with at least one box in the binding-dominated regime and at least one box in the synthesis-dominated regime; and (iii) a pure synthesis-dominated regime, where all boxes are in the synthesis-dominated regime. In all three regimes, the trap-binding efficiency is higher in systems with a faster diffusion, a lower synthesis of new proteins, or a higher trap binding affinity. A lower synthesis of new
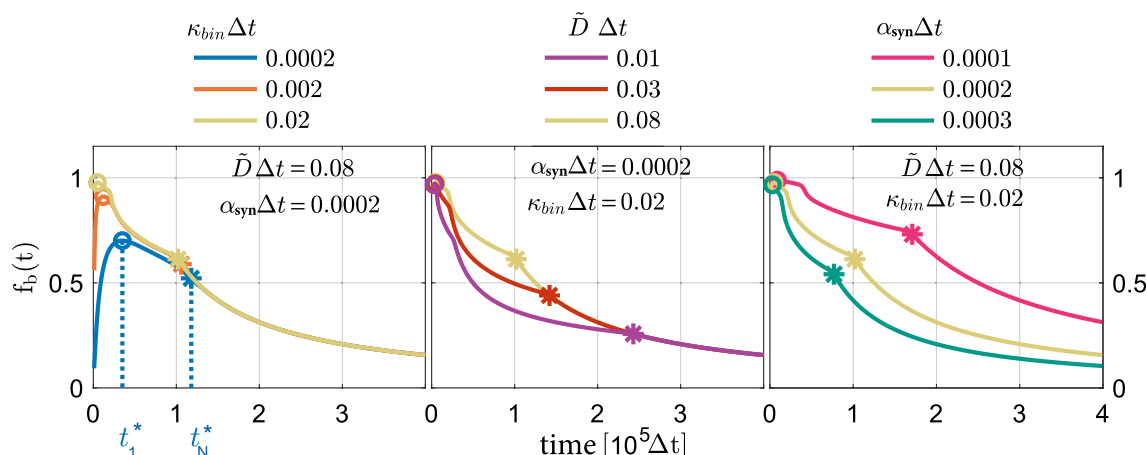
Fig. 4 Trap-binding efficiency $f_b(t)$ in 100-box systems with synthesis rate α_{syn}, binding affinity κ_{bin} and diffusion rate $\widetilde{D}$. Time-evolution of the trap-binding efficiency $f_b(t)$ for a 100-box system with synthesis of proteins restricted to boxes 12 to 19 and system parameters as indicated. Circles and stars mark the transition from the pure binding-dominated to the mixed regime at t_1^* and from the mixed to the pure synthesis-dominated regime at t_{100}^*, respectively

proteins, a faster diffusion, or a smaller binding affinity of proteins and traps attenuate the accumulation of trapped proteins in the proximity of the synthesis sites and, thus, cause a delay of the transition from the pure binding-dominated to the mixed regime. A transition from the mixed to the pure synthesis-dominated regime occurs when the box that is the most distant from the synthesis sites enters the synthesis-dominated regime. Therefore, this transition is delayed by slower diffusion, weaker binding, and lower synthesis.

Inhomogeneities in the trap density modulate the distribution of bound molecules

So far, we discussed systems with a homogeneous trap density, i.e, with a location-independent maximal number of bound proteins per box $N_{i,b}^{max} = N_b^{max}$ for all boxes $i = 1, ..., M$. In this paragraph, we investigate how inhomogeneities in the trap density modulate the dynamics and the localization of trap binding. In an experimental setup, these inhomogeneities may arise from an intended trap patterning, an incomplete coating of the reaction chamber or microfluidic chip with traps, or from impurities in the reaction solution that block the traps in an unspecific manner before proteins can bind. In Fig. 5 we plotted distributions of bound proteins at specific times for binding affinity $\kappa_{bin} = 2e-4.\Delta t^{-1}$, diffusion rate $\widetilde{D} = 8.3e-2.\Delta t^{-1}$, synthesis rate $\alpha_{syn} = 2e-4.\Delta t^{-1}$ and the trap pattern indicated by the blue solid line. The distribution of bound proteins may reflect the trap pattern already at early times when most of the traps are still unoccupied because the probability that a protein gets trapped is reduced in areas with a lower trap density.
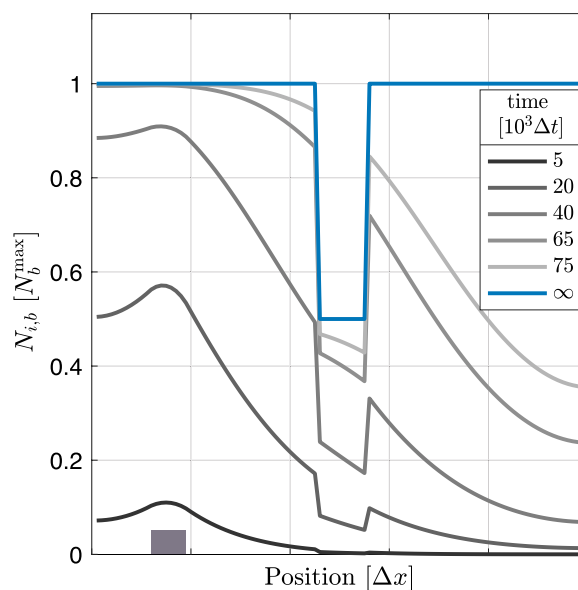

Fig. 5 Inhomogeneities in the trap pattern of a 100-box system can be reflected by the distributions of bound proteins long before the traps get saturated. Local numbers of bound proteins $N_{i,b}(t)$ for binding affinity $\kappa_{bin} = 2e-4.\Delta t^{-1}$, diffusion rate $\widetilde{D} = 8.3e-2.\Delta t^{-1}$, synthesis rate $\alpha_{syn} = 2e-4.\Delta t^{-1}$, and specific points in time as indicated in the figure legend. Numbers are normalized to the overall largest value of the local maximal numbers $N_{i,b}^{max}$ of bound proteins, i.e., $N_b^{max} = max(N_{i,b}^{max} : i = 1, ..., 100)$. The local maximal numbers $N_{i,b}^{max}$ are equivalent to the local numbers of traps (solid blue line, "∞"), i.e., the inhomogeneous trap pattern. The gray box indicates the area of protein synthesis

How clearly the trap pattern is imprinted in the distribution of bound proteins at a given point in time can differ for binding reaction-limited systems (with relatively

high diffusion constants and low binding affinities) and diffusion-limited systems (with relatively low diffusion constants and high binding affinities). For the trap pattern in Fig. 5, the distributions of bound proteins reflect the inhomogeneous trap density already at early time points if the system is binding reaction-limited and not diffusion-limited.

Systems with spatially restricted binding domains

For certain applications or for technical reasons, binding of proteins to traps might be restricted to a specific region within the chip surface. We model such systems by setting the maximal number of bound proteins $N_{i,b}^{max} = 0$ for all boxes i except for $i = i_{trap} = 15 + k_{dist}$, ..., $15 + k_{dist} + k_{width}$, where k_{dist} indicates the distance

between the synthesis sites and the trap region, and k_{width} refers to the number of consecutive boxes that define the trap region. We set $k_{width} = 8$ and $N_{i_{trap},b}^{max} = N_b^{max} > 0$. Figure 6 shows total numbers and exemplary distributions of bound proteins at several points in time and for different trap positions. Proteins that diffuse faster and are less easily captured by traps have a greater chance to spread throughout the whole reaction chamber such that the number of proteins in solution gets more homogeneous. Thus, in these systems the total number of bound proteins as a function of time depends much less on the distance k_{dist} between synthesis sites and traps compared to systems with slowly diffusing proteins with a high trap-binding affinity. For a system with known synthesis rate
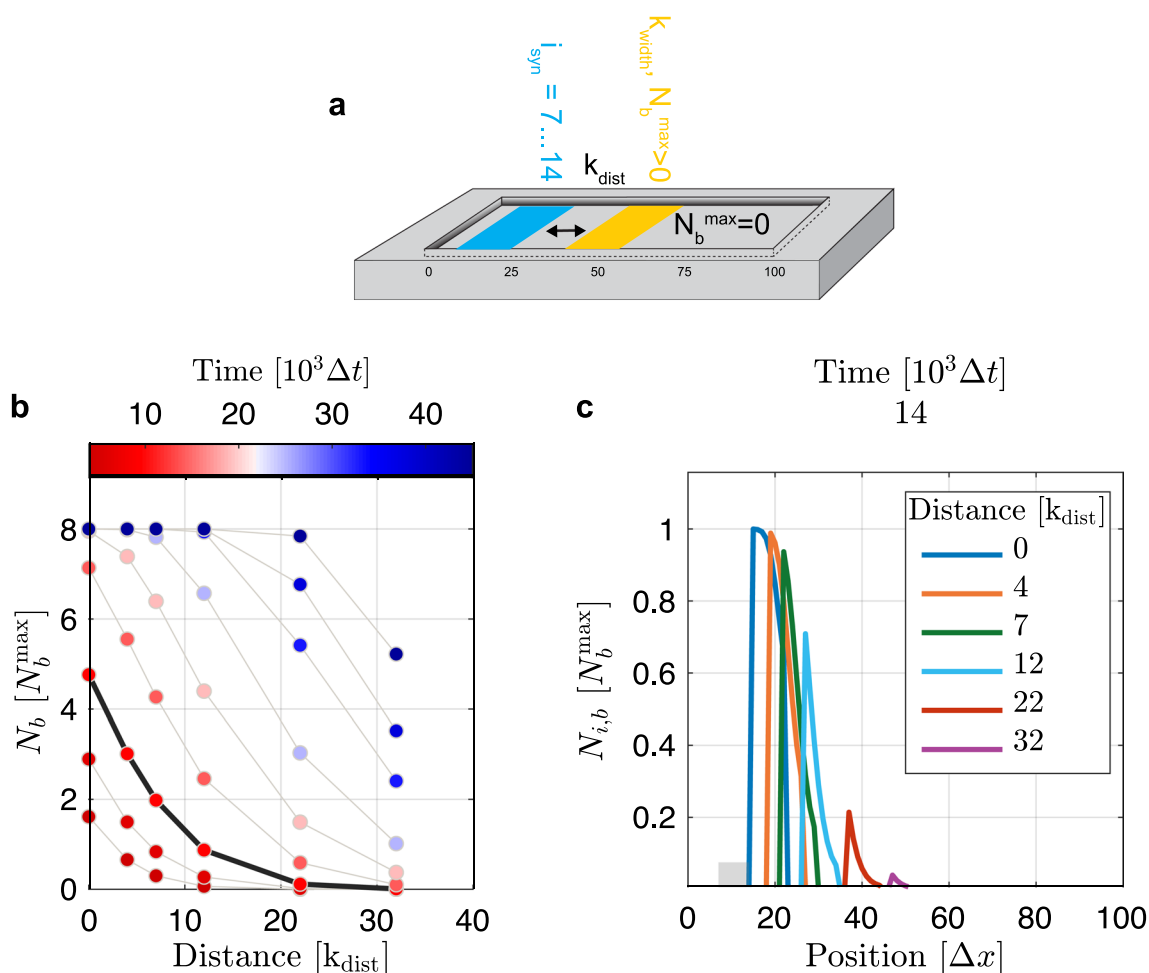
Fig. 6 Binding in spatially restricted domains. **a** The system consists of 100 boxes where binding is restricted to 8 consecutive boxes i_{trap} such that $N_{i,b}^{max} = N_b^{max}$ for $i = i_{trap}$ and $N_{i,b}^{max} = 0$ elsewhere. Synthesis sites are at boxes $i_{syn} = 7, ..., 14$ (blue) and boxes with traps (yellow) are located in a specific distance k_{dist} to the synthesis site as indicated. Further system parameters are binding affinity $\kappa_{bin} = 2e - 2 . \Delta t^{-1}$, diffusion rate $\widetilde{D} = 9.8e - 3 . \Delta t^{-1}$, and synthesis rate $\alpha_{syn} = 2e - 4 . \Delta t^{-1}$. **b** Total number of bound proteins in the biochip at different time points and distances between synthesis sites and traps. The bold line corresponds to the time point in (**c**), which shows the distributions of bound proteins for different trap distances k_{dist}

and diffusion rate, these differences in the trap-position dependence could be exploited to determine an unknown binding affinity κ_{bin}: Using the binding affinity as the fit parameter, the computed total number of bound proteins could be fitted to the measured amount of bound proteins, either at a specific time point for different trap positions or at several time points for a specific trap position.

Experimental validation: protein synthesis, diffusion, and trap binding in a quasi-twodimensional silicon-based compartmental system

In this section, we compare experimental data to predictions from our computational 100-box system to validate our modeling approach. The experimental setup has been described previously [26]. Briefly, we used compartments of dimensions $200\,\mu\mathrm{m}$ width, $1000\,\mu\mathrm{m}$ length, and $2\,\mu\mathrm{m}$ height. First, we uniformly immobilized DNA along the width of the compartment, close to one edge. Here, we used DNA encoding the bacteriophage protein gp10 as a model protein. Antibody-based traps (tagged protein gp11-HA) were then immobilized all along the compartments, excluding the area of the DNA. After washing all non-immobilized bio-molecules, we filled the compartments with a cell-free reaction solution and sealed them. Incubating the compartments for 2–3 h enabled the synthesis, diffusion, and trapping of gp10 from the immobilized DNA to the traps to occur.

First, we studied the distribution of trapped gp10 for different fractions of coding sequences in the DNA brush. The protein synthesis rate can be controlled by the amount of coding DNA provided that all components required for protein synthesis are added in excess. Figure 7 shows that for a higher synthesis rate indeed a larger part of the chip is covered with gp10 120 min after protein synthesis has started and that our computational 100-box model reproduces the experimental data.
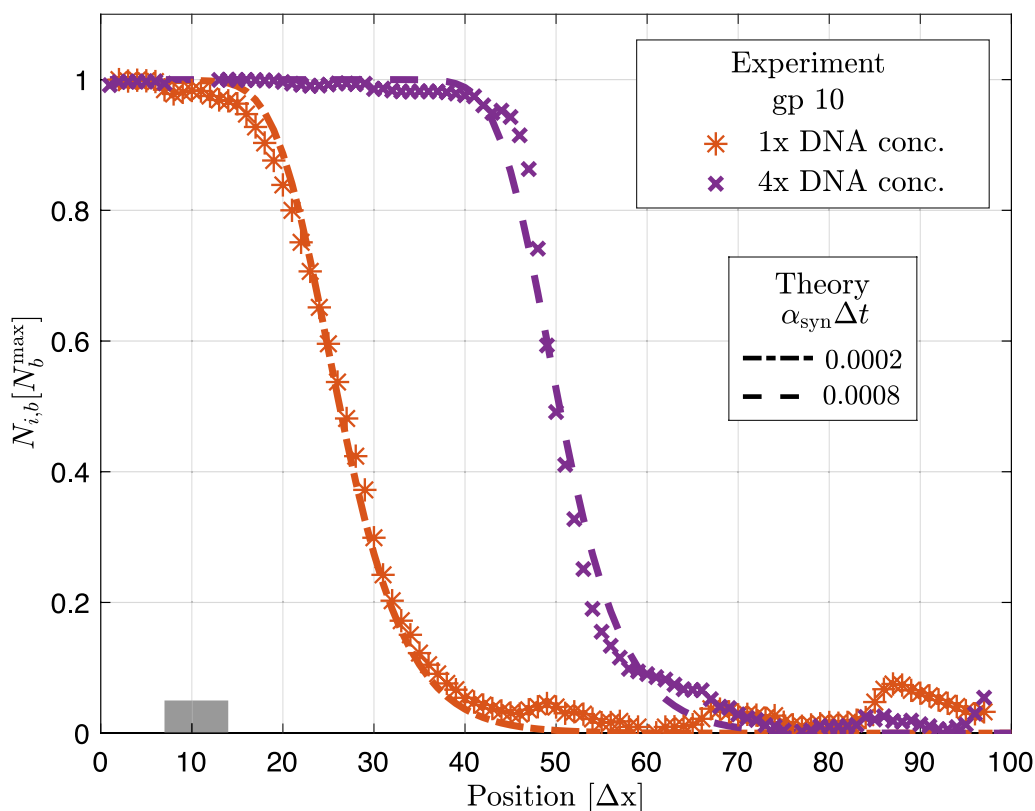

Fig. 7 Experimental validation of the 100-box model. Fluorescence signals of trap-bound gp10 synthesized in a quasi-twodimensional, silicon-based expression system recently introduced [26] (symbols) and corresponding simulation results at $t = 21 \cdot 10^3 \Delta t$ (dashed lines). In the experiment, the gp10 synthesis rate is roughly proportional to the amount of gp10-coding DNA sequences. Simulations were performed in the 100-box system as described in the main text with synthesis sites being located at boxes $i_{\mathrm{syn}} = 7, ..., 14$, binding affinity $\kappa_{\mathrm{bin}} = 2e - 3 \, .\Delta t^{-1}$, diffusion rate $\widetilde{D} = 0.045 \, .\Delta t^{-1}$ and synthesis rate as indicated. Increasing the amount of gp10-coding DNA sequences on the biochip by a factor of 4 corresponds to an increase in the simulation parameter α_{syn} by roughly the same fold change
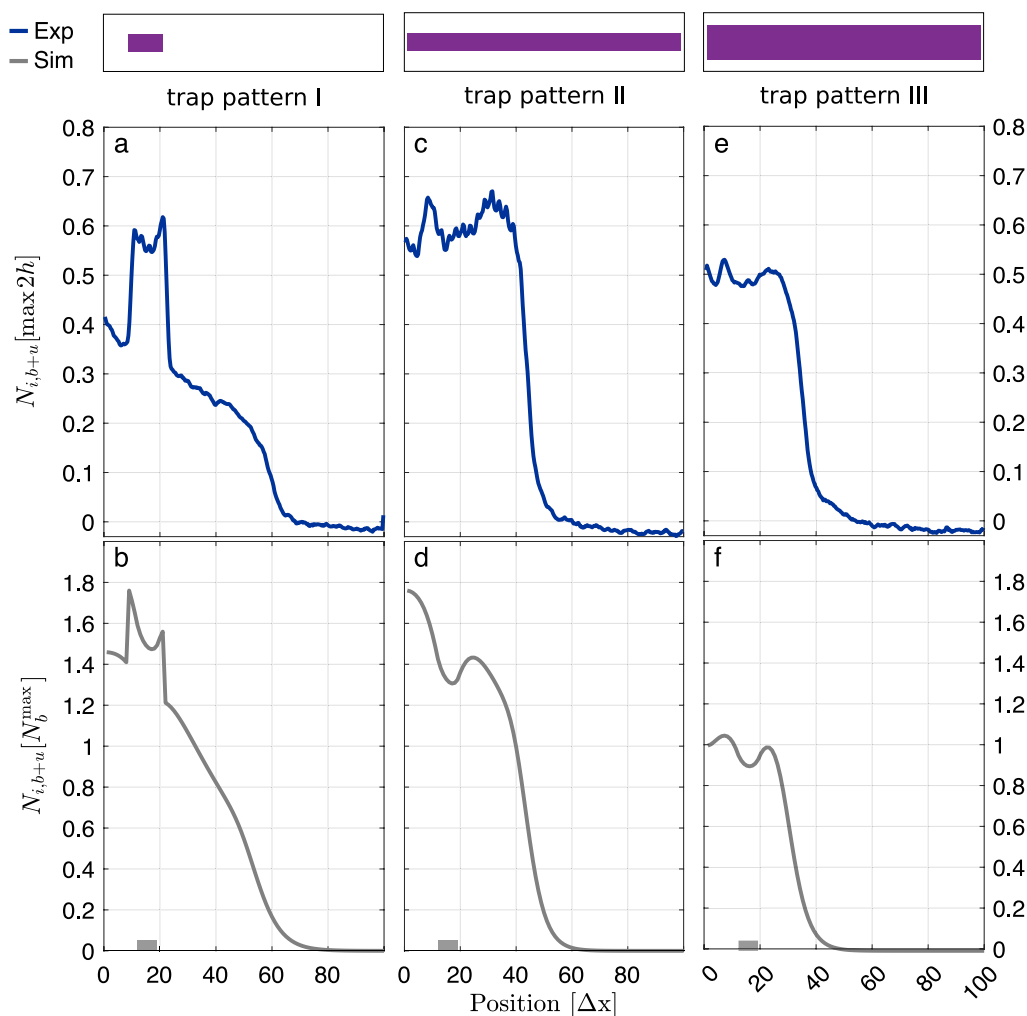
Fig. 8 Qualitative comparison of experimental results and simulations for different trap patterns. **a**, **c**, **e** Fluorescence signals of bound and unbound ($b + u$) GFP synthesized in a quasi-twodimensional, silicon-based expression system recently introduced [26] for different trap patterns (purple bars) 1 h after start of synthesis. Signal normalized to the respective maximum value measured after 2 h. **b**, **d**, **f** Simulations of GFP synthesis, diffusion and binding in a 100-box system with synthesis sites being located at boxes $i_{syn} = 12, ..., 19$ (gray boxes), trap patterns as in **a**, **c** and **e**, respectively, and at simulation time $t = 39 \cdot 10^3 \Delta t$. Number of bound and unbound GFP molecules is normalized to the respective maximal number of traps per box

Second, we checked whether the 100-box model correctly predicts the influence of trap patterns on the distribution of proteins within the chip and their binding to the traps. Instead of gp10, which requires post-staining for visualization [26], we used GFP as a model protein and reporter. GFP is a fluorescent protein and can therefore be monitored both as a freely diffusing molecule and when bound to a trap. We performed experiments with three different trap distributions, see Fig. 8: for the first trap distribution, traps were only present in proximity to the DNA brushes; for the second, traps were distributed homogeneously along the chip; for the third, the area covered by traps and, thus, the total number of traps was approximately doubled compared to the second. After start of synthesis, GFP fills both the traps and the reaction solution of the chip. The trap pattern has an impact on the distribution of GFP because the traps effectively slow down the spread of GFP within the chip, see Fig. 8. The strong correlation between simulation and experimental results suggests that the computational model provides an adequate

description of protein synthesis, diffusion, and binding in compartmentalized systems introduced in [26].

Conclusion

Recently, we introduced quasi-twodimensional, silicon-based systems for the in-vitro synthesis and assembly of bionanomachines consisting of multiple protein subunits, where we introduced a model to examine different modes of protein assembly [26]. Here, we discussed a simplified version of these systems, in which a single protein subunit is synthesized either everywhere or in a spatially defined region of the chip and is later bound to the chip surface by molecular traps. The dynamics of these simplified systems are governed by the synthesis rate α_{syn}, the diffusion constant D of the protein, and the trap-binding affinity κ_{bin}. We studied the influence of these parameters on the system's trap-binding efficiency $f_b(t)$, which is defined as the time-dependent fraction of proteins that are bound to traps, see Eq. (4). In particular, we took into account that the trap density is limited. We found that systems with spatially homogeneous synthesis of proteins undergo a transition from a binding-dominated to a synthesis-dominated regime after protein synthesis has started: The first regime is characterized by a high fraction of proteins that get bound to traps, whereas in the second regime a high fraction of traps is occupied and the systems fills up with unbound proteins. Thus, the trap-binding efficiency increases in the binding-dominated regime and decreases in the synthesis-dominated regime. A spatial confinement of protein synthesis introduces more complexity to the binding dynamics: Systems with a spatially inhomogeneous synthesis that are not well-mixed also exhibit a pure binding-dominated and a pure synthesis-dominated regime, but these are separated by a mixed regime. In the mixed regime, local binding-dominated coexist with local synthesis-dominated regimes. Spatial inhomogeneities of protein synthesis are not the only way to modulate the dynamics of protein spreading and trap binding in the biochip. Further complexities can be introduced by using non-uniform distributions of traps on the surface of the synthesis systems. Depending on the system parameters, the spatial distribution of bound proteins is shaped by the trap pattern long before all traps get saturated. Therefore, designing non-uniform trap distributions provides a possibility to control both the temporal and the spatial dependence of protein-trap binding. Furthermore, we studied systems with spatially restricted binding where traps are only present in a distinct, narrow region on the surface. We showed

that the system parameters determine the impact of the distance between synthesis and trap sites on the protein-trap binding dynamics. This parameter-dependence could potentially open up a new path to infer for example unknown binding affinities from systems with known synthesis rates and diffusion constants: To this end, the model presented in this work might be fitted to the fluorescence time courses of trapped proteins measured in a series of experiments with increasing distances between synthesis and trap sites. Finally, we performed experiments to validate our modeling approach. We used single-gene versions of our silicon-based compartmental system introduced in [26] to synthesize the proteins gp10 and GFP for different synthesis rates (gp10) and trap patterns (GFP). Simulations based on the here defined diffusion system with limited trap density reproduce the experimental findings, which indicates that the presented model is an adequate description of protein synthesis, diffusion and binding in quasi-twodimensional biochips. Overall, our work broadens the theoretical understanding of synthesis, diffusion, and binding processes in compartmental systems. This advancement enables improved control over directed molecular self-assembly, facilitating the fabrication of nanomachines for applications in synthetic biology or nanotechnology (Additional file 1, 2).

Methods

Time-dependent number of bound proteins in the one-box system

We solved the differential Eqs. (1) and (2) using the software Wolfram Mathematica. The number $N_b(t)$ of bound proteins relative to its maximal value N_b^{max} is given by

$$\frac{N_b(t)}{N_b^{\text{max}}} = 1 - \eta \frac{e^{-\lambda^2 C_\zeta(t)^2}}{e^{-\lambda^2 C_\zeta(0)^2} + N_b^{\text{max}} \eta \lambda \sqrt{\pi} \left(\text{Erf}\left[\lambda C_\zeta(t) \right] - \text{Erf}\left[\lambda C_\zeta(0) \right] \right)} \tag{15}$$

with $\text{Erf}(z)$ indicating the Gauss error function,

$$\lambda \equiv \sqrt{\frac{\kappa_{\text{bin}}}{2\alpha_{\text{syn}}}}, \quad C_\zeta(t) \equiv -N_b^{\text{max}}$$
$$+ N_b(0) + N_u(0) + \alpha_{\text{syn}} t, \quad \eta \equiv \left(1 - \frac{N_b(0)}{N_b^{\text{max}}} \right),$$

and all further symbols defined in the main text.

Experimental methods

The experimental setup was previously described in [26]. For visualization of on-chip dynamic expression, Fig. 8, the chip was placed in an incubator

Förste *et al. Microbial Cell Factories*   (2023) 22:244

Page 12 of 13

chamber (Bold-line stage top incubator, Okolab) installed on the microscope and set to 30 °C. Fluorescent images were obtained using an AxioObserver Z1 inverted microscope with a motorized stage (Zeiss), 10x/0.3 MPlanFL N (Olympus) Objective. Illumination was performed using a Colibri2 LED illumination system equipped with a 470 nm LED module (Zeiss) and filter sets 38 HE (Zeiss, excitation 470/40 nm, dichroic mirror 495 nm, emission 525/50 nm). Images were captured using an iXon Ultra CCD camera (Andor Technology, Belfast, UK). Chip alignment and multi-image time series acquisition was performed using the Zeiss ZEN 2012 software. Images were analyzed as previously described in [26].

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12934-023-02237-5.

Additional file 1: Additional parameters and trap distributions.

Additional file 2: Movie S1. Dynamics of GFP expression, diffusion and binding for different trap patterns. Montage of three wells, each with different trap pattern density, as explained in Fig. 8. Expression originates from immobilized DNA at ~150 μm from the left side of the well. Scale bar 100 μm. Images were acquired every 3 minutes.

Author contributions
SF: contributed to software development, conducted formal analysis, performed investigations, participated in writing, reviewing, and editing the manuscript, and created visualizations. OV: contributed to the methodology development, conducted validation, participated in writing, reviewing, and editing the manuscript, and created visualizations. SSD: contributed to the methodology development, participated in reviewing, and editing the manuscript. RHBZ: contributed to the methodology development, provided resources, and acquired funding. RL: contributed to the conceptualization of the study, provided resources, participated in writing, reviewing, and editing the manuscript, and acquired funding. SR: contributed to the conceptualization, conducted formal analysis, wrote the original draft, participated in reviewing, and editing the manuscript, created visualizations, and provided supervision. All authors read and approved the final manuscript.

Availability of data and materials
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

References
1. Gazit E. Reductionist approach in peptide-based nanotechnology. Annu Rev Biochem. 2018;20(87):533–53. https://doi.org/10.1146/annurev-biochem-062917-012541.
2. Lampel A, Ulijn RV, Tuttle T. Guiding principles for peptide nanotechnology through directed discovery. Chem Soc Rev. 2018;47(10):3737–58. https://doi.org/10.1039/c8cs00177d.
3. Kobayashi N, Arai R. Design and construction of self-assembling supramolecular protein complexes using artificial and fusion proteins as nanoscale building blocks. Curr Opin Biotechnol. 2017;46:57–65. https://doi.org/10.1016/j.copbio.2017.01.001.
4. Blanchard AT, Salaita K. Emerging uses of DNA mechanical devices. Science. 2019;365(6458):1080–1. https://doi.org/10.1126/science.aax3343.
5. Madsen M, Gothelf KV. Chemistries for DNA nanotechnology. Chem Rev. 2019;119(10):6384–458. https://doi.org/10.1021/acs.chemrev.8b00570.
6. Kopperger E, List J, Madhira S, Rothfischer F, Lamb DC, Simmel FC. A self-assembled nanoscale robotic arm controlled by electric fields. Science. 2018;359(6373):296–301. https://doi.org/10.1126/science.aao4284.
7. Keya JJ, Suzuki R, Kabir AM, Inoue D, Asanuma H, Sada K, Hess H, Kuzuya A, Kakugo A. DNA-assisted swarm control in a biomolecular motor system. Nat Commun. 2018;9(1):453. https://doi.org/10.1038/s41467-017-02778-5.
8. Seeman NC, Sleiman HF. DNA nanotechnology. Nat Rev Mater. 2017;3(1):1–23. https://doi.org/10.1038/natrevmats.2017.68.
9. Zhang G, Surwade SP, Zhou F, Liu H. DNA nanostructure meets nanofabrication. Chem Soc Rev. 2013;42(7):2488–96. https://doi.org/10.1039/c2cs35302d.
10. Besford QA, Cavalieri F, Caruso F. Glycogen as a building block for advanced biological materials. Adv Mater. 2020;32:1904625. https://doi.org/10.1002/adma.201904625.
11. Miao T, Wang J, Zeng Y, Liu G, Chen X. Polysaccharide-based controlled release systems for therapeutics delivery and tissue engineering: from bench to bedside. Adv sci. 2018;5(4):1700513. https://doi.org/10.1002/advs.201700513.
12. Feng X, Du C, Li J. Molecular assembly of polysaccharide-based microcapsules and their biomedical applications. Chem Rec. 2016;16(4):1991–2004. https://doi.org/10.1002/tcr.201600051.
13. Ellis E, Moorthy S, Chio WI, Lee TC. Artificial molecular and nanostructures for advanced nanomachinery. Chem Commun. 2018;54(33):4075–90. https://doi.org/10.1039/c7cc09133h.
14. Kim K, Guo J, Liang Z, Fan D. Artificial micro/nanomachines for bioapplications: biochemical delivery and diagnostic sensing. Adv Funct Mater. 2018;28(25):1705867. https://doi.org/10.1002/adfm.201705867.
15. Rudorf S, Rädler JO. Self-assembly of stable monomolecular nucleic acid lipid particles with a size of 30 nm. J Am Chem Soc. 2012;134(28):11652–8. https://doi.org/10.1021/ja302930b.
16. Qian L, Winfree E. Scaling up digital circuit computation with DNA strand displacement cascades. Science. 2011;332(6034):1196–201. https://doi.org/10.1126/science.1200520.
17. Lin L, Yang C, Ma M, Ma S, Yan H. A clock synchronization method for molecular nanomachines in bionanosensor networks. IEEE Sensors J. 2016;16(19):7194–203. https://doi.org/10.1109/JSEN.2016.2591823.
18. Baroncini Massimo, Casimiro Lorenzo, de Vet Christiaan, Groppi Jessica, Silvi Serena, Credi Alberto. Making and operating molecular machines: a multidisciplinary challenge. ChemistryOpen. 2018;7(2):169–79. https://doi.org/10.1002/open.201700181.
19. Daube SS, Bar-Ziv RH. Protein nanomachines assembly modes: cell-free expression and biochip perspectives. Wiley Interdiscip Rev Nanomed Nanobiotechnol. 2013;5(6):613–28. https://doi.org/10.1002/wnan.1234.
20. Jacobs WM, Frenkel D. Self-assembly of structures with addressable complexity. J Am Chem Soc. 2016;138(8):2457–67. https://doi.org/10.1021/jacs.5b11918.

Förste *et al. Microbial Cell Factories*   (2023) 22:244

Page 13 of 13

21. Whitelam S, Jack RL. The statistical mechanics of dynamic pathways to self-assembly. Annu Rev Phys Chem. 2015;66(1):143–63. https://doi.org/10.1146/annurev-physchem-040214-121215.

22. Gartner FM, Graf IR, Wilke P, Geiger PM, Frey E. Stochastic yield catastrophes and robustness in self-assembly. Elife. 2019. https://doi.org/10.7554/eLife.51020.

23. Grzybowski BA, Fitzner K, Paczesny J, Granick S. From dynamic self-assembly to networked chemical systems. Chem Soc Rev. 2017;46(18):5647–78. https://doi.org/10.1039/c7cs00089h.

24. Murugan A, Zou J, Brenner MP. Undesired usage and the robust self-assembly of heterogeneous structures. Nat Commun. 2015;6(1):6203. https://doi.org/10.1038/ncomms7203.

25. Hongzhou Gu, Chao Jie, Xiao Shou-Jun, Seeman Nadrian C. A proximity-based programmable DNA nanoscale assembly line. Nature. 2010;465(7295):202–5. https://doi.org/10.1038/nature09026.

26. Vonshak O, Divon Y, Förste S, Garenne David, Noireaux V, Lipowsky R, Rudorf S, Daube SS, Bar-Ziv RH. Programming multi-protein assembly by gene-brush patterns and two-dimensional compartment geometry. Nat Nanotechnol. 2020;15(9):783–91. https://doi.org/10.1038/s41565-020-0720-7.

27. Vishwanath S, Sukhwal A, Sowdhamini R, Srinivasan N. Specificity and stability of transient protein-protein interactions. Curr Opin Struct Biol. 2017;44:77–86. https://doi.org/10.1016/j.sbi.2016.12.010.

Publisher's Note