Politecnico di Torino

ScuDo
Scuola di Dottorato ~ Doctoral School
WHAT YOU ARE, TAKES YOU FAR

Doctoral Dissertation
Doctoral Program in Doctoral Program in Physics (*XXXVI*cycle)

Competition and band-pass filtering in biology

modelling and experimental validation in membrane trafficking and post-transcriptional gene regulation

By

Elsi Ferro

******

Supervisor(s):
Dr. Carla Bosia, Supervisor

Doctoral Examination Committee:
Prof. Alejandra Ventura , Referee, Universidad de Buenos Aires
Prof. Matteo Osella, Referee, Università di Torino

Politecnico di Torino
2024

Declaration

I hereby declare that the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Elsi Ferro
2024

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

Abstract

Cells have often been found out to exploit mathematical features of biochemical networks to encode biological information. In this thesis we investigate two mechanisms that can shape the behavior of cells by relying on mathematical properties: competition and frequency preference.
Competition constitutes a regulatory mechanism thought to underlie the behavior of many biological systems, from the molecular to the ecosystem level. Here we propose competition as a mechanism able to generate mutual segregation between molecular species, a well-known feature of membrane proteins involved in the trafficking of material across the cell, whose origin remains poorly understood. We thus formulate several different models of proteins competing for a shared resource using Ordinary Differential Equations (ODEs) and stochastic simulations. By formulating predictions on protein mean values and mutual fluctuations, we find that their amounts can be anticorrelated, thereby suggesting that competition can lead to segregation between protein species. For experimental validation, wet-lab experiments are performed by our collaborators to track the dynamics of two proteins - Rab5 and Rab11 - on the membrane of cellular organelles by fluorescence microscopy imaging. By analyzing these image data, we validate the qualitative mean value and correlation behaviors predicted by one of our models, thereby confirming that competition for a shared resource might underlie mutual protein segregation.
On the other side, we investigate frequency preference in the epigenetic context of post-transcriptional regulation - where short RNAs called microRNAs (miRNAs) repress gene activity by inhibiting RNAs.
Frequency preference - or band pass filtering - is the ability of a system to respond to certain frequencies of an input signal while filtering out others. While some mechanisms of genetic regulation have been shown to exhibit band-pass filtering properties, little effort has been devoted to unveil the role of periodicity in the interaction between miRNAs and RNAs. Here we theoretically and experimentally address

the frequency preference properties of this interaction by using ODE modelling and wet-lab experiments. On the theoretical side, we mathematically model periodic pulses of miRNA synthesis and we analyze the resulting RNA repression as a function of frequency. We find that this repression can be maximal for a narrow range of frequencies, thereby exhibiting a band-pass filtering behaviour. Furthermore, we show how this property is tied to the relative kinetics of target and repressor, thus providing suggestions for experimental validation. In parallel, we develop and test an experimental setup aimed at investigating band-pass filtering in live cells using a synthetic miRNA-RNA network.

Contents

List of Figures

List of Tables

Chapter 1

Quantitative biology: at the intersection between physics and biology

The first contact between physics and biology occurred when DNA, the fundamental molecule carrying genetic information in all living beings, was discovered by Watson and Crick in 1953 [6]. X-ray diffraction images of this molecule, composed of two twisted strands that form the shape of a helix, were captured by Rosalind Franklin and Maurice Wilkins, thus helping to solve the mystery of its geometrical structure [7, 8]. Similar to X-ray diffraction, many tools based on physical principles have found application in biology and medicine.

The vast world of microscopy constitutes a striking example of how living matter can be investigated using physics. Microscopy techniques based on optical principles, such as confocal microscopy, are nowadays widely used for investigating cell functions, whereas approaches relying on particle beams such as scanning electron microscopy (SEM) are at the base of studies unveiling cellular and sub-cellular structures [9].

On the medical side, physicists contributed to develop imaging techniques such as computer assial tomography (CAT) and nuclear magnetic resonance (NMR), that revolutionized the field of diagnostics [10].

However, the use of physics for the development of tools in experimental biology did not yet imply a common ground of investigation for the two sciences. For most of the last century, biology was viewed as too complex and disordered to be mathematically

modeled, and because of its unpredictability it was not considered of interest by physicists. Indeed, the behaviour of living matter is typically affected by a high number of variables [11]. Biological systems, such as a living cell or the human brain, are constituted by a great number of intricately interconnected elements, and thus pose the challenging issue of simplifying a complex system into its fundamental parts [12].

A complex system is characterized by the emergence of a collective behaviour from the individual behaviours of its numerous components. Such global behaviour cannot be explained by the sole composition of individual component dynamics, thus constituting a property of the system as a whole. Consider for instance a population of migrating cells: it is an out-of-equilibrium system composed by single cells, each one with its own dynamics, interacting with one another and with the surrounding environment. However, if we observe the system as a whole, we notice a global direction of motion, with the population maintaining a coherent global shape - a property that does not result from the mere dynamics of its single components [13]. Similar examples can be found also at different scales: the cohesive motion of bird flocks or the schooling of fish are emergent behaviours of collective and coordinated movement [14]. If we look at smaller length scales, we find molecular motors within single cells, that often need to work together in a coordinated manner: cell division requires the cooperation of multiple motor proteins to segregate chromosomes - a process that involves intricate signaling mechanisms [15].

The problem of connecting macroscopic (e.g. bird flock motion) and microscopic (e.g. single bird motion) behaviours in a complex system is the typical aim of statistical physics [16]. This framework considers the laws governing the dynamics of single system components and derives macroscopical features through a probabilistic approach, thus avoiding the costly task of solving the equations of motion for each single component. For instance, while classical thermodynamics considers a gas as characterized by macroscopic variables such as volume, pressure and temperature, statistical thermodynamics views it as a system composed by a great number of particles - each one of them following Newton's laws and quantum mechanics - and derives classical macroscopic laws by applying probability theory to the equations of motion.

Statistical physics has played a fundamental role in the approachment of physics to biology, as it has allowed to build mathematical models that connect underlying microscopic mechanisms to the macroscopic behaviours of molecules, cells, and

organisms [17]. In this framework, disorder and heterogeneity become themselves emergent behaviours of interest, rather than posing a limitation to scientific investigation.

Ultimately, the last decades saw an increase in the amount of biological data, which encouraged physicist to model systems with unprecedented complexity: while tracking a few molecular species was considered state-of-the-art until not long ago, recent breakthroughs in sequencing, proteomics and microscopy techniques revolutionized the accessibility to experimental data. This data availability called for more quantitative approaches of analysis and opened the possibility to uncover new mechanisms governing the behaviour of complex living organisms. As a consequence, physicists started modelling biological systems composed by progressively larger numbers of components and covering diverse length scales: topics range from information processing in genetic networks [18] and metabolism organization [19], to ecosystem dynamics [20].

Such quantitative models do not merely fit experimental data, but aim to make predictions about biological mechanisms. This approach, often called "quantitative biology", is an established interdisciplinary field of research where physics and biology continuously connect, both in the techniques employed and in the scientific questions addressed. As the physicist William Bialek stated, one of the challenges of modern science lies in the "development of a theoretical physics of biological systems that reaches the level of predictive power that has become standard in other areas of physics" [21].

This thesis falls under the framework of quantitative biology and pursues the following frame: (i) we specify the principles that delineate a specific question at a systemic level; (ii) we represent these principles using a mathematical formalism; (iii) we formulate quantitative predictions of our mathematical theory relying on experimentally accessible observables; (iv) we identify the experimental method that enables to measure such observables; (v) we verify our predictions through experiments.

To ensure tightly controlled experiments when dealing with a complex biological system, we borrow strategies from experimental physics: we fix all variables of the system and we choose a single "free variable" to be tuned to perturb the system. Next, we analyze the system's response to this variation by making measurements that correspond to our theoretical observables. Thus, we test the mathematical model by comparing theoretical predictions to empirical results in order to change the

potentially disproved hypothesis.
This thesis investigates two different biological issues with a physical approach. Due to the significant difference between the two topics, the thesis is composed by two separate parts. On one hand, we study how competition for limited resources shapes the interaction of two molecules that play a role in a fundamental cellular function - membrane trafficking. Competition is a widespread concept in physics, involving to two or more interacting physical entities. It can refer for instance to competing forces, such as gravitational attraction [22], or to thermodynamic effects arising when multiple processes or interactions within a system compete for limited energy or resources [23]. Similarly, biology often involves systems with multiple competing entities, which may be molecules [24], organisms [25] and species [26]. Likewise, competition among such entities plays a role in shaping the behavior and outcomes of these complex systems, leading to emergent properties and patterns.
The second project investigates how the periodic synthesis of a regulatory molecule - the microRNA - affects a gene's activity. In particular, we study effects driven by synthesis frequency, addressing whether gene activity regulation might be subject to mechanisms of band-pass filtering. Frequency-dependent phenomena hold a significant role in many areas of physics, ranging from light diffraction [27] and mechanical resonance [28] to electric band-pass filtering [29]. Similarly, the frequency of periodic inputs affects the behaviour of some biological systems, such as genes regulating circadian rythms [30] or photoreceptors of the human eye [31].
Our theoretical framework models biochemical systems composed by interacting molecules using Ordinary Differential Equations (ODEs) and stochastic simulation algorithms, whereas our experimental approach relies on typical wet-lab measurements of cellular and molecular biology. Chapter 2 is devoted to summarize the fundamentals of our theoretical strategies.
Subsequently, chapter 3 will give a broad introduction exploring the concept of competition - a pivotal feature of both systems investigated in this thesis - bringing forward examples from both physical and biological systems.
Parts I and II are devoted to illustrate the two separate research projects.
Part I, composed by chapters 4, 5, 6 and 7, illustrates the first project concerning competition between Rab proteins involved in membrane trafficking. In particular, chapter 4 provides the necessary backgroud knowledge on membrane trafficking processes and the role hereby played by the Rab protein family; chapter 5 illustrates theoretical modelling and formulates predictions; chapter 6 explains experimental

activity and results; eventually, chapter 7 adds further theoretical insights together with their experimental validation.
Part II, composed by chapters 8, 9, 10 and 11, concerns the second project, which investigates frequency preference in microRNA-mediated regulation. In particular, chapter 8 introduces the topics of frequency preference and band-pass filtering, whereas chapter 9 provides backround knowledge concerning microRNAs and their functions. Eventually, chapters 10 and 11 explore respectively the theoretical and experimental investigations, along with corresponding results.

Chapter 2

Theoretical modelling of biochemical systems: Ordinary Differential Equations and stochastic simulation algorithms

Systems composed by multiple interacting molecules can be viewed as networks, where molecules are represented by nodes and interactions amongst them are indicated as arrows [32]:

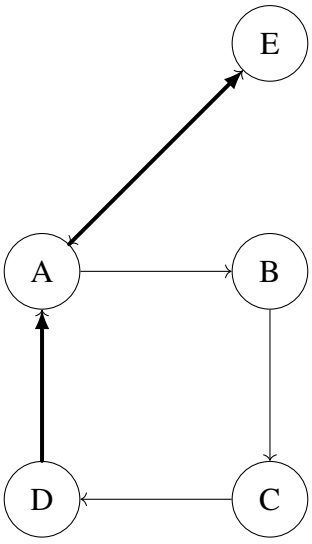


Fig. 2.1 Example of a simple network of interacting molecules.

In this framework, each arrow represents a positive (pointed arrow) or a negative (flat-headed arrow) interaction [32], but does not specify the actual biochemical reactions implicitly involved. For instance, let A and B be two different protein species in the network depicted in Figure 2.1. Then, the positive arrow between nodes A and B might indicate many kinds of biochemical interactions: for instance, A might cause the production of B by recruiting the molecular machinery necessary for its translation, or as well it could activate the production of B by binding to its gene's promoter and initiating its transcription [32]. Thus, an explicit description of biochemical reactions involved in each interaction is required to predict a network's behaviour.

However, a biochemical reaction can be described with lower or higher resolution. As an example [33], the production of an mRNA species can be described with high resolution by considering multiple steps, such as: (i) transcription factor binding to the specific DNA promoter site; (ii) recruitment of RNA polymerase and other transcriptional machinery; (iii) binding of RNA polymerase to the promoter region to initiate transcription; (iv) synthetization of an RNA transcript from the DNA template strand; (v) elongation of RNA transcript; (vi) transcription termination, which results in a complete mRNA transcript. This description would require considering at least 4 different variables (i.e. molecular species) and 6 distinct biochemical reactions.

Since biochemical networks often involve a great number of biochemical reactions, these need to be simplified and described with lower resolution in order to reduce the number of variables and parameters. The six steps of transcriptional regulation can be for instance condensed in a single-step process, where the production of an mRNA transcript simply results from the binding of a transcription factor to DNA, neglecting intermediate reactions.

Many elementary biochemical reactions such as decomposition reactions, bimolecular association reactions, and enzyme-substrate reactions, are demonstrated to follow the Mass Action Law [34]. This principle states that the rate of a reaction is directly proportional to the product of the concentrations of the reactants.

Let's consider for instance a simple reaction where a single reactant species A is degraded with rate constant k:

$$A \xrightarrow{k} 0$$

Assuming the Mass Action principle, the reaction rate V will be equal to $V = k \cdot [A]$, where $[A]$ is the contrentration of molecular species A. . The same holds for

dissociation reactions such as the following, where a molecular species A dissociates into two species B and C:

$$A \xrightarrow{k} B + C$$

These reactions fall under the category of First-order kinetics, where reaction rates depend linearly on a reactant's concentration.
Conversely, the inverse biochemical reaction where two species associate resulting in a molecular complex follows Second-order kinetics:

$$B + C \xrightarrow{k} A$$

Here, according to the Mass Action Law, the rate of reaction is proportional to the product of the two reactants concentrations: $V = k \cdot [B] \cdot [C]$, and thus depends on a squared concentration.
However, if we consider more articulate biochemical reactions, the Mass Action principle can lead to more complex dependencies from molecule concentrations. Take for instance the typical catalytic reaction, where E represents an enzyme, S a substrate molecule and P the product molecule:

$$\mathrm{E} + \mathrm{S} \underset{\mathrm{k_{off}}}{\overset{\mathrm{k_{on}}}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{\mathrm{k_{cat}}} \mathrm{E} + \mathrm{P} \tag{2.1}$$

Reaction rates k_{on}, k_{off} and k_{cat} represent respectively the rate of E and S binding, the rate of their unbinding, and the rate of catalysis (i.e. the dissociation that leads to product formation and release of the enzyme).
In this biochemical system, the rate of P production has been demonstrated to depend on the substrate concentration $[S]$ following the Michaelis-Menten function [35]:

$$V = \frac{V_{max} \cdot [S]}{K_M + [S]} \tag{2.2}$$

where V_{max} represents the maximum rate achieved for maximal substrate concentration and K_M, named as "Michaelis-Menten constant", is the combination of reaction rate constants $K_M = \frac{k_{off} + k_{cat}}{k_{on}}$.
Thus, as the substrate is consumed during the progression of catalysis, the rate of P production is reduced in a nonlinear way, reaching null values in the absence of substrate. This kind of kinetics has been observed not only in the strict context of

catalytic reactions, but also in other kinds of interactions where the reaction rate depends nonlinearly on some reactant's concentration.
More complex kinetics include for instance Hill kinetics [36], used to describe binding reactions involving molecules with multiple binding sites, where binding events become more likely as binding sites are progressively occupied (i.e. cooperative binding).
In a biochemical network containing multiple reactions, kinetics plays a key role in shaping the global system's behaviour: distinct kinds of kinetics imply different relationships between nodes. These relationships can be mathematically modelled by considering a biochemical network as a dynamical system whose variables correspond to nodes, with values typically referring to molecule numbers or concentrations. Since molecules are individual objects, the quantitative measure of molecular content is in principle discrete, and its temporal evolution would thus require discrete updates. Moreover, biochemical reactions are stochastic events and thus demand for a probabilistic modelling approach, where reaction rates become reaction probabilities and molecule amounts are replaced by their probability distributions.
The Chemical Master Equation (CME) can be used to model the time evolution of the network state's probability distribution (i.e. the joint distribution of network variables) [37].
Take for instance a simple network involving three molecular species A, B, and C. Let's assume that the three species are involved in a bimolecular association reaction, where A and B reversibly bind to form the molecular complex C with a certain rate constant k_{on}, and unbind with rate constant k_{off}:

$$\mathrm{A} + \mathrm{B} \underset{\mathrm{k_{off}}}{\overset{\mathrm{k_{on}}}{\rightleftharpoons}} \mathrm{C} \tag{2.3}$$

Let A, B, and C be the respective molecule numbers of each species. The state of the system - defined by such molecule numbers - is subject to transitions due to the binding and unbinding reactions, each transition occurring with a certain probability. Assuming the validity of the Mass Action Law and assuming that, at each time step, the transition probabilities depend only on the current state of the system (i.e., that the network's time evolution is a Markov process [38]):

(i) the transition from (A,B,C) to $(A-1,B-1,C+1)$, representing the binding reaction, can occur with probability $k_{on} \cdot A \cdot B \cdot P(A,B,C)$.

(ii) the transition from (A,B,C) to $(A+1,B+1,C-1)$, representing the unbinding reaction, can occur with probability $k_{off} \cdot C \cdot P(A,B,C)$.

Thus, the system's CME can be written as:

$$\frac{dP(A,B,C)(t)}{dt} = k_{\text{on}}(A+1)(B+1)P(A+1,B+1,C-1)(t) \tag{2.4}$$

$$+ k_{\text{off}}(C+1)P(A-1,B-1,C+1)(t) \tag{2.5}$$

$$- (k_{\text{off}}C + k_{\text{on}}AB)P(A,B,C)(t) \tag{2.6}$$

and its solution provides the temporal evolution of the probability distribution of the network's state, $P(A,B,C)(t)$.
However, solving the CME might be challenging or unfeasible even for simple reaction networks [39]. For this reason, the time evolution of biochemical networks is often modelled with stochastic simulation algorithms that attempt to approximate the exact probability distributions provided by the CME. The most common algorithm employed to simulate biochemical networks is the stochastic Gilliespie algorithm [40]. This algorithm - similar to other stochastic algorithms - treats biochemical reactions as discrete events occurring at random times. The state of the system, given by the number of molecules of each species, is updated when one of these reactions occurs. The algorithm, which will be adopted in this thesis to simulate the dynamics of a network of interacting proteins, works as follows:

(i) Probabilities a_i are computed for each reaction i according to the Mass Action Law. If we consider a network formed by three molecular species, whose molecule numbers are A, B and C, these would correspond to the probability of binding, $a_1 = k_{on} \cdot A \cdot B$, and the probability of unbinding, $a_2 = k_{off} \cdot C$.
(ii) The time until the next reaction occurs is computed as $t_{\text{next}} = \frac{1}{a_{tot}} \ln\left(\frac{1}{r_1}\right)$ where r_1 is a random uniform number between 0 and 1, and a_{tot} is the sum of all reaction probabilities, thus for the network composed by A, B and C: $a_{tot} = a_1 + a_2 = k_{on} \cdot A \cdot B + k_{off} \cdot C$.
(iv) The next occurring reaction i is the one that satisfies: $i = \text{argmin}\left(\sum_{j=1}^{i} \frac{a_j}{a_0} > r_2\right)$,

where r_2 is another random number drawn uniformly between 0 and 1. This guarantees that the likelihood of each reaction to be chosen reflects its biochemical rate constant, as well as the concentrations of the species involved, statistically converging to the Mass Action rates.
Thanks to algorithms such as the Gilliespie, it is possible to estimate the probability distribution of a biochemical network's state without facing the solution of the CME. Nevertheless, the number of molecules of each species involved in a biochemical system - such as the cell - may be large, leading to small stochastic fluctuations and thereby allowing to neglect deviations from mean molecule amounts, adopting the so-called "mean-field approach". In this condition, continuous variables can be adopted to describe molecular species, and their temporal evolution can be described deterministically, using Ordinary Differential Equations (ODEs). It is indeed demonstrated that stochastic algorithms such as the Gillespie will converge to mean values predicted by ODEs if molecule amounts are large enough, correlations between species are negligible, and the system is well-mixed.
Under these assumptions, the network with binding and unbinding involving species A, B and C can be modelled with the following ODE system:

$$\frac{dA}{dt} = -k_{\text{on}} \cdot A \cdot B + k_{\text{off}} \cdot C \tag{2.7}$$

$$\frac{dB}{dt} = -k_{\text{on}} \cdot A \cdot B + k_{\text{off}} \cdot C \tag{2.8}$$

$$\frac{dC}{dt} = k_{\text{on}} \cdot A \cdot B - k_{\text{off}} \cdot C \tag{2.9}$$

Many ODE systems such as this one may not be solvable analytically, and thus require numerical integration methods such as the Runge-Kutta algorithm [41]. In this thesis we adopt an ODE-based approach to estimate the dynamics of mean molecule values. On the other hand, the Gilliespie algorithm will be employed to obtain probability distributions of molecule amounts necessary for computing statistical moments such as the Pearson correlation.

Chapter 3

Competition

Competition mechanisms are ubiquitous in our universe. The mere interaction of two or more entities with the environment - let these entities be particles as well as living organisms - implies competition.
The idea of competition has been an integral part of physical thinking since its inception. In the early days of physics, this concept can be found in the study of celestial mechanics: Kepler's first law, also known as the law of ellipses, described the shape of planetary orbits as ellipses with the Sun at one of the foci - a behaviour that later found its explanation in the competition between a planet's inertia and its gravitational attraction to the Sun [22].
Throughout the last century, competition has been involved in the most diverse fields of physics, from interactions governing astrophysical phenomena [42], to those that influence the phases of condensed matter systems [43], or even shape quantum states [44]. The competitive interplay between different interaction timescales was even observed to influence the arrangement of particles in materials, affecting the atomic and molecular structure of matter [45].
Conversely, the biological notion of competition has been limited for a long time to its role within the Darwinian framework of evolutionary biology: the principle that species or organisms within a population compete for limited resources, and that selective pressure drives adaptation mechanisms to enhance competitiveness, is well-known to underlie the diversity of life on Earth. Studies spanning from the 18th century to the present day explored evolutionary competition through diverse lines of evidence, including comparative anatomy [46], molecular biology [47], artificial selection experiments [48], and direct observations of natural selection in

contemporary species [49]. The Lotka-Volterra equations describe one of the most well-known mathematical models of intraspecific competition [50, 51].
However, as physics and biology started their approachment, a number of studies began to quantitatively investigate competition beyond the evolutionary field. Among these, a cornerstone study by Chodwdhury and collaborators explored competition and cooperation among biological motors within cells using statistical physics [52]. This kind of studies marked a shift towards the realm of quantitative biology, demonstrating how competition mechanisms are not limited to either subatomic particles or astronomical scales, but also extend to the microscopic world of cellular processes. Starting from the early 2000s, both physics and biology journals saw an increase of publications exploring competition in cellular and molecular biology with the use of mathematical modelling. Topics included competition among stem cells [53] or cell competition underlying tumor development [54], competition among yeast or bacterial strains [55], and competition between different kinds of molecules ranging from receptors to nucleic acids [56, 57].
Since this thesis focuses on a biomolecular level, in the next section we will briefly give some examples of ODE modelling and stochastic simulation of competitive mechanisms in this context.

3.1 Competition in biochemical systems

Any biochemical reaction implies the exploitation of some resource - let that be an enzyme, a molecular co-factor or simply energy. For instance, some biochemical reactions require the binding of a ligand to a molecule possessing only one binding site - an event that reversibly or irreversibly sequestrates the molecule from the environment, making it unaccessible to other potential binders [58]. Let's now consider the situation where two or more ligand species require binding to the same molecule to perform their biological tasks; whenever the number of such molecules is limited, competition will arise spontaneously, favouring the biological function of the ligand most capable of binding. In this scenario, the shared molecule can be viewed as a limited resource, whose availability to each ligand species depends not only on the ligand's binding ability, but also on that of competitor ligands. This simple network introduces a hidden mechanism that connects competitor ligands even without them physically interacting.

A well-known situation where two ligands compete for binding to the same molecular species is the competition between epinephrine and acetylcholine for binding to the receptors on the surface of heart muscle cells [59]. The heart is regulated by the autonomic nervous system, which has two branches: the sympathetic nervous system - that releases the adrenaline neurotransmitter during stressful situations - and the parasympathetic nervous system - responsible for calming the body down after stress by releasing acetylcholine. Both neurotransmitters function by binding to a specific kind of receptor present on heart muscle cells, thus giving rise to a competition mechanism. While the binding of adrenaline to the receptor initiates a signalling cascade that leads to increased heart rate and contractility, the binding of acetylcholine conversely initiates a pathway leading to reduced heart rate and contractility. Thus, the physiological outcome - aimed at maintaining the heart rate within an adaptive range - depends on the respective amounts of adrenaline and acetylcholine ligands able to bind to the receptor.
One of the simplest mathematical models that describes ligand-ligand competition for a generic receptor - widely used in pharmacology and biochemistry - is represented by the the following reactions [58]:

$$\mathrm{R} + \mathrm{A} \underset{\mathrm{k_{off_A}}}{\overset{\mathrm{k_{on_A}}}{\rightleftharpoons}} \mathrm{RA} \tag{3.1}$$

$$\mathrm{R} + \mathrm{B} \underset{\mathrm{k_{off_B}}}{\overset{\mathrm{k_{on_B}}}{\rightleftharpoons}} \mathrm{RB} \tag{3.2}$$

where *R* represents the receptor molecule, *A* and *B* represent the two competing ligands, *RA* and *RB* represent respective molecular complexes formed with the receptor, and rates k_{on} and k_{off} are the rate constants of binding and unbinding reactions relative to each ligand. Assuming that reactions follow the Mass Action Law, the system's time evolution can be modelled with the following ODEs:

$$\frac{d[R]}{dt} = -k_{on_A}[R][A] + k_{off_A}[RA] - k_{on_B}[R][B] + k_{off_B}[RB] \tag{3.3}$$

$$\frac{d[A]}{dt} = -k_{on_A}[R][A] + k_{off_A}[RA] \tag{3.4}$$

$$\frac{d[RA]}{dt} = k_{on_A}[R][A] - k_{off_A}[RA] \tag{3.5}$$

$$\frac{d[B]}{dt} = -k_{on_B}[R][B] + k_{off_B}[RB] \tag{3.6}$$

$$\frac{d[RB]}{dt} = k_{on_B}[R][B] - k_{off_B}[RB] \tag{3.7}$$

where concentrations are indicated for each molecular species enclosed in square brackets.
This model allows to calculate the fraction of receptor bound to each competing ligand at the steady state, by solving $\frac{d[S]}{dt} = 0$ for each species S. From the third and the fifth equations, we express $[RA]$ in terms of $[R]$ and $[A]$, as well as $[RB]$ in terms of $[R]$ and $[B]$:

$$[RA]_{ss} = \frac{k_{on_A}[R]_{ss}[A]_{ss}}{k_{off_A}} \tag{3.8}$$

$$[RB]_{ss} = \frac{k_{on_B}[R]_{ss}[B]_{ss}}{k_{off_B}} \tag{3.9}$$

In this system the total receptor concentration (comprising bound and unbound concentrations) is conserved and thus always equal to a constant $[R_{TOT}] = [R] + [RA] + [RB]$. We can therefore obtain equilibrium fractions of receptor bound to each ligand, F_A and F_B, dividing by $[R_{TOT}]$:

$$F_A = \frac{[RA]_{ss}}{[R_{TOT}]} = \frac{\frac{k_{on_A}[R]_{ss}[A]_{ss}}{k_{off_A}}}{[R_{TOT}]} \qquad \text{Fraction of receptor bound to ligand } A \tag{3.10}$$

$$F_B = \frac{[RB]_{ss}}{[R_{TOT}]} = \frac{\frac{k_{on_B}[R]_{ss}[B]_{ss}}{k_{off_B}}}{[R_{TOT}]} \qquad \text{Fraction of receptor bound to ligand } B \tag{3.11}$$

which depend on the binding and unbinding rate constants, and where equilibrium concentrations $[A]_{ss}$, $[B]_{ss}$ and $[R]_{ss}$ will depend on the respective relative concentrations of ligands and receptor. As one can observe, these fractions are proportional to $\frac{k_{on}}{k_{off}}$, indicating that a ligand with a higher binding-to-unbinding rate constant ratio has a binding advantage over competitors. In the adrenaline-acetylcholine competition, this would advantage the signalling pathway triggered by the most competitive ligand, leading to the prevalence of either sympathetic or parasympathetic functions.

Similar competitive scenarios have been found to govern different cellular functions. In growth factor signaling, competition among ligands can determine whether cells undergo proliferation, differentiation, or apoptosis [60]. Moreover, competition can influence cell fate decisions: competing ligands can dictate differentiation of cells into one cell type or another [61, 62]. For instance, in neural development, a cell can become a neuron or a glial cell based on competition for receptor binding [63]. The immune system is also affected by competitive mechanisms: competition among antigens for binding to receptors can determine the strength and specificity of the immune response, leading to immune reactions against specific targets [64–66]. Ligand competition can also regulate gene expression: winning ligands can lead to the activation of specific genes over others [67].

To our particular interest, competition has recently been proposed to shape post-transcriptional regulation - the biochemical system under study in the second project of this thesis. Indeed, competitive interactions among post-transcriptional regulators (miRNAs) and among their targets have been predicted to generate correlations between distinct RNA species [57], thus constituting a mechanism able to simultaneously tune multiple molecules. Therefore, although our project dealing with post-transcriptional regulation does not primarily focus on competition, we will as well study specific cases where this feature might affect the system's properties.

At the same time, the influence of competitive mechanisms remains poorly understood within some other cellular contexts. One of these is membrane trafficking, that is the process by which molecules are distributed throughout the cell and released to or internalised from the extracellular space [68]. In this context, the first project of this thesis proposes competition as a novel driving mechanism, showing that it might underlie the protein segregation events necessary for the proper functioning of trafficking processes.

The implications of competition in the two biochemical system - post-transcriptional

regulation and membrane trafficking - will be analyzed more in detail in each project's dedicated section.

Part I

Competition as a segregation mechanism in membrane trafficking

Chapter 4

Rab proteins in membrane trafficking

4.1 Membrane trafficking

The flow of material between the inside and the outside of the cell and among different locations inside the cell is necessary for its proper functioning. Once proteins, lipids and carbohydrates are synthesized, they are transported from the endoplasmatic reticulum to the Golgi, and from there they are directed towards different destinations or expelled from the cell through its membrane [68]. Conversely, molecules that cannot be produced by the cell itself - for instance vitamins - are captured from the extracellular space and internalized [69]. These processes guarantee the balance between the uptake and the release of nutrients, and are thus crucial for the cell's physiological health [70].

The ensemble of pathways involved in this flow of molecules is referred to as "membrane trafficking" [68] (see Figure 4.1). Among the most important membrane trafficking pathways are the endocytosis - comprising all processes that follow the internalization of molecules until their delivery - and the exocytosis - that conversely leads to to the secretion of molecules out of the cell.

In general, the transport of cargoes requires the formation of a motile transport intermediate: many pathways exploit lipid vesicles as means of trasport [71]. These vesicles are originated at the cell membrane: the membrane is deformed to generate a curved structure - a process called "vesicle budding"; next, the association of specific proteins at the neck of the budding vesicle allows its detachment from the membrane [72]. Once a vesicle is split from the membrane, it is able to reach its target desti-

nation by moving inside the cell. Vesicle motility occurs along the cytoskeleton on microtubules, thanks to specific molecular motors [73]. At destination, facilitated by the key protein family SNARE, vesicles will anchor their acceptor compartments - a mechanism referred to as "vesicle tethering" - and fuse with their membrane to release cargo [1]. This membrane plasticity is fundamental for the continuously occurring exchange of material: vesicles are able to fuse with other similar vesicles or with cellular organelles, and even split into smaller vesicles by fission [74].
Vesicle types can be classified based on the type of cargo they carry, the pathway where they belong and the type of proteins that characterize their surface. Indeed, which proteins cover vesicle surfaces depends on the trafficking stage that the vesicle is going through [75]. For instance, clathrin-coated vesicles transport cargo from the plasma membrane to specific compartments and to the Golgi [76], whereas the next transport stages are performed by vesicles covered with different coating proteins, such as COPI [77]. In turn, acceptor compartments expose specific proteins on their membranes, that act as signals allowing vesicles to recognize their correct destination [1].
Among compartments, vesicles called endosomes play a crucial role in trafficking by acting as sorting stations at crossroads between different pathways: inside them, different cargoes are sorted, as each one has to be delivered to its specific destination[78].
Endosomes are formed by the fusion of multiple vesicles along the endocytic route, reaching a diameter of approximately 0.5 μm. They are composed of regions of thin tubular extensions (~ 60 *nm* diameter) and large vesicles (~ 400 *nm* diameter) that confer them a multi-vesicular structure [78]. These distinct sub-domains are thought to be functionally determinant: sorting occurs through the localization of different cargoes to these very domains [77]. For instance, cargo targeted as waste to be recycled localizes on the endosome's tubular domains [77]. Tubules subsequently concentrate the recycling cargo for efficient transport out of the endosome, to deliver it to the so-called "endosomal recycling compartment". Cargoes that arrive to this compartment mostly reach the plasma membrane within minutes, to be subsequently expelled from the cell [78].

These membrane trafficking processes require the contribution of different molecule types, including the mentioned SNARE family and the various vesicle coating proteins. However, the Rab protein family plays perhaps the most crucial

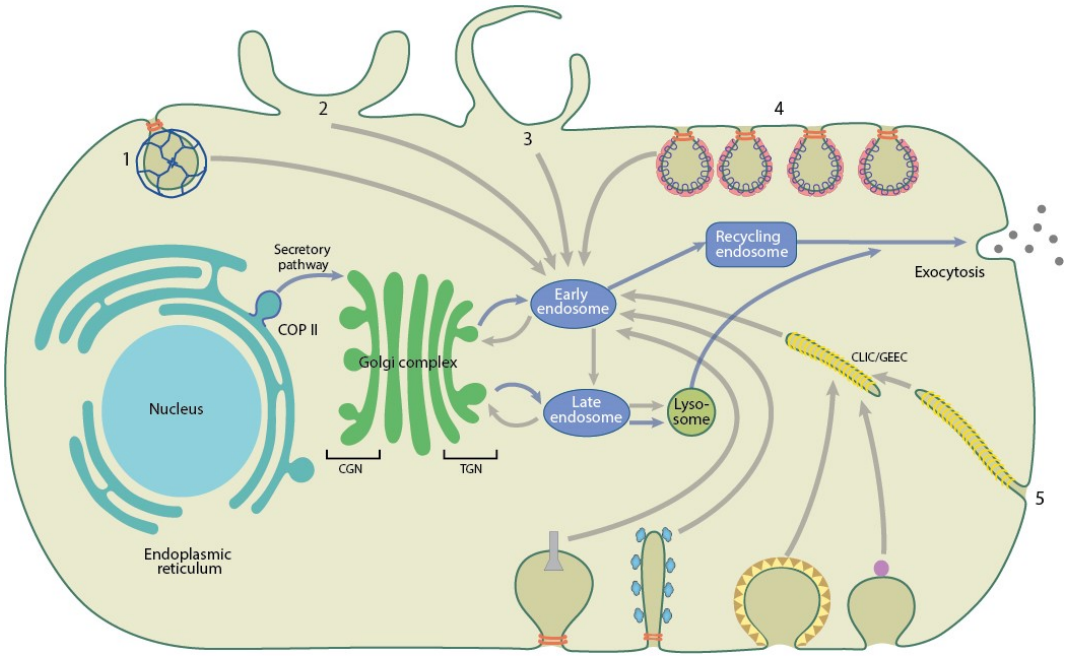


Fig. 4.1 Overview of cellular membrane trafficking pathways (adapted from: www.mechanobio.info).

role. These molecules were found to regulate all key trafficking steps, from the formation of vesicles to their direction to destination [79]. Their main roles in membrane trafficking pathways are schematized in Figure 4.2. The next sections are thus dedicated to characterize these proteins and their functions.

4.2 Rab proteins in membrane trafficking

The Rab family plays a central role in membrane trafficking, and is fundamental for the eukaryotic cell. Its importance is supported by its high conservation through evolution. Along their evolutionary path, these proteins underwent mutations with high plasticity: as eukaryotes diverged, they gained or lost members of the Rab family, generating thousands of variants among the different analyzed genomes [1]. The number of Rab variants can range from the 11 family members of the Saccharomyces Cerevisiae budding yeast to the 66 variants encoded in the human genome. Since no other protein family involved in intracellular transport has undergone such an expansion, it is thought that mutations in Rab proteins are the primary reason behind the differences between cellular organelles and compartments among species [1].

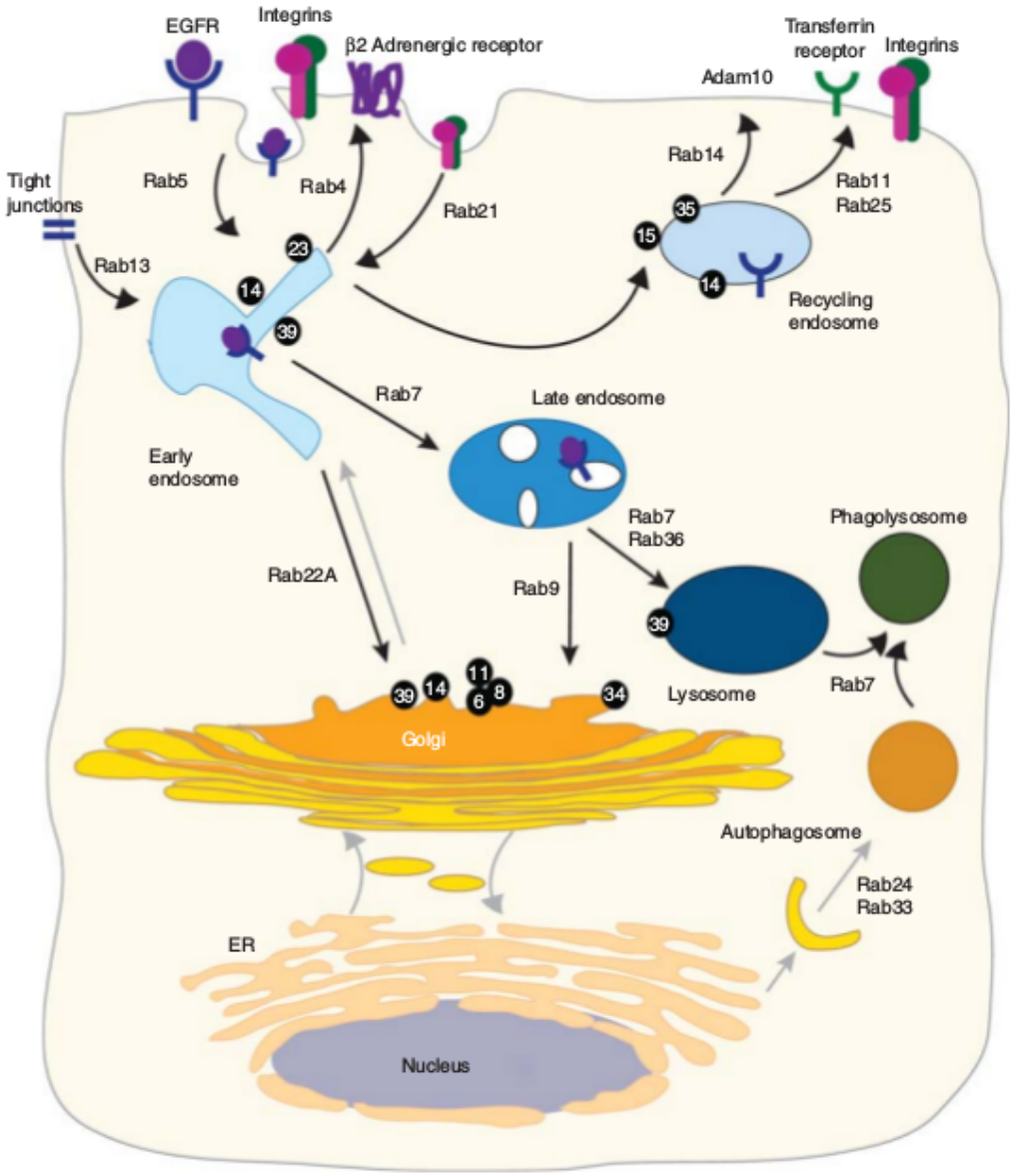


Fig. 4.2 Overview of Rab GTPases in the endocytic and recycling pathways (adapted from [1])

Consistently, in vivo studies in liver showed that their depletion causes a dramatic reduction in the number of endosomes, leading to endocytosis malfunctions [80, 81]. Because of this connection to trafficking defects, many members of the Rab family have been linked to tumour development and degenerative diseases [82, 83].
Rab proteins are part of a broader group of molecules - called GTPases - that behave like molecular switches: they continuously cycle between an active and an inactive state (see Figure 4.3). When active, they are bound to guanosine triphosphate (GTP), whereas when in their inactive state they are bound to guanosine diphosphate (GDP) [84]. The transition between these two states occurs either by hydrolysis, that turns GTP into GDP, or by nucleotide exchange, that vice versa converts GDP into GTP. These biochemical reactions are catalyzed by specific enzymes: Guanine Nucleotide Exchange Factors (GEFs) catalyze nucleotide exchange leading to Rab activation, whereas GTPase-Activating Proteins (GAPs) catalyze hydrolysis, causing Rab deactivation [84].

$$\mathrm{Rab}^{\mathrm{GDP}} \underset{\mathrm{GAP}}{\overset{\mathrm{GEF}}{\rightleftharpoons}} \mathrm{Rab}^{\mathrm{GTP}}$$

When silent, Rab proteins are kept in their inactive state (GDP-bound) in the cytosol, bound to molecular factors - the Rab GDP Dissociation Inhibitors (GDIs) - that prevent them from binding to compartmental membranes [85]. When needed, they are recruited on a compartment's membrane, where they are activated by GEFs. The conformational change caused by these enzymes makes Rabs unaccessible to GDI and thus stably bound to the membrane. Once active (bound to GTP), Rabs are able to bind effector molecules to execute their biological functions [84]. When their functions are terminated, hydrolysis by GAPs deactivates Rab proteins, reverting them to their GDP-bound state and making them once again accessible to GDI binding for storage in the cytosol.

The interconversion between GDP and GTP is the distinctive reaction that allows all GTPases to function: GTP is a high-energy molecule which serves as a resource for many cellular processes, including those of trafficking. It is indeed observed that the cycling between GDP- and GTP-bound forms underlies all functions exerted by Rab proteins in membrane trafficking events. For instance, GTP-bound Rabs are able to recruit specific effector molecules to initiate vesicle budding or to facilitate

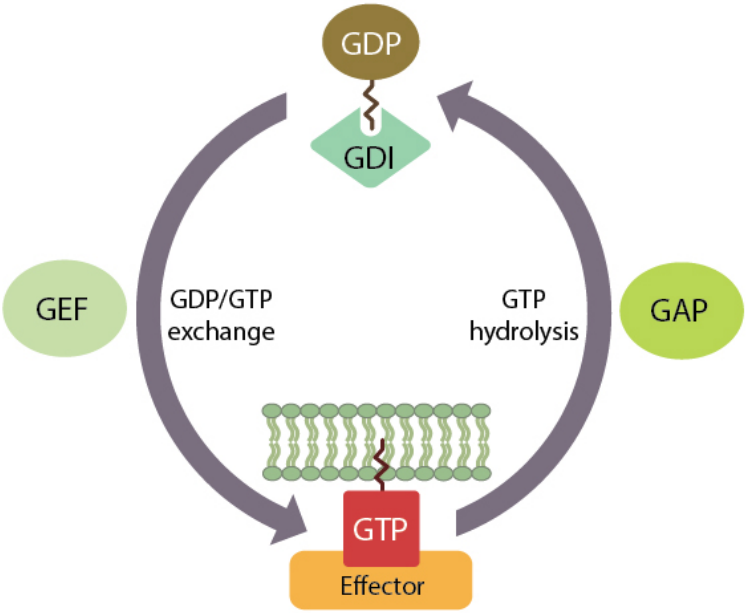


Fig. 4.3 Schematic of the Rab protein activation cycle. GEF enzymes mediate Rab activation by catalyzing GDP/GTP exchange; GAP enzymes mediate Rab deactivation by catalyzing GTP hydrolysis.

the fusion of vesicles with target compartments.

Through their cycling behavior, Rab proteins ensure the spatiotemporal regulation of all sequential stages of trafficking. Indeed, distinct members of the Rab family typically function one after the other in a sequence of trafficking steps. To coordinate their functions, these proteins typically use a mechanism referred to as "effector handoff", where a Rab species recruits GEFs, GAPs or other molecules needed by the next Rab along the pathway. For instance, a Rab can trigger activation of the next by recruiting its GEF; in turn, the newly activated Rab can recruit a GAP that terminates the action of the previous [1].

These interactions between GEFs, GAPs and effectors give rise to cascades involving multiple Rab species, generating crosstalk between distinct Rab proteins [86]. An example concerns the activities of Rab5 and Rab7, which are notoriously coupled by multivalent effector molecules capable of binding both species [1]. Similarly, Rab4 and Rab11 are known to cooperate through shared effectors along the route that leads to the outflow of waste - the so-called recycling pathway [87]. However, several effectors that couple sequentially acting Rab proteins still need to be identified to fully understand how cargo flows between different compartments.

The endocytic pathway is to date the best characterized trafficking route in terms of Rab protein functions. The first proteins found to participate in this pathway in mammalian cells were Rab5 and Rab7 [88]. Along this route, the two molecules regulate key steps in cargo internalization, uptake and transport to latter compartments [1]. While Rab5 is found on endosomes that operate in the first stage of endocytosis, i.e. early endosomes, Rab7 is detected during later endocytic stages on the so-called late endosomes. Interestingly, the two Rab species are thought to act as identity markers, making the two kinds of compartments recognizable by the trafficking machinery [2].
Following endocytosis, cargoes stored in endosomes are sorted according to their different destinations, with Rab4 and Rab11 as regulators [87]. However, how Rabs contribute to the various branchpoints of this route is only partially understood at present [89].
Remarkably, it was noticed that Rab proteins tend to have a narrower distribution on the surface of compartments with respect to other proteins involved in membrane trafficking [1]. Moreover, multiple evidences by Zerial and coworkers showed that each intracellular location appears characterized by distinct members of the Rab family, thus making these proteins the first true molecular markers for different compartments involved in trafficking [90]. Not only different Rabs are enriched in different compartments, but also in unique domains of the same compartment where specific functions take place. For instance, endosomes harbor Rab5, Rab4 and Rab11 in distinct and sometimes overlapping regions that play distinct roles in endosome maturation, cargo sorting and recycling [91]. In sum, Rab proteins form patterns that enable the unambiguous identification of target membranes. This interesting phenomenon is referred to as "membrane identity" [1]. Yet, the mechanisms that allow these patterns to be correctly localized and formed are still unclear, and the currently proposed hypotheses are brought forward in the next section.

4.3 Rab proteins and membrane identity

The current idea is that Rabs define the identity of compartmental membranes, making them recognisable in the busy trafficking scenario [92]. However, also other GTPases and specific phosphoinositide lipids have been proposed to specify membrane identity [93]. It seems likely that a single factor cannot explain by itself how

membrane identity is achieved. In vitro studies currently support the hypothesis of cooperativity, as it has emerged that Rab and SNARE machineries of early endosomes are both necessary for membrane fusion and cargo delivery [1].
Nevertheless, the issue of how inactive Rabs, initially stored in the cytosol, are targeted to specific compartments or even sub-domains for activation is still unsolved. It was proposed that a structural feature of the Rab molecule might provide a targeting mechanism, but later it was demonstrated that this alone does not provide a sufficient condition, and that multiple regions of the molecule contribute to the localization of Rabs on precise membrane locations [94, 93].
An alternative hypothesis explaining the localized activation of Rab proteins relies on GEF enzymes. Some studies showed that the presence of these enzymes is sufficient to target a Rab to a specific membrane location within the cell [95, 96]. Similarly, it was also observed that GDI alone is able to extract Rabs from a lipid bilayer (i.e. the common structure of compartmental membranes) [93].
Other mechanisms can determine Rab localization by acting on their activation cycle. For instance, covalent modification such as phosphorylation or some bacterial pathogens able to modify these proteins (e.g. *Legionella*) can alter the GDP-GTP cycle and thus domain formation on membranes [1].
Crucial for precise membrane identification is that domains formed by different Rab GTPases are maintained tightly separated, either in space or in time. Diverse potential mechanisms able to generate this reciprocal exclusivity have been proposed. In general, a prerequisite for the spatial or temporal confinement of distinct Rabs on organelle membranes is that a Rab species affects the accumulation of its neighbour. Often, these interactions rely on mutual inhibition exerted by a Rab on its cognate species [1]. For instance, inhibition between distinct Rabs has been shown to underlie the mutual exclusivity of Rab5 and Rab7 necessary to maintain the identity of early and late endosomes in the endocytic pathway [3]. At the first stages of the pathway, endosomes are Rab5-positive, whereas Rab7 is absent. Through repetitive fusion and fission events, endocytosed cargo is transferred into these endosomes, which become progressively fewer and larger while moving towards the cell nucleus. At this stage - reached in 10-15 minutes from the pathway initiation - the concentration of Rab5 reaches a peak which is soon followed by its complete loss and replacement by Rab7. This change of membrane identity marks the transition from early to late endosomes and is referred to as "Rab conversion" [1].

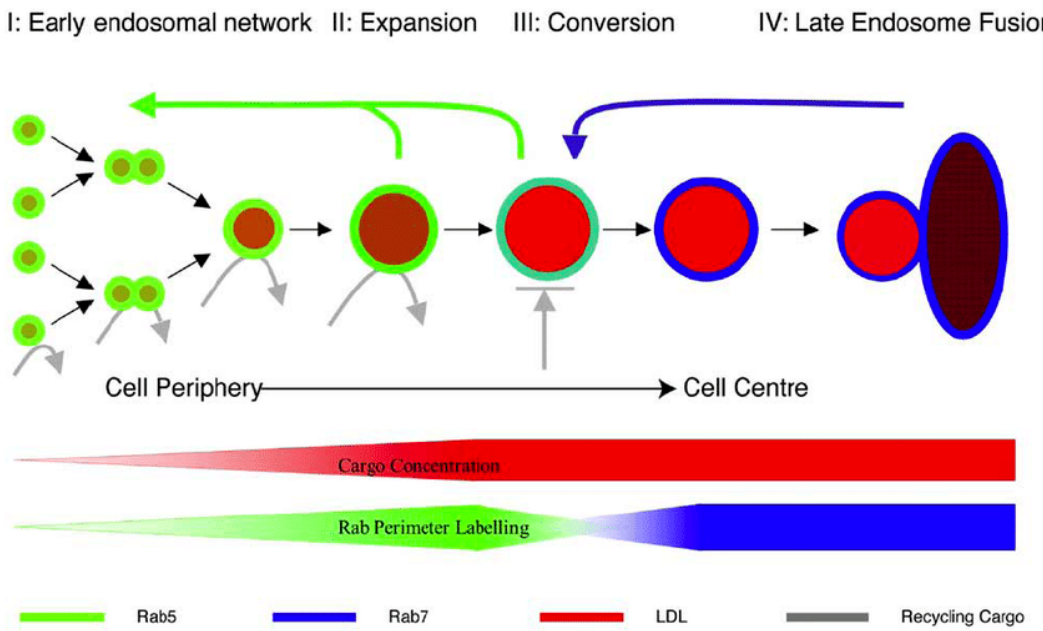


Fig. 4.4 Progression from Rab5-positive early endosomes to Rab7-positive late endosomes (adapted from: [2])

Zerial and co-workers studied the interaction between Rab5 and Rab7 to uncover the mechanism that allows their temporal segregation [3]. The authors predicted temporal concentrations of the two proteins using ODE modelling. By analysing the stability of different models, they were able to identify the best mechanism to explain experimental data collected on early and late endosomes: they proposed the so-called "cut-out switch" (see scheme in Figure 4.5), a mechanism where Rab5 initially triggers the activation of Rab7, whose progressive increase above a threshold eventually inhibits Rab5 activity leading to its disassembly from the membrane. Their analysis showed that the system displays a bistable behaviour: Rab5 accumulation is followed by a sudden transition at which Rab5 is completely lost and replaced by Rab7.

However, specificity and mutual exclusivity have been as well observed in some cases where the Rab proteins involved have not been found connected by straighforward interactions. Experiments conducted by our collaborators evidenced that Rab5 and Rab11, two Rab GTPases with no known connecting interactions, are found on the same kind of compartment - the recycling endosome - where they tend to form spatially exclusive domains. In this thesis we propose an indirect

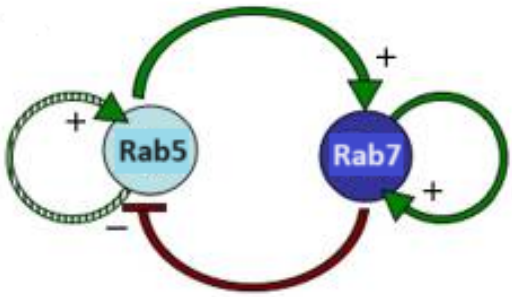


Fig. 4.5 Schematic of the cut-out switch model developed by Zerial and coworkers for Rab domain identity generation and endosome conversion from Rab5 to Rab7 (adapted from: [3]).

relationship relying on competition as a potential mechanism able to generate this mutual segregation, which will be illustrated in the next chapter.

Chapter 5

The Rab5-Rab11 problem: theoretical models and predictions

5.1 Rab5 and Rab11 in literature

Among Rab GTPases, Rab5 has long been known for its crucial role in the endocytic pathway [97]. A substantial body of literature highlights Rab5's localization on early endosomes, where it orchestrates the fate of cargo molecules by interacting with effectors and other factors.
On the other hand, Rab11 is known to control the recycling pathway [98]. This protein has been observed to act on the membrane of recycling endosomes, where cargo molecules are sorted to send waste out of the cell.
However, early and recycling endosomes are at the crossroad of several transport routes and their distinction is mainly based on the flow direction of cargo molecules [87]. Therefore, despite the rare co-occurrence of Rab5 and Rab11 on endosomes, Rab5 and Rab11 pathways are thought to be linked.
Moreover, Vantaggiato and coworkers demonstrated that the two Rab proteins both interact with an effector molecule present on intracellular membranes called Zfyve26 [99]. Specifically, they showed that Zfyve26 mutations affect the membrane binding and activity of both GTPases. Thus, although evidence of a direct interaction between the two GTPases is lacking, these findings suggest that Rab5 and Rab11 may be indirectly linked through interaction with their shared effector Zfyve26.
Therefore, through theoretical modelling we formulate the hypothesis that both Rab5

and Rab11 might be able to bind to Zfyve26. In this scenario Zfyve26 is a resource shared by Rab5 and Rab11 but can only bind one of the two GTPases at a time, thus generating competition. Such competition hypothesis will be developed through the formulation of five different models of microscopic interaction illustrated in the next sections. For each model, we study mean values and reciprocal fluctuations of the two Rabs as functions of the amout of shared resource Zfyve26.

5.2 Mathematical models

5.2.1 Model assumptions

All our mathematical models rely on a few assumptions based on preexisting biological evidence:

- The total number of molecules in the system is constant, i.e. the system is closed. This assumption relies on the knowledge that the temporal scale of GTPase signalling reactions is negligible with respect to the time needed by the cell to synthesize new GTPase molecules or degrade them [100, 101]. Therefore, our mathematical models do not consider biochemical reactions where Rab proteins are created or degraded, and instead rely on the conservation of the same global pool of Rab molecules.
- Both Rab5 and Rab11 undergo the activation cycle common to all GTPases:
  - The Rab is in its inactive (Rab^{GDP}) state in the cytosol;
  - Rab^{GDP} is recruited to the membrane;
  - Rab^{GDP} is activated on the membrane by a GEF enzyme specific to the Rab species, which converts GDP into GTP by nucleotide exchange;
  - The active Rab (Rab^{GTP}) is inactivated by GAP enzymes specific to the Rab species, which covert GTP into GDP by hydrolysis;
  - Rab^{GDP} dissociates from the membrane, and is transferred back to the cytosol.

  Note that these steps imply that the inactive (Rab^{GDP}) GTPases can shuttle between the cytoplasm and the endocytic membrane, whereas the active

(Rab^{GTP}) GTPases are restricted to the membrane: the nucleotide exchange mediated by GEFs and the nucleotide hydrolysis mediated by GAPs occur only at the endocytic membrane. This assumption stems from the knowledge that Rab activation is closely related to their association with membranes [84].

- Zfyve26 is modelled as a monovalent Rab effector (i.e. possessing a single GTPase binding site) thus generating competition between Rab5 and Rab11 for binding. This feature, which ensures specific one-to-one interaction with a Rab protein, is shared by many Rab effectors [102].

- Zfyve26 binds to $Rab5^{GTP}$ [99], and the resulting molecular complex, i.e. $Rab5^{GTP}:Zfyve26$, mediates activation of further Rab5 molecules, i.e. Zfyve26 mediates an activatory Rab5 feedback loop. This modelling choice relies on the evidence that many Rab GTPases, including Rab5 [103], form domains on compartmental membranes thanks to positive self-loop mechanisms mediated by effectors.

- Zfyve26 binds either to the inactive ($Rab11^{GDP}$) or the active ($Rab11^{GTP}$) form of Rab11 and the resulting molecular complex induces either Rab11 deactivation or activation. We choose to explore these four possibilities because there does not exist more articulate information than the unspecific evidence of Rab11 interaction with Zfyve26 [99].

- The system is well-mixed, meaning that molecule diffusion on the endocytic membrane is fast with respect to reactions involving the two GTPases. This assumption allows us to model reactions occurring on the endosomal membrane using Ordinary Differential Equations independent from space.

The next section will illustrate the development of mathematical models based on these assumptions.

5.2.2 Model formulation

All possible reactions considered in our models, with each model containing a subset of such reactions, can be summarized as follows:

- Binding to the membrane and unbinding from the membrane, occurring at rates k_{in} and k_{out}, are represented by reactions:

$$\mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Cyt}} \xrightarrow[\mathrm{k_{in}}]{} \mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}}$$
$$\mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}} \xrightarrow[\mathrm{k_{out}}]{} \mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Cyt}}$$

  where *Rab* indicates any of the two GTPases (Rab5 or Rab11).

- Reactions describing the Rab activity cycle, i.e. activation by GEFs and deactivation by GAPs, are represented as:

$$\mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}} \xrightarrow[\mathrm{k_{GEF}}]{} \mathrm{Rab}^{\mathrm{GTP}}_{\mathrm{Mem}}$$
$$\mathrm{Rab}^{\mathrm{GTP}}_{\mathrm{Mem}} \xrightarrow[\mathrm{k_{GAP}}]{} \mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}}$$

  where parameters k_{GEF} and k_{GAP} represent rates of Rab activation and deactivation.

- Reactions describing the interaction with Zfyve26 (i.e. effector binding and unbinding, either with the active or the inactive Rab) read as follows:

$$\mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}} + \mathrm{E} \xrightarrow[\mathrm{k_A}]{} \mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}} : \mathrm{E}$$
$$\mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}} : \mathrm{E} \xrightarrow[\mathrm{k_D}]{} \mathrm{Rab}^{\mathrm{GDP}}_{\mathrm{Mem}} + \mathrm{E}$$

  or

$$\mathrm{Rab}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E} \xrightarrow[\mathrm{k_{A'}}]{} \mathrm{Rab}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E}$$
$$\mathrm{Rab}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E} \xrightarrow[\mathrm{k_{D'}}]{} \mathrm{Rab}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E}$$

  where E represents Zfyve26, $RabGDP : E$ represents a Rab^{GDP}-Zfyve26 molecular complex, whereas $Rab^{GTP} : E$ represents a Rab^{GTP}-Zfyve26 molecular complex. k_A and k_D describe respectively the rates of RabGDP-Zfyve26

complex formation and dissociation, whereas k'_A and k'_D represent respectively the rates of Rab^{GTP}-Zfyve26 complex formation and dissociation. Each Rab is assumed to undergo only a subset of such reactions, that depends on the kind of interaction with Zfyve26 assumed. The different possibilities considered give rise to the different models described hereafter.

Competitive models I and II

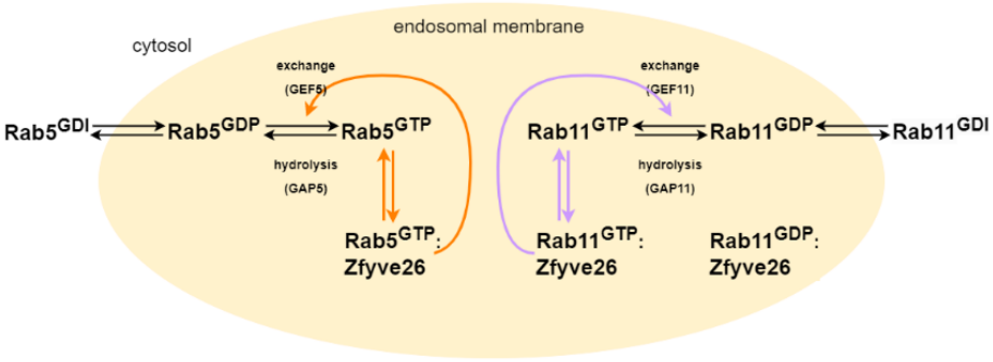


Fig. 5.1 Schematization of reactions in the competitive model ComI.

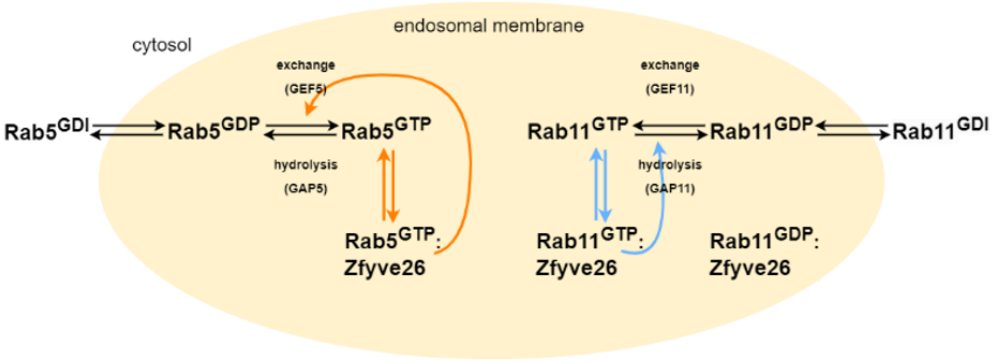


Fig. 5.2 Schematization of reactions in the competitive model ComII.

While sharing the positive Rab5 self-loop mediated by $Rab5^{GTP} : Zfyve26$ common to all models, the first two models ComI and ComII hypothesize that Rab11 binds to Zfyve26 when active, thus implying the following biochemical reactions for the two GTPases:

$$\begin{aligned}
\mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Cyt}} &\xrightarrow[\mathrm{k_{in5}}]{} \mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{k_{out5}}]{} \mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Cyt}} \\
\mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{k_{GEF5}}]{} \mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} \\
\mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{k_{GAP5}}]{} \mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E} &\xrightarrow[\mathrm{k_{A5'}}]{} \mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E} \\
\mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E} &\xrightarrow[\mathrm{k_{D5'}}]{} \mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Cyt}} &\xrightarrow[\mathrm{k_{in}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{k_{out}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Cyt}} \\
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{k_{GEF}}]{} \mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} \\
\mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{k_{GAP}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E} &\xrightarrow[\mathrm{k_{A'}}]{} \mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E} \\
\mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E} &\xrightarrow[\mathrm{k_{D'}}]{} \mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E}
\end{aligned}$$

and the corresponding system of ODEs:

$$\frac{d[Rab5^{GDP}_{Cyt}]}{dt} = k_{out5}[Rab5^{GDP}_{Mem}] - k_{in5}[Rab5^{GDP}_{Cyt}] \tag{5.1}$$

$$\frac{d[Rab5^{GDP}_{Mem}]}{dt} = k_{in5}[Rab5^{GDP}_{Cyt}] - k_{out5}[Rab5^{GDP}_{Mem}] + k_{GAP5}[Rab5^{GTP}_{Mem}] - k_{GEF5}[Rab5^{GDP}_{Mem}] \tag{5.2}$$

$$\frac{d[Rab5^{GTP}_{Mem}]}{dt} = k_{GEF5}[Rab5^{GDP}_{Mem}] - k_{GAP5}[Rab5^{GTP}_{Mem}] - k_{(A5')}[Rab5^{GTP}_{Mem}][E] + k_{(D5')}[Rab5^{GTP}_{Mem} : E] \tag{5.3}$$

$$\frac{d[Rab5^{GTP}_{Mem} : E]}{dt} = k_{(A5')}[Rab5^{GTP}_{Mem}][E] - k_{(D5')}[Rab5^{GTP}_{Mem} : E] \tag{5.4}$$

$$\frac{d[Rab11^{GDP}_{Cyt}]}{dt} = k_{out11}[Rab11^{GDP}_{Mem}] - k_{in11}[Rab11^{GDP}_{Cyt}] \tag{5.5}$$

$$\frac{d[Rab11^{GDP}_{Mem}]}{dt} = k_{in11}[Rab11^{GDP}_{Cyt}] - k_{out11}[Rab11^{GDP}_{Mem}] + k_{GAP11}[Rab11^{GTP}_{Mem}] - k_{GEF11}[Rab11^{GDP}_{Mem}] \tag{5.6}$$

$$\frac{d[Rab11^{GTP}_{Mem}]}{dt} = k_{GEF11}[Rab11^{GDP}_{Mem}] - k_{GAP11}[Rab11^{GTP}_{Mem}] - k_{(A11')}[Rab11^{GTP}_{Mem}][E] + k_{(D11')}[Rab11^{GTP}_{Mem} : E] \tag{5.7}$$

$$\frac{d[Rab11^{GTP}_{Mem} : E]}{dt} = k_{(A11')}[Rab11^{GTP}_{Mem}][E] - k_{(D11')}[Rab11^{GTP}_{Mem} : E] \tag{5.8}$$

$$\frac{d[E]}{dt} = k_{D5'}[Rab5^{GTP}_{Mem} : E] - k_{A5'}[Rab5^{GTP}_{Mem}][E] + k_{D11'}[Rab11^{GTP}_{Mem} : E] - k_{A11'}[Rab11^{GTP}_{Mem}][E] \tag{5.9}$$

where the parameter k_{GEF5} is made dependent on the amount of $Rab5^{GTP}$: $Zfyve26$ molecular complex as a Hill function, to implement the assumed activatory self-loop mediated by the effector:

$$k_{GEF5} = k_{GEF5b} + k_{GEF5MAX}\left(\frac{(Rab5^{GTP}_{Mem} : E)^n}{((Rab5^{GTP}_{Mem} : E)^n + k^n_{GEF5M})}\right) \tag{5.10}$$

whereas the Rab5 deactivation rate k_{GAP5} is constant:

$$k_{GAP5} = k_{GAP5b} \tag{5.11}$$

here k_{GEF5b} and k_{GAP5b} respectively represent constant basal rates of GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GEF5MAX}$ is the maximal rate of Rab5 activation mediated by the effector, approached as the level of Rab5-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GEF5M}, determines the number of complex molecules required to reach half of the maximum rate $k_{GEF5MAX}$.

However, the two models differ in the action exerted by the $Rab11^{GTP}$:Zfyve26 complex, activatory for Competitive model I and inhibitory for Competitive model II (see schemes 5.1, 5.2). This activation/inhibition is modelled by making either the Rab11 activation rate kGEF11 or its deactivation kGAP11 dependent on the amount of $Rab11^{GTP} : Zfyve26$ complex.

For model ComI, where the complex fulfils an activatory function, the rate k_{GEF11} is a Hill function of $Rab11^{GTP} : Zfyve26$ whereas k_{GAP11} is kept constant and corresponding to the basal deactivation rate:

$$k_{GEF11} = k_{GEF11b} + k_{GEF11MAX}\left(\frac{(Rab11^{GTP}_{Mem} : E)^n}{((Rab11^{GTP}_{Mem} : E)^n + k^n_{GEF11M})}\right) \tag{5.12}$$

$$k_{GAP11} = k_{GAP11b} \tag{5.13}$$

here k_{GEF11b} and k_{GAP11b} respectively represent constant basal rates Rab11 of GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GEF11MAX}$ is the maximal rate of Rab11 activation mediated by the effector, approached as the level of Rab11-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GEF11M}, determines the number of complex molecules required to reach half of the maximum rate $k_{GEF11MAX}$.

On the other hand, for model ComII, where the complex causes Rab11 deactivation, it is the rate k_{GAP11} that depends on $Rab11^{GTP} : Zfyve26$ as a Hill function while k_{GEF11} is kept constant and corresponding to the basal activation rate:

$$k_{GAP11} = k_{GAP11b} + k_{GAP11MAX} \left(\frac{(Rab11^{GTP}_{Mem} : E)^n}{((Rab11^{GTP}_{Mem} : E)^n + k^n_{GAP11})} \right) \tag{5.14}$$

$$k_{GEF11} = k_{GAP11b} \tag{5.15}$$

here k_{GEF11b} and k_{GAP11b} respectively represent constant basal rates Rab11 of GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GAP11MAX}$ is the maximal rate of Rab11 activation mediated by the effector, approached as the level of Rab11-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GAP11M} determines the number of complex molecules required to reach half of the maximum rate $k_{GAP11MAX}$.

Competitive model III and IV

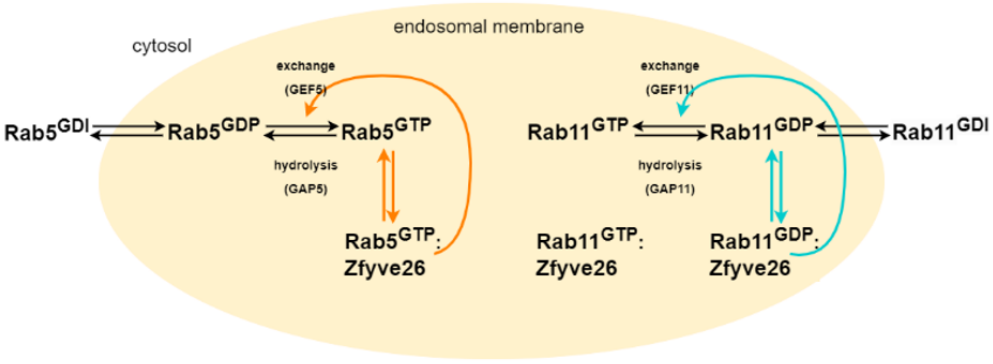


Fig. 5.3 Schematization of reactions in the competitive model ComIII.

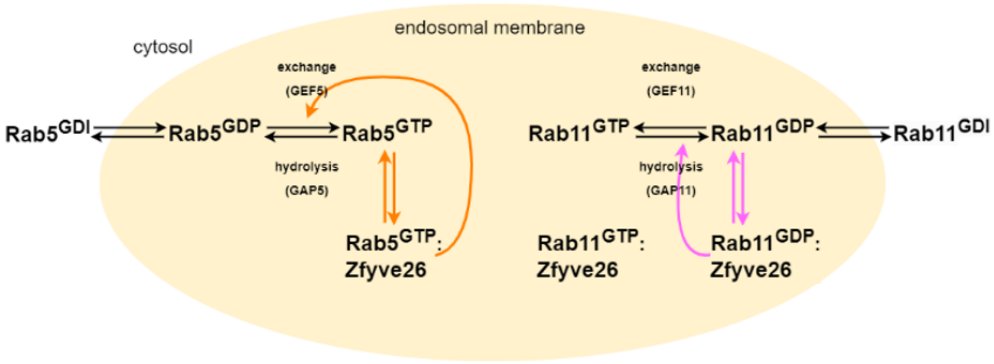


Fig. 5.4 Schematization of reactions in the competitive model ComIV.

The third and fourth models assume that Rab11 binds to Zfyve26 when inactive, together with the Rab5 activatory self-loop mediated by $Rab5^{GTP} : Zfyve26$ shared by all models, and thus imply the following biochemical reactions:

$$\begin{aligned}
\mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Cyt}} &\xrightarrow[\mathrm{K_{in5}}]{} \mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{K_{out5}}]{} \mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Cyt}} \\
\mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{K_{GEF5}}]{} \mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} \\
\mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{K_{GAP5}}]{} \mathrm{Rab5}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E} &\xrightarrow[\mathrm{K_{A5'}}]{} \mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E} \\
\mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} : \mathrm{E} &\xrightarrow[\mathrm{K_{D5'}}]{} \mathrm{Rab5}^{\mathrm{GTP}}_{\mathrm{Mem}} + \mathrm{E}
\end{aligned}$$

$$\begin{aligned}
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Cyt}} &\xrightarrow[\mathrm{K_{in}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{K_{out}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Cyt}} \\
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{K_{GEF}}]{} \mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} \\
\mathrm{Rab11}^{\mathrm{GTP}}_{\mathrm{Mem}} &\xrightarrow[\mathrm{K_{GAP}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} \\
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} + \mathrm{E} &\xrightarrow[\mathrm{K_{A}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} : \mathrm{E} \\
\mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} : \mathrm{E} &\xrightarrow[\mathrm{K_{D}}]{} \mathrm{Rab11}^{\mathrm{GDP}}_{\mathrm{Mem}} + \mathrm{E}
\end{aligned}$$

and the corresponding system of ODEs:

$$\frac{d[Rab5^{GDP}_{Cyt}]}{dt} = k_{out5}[Rab5^{GDP}_{Mem}] - k_{in5}[Rab5^{GDP}_{Cyt}] \tag{5.16}$$

$$\begin{aligned}\frac{d[Rab5^{GDP}_{Mem}]}{dt} &= k_{in5}[Rab5^{GDP}_{Cyt}] - k_{out5}[Rab5^{GDP}_{Mem}] \\ &\quad + k_{GAP5}[Rab5^{GTP}_{Mem}] - k_{GEF5}[Rab5^{GDP}_{Mem}]\end{aligned} \tag{5.17}$$

$$\begin{aligned}\frac{d[Rab5^{GTP}_{Mem}]}{dt} &= k_{GEF5}[Rab5^{GDP}_{Mem}] - k_{GAP5}[Rab5^{GTP}_{Mem}] \\ &\quad - k_{(A5')}[Rab5^{GTP}_{Mem}][E] + k_{(D5')}[Rab5^{GTP}_{Mem} : E]\end{aligned} \tag{5.18}$$

$$\frac{d[Rab5^{GTP}_{Mem} : E]}{dt} = k_{(A5')}[Rab5^{GTP}_{Mem}][E] - k_{(D5')}[Rab5^{GTP}_{Mem} : E] \tag{5.19}$$

$$\frac{d[Rab11^{GDP}_{Cyt}]}{dt} = k_{out11}[Rab11^{GDP}_{Mem}] - k_{in11}[Rab11^{GDP}_{Cyt}] \tag{5.20}$$

$$\begin{aligned}\frac{d[Rab11^{GDP}_{Mem}]}{dt} &= k_{in11}[Rab11^{GDP}_{Cyt}] - k_{out11}[Rab11^{GDP}_{Mem}] \\ &\quad + k_{GAP11}[Rab11^{GTP}_{Mem}] - k_{GEF11}[Rab11^{GDP}_{Mem}] \\ &\quad - k_{A11}[Rab11^{GDP}_{Mem}][E] + k_{(D11)}[Rab11^{GDP}_{Mem} : E]\end{aligned} \tag{5.21}$$

$$\frac{d[Rab11^{GTP}_{Mem}]}{dt} = k_{GEF11}[Rab11^{GDP}_{Mem}] - k_{GAP11}[Rab11^{GTP}_{Mem}] \tag{5.22}$$

$$\frac{d[Rab11^{GDP}_{Mem} : E]}{dt} = k_{(A11)}[Rab11^{GDP}_{Mem}][E] - k_{(D11)}[Rab11^{GDP}_{Mem} : E] \tag{5.23}$$

$$\begin{aligned}\frac{d[E]}{dt} &= k_{D5'}[Rab5^{GTP}_{Mem} : E] - k_{A5'}[Rab5^{GTP}_{Mem}][E] \\ &\quad + k_{(D11)}[Rab11^{GDP}_{Mem} : E] - k_{A11}[Rab11^{GDP}_{Mem}][E]\end{aligned} \tag{5.24}$$

As in all models, the parameter k_{GEF5} is modelled as dependent on the amount of $Rab5^{GTP} : Zfyve26$ molecular complex as a Hill function, to implement the assumed activatory self-loop mediated by the effector:

$$k_{GEF5} = k_{GEF5b} + k_{GEF5MAX}\left(\frac{(Rab5^{GTP}_{Mem} : E)^n}{((Rab5^{GTP}_{Mem} : E)^n + k^n_{GEF5M})}\right) \tag{5.25}$$

whereas the Rab5 deactivation rate k_{GAP5} is kept constant:

$$k_{GAP5} = k_{GAP5b} \tag{5.26}$$

where k_{GEF5b} and k_{GAP5b} respectively represent constant basal rates of GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GEF5MAX}$ is the maximal rate of Rab5 activation mediated by the effector, approached as the level of Rab5-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GEF5M} determines the number of complex molecules required to reach half of the maximum rate $k_{GEF5MAX}$.

However, the two models differ in the action exerted by the Rab11GDP:Zfyve26 complex, activatory for Competitive model III and inhibitory for Competitive model IV (see schemes 5.3, 5.4). This activation/inhibition is modelled by making either the Rab11 activation rate k_{GEF11} or its deactivation k_{GAP11} dependent on the amount of $Rab11^{GDP} : Zfyve26$ complex.
For model ComIII, where the complex fulfils an activatory function, the rate k_{GEF11} is a Hill function of Rab11GDP:Zfyve26 whereas k_{GAP11} is kept constant and equal to basal deactivation:

$$k_{GEF11} = k_{GEF11b} + k_{GEF11MAX}\left(\frac{(Rab11^{GDP}_{Mem} : E)^n}{((Rab11^{GDP}_{Mem} : E)^n + k^n_{GEF11M})}\right) \tag{5.27}$$

$$k_{GAP11} = k_{GAP11b} \tag{5.28}$$

where k_{GEF11b} and k_{GAP11b} respectively represent constant basal rates of Rab11 GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GEF11MAX}$ is the maximal rate of Rab11 activation mediated by the effector, approached as the level of Rab11-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GEF11M} determines the number of complex molecules required to reach half of the maximum rate $k_{GEF11MAX}$.

Conversely, in model ComIV, where the complex causes Rab11 deactivation, it is the rate k_{GAP11} that depends on $Rab11^{GDP} : Zfyve26$ as a Hill function while k_{GEF11} is kept constant and corresponding to a basal activation:

$$k_{GAP11} = k_{GAP11b} + k_{GAP11MAX}\left(\frac{(Rab11^{GDP}_{Mem} : E)^n}{((Rab11^{GDP}_{Mem} : E)^n + k^n_{GAP11})}\right) \tag{5.29}$$

$$k_{GEF11} = k_{GAP11b} \tag{5.30}$$

where k_{GEF11b} and k_{GAP11b} respectively represent constant basal rates Rab11 of GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GAP11MAX}$ is the maximal rate of Rab11 activation mediated by the effector, approached as the level of Rab11-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GAP11M} determines the number of complex molecules required to reach half of the maximum rate $k_{GAP11MAX}$.

Non-competitive model Cas

Additional to the four competitive models ComI-IV, we developed a non-competitive model to be adopted as our null hypothesis, where Rab11 is not allowed to bind to Zfyve26. Thus, this model shares the Rab5 positive self-loop mediated by $Rab5^{GTP} : Zfyve26$ common to all models, but lacks competition between the two GTPases for the effector. Here, the interaction between Rab5 and Rab11 relies on a typical mechanism of GTPase interaction [104–107] where one Rab in complex with the effector affects the activity of the other Rab species. It is indeed the $Rab5^{GTP} : Zfyve26$ complex that triggers deactivation of Rab11. This mechanism is commonly referred to as "GTPase cascade", hence our model name Cas, and is described by the following set of equations:

$$\frac{d[Rab5^{GDP}_{Cyt}]}{dt} = k_{out5}[Rab5^{GDP}_{Mem}] - k_{in5}[Rab5^{GDP}_{Cyt}] \tag{5.31}$$

$$\frac{d[Rab5^{GDP}_{Mem}]}{dt} = k_{in5}[Rab5^{GDP}_{Cyt}] - k_{out5}[Rab5^{GDP}_{Mem}] + k_{GAP5}[Rab5^{GTP}_{Mem}] - k_{GEF5}[Rab5^{GDP}_{Mem}] \tag{5.32}$$

$$\frac{d[Rab5^{GTP}_{Mem}]}{dt} = k_{GEF5}[Rab5^{GDP}_{Mem}] - k_{GAP5}[Rab5^{GTP}_{Mem}] - k_{(A5')}[Rab5^{GTP}_{Mem}][E] + k_{(D5')}[Rab5^{GTP}_{Mem} : E] \tag{5.33}$$

$$\frac{d[Rab5^{GTP}_{Mem} : E]}{dt} = k_{(A5')}[Rab5^{GTP}_{Mem}][E] - k_{(D5')}[Rab5^{GTP}_{Mem} : E] \tag{5.34}$$

$$\frac{d[Rab11^{GDP}_{Cyt}]}{dt} = k_{out11}[Rab11^{GDP}_{Mem}] - k_{in11}[Rab11^{GDP}_{Cyt}] \tag{5.35}$$

$$\frac{d[Rab11^{GDP}_{Mem}]}{dt} = k_{in11}[Rab11^{GDP}_{Cyt}] - k_{out11}[Rab11^{GDP}_{Mem}] + k_{GAP11}[Rab11^{GTP}_{Mem}] - k_{GEF11}[Rab11^{GDP}_{Mem}] \tag{5.36}$$

$$\frac{d[Rab11^{GTP}_{Mem}]}{dt} = k_{GEF11}[Rab11^{GDP}_{Mem}] - k_{GAP11}[Rab11^{GTP}_{Mem}] \tag{5.37}$$

$$\frac{d[E]}{dt} = k_{D5'}[Rab5^{GTP}_{Mem} : E] - k_{A5'}[Rab5^{GTP}_{Mem}][E] \tag{5.38}$$

$$\tag{5.39}$$

Note that in this system only Rab5 is allowed to bind to the effector, thus reactions of Rab11 binding and unbinding from Zfyve26 are not considered and the Rab11-Zfyve26 complex does not appear in the equations.
Similar to other models, the parameter k_{GEF5} is made dependent on the amount of $Rab5^{GTP} : Zfyve26$ molecular complex as a Hill function, to implement the assumed activatory self-loop mediated by the effector:

$$k_{GEF5} = k_{GEF5b} + k_{GEF5MAX}\left(\frac{(Rab5^{GTP}_{Mem} : E)^n}{((Rab5^{GTP}_{Mem} : E)^n + k^n_{GEF5M})}\right) \tag{5.40}$$

whereas the Rab5 deactivation rate k_{GAP5} is constant:

$$k_{GAP5} = k_{GAP5b} \tag{5.41}$$

here k_{GEF5b} and k_{GAP5b} respectively represent constant basal rates of GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GEF5MAX}$ is the maximal rate of Rab5 activation mediated by the effector, approached as the level of Rab5-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GEF5M} determines the number of complex molecules required to reach half of the maximum rate $k_{GEF5MAX}$.

However, opposed to the four competitive models, the Rab5GTP:Zfyve26 complex does not only trigger Rab5 activation but also causes Rab11 deactivation, making also k_{GAP11} a Hill function dependent on the same complex $Rab5^{GTP} : Zfyve26$:

$$k_{GAP11} = k_{GAP11b} + k_{GAP11MAX}\left(\frac{(Rab5^{GTP}_{Mem} : E)^n}{((Rab5^{GTP}_{Mem} : E)^n + k^n_{GAP11M})}\right) \tag{5.42}$$

whereas the Rab11 activation rate k_{GEF11} is constant and equal to its basal value:

$$k_{GEF11} = k_{GEF11b} \tag{5.43}$$

where k_{GEF11b} and k_{GAP11b} respectively represent constant basal rates of GDP/GTP exchange and hydrolysis, whereas the additive terms represent Hill-like functions of the amount of complex. $k_{GAP11MAX}$ is the maximal rate of Rab11 deactivation mediated by the effector, approached as the level of Rab5-Zfyve26 complex increases. The Michaelis-Menten constant, k_{GAP11M} determines the number of complex molecules required to reach half of the maximum rate $k_{GAP11MAX}$.

5.3 Model parameter estimation

Baseline parameter values were chosen in agreement with orders of magnitude reported in literature. The basal Rab5 GEF activity rate (i.e. the rate of Rab5 activation

by its GEF) was found to be of the order $10^{-2}s^{-1}$ in vitro [108]. We thus set the basal Rab5 activation rate, k_{GEF5b}, to $10^{-2}s^{-1}$. Due to the absence of similar data on Rab11, the basal Rab11 activation rate was set to an identical value for symmetry: $k_{GEF11b} = 10^{-2}s^{-1}$.
Since rate constants associated to Zfyve26-mediated Rab activation are unknown, the maximum value of the Hill function $k_{GEF5MAX}$ was set to the same value of the basal rate of GDP/GTP exchange, that is to $10^{-2}s^{-1}$. In the case of models implying positive modulation of Rab11 activity by Zfyve26, the corresponding maximum rate $k_{GEF11MAX}$ was set to a value identical to $k_{GEF5MAX}$. A value of $10^{-2}s^{-1}$ was measured for Rab5 basal hydrolysis rate [109]. We thereby set k_{GAP5b} to $10^{-2}s^{-1}$ and we used an identical value for Rab11, k_{GAP11b}=$10^{-2}s^{-1}$. For models implying negative regulation of Rab11 activity by Zfyve26, we set the maximum of the Hill function $k_{GAP11MAX}$, to the same value of the basal hydrolysis rate, that is to $10^{-2}s^{-1}$.
Remaining parameters relative to Hill functions were set by choosing realistic orders of magnitude: Michaelis-Menten constants of Hill functions, representing half of the amount of Rab-effector complex required to reach maximum feedback rates, were all set to the same arbitrary value of 50 molecules. Exponents of activatory Hill functions were set to $n = 2$ in order to account for cooperativity, as it has been proposed that the recruitment of GEFs by Rab effectors generates cooperative Rab activation [110]. For symmetry, exponents of Hill functions describing k_{GAPMAX} rates were set to the same value.
Rates of Rab binding and unbinding from the effector Zfyve26 were arbitrarily set to $10^{-}3$ for both Rab5 and Rab11.
The rate of Rab5 membrane binding was measured to be of the order of $10^{-2}s^{-1}$ in vitro [111]. We thereby set both parameters k_{in5} and k_{in11} to $10^{-2}s^{-1}$. Since precise kinetic measurements of GDI-mediated Rab membrane extraction are lacking, we set rates k_{out5} and k_{out11} to the same baseline value of $10^{-2}s^{-1}$.

5.4 Model screening for Rab5 and Rab11 mean values

For each model, we address the Zfyve26-dependent behaviour of Rab5 and Rab11 mean numbers of molecules at the steady state. In particular, because we are interested in observables susceptible of experimental validation, we formulate qualitative

predictions on two variables of the system that can be assessed experimentally: the total number of Rab molecules on the membrane and the number of active molecules for both Rab5 and Rab11.
Note that for the single Rab species, the numer of active Rab molecules is given by the sum of free active Rabs, i.e. Rab^{GTP}_{Mem}, and - if the model implies that the active Rab forms a complex with Zfyve26 - active Rabs in complex with the effector, i.e. $Rab^{GTP}_{Mem} : Zfyve26$:

$$Rab^{GTP} = Rab^{GTP}_{Mem} + Rab^{GTP}_{Mem} : Zfyve26 \tag{5.44}$$

whereas the total number of molecules on the membrane is given by the sum of free inactive Rabs (i.e. Rab^{GDP}_{Mem}), free active Rabs (i.e. Rab^{GTP}_{Mem}), and either inactive or active Rabs in complex with the effector depending on the kind of complex implied by the model (i.e. either $Rab^{GDP}_{Mem} : Zfyve26$ or $Rab^{GTP}_{Mem} : Zfyve26$):

$$Rab_{Mem} = Rab^{GDP}_{Mem} + Rab^{GTP}_{Mem} + Rab^{GTP}_{Mem} : Zfyve26 \tag{5.45}$$

or

$$Rab_{Mem} = Rab^{GDP}_{Mem} + Rab^{GTP}_{Mem} + Rab^{GDP}_{Mem} : Zfyve26 \tag{5.46}$$

The total number of molecules for each GTPase, which is conserved in our system, is the sum of all forms - inactive and active, both free and in complex with the effector:

$$Rab_{TOT} = Rab^{GDP}_{Cyt} + Rab^{GDP}_{Mem} + Rab^{GTP}_{Mem} + Rab^{GTP}_{Mem} : Zfyve26 \tag{5.47}$$

or

$$Rab_{TOT} = Rab^{GDP}_{Cyt} + Rab^{GDP}_{Mem} + Rab^{GTP}_{Mem} + Rab^{GDP}_{Mem} : Zfyve26 \tag{5.48}$$

5.4.1 Mean Rab5 and Rab11 values as functions of Zfyve26 and kGEF/kGAP parameters

To compute GTPase mean values as functions of the effector amount, we solved each model's ODE system at the steady state with baseline parameter values, using identical total amounts of Rab5 and Rab11 ($Rab5_{TOT} = 200$ molecules and $Rab11_{TOT} = 200$ molecules) and different initial Zfyve26 values ranging from 0 to 400 molecules. We thus computed the resulting amount of active Rab, i.e Rab^{GTP}, and the total amount of Rab on the membrane, i.e. Rab_{Mem}, for both Rab5 and Rab11 as described above.
To account for different strengths of Zfyve26 action, leading to weaker or stronger activatory/inhibitory action of the effector, we computed the same numbers of molecules using different maximum values of the Hill functions describing *kGEF* and *kGAP* rates. To this purpose, all parameters were kept fixed to baseline values except for $k_{GEF5MAX}$ - the maximum Zfyve26-mediated Rab5 activation rate - and either $k_{GEF11MAX}$ or $k_{GAP11MAX}$, depending on whether the model implies an activatory or inhibitory Zfyve26 action on Rab11. The varying parameter was assigned values 10^{-3}, 10^{-2}, 10^{-1}, 10^{0}.

We found that all models (ComI-IV and Cas) predict an increase of $Rab5^{GTP}$ as a function of Zfyve26 regardless of k_{GEFMAX} and k_{GAPMAX} parameter values (Figures 5.5, 5.6, 5.7, 5.8, 5.9, left panels), consistent with the presence of an activatory Rab5 self-loop mediated by Zfyve26. Similarly, all models generate an increasing total Rab5 amount on the membrane, $Rab5_{Mem}$, although less significant (Figures 5.5, 5.6, 5.7, 5.8, 5.9, right panels).
For all models, the steepness of increase of both $Rab5^{GTP}$ and $Rab5_{Mem}$ resulted mostly sensitive to changes in $k_{GEF5MAX}$ and less in $k_{GEF11MAX}$ or $k_{GAP11MAX}$. In particular, higher $k_{GEF5MAX}$ values resulted in steeper $Rab5_{Mem}$ and $Rab5^{GTP}$ increases, coherently with the strenghtened activatory self-loop.
On the contrary, the distinct models generated different Zfyve26-dependent predictions for the amounts of active Rab11 molecules and the total number of Rab11 molecules on the membrane.
In particular, models ComI and ComII both predict a $Rab11^{GTP}$ rise for increasing number of Zfyve26 molecules, and analogous results are obtained for $Rab11_{Mem}$ (Figures 5.5, 5.6). Moreover, for model ComI this rise results steeper for higher values of $k_{GEF11MAX}$, consistently with a stronger Rab11 activation mediated by

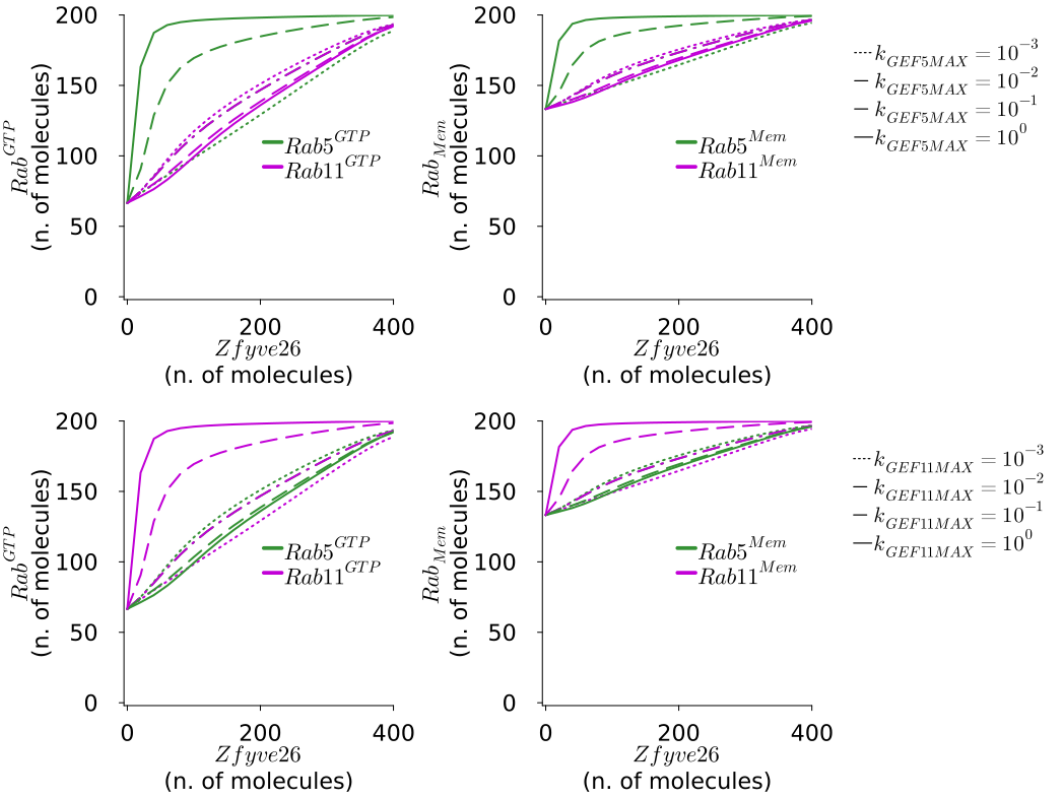


Fig. 5.5 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComI. Top panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GEF5MAX}$. Lines with different dash styles are referred to different values of the parameter. Bottom panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GEF11MAX}$. Lines with different dash styles are referred to different values of the parameter.

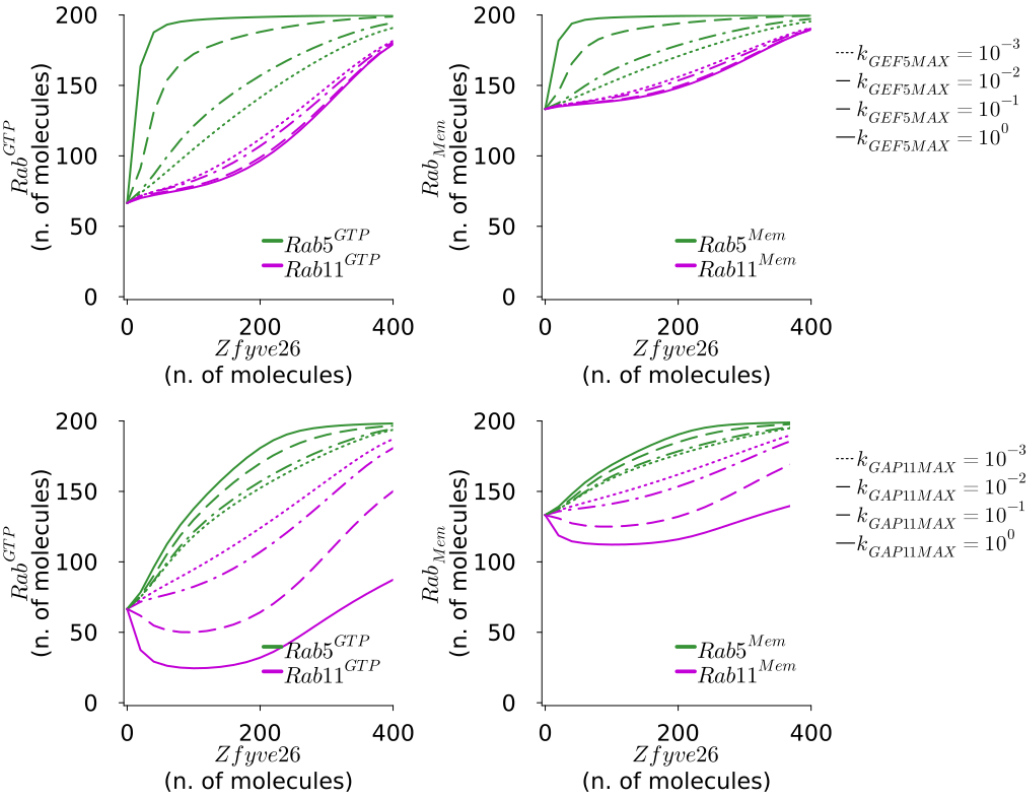


Fig. 5.6 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComII. Top panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GEF5MAX}$. Lines with different dash styles are referred to different values of the parameter. Bottom panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GAP11MAX}$. Lines with different dash styles are referred to different values of the parameter.

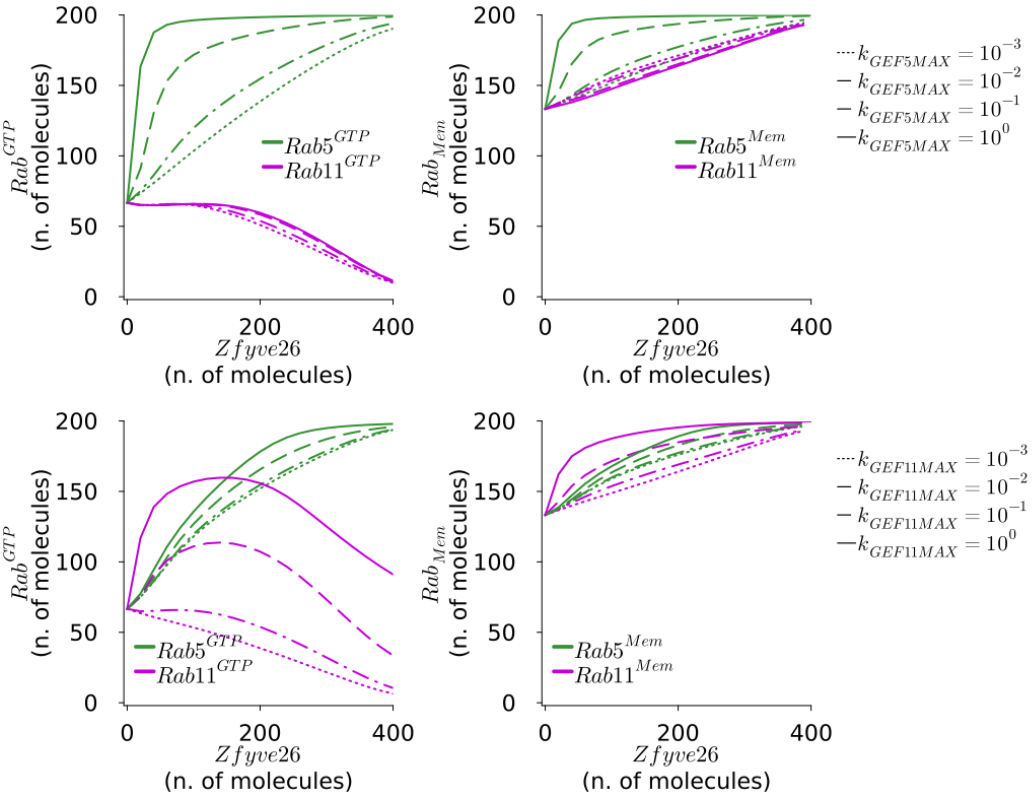


Fig. 5.7 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComIII. Top panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GEF5MAX}$. Lines with different dash styles are referred to different values of the parameter. Bottom panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GEF11MAX}$. Lines with different dash styles are referred to different values of the parameter.

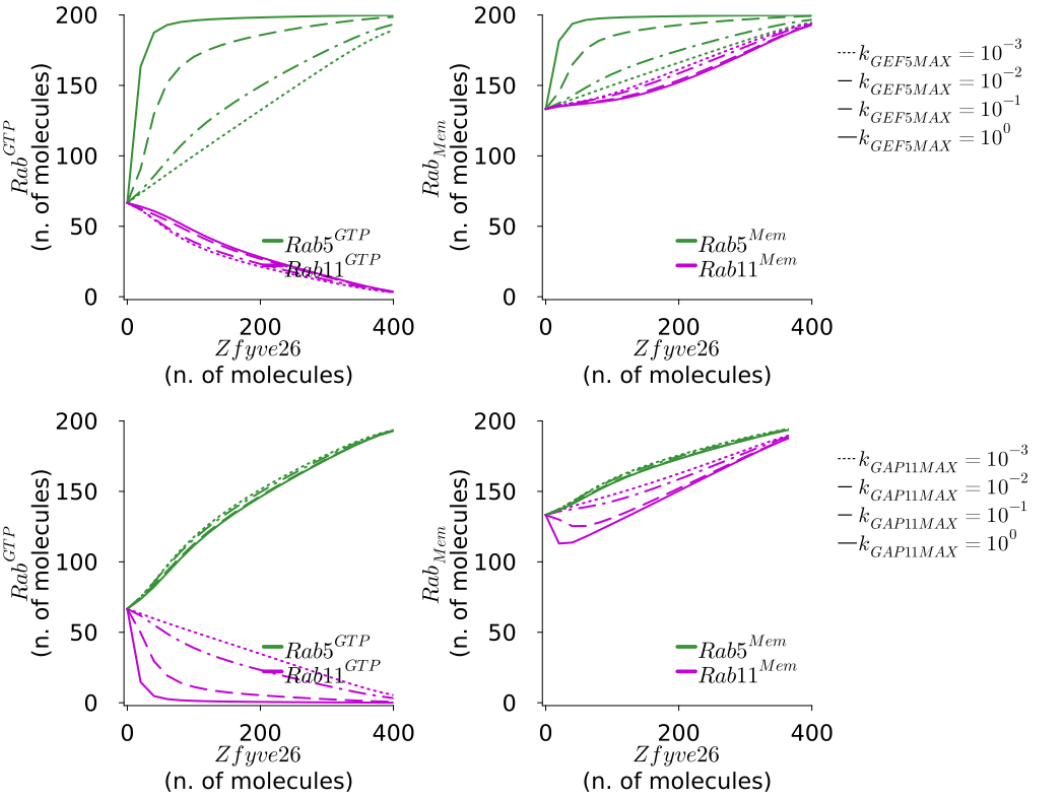


Fig. 5.8 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComIV. Top panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GEF5MAX}$. Lines with different dash styles are referred to different values of the parameter. Bottom panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GAP11MAX}$. Lines with different dash styles are referred to different values of the parameter.

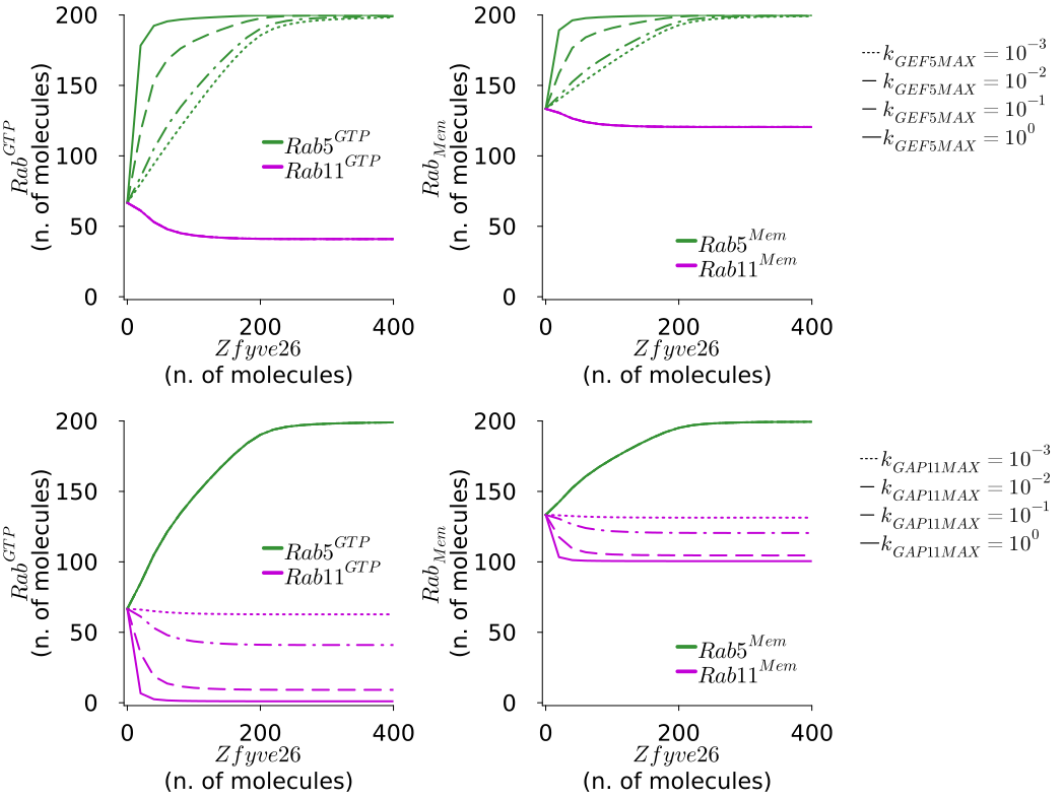


Fig. 5.9 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model Cas. Top panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GEF5MAX}$. Lines with different dash styles are referred to different values of the parameter. Bottom panels represent $Rab5^{GTP}$ and $Rab11^{GTP}$ (left), $Rab5_{Mem}$ and $Rab11_{Mem}$ (right) as functions of Zfyve26, obtained for different values of $k_{GAP11MAX}$. Lines with different dash styles are referred to different values of the parameter.

the effector and a consequently enhanced Rab11 accumulation on the membrane. Conversely, milder changes of steepness were obtained by varying $k_{GEF5MAX}$.
Note instead that the Zfyve26-dependent $Rab11^{GTP}$ increasing trend predicted by model ComII is lost for the highest considered $k_{GAP11MAX}$ value (Figure 5.6, bottom left panel). However, such value of $k_{GAP11MAX}$ falls beyond realistic rates of hydrolysis and thus results as biologically unlikely [].
Opposite to the first two models, models Cas, ComIII and ComIV generate a decrease of $Rab11^{GTP}$ as a function of the number of Zfyve26 molecules (Figures 5.7, 5.8, 5.9). In particular, in model ComIV, $Rab11^{GTP}$ displays a steeper decrease for higher values of $k_{GAP11MAX}$, coherently with a stronger Zfyve26-mediated deactivation, whereas showing a weaker response to changes in $k_{GEF5MAX}$. Instead, note that the Zfyve26-dependent $Rab11^{GTP}$ decreasing trend predicted by model ComIII is disrupted for the highest adopted value of $k_{GEF11MAX}$ (10^0)(Figure 5.7, bottom left panel). However, such value of $k_{GEF11MAX}$ falls beyond realistic rates of GDP/GTP exchange and thus results as biologically unlikely [108, 109].
Only model Cas predicts a reduction of $Rab11_{Mem}$ as a function of Zfyve26, whereas the two competitive models ComIII and ComIV show instead a net increase of $Rab11_{Mem}$. For all three models, the $Rab11_{Mem}$ growth/reduction resulted steeper for higher values of $k_{GAP11MAX}$ and mildly affected by changes in $k_{GEF5MAX}$.
In conclusion, we observed that all models predict an increase in Rab5 activity and abundance as functions of the amount of effector. Rab11 abundance and activity are instead predicted to either decrease or increase depending on the hypothesized interaction with the effector: all competitive models predict a Zfyve26-dependent growth of Rab11 abundance opposed to the Cas model, which generates a decreasing $Rab11_{Mem}$ amount; whereas Rab11 activity is predicted to increase with Zfyve26 by models ComI and ComII and to decrease by models ComIII, ComIV and Cas.

5.4.2 Mean Rab5 and Rab11 values as functions of Zfyve26 and the relative Rab5/Rab11 pool size

Because all models are described by coupled ODEs where Rab5 and Rab11 are connected either direcly (Cas model) or indirectly (ComI-IV models), we expect reciprocal Rab5 and Rab11 amounts to affect the Zfyve26-dependent behaviour of the two proteins. Moreover, since GTPase amounts on the endosomal membrane can vary significantly, the robustness of Zfyve26-dependent qualitative predictions to

changes in the relative amount of the two competing GTPases needs to be verified. We thus computed mean values of Rab5 and Rab11 as functions of both the number of effector molecules and the ratio between Rab5 and Rab11 molecule numbers fed to the system.

To this purpose, we used baseline parameter values, Zfyve26 values ranging from 0 to 400 molecules, and a total pool of 400 Rab molecules that was partitioned into progressively increasing $Rab5_{TOT}/Rab11_{TOT}$ ratios ranging from 10^{-1} to 10^{1}. We computed total Rab5 and Rab11 membrane amounts, $Rab5_{Mem}$ and $Rab11_{Mem}$, and $Rab5^{GTP}$ and $Rab11^{GTP}$ amounts as functions of both the number of effector molecules and the $Rab5_{TOT}/Rab11_{TOT}$ ratio.

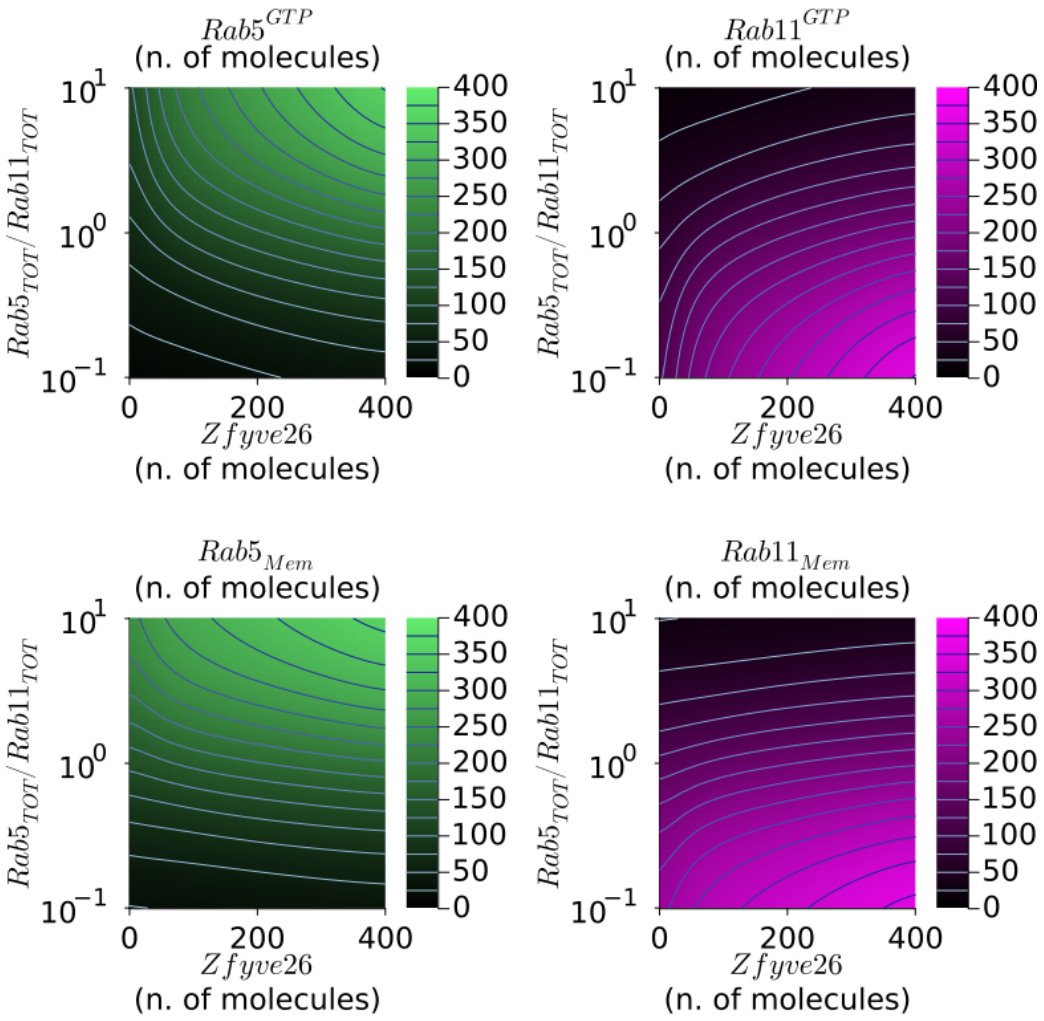


Fig. 5.10 Heatmaps of $Rab5^{GTP}$ (top left), $Rab11^{GTP}$ (top right), $Rab5_{Mem}$ (left) and $Rab11_{Mem}$ as functions of Zfyve26 and the relative GTPase amount $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComI. Curves represent points of constant molecule level.

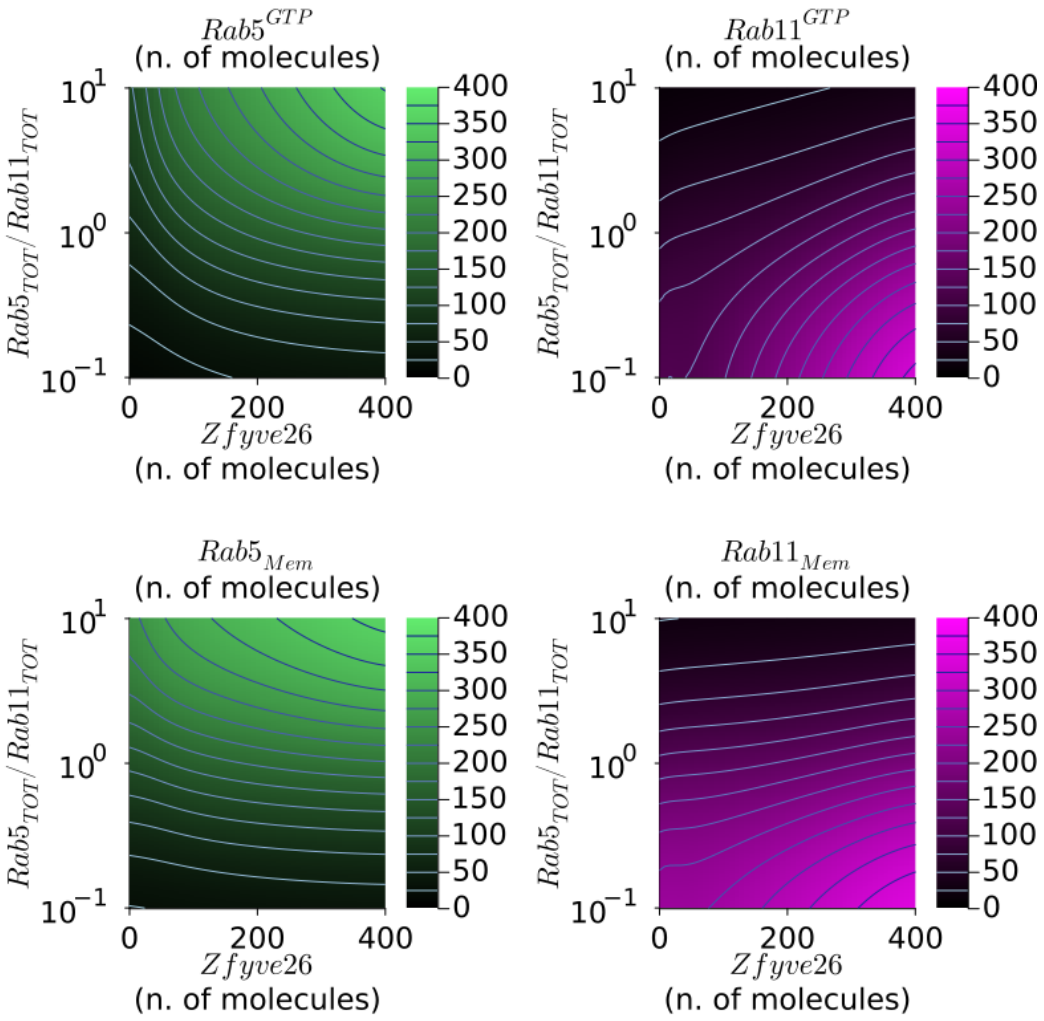


Fig. 5.11 Heatmaps of $Rab5^{GTP}$ (top left), $Rab11^{GTP}$ (top right), $Rab5_{Mem}$ (left) and $Rab11_{Mem}$ as functions of Zfyve26 and the relative GTPase amount $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComII. Curves represent points of constant molecule level.

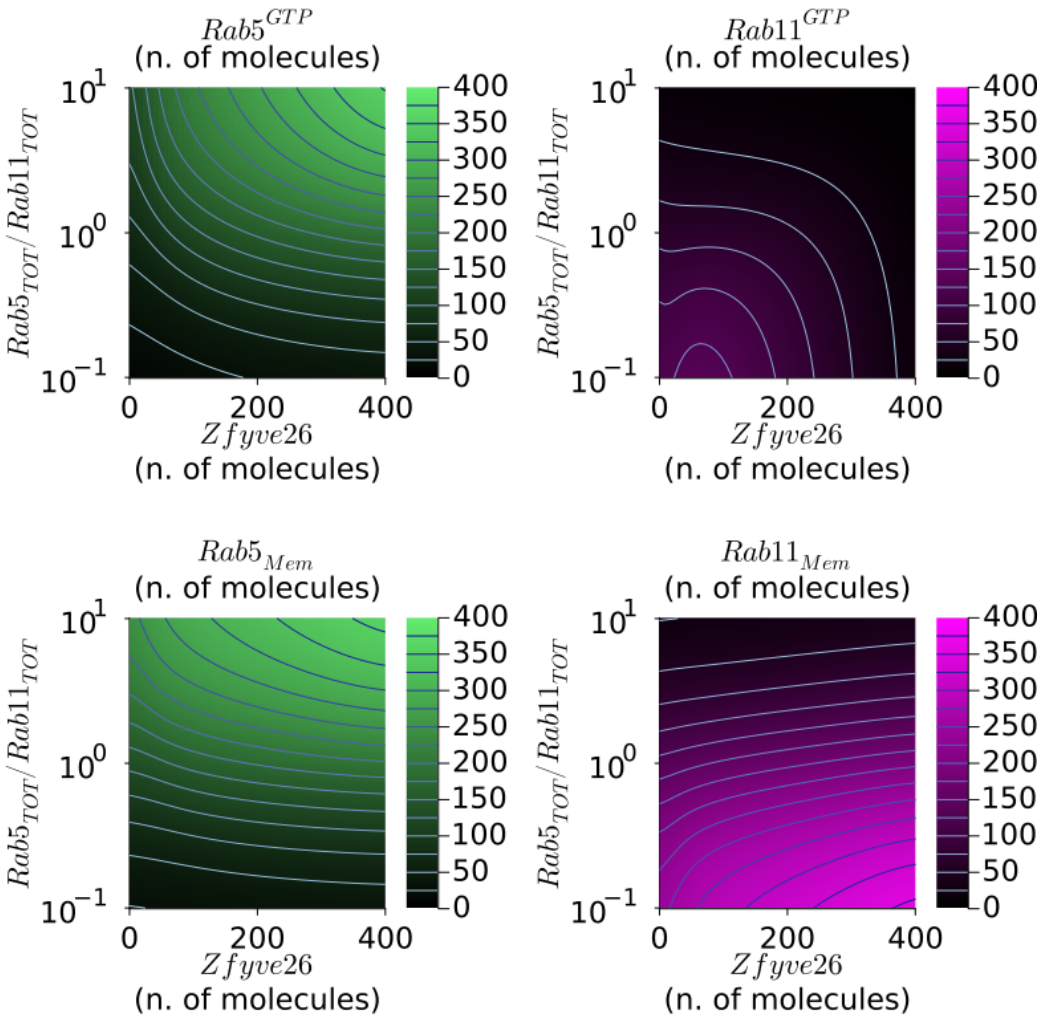


Fig. 5.12 Heatmaps of $Rab5^{GTP}$ (top left), $Rab11^{GTP}$ (top right), $Rab5_{Mem}$ (left) and $Rab11_{Mem}$ as functions of Zfyve26 and the relative GTPase amount $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComIII. Curves represent points of constant molecule level.

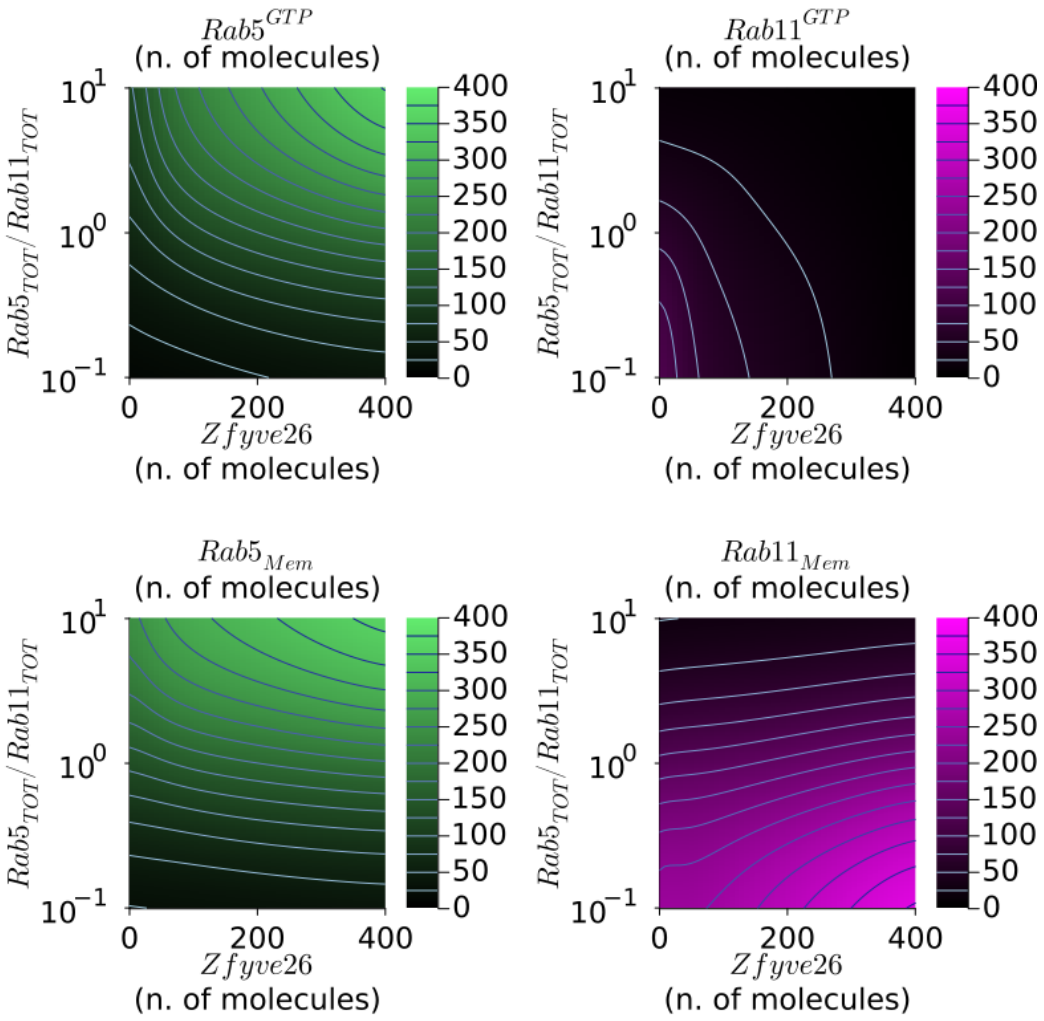


Fig. 5.13 Heatmaps of $Rab5^{GTP}$ (top left), $Rab11^{GTP}$ (top right), $Rab5_{Mem}$ (left) and $Rab11_{Mem}$ as functions of Zfyve26 and the relative GTPase amount $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComIV. Curves represent points of constant molecule level.

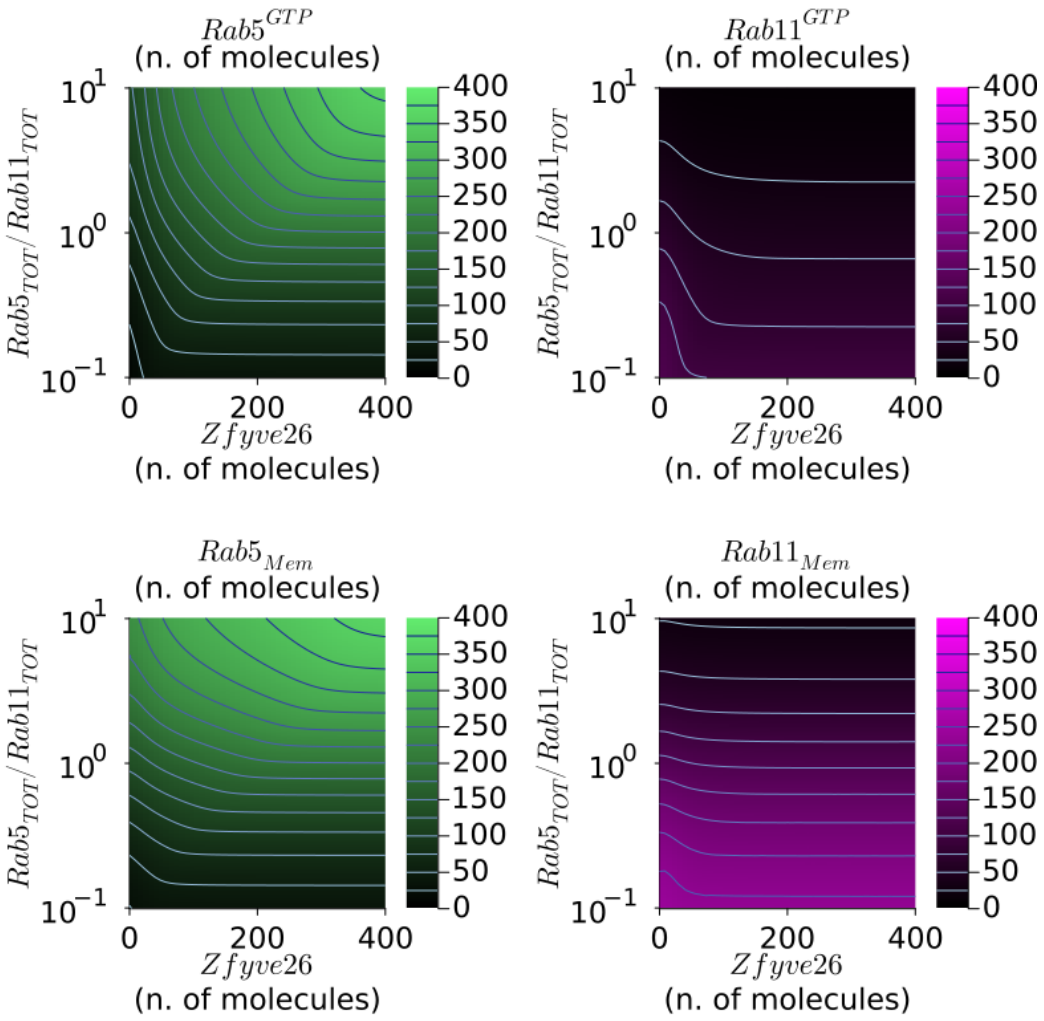


Fig. 5.14 Heatmaps of $Rab5^{GTP}$ (top left), $Rab11^{GTP}$ (top right), $Rab5_{Mem}$ (left) and $Rab11_{Mem}$ as functions of Zfyve26 and the relative GTPase amount $Rab5_{TOT}/Rab11_{TOT}$ predicted by model Cas. Curves represent points of constant molecule level.

For all models, computed predictions show that, for any fixed Zfyve26 amount, increasing the total number of molecules of a GTPase over the other leads it to both increased activity and increased abundance on the membrane (Figures 5.10, 5.11, 5.12, 5.13, 5.14). Take for instance model ComI: Rab5 activity and abundance result higher if Rab5 is numerically prevalent over Rab11, whereas the opposite occurs if conversely Rab11 prevails over Rab5 (Figure 5.10). Similar observations hold for the remaining models, both competitive (ComII-IV) and non-competitive (Cas). Overall, qualitative Zfyve26-dependent predictions are not disrupted by changes in

the relative Rab5-Rab11 amount present in the system.
For further clarification, we report Rab^{GTP} and Rab_{Mem} curves for both Rab5 and Rab11 as functions of Zfyve26 obtained for three different values of $Rab5_{TOT}/Rab11_{TOT}$ (0.3, 1.0, 3.0) picked from the explored range (Figures 5.15, 5.16, 5.17, 5.18, 5.19). For each model, the three curves display concordant qualitative trends as functions of the effector, thus showing that Zfyve26-dependent predictions are qualitatively maintained when Rab5 and Rab11 amounts are unbalanced.

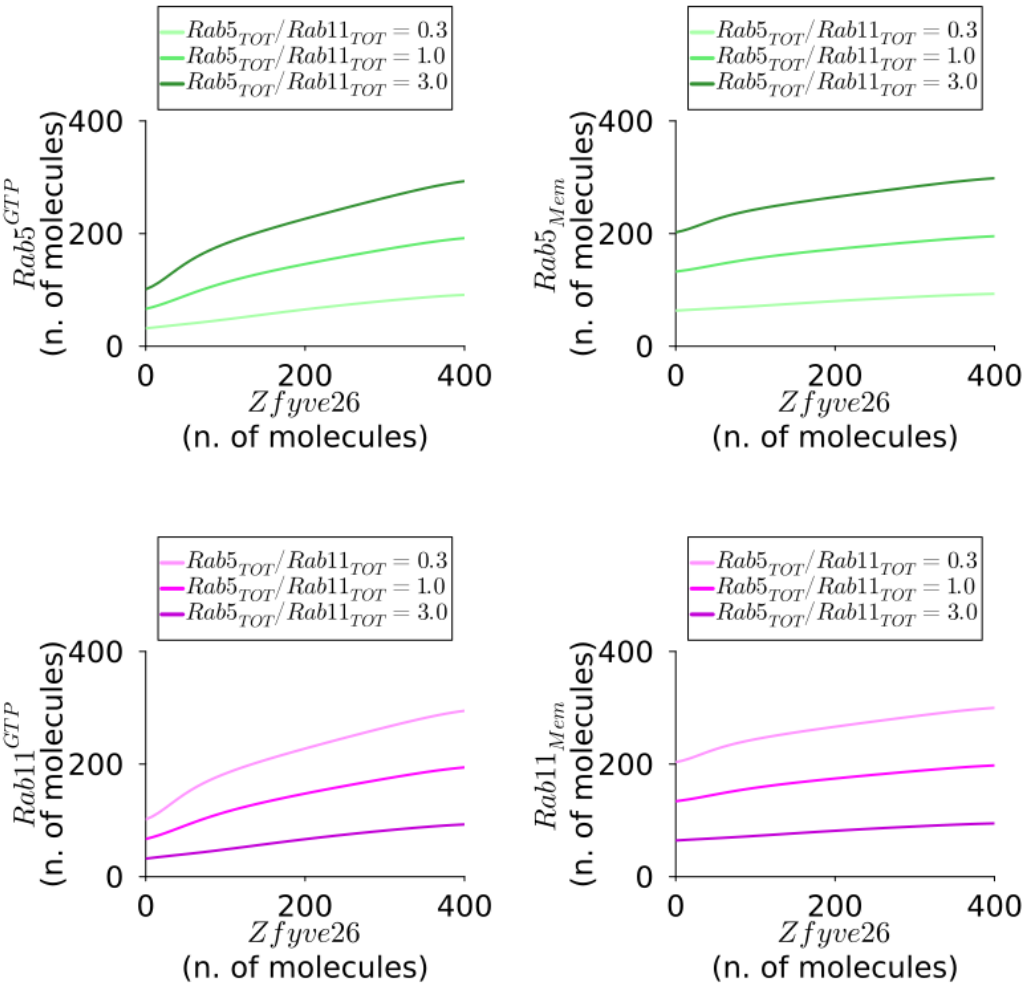


Fig. 5.15 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComI. Top panels represent $Rab5^{GTP}$ (left) and $Rab5_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of green correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$. Bottom panels represent $Rab11^{GTP}$ (left) and $Rab11_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of purple correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$.

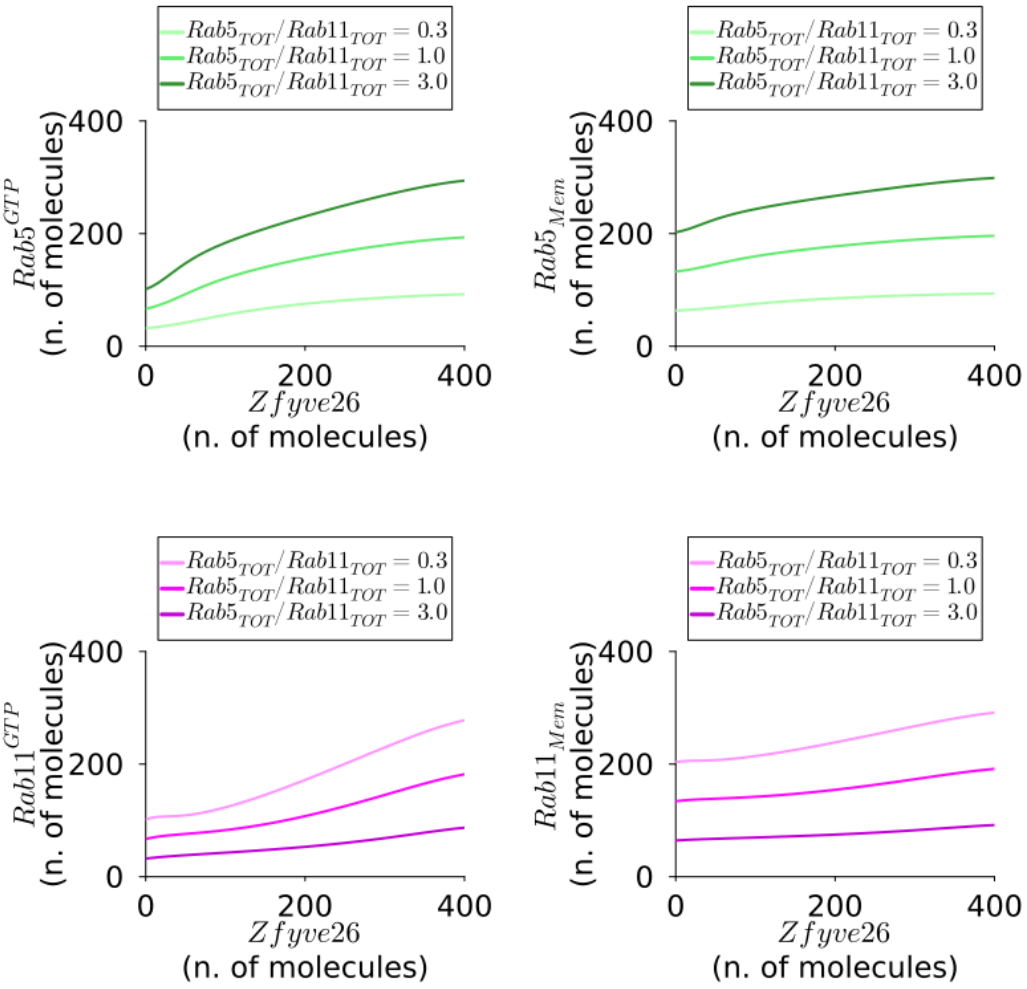


Fig. 5.16 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComII. Top panels represent $Rab5^{GTP}$ (left) and $Rab5_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of green correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$. Bottom panels represent $Rab11^{GTP}$ (left) and $Rab11_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of purple correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$.

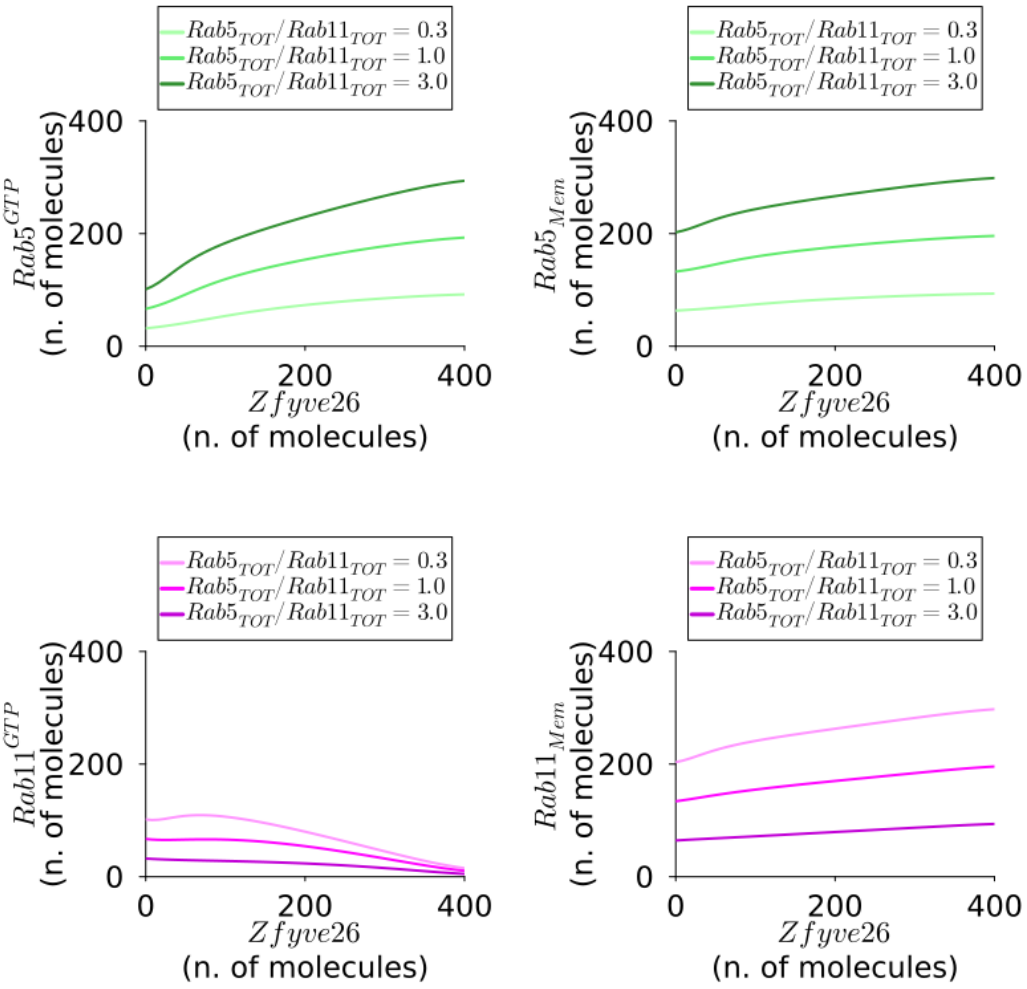


Fig. 5.17 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComIII. Top panels represent $Rab5^{GTP}$ (left) and $Rab5_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of green correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$. Bottom panels represent $Rab11^{GTP}$ (left) and $Rab11_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of purple correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$.

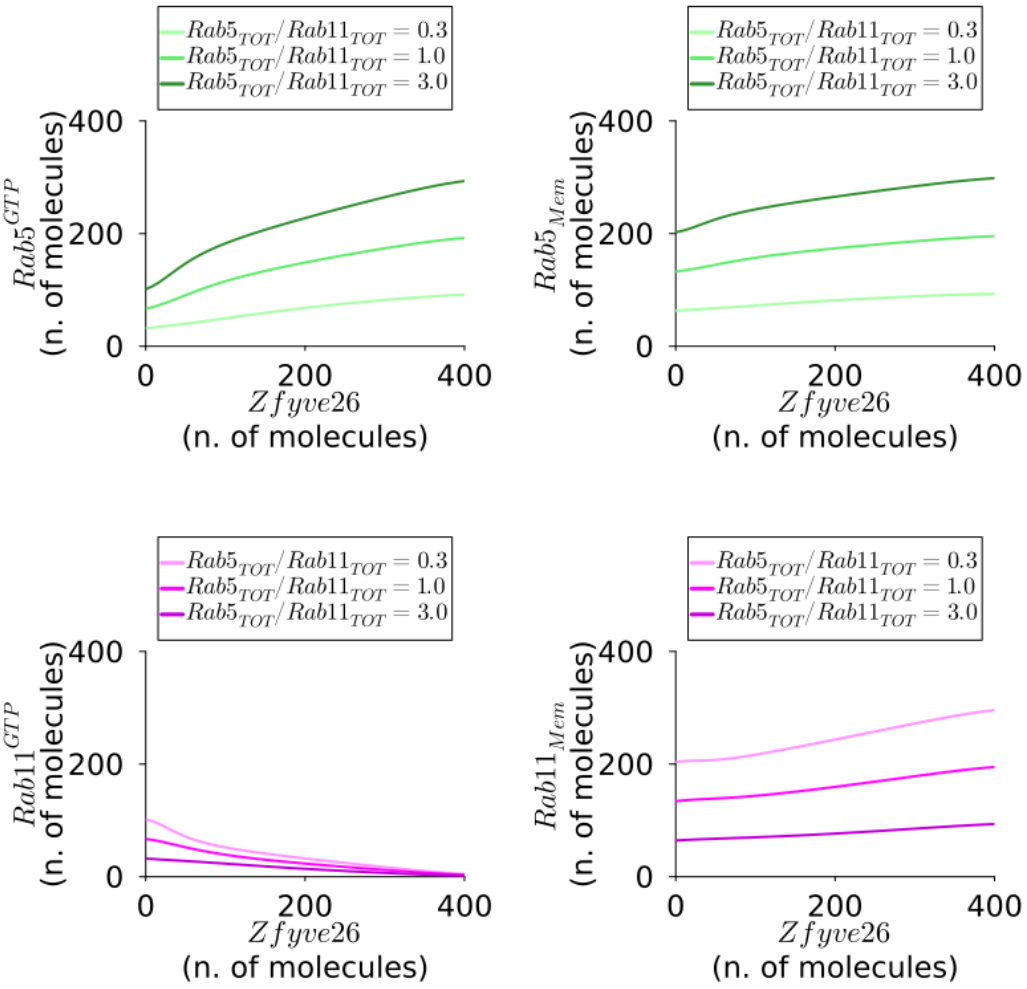


Fig. 5.18 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model ComIV. Top panels represent $Rab5^{GTP}$ (left) and $Rab5_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of green correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$. Bottom panels represent $Rab11^{GTP}$ (left) and $Rab11_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of purple correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$.

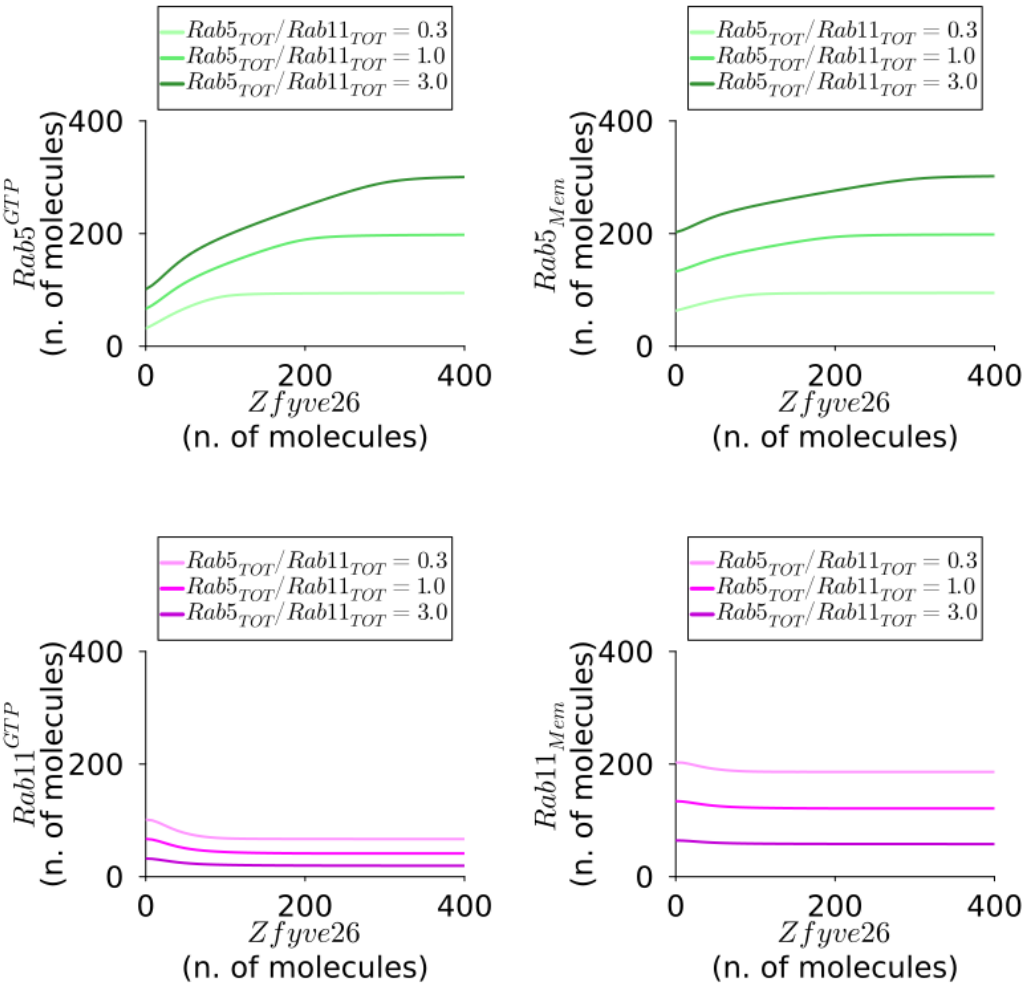


Fig. 5.19 Rab5 and Rab11 membrane amounts and active amounts as functions of Zfyve26 predicted by model Cas. Top panels represent $Rab5^{GTP}$ (left) and $Rab5_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of green correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$. Bottom panels represent $Rab11^{GTP}$ (left) and $Rab11_{Mem}$ (right) as functions of Zfyve26. Curves with different shades of purple correspond to different values of $Rab5_{TOT}/Rab11_{TOT}$.

5.5 Model screening for Rab5 and Rab11 correlations

Since some of our models generated qualitatively similar predictions for Rab5 and Rab11 mean values, we addressed mutual fluctuations of the two proteins as functions of Zfyve26 using stochastic simulations. As illustrated in the introduction, this

choice was motivated by previous theoretical evidence suggesting that the amounts of molecules competing for a shared resource can be correlated, and that such correlations depend on the total amount of resource available.
Consistent with our experimental constraints (see Chapter 6), we addressed correlations susceptible of live-cell measurements, i.e. steady state correlations between the following pairs of variables describing either the membrane abundance or the activity of one of the two Rabs:

- $\rho(Rab5_{Mem}, Rab11_{Mem})$
- $\rho(Rab11_{Mem}, Rab5^{GTP})$
- $\rho(Rab5_{Mem}, Rab5^{GTP})$
- $\rho(Rab11_{Mem}, Rab11^{GTP})$

To assess the dependence of such correlations on both the effector amount and the relative amount of the two proteins, we computed them as functions of both Zfyve26 and the ratio between amounts of the two Rabs present in the system, i.e. $Rab5_{TOT}/Rab11_{TOT}$.

5.5.1 Rab5-Rab11 correlations as functions of Zfyve26 and relative Rab5/Rab11 pool size

We performed stochastic simulations of each model using the Gillespie algorithm, for different Zfyve26 values ranging from 0 to 400 molecules , with a total GTPase pool fixed to 400 molecules and $Rab5_{TOT}/Rab11_{TOT}$ ranging from 10^{-1} to 10^{1}. To account for the potential variability of Zfyve26 amount on the endosomal membrane deriving from extrinsic factors, we added noise on the number of effector molecules fed to the system. To this purpose, at each simulation we drew the initial value of Zfyve26 from a Gaussian distribution centered on its mean value (coefficient of variation = 0.05). For each simulation, the number of Rab proteins on the membrane (Rab_{Mem}) and the number of active Rab proteins (Rab^{GTP}) were collected at the steady state for both Rab5 and Rab11. Thus, after performing 10^4 simulations for each Zfyve26 mean value, we computed correlations between the pairs of variables defined above, as:

$$\rho(a,b) = \frac{\sum_{i=1}^{n}(a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a})^2}\sqrt{\sum_{i=1}^{n}(b_i - \bar{b})^2}} \tag{5.49}$$

where a and b are the two variables representing either Rab abundances or activities, i identifies the simulation, and n indicates the total number of simulations performed.

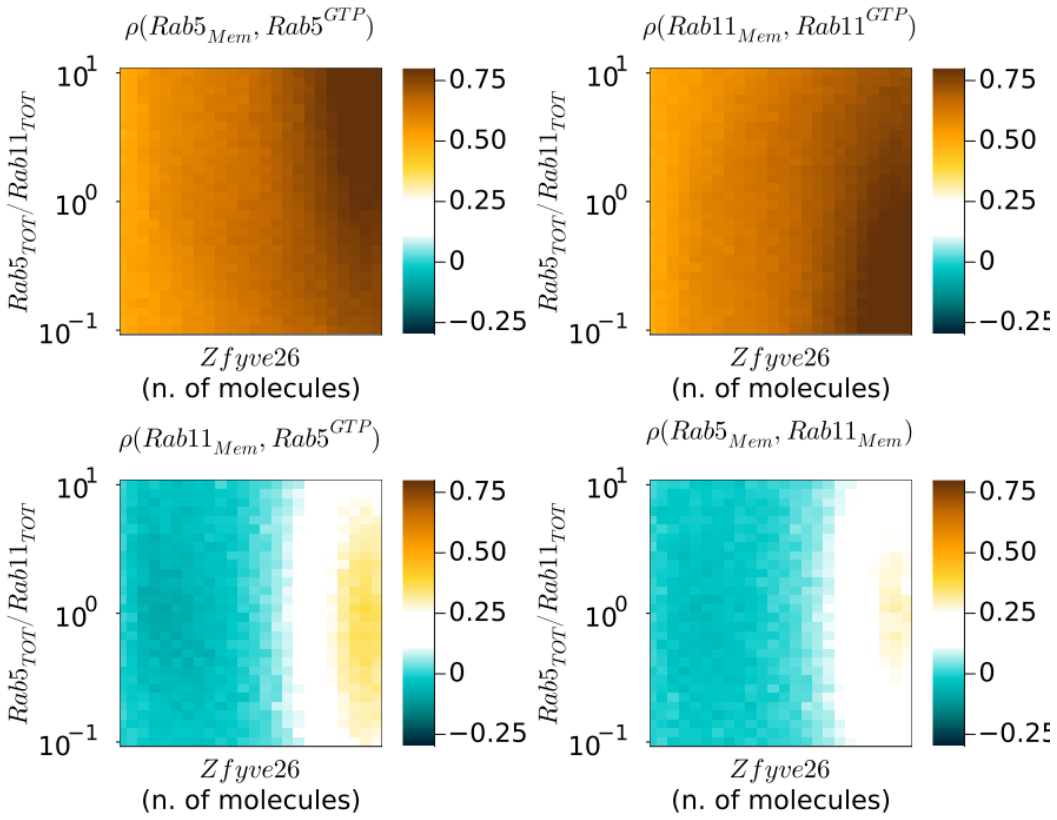


Fig. 5.20 Correlations between Rab membrane abundances and activities as functions of Zfyve26 and $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComI: $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (top left panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (top right panel), $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (bottom right panel).

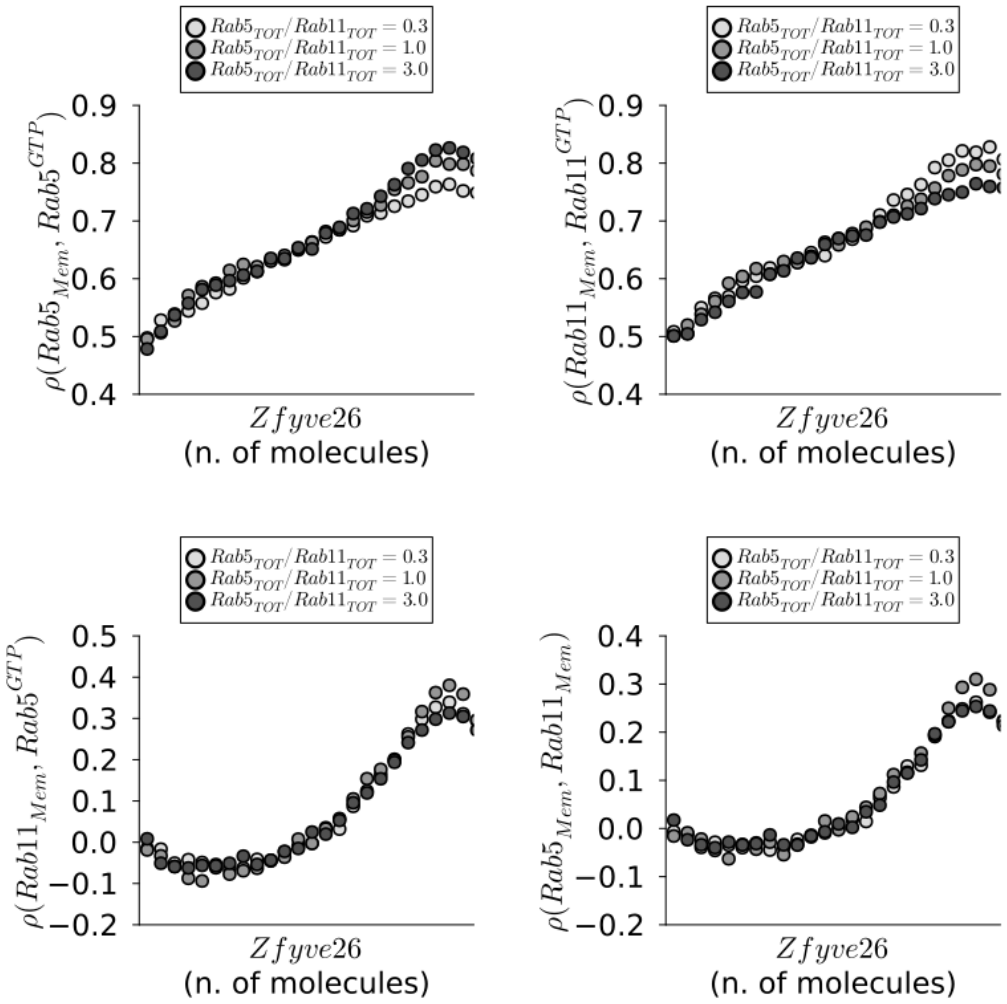


Fig. 5.21 Correlations between Rab membrane abundances and activities as functions of Zfyve26, for three different values of $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComI: $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (top left panel), $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (top right panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (bottom right panel).

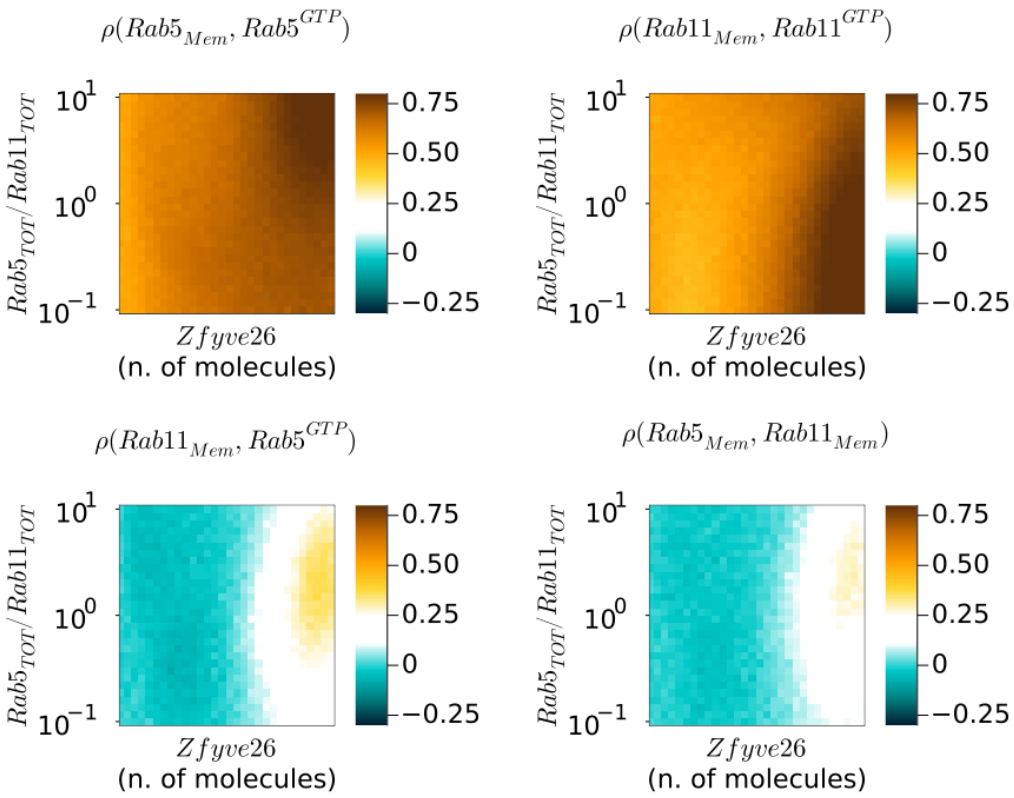


Fig. 5.22 Correlations between Rab membrane abundances and activities as functions of Zfyve26 and $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComII: $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (top left panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (top right panel), $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (bottom right panel).

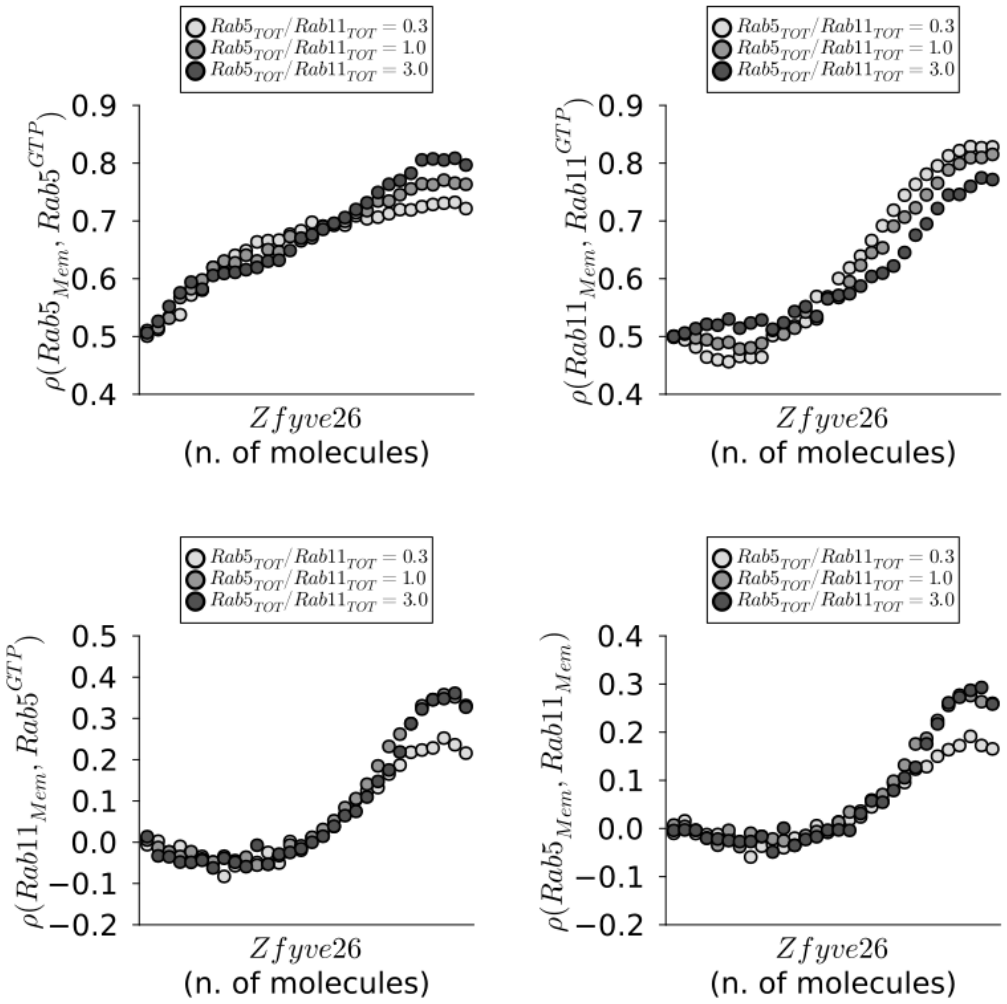


Fig. 5.23 Correlations between Rab membrane abundances and activities as functions of Zfyve26, for three different values of $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComII: $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (top left panel), $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (top right panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (bottom right panel).

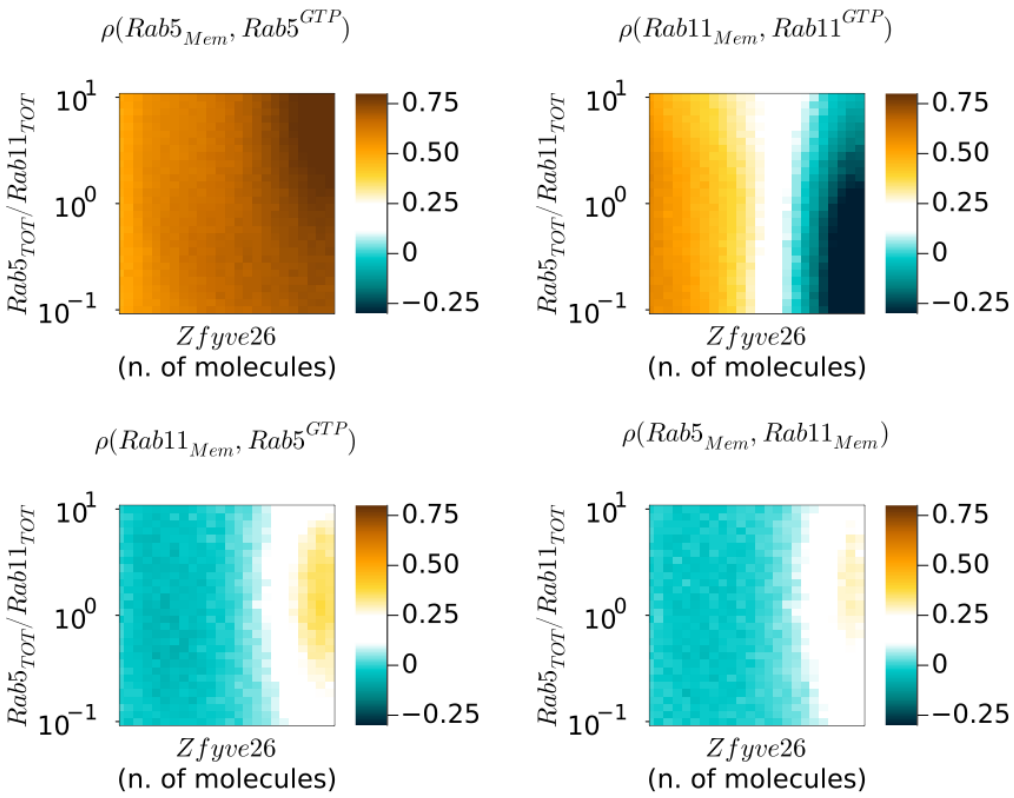


Fig. 5.24 Correlations between Rab membrane abundances and activities as functions of Zfyve26 and $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComIII: $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (top left panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (top right panel), $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (bottom right panel).

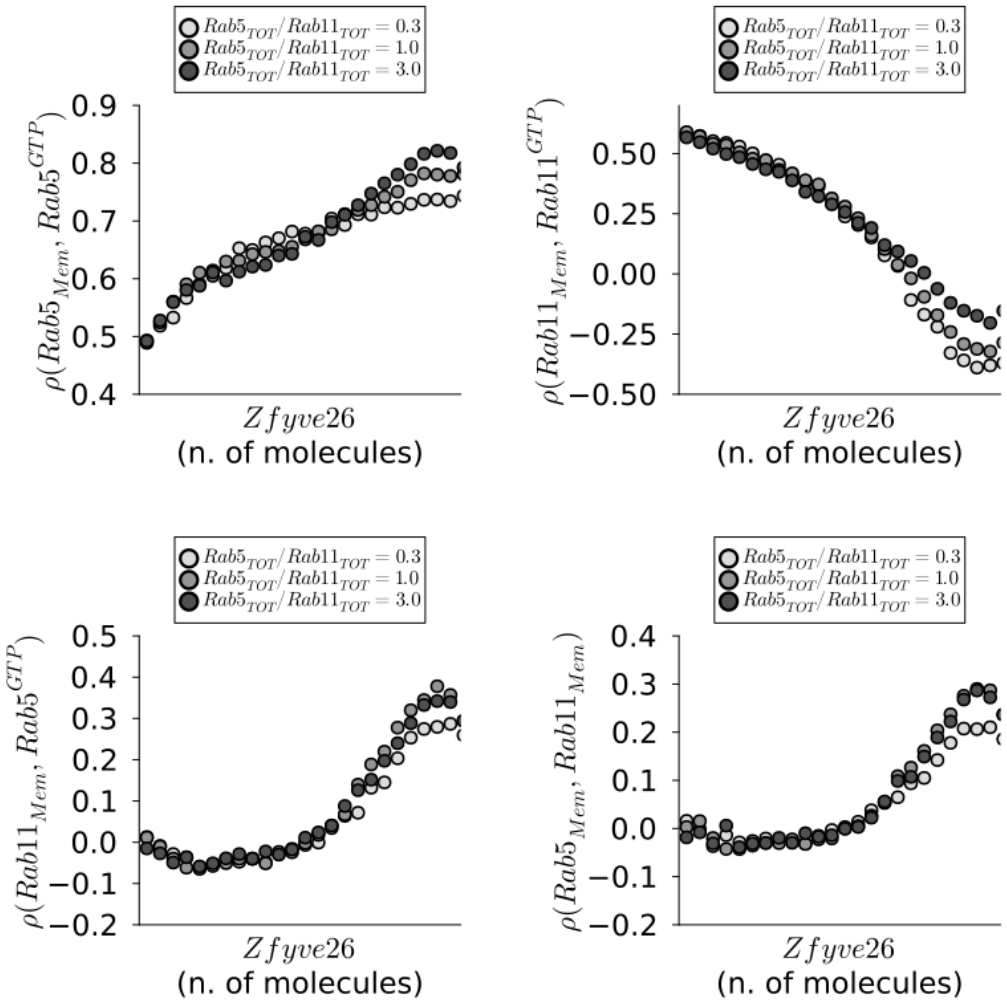


Fig. 5.25 Correlations between Rab membrane abundances and activities as functions of Zfyve26, for three different values of $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComIII: $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (top left panel), $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (top right panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (bottom right panel).

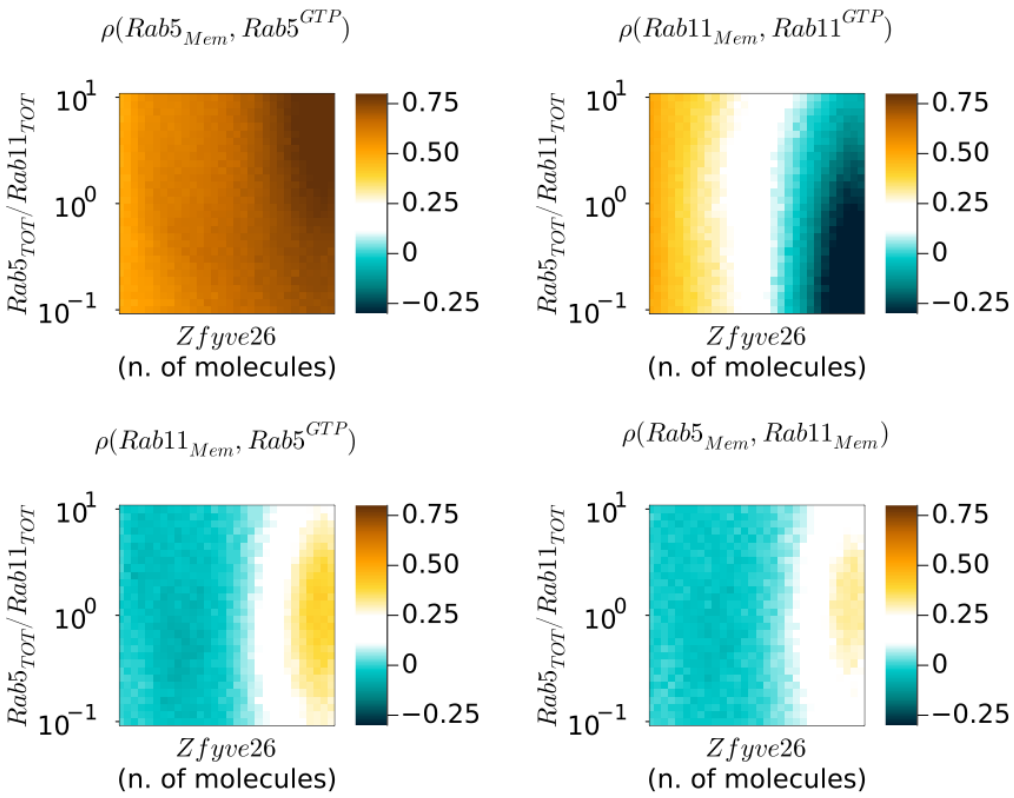


Fig. 5.26 Correlations between Rab membrane abundances and activities as functions of Zfyve26 and $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComIV: $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (top left panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (top right panel), $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (bottom right panel).

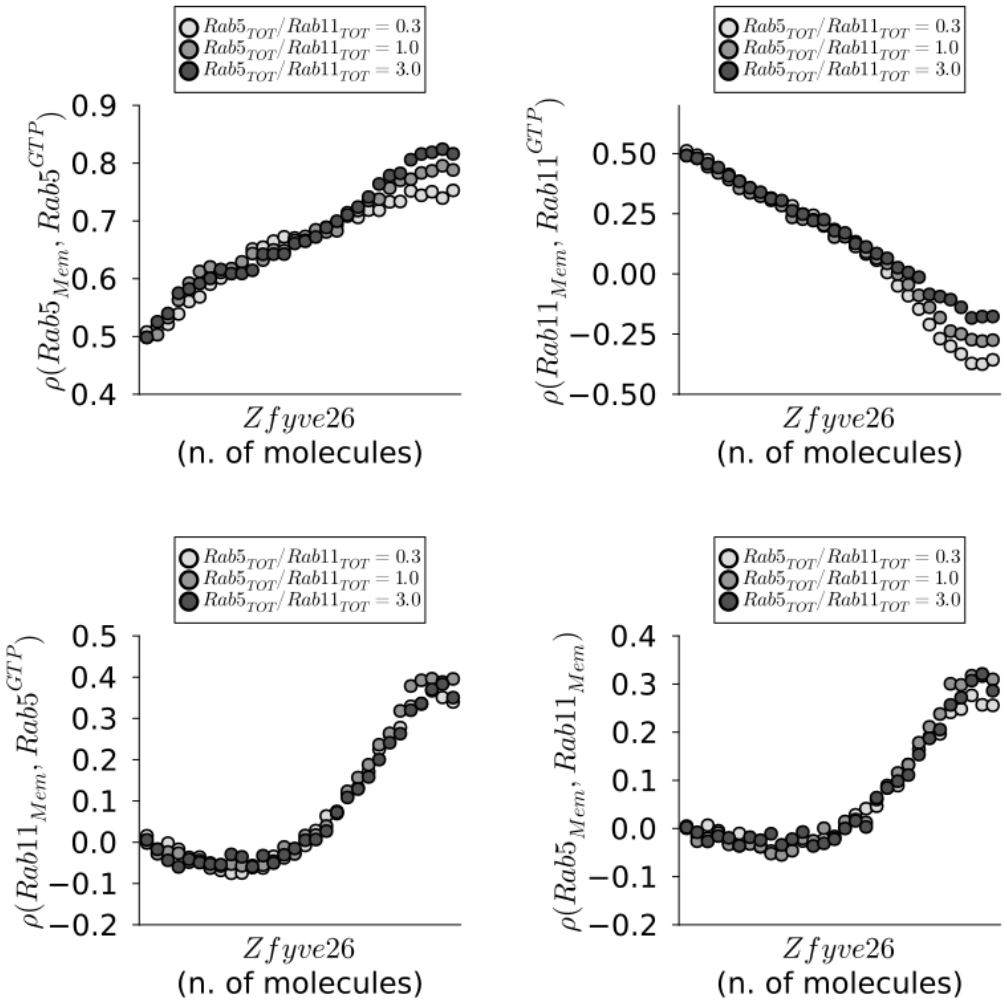


Fig. 5.27 Correlations between Rab membrane abundances and activities as functions of Zfyve26, for three different values of $Rab5_{TOT}/Rab11_{TOT}$ predicted by model ComIV: $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (top left panel), $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (top right panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (bottom right panel).

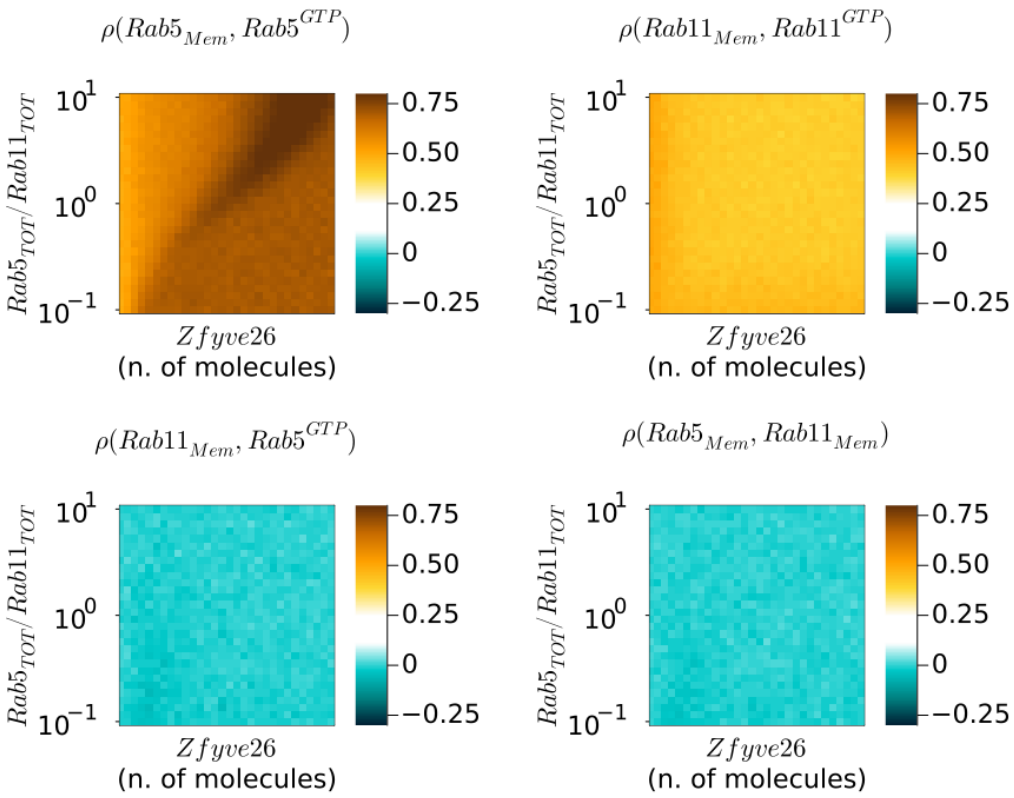


Fig. 5.28 Correlations between Rab membrane abundances and activities as functions of Zfyve26 and $Rab5_{TOT}/Rab11_{TOT}$ predicted by model Cas: $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (top left panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (top right panel), $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (bottom right panel).

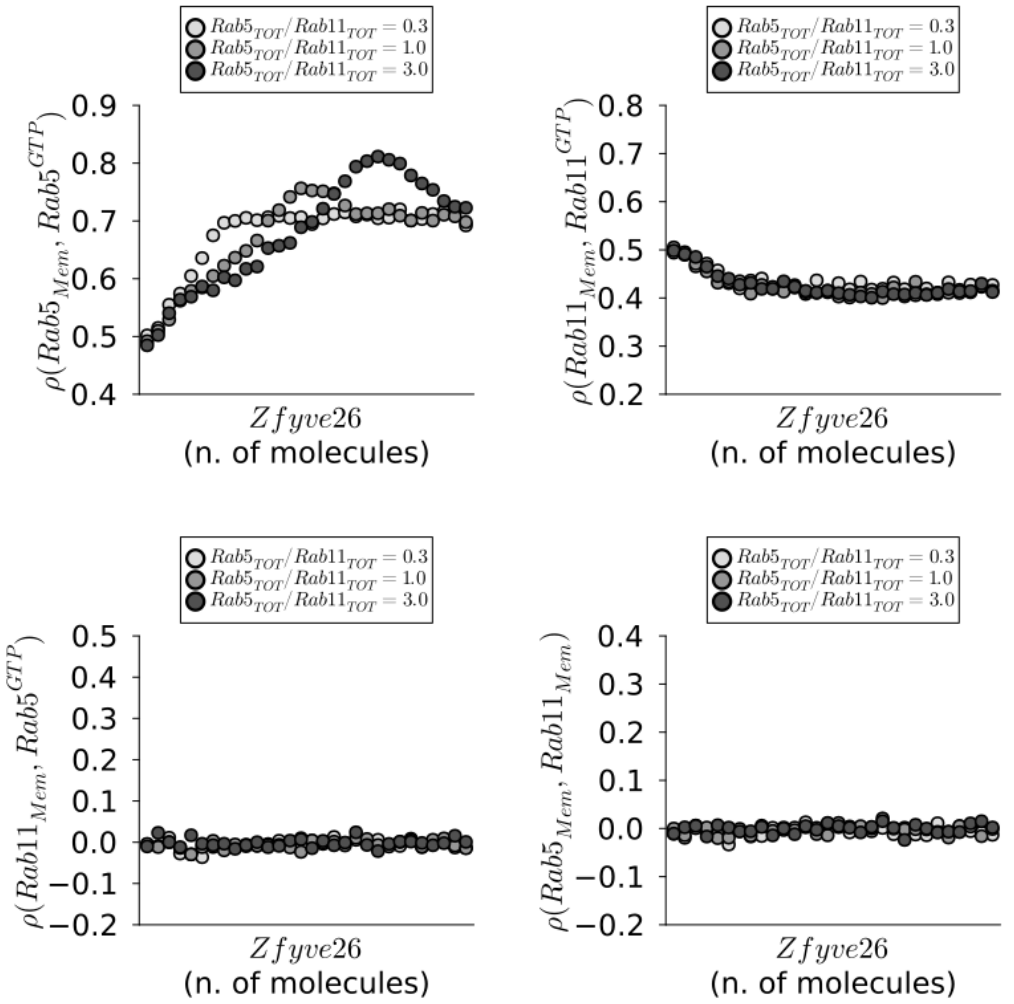


Fig. 5.29 Correlations between Rab membrane abundances and activities as functions of Zfyve26, for three different values of $Rab5_{TOT}/Rab11_{TOT}$ predicted by model Cas: $Rab5_{Mem}$-$Rab5^{GTP}$ correlation (top left panel), $Rab11_{Mem}$-$Rab11^{GTP}$ correlation (top right panel), $Rab11_{Mem}$-$Rab5^{GTP}$ correlation (bottom left panel), and $Rab5_{Mem}$-$Rab11_{Mem}$ correlation (bottom right panel).

The correlation between the number of Rab5 molecules on the membrane and the number of active Rab5 molecules, i.e. $\rho(Rab5_{Mem}, Rab5^{GTP})$, resulted positive and increasing with Zfyve26 for all models (Figures 5.20, 5.22, 5.24, 5.26, 5.28, bottom left panel). This finding is consistent with the presence of a positive Rab5 self-loop mediated by the $Rab5_{GTP} : Zfyve26$ complex: the highest the effector level, the more Rab5 molecules are active and bound to Zfyve26, and are thus retained on the membrane unable to leave. Therefore, activity fluctuations are associated to concor-

dant fluctuations in the total count of Rab5 molecules present on the membrane.
On the contrary, the resulting correlation between the amount of Rab11 on the endosomal membrane and its activity, i.e. $\rho(Rab11_{Mem}, Rab11^{GTP})$, differs between models (Figures 5.20, 5.22, 5.24, 5.26, 5.28, bottom right panel). Only models ComI and ComII generate an increasing correlation $\rho(Rab11_{Mem}, Rab11^{GTP})$ as a function of Zfyve26, similar to $\rho(Rab5_{Mem}, Rab5^{GTP})$ (Figures 5.20, 5.22, bottom right panel). Indeed, similar to the case of Rab5, these models imply that Rab11 interacts with Zfyve26 while active, by forming a $Rab11_{GTP} : Zfyve26$ complex, and thus lead to concordant fluctuations in the number of Rab11 molecules on the membrane and its active amount for high effector levels.
The same correlation $\rho(Rab11_{Mem}, Rab11^{GTP})$ results instead decreasing with Zfyve26 from models ComIII, ComIV and Cas (Figures 5.24, 5.26, 5.28, bottom right panel) Specifically, while the non-competitive model (Cas) predicts a less significant correlation decrease, the two competitive models ComIII and ComIV generate a switch from positive to negative $\rho(Rab11_{Mem}, Rab11^{GTP})$ values for increasing Zfyve26. This is coherent with the Rab11-effector interaction implied by the two models, where Zfyve26 binds to the inactive form of Rab11: a high effector level leads the majority of Rab11 molecules to be retained on the membrane in their inactive form, negatively coupling fluctuations in Rab11 activity and abundance.
Overall, Zfyve26-dependent correlations involving the membrane abundance and the activity of the single Rab species are predicted to increase by models where the effector interacts with the active form of the protein, and conversely to decrease if the effector interacts with the inactive form of the protein.
On the other hand, correlations involving activities/abundances of the two distinct species, i.e. $\rho(Rab11_{Mem}, Rab5^{GTP})$ and $\rho(Rab5_{Mem}, Rab11_{Mem})$, follow qualitatively similar trends as functions of the effector level for all competitive models (ComI-IV), presenting a negative minimum for low Zfyve26 and increasing to positive values for high Zfyve26 (Figures 5.20, 5.22, 5.24, 5.26, top panels). Indeed, when the amount of effector is limited, the binding of each GTPase to Zfyve26 is mutually exclusive, and thus the two Rab amounts are subject to opposite fluctuations. Such effect fades out if the amount of Zfyve26 is enough to exhaust the entire GTPase pool, as suggested by the progressive correlation increase. These correlations are particularly significant in the context of competition between Rab5 and Rab11 for binding to the shared resource Zfyve26: our results agree with previous theoretical predictions [57, 112, 113] showing that a limiting resource can cause

correlation between competitors, and that increasing the amount of resource disrupts this coupling. In line with this observation, the only model lacking competition, i.e. model Cas, generates $\rho(Rab11_{Mem}, Rab5^{GTP})$ and $\rho(Rab5_{Mem}, Rab11_{Mem})$ correlations constantly fluctuating around zero (Figure 5.28, bottom panels).
In general, qualitative correlation behaviours as function of Zfyve26 are maintained for different Rab5-Rab11 relative pool sizes. For clarity, correlations as functions of the effector level are also reported for three different values of $Rab5_{TOT}/Rab11_{TOT}$ picked from the explored range (0.3, 1.0, 3.0) (Figures 5.21, 5.23, 5.25, 5.27, 5.29), to show that Zfyve26-dependent trends result similar regardless of the relative amount of the two GTPases.

5.6 Summary of theoretical predictions

To qualitatively summarize theoretical predictions generated by all of our mathematical models, we report them in the following table. This table can be used for comparing predictions with experimental results illustrated in the next chapter.

	ComI	ComII	ComIII	ComIV	Cas
$Rab5_{Mem}$					
$Rab11_{Mem}$					
$Rab5^{GTP}$					
$Rab11^{GTP}$					
$\rho(Rab5_{Mem}, Rab11_{Mem})$					
$\rho(Rab5^{GTP}, Rab11_{Mem})$					
$\rho(Rab5_{Mem}, Rab5^{GTP})$					
$\rho(Rab11_{Mem}, Rab11^{GTP})$					

Increase
Decrease
None

Fig. 5.30 Table of qualitative theoretical predictions on Rab5 and Rab11 membrane abundances, activities and correlations as functions of the Zfyve26 effector level.

Chapter 6

Experimental investigation of Rab5 and Rab11 mean values and correlations

The idea behind the experiments of this project is investigating the hypothesis of competition between Rab5 and Rab11 by verifying theoretical predictions dependent on the number of effector molecules. If the hypothesis of Rab5 and Rab11 competing for the shared effector is valid, we expect an increase in the amount of Zfyve26 to loosen competition between the two proteins and to affect their abundances, activities and mutual fluctuations consistently with our predictions. Moreover, since our distinct competitive models hypothesize different kinds of interactions with Zfyve26, we aim at comparing their different predicted results with experimental data to understand which Rab form - active or inactive - interacts with the effector and whether it fulfils an activatory or inhibitory action.
To address whether predicted Zfyve26-dependent changes are recovered by experiments, our collaborators modulate the amount of effector available in the cell and we quantify mean Rab amounts and correlations on the endocytic membrane using fluorescence microscopy images. Additionally, photon counting experiment are performed to estimate the relative amount of the two GTPases on the membrane.

6.1 Introduction to experimental techniques

6.1.1 Fluorescence microscopy

Experimental biology makes extensive use of microscopy to follow the behaviour of molecules in space and time. In general, the real-time imaging of processes occurring inside a live cell requires modifying the molecules of interest to make them visually detectable. While there exists a variety of microscopy principles that allow a static visualization of non-living samples, fluorescence microscopy represents to date the most widely exploited imaging tool compatible with life [114]. Indeed, the use of specific excitation and emission wavelengths minimizes its toxicity, and the availability of non-invasive fluorescent markers avoids disrupting cellular processes [114]. Moreover, the principles employed to mark molecules often exploit mechanisms endogenous to life itself, without affecting cell viability. For instance, a typical technique relies on inducing the cell itself to synthesize fluorescent versions of the molecule of interest: cells fed with artificial DNA that codes for a fluorescent protein are naturally forced to produce that protein [115]. Once the fluorescent protein is produced by the cell, its spatial and temporal dynamics can be visualized. For instance, let's consider a molecule involved in membrane trafficking, such as a Rab protein: this molecule can be visualized by introducing into cells a DNA sequence carrying the Rab protein gene fused to a gene encoding a fluorescent protein. This enables for instance the visualization of Rab domains on the membrane of intracellular compartments and organelles. Figure 6.1 reports an example of a fluorescence microscopy image of Rab5 acquired by our collaborators on the membrane of a recycling endosome.

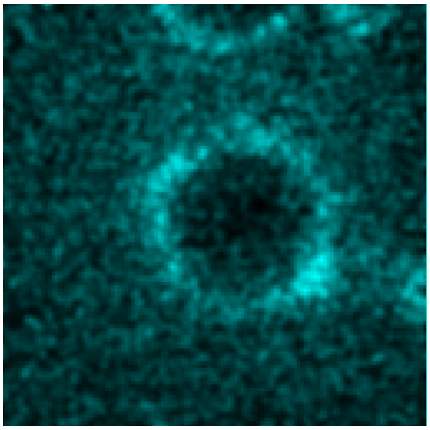


Fig. 6.1 Example of Rab5 imaged on the membrane of recycling endosomes by fluorescence microscopy.

The next section illustrates the experimental approaches that allow to synthesize the artificial DNA sequence carrying the gene of interest, and introduce it into live cells for fluorescence microscopy imaging.

6.1.2 Plasmid DNA cloning and transient cell transfection

Artificial DNAs commonly used to introduce a gene of interest into cells are circular fragments derived from bacterial DNA called "plasmids". Plasmids carrying the genes of interest - e.g. the genes coding for a molecule of interest and a fluorescent protein - often need to be generated by modifying previously existent DNA with cloning techniques: the use of specific enzymes and Polymerase Chain Reaction (PCR) allows "cutting, copying and pasting" DNA fragments to build the desired plasmid sequence before feeding the plasmid to cells.
The principle of introducing exogenous nucleic acids into cells is referred to as "cell transfection". This technique exploits chemical properties of the cell membrane to facilitate the entry of genetic material: nucleic acids are associated to positively charged lipids. In this way, DNA can then be internalized by the negatively charged cell membrane.
Following transfection, the introduced nucleic acids remain in the cell for a period of days at longest, and thus does not become stably part of the cell's genome and cannot be inherited by daughter cells. Therefore, transient transfection experiments are suitable for the observation of cellular processes that span up to few days. During

this period of time, the cell synthesizes molecules genetically encoded in the plasmid. Hence, this approach suits the visualization of Rab proteins on the membrane of endosomes - an event associated to recycling processes, which span minutes [1].

6.1.3 Fluorescence resonance energy transfer (FRET)

FRET is a physical process by which energy is transferred nonradiatively from an excited molecular fluorophore - called "donor" - to another fluorophore - called "acceptor" - by means of intermolecular long-range dipole-dipole coupling [116]. The efficiency of this energy transfer depends on the overlap between the emission spectrum of the donor and the absorption spectrum of the acceptor (Figure 6.2), and on the distance between donor and acceptor (Figure 6.3). Specifically, FRET efficiency is inversely proportional to the sixth power of the distance between donor and acceptor. In particular, FRET is highly efficient if the donor and the acceptor are positioned within the Förster radius, i.e. the distance at which half the excitation energy of the donor is transferred to the acceptor (typically 3-6 nm) [116]. Thanks to this property, FRET provides a sensitive tool for the measurement of molecular proximity at at ångstrom distances (10-100 ångstrom).

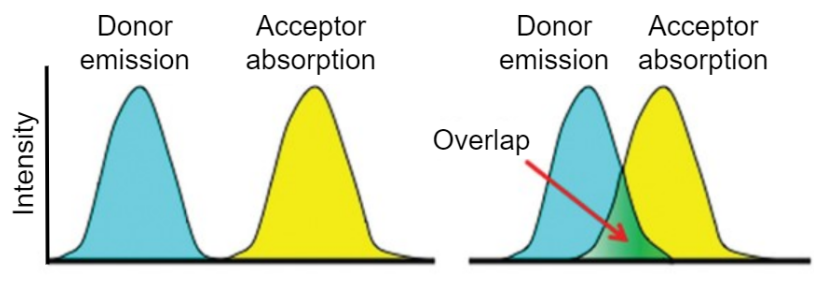


Fig. 6.2 Overlap between emission spectrum of the donor and absorption spectrum of the acceptor. Significant overlap (shaded region) between the donor emission and the acceptor excitation is one of the requirements for a FRET molecule pair (adapted from: Khandria and Senes, Methods in Molecular Biology 2013)

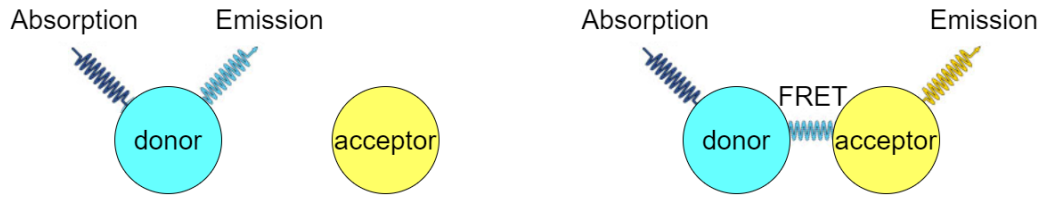


Fig. 6.3 Schematic of the FRET working principle. FRET is highly sensitive to the distance between donor and acceptor fluorophores.

FRET is widely employed for the detection of molecular events of interest, such as interactions between proteins or changes in molecule conformation. In our experiments, it can be exploited to detect protein activation. Indeed, donor and acceptor fluorophores can be connected to a protein in such a way that the conformational change brings the two close together, and FRET can occur. For instance, in the FRET probes used in our experiments, donor and acceptor are connected to a Rab protein. When the Rab protein is inactive, the two fluorophores are far enough from each other that no FRET can occur; thus, excitation of the donor results in fluorescence emission by the donor itself. Conversely, Rab activation causes conformational changes that bring the two fluorophores close together; thus FRET occurs and leads to fluorescence emission by the acceptor.
Specifically, FRET biosensors used in our experiments are constituted by the Rab protein fused to a donor mECFP (cyan fluorescent protein) and an acceptor YFP (yellow fluorescent protein) attached by a linker (Rab11 FRET probe AS-Rab11 in Figure 6.4):

- When the Rab protein is inactive, excitation of the donor mECFP results in emission of its constitutive cyan fluorescence (mECFP channel); if conversely the Rab protein becomes active, excitation of mECFP leads to emission of yellow fluorescence by the acceptor YFP due to FRET (FRET channel). Thus, excitation of mECFP provides a proxy of Rab activity.
- Regardless of FRET, excitation of YFP results in emission by YFP itself (YFP channel), and thus provides a proxy of total Rab abundance.

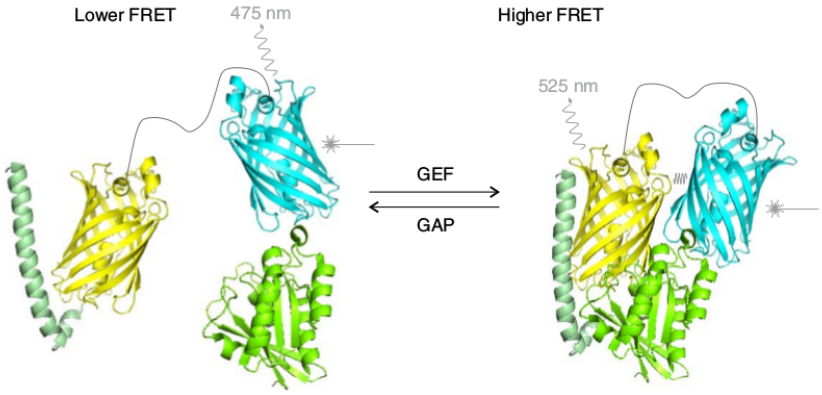


Fig. 6.4 Tridimensional representation of the FRET probe AS-Rab11 (Rab11 activity sensor) in its inactive (bound to GDP, left) or active conformation (bound to GTP, right). Yellow and light blue β-barrels represent yellow- and cyan-emitting fluorescent proteins, respectively. Sea-green α-helix, black line, and green β structures indicate respectively the Rab11-GTP-binding domain, the linker domain and the Rab11 protein (adapted from [4]

This approach is called sensitized emission FRET and is among the most common forms of FRET microscopy. Sensitized emission FRET relies on measuring the FRET signal and the CFP signal by fluorescence imaging [116], and adopting the so-called "FRET ratio" as indicator of FRET efficiency:

$$FRETratio = \frac{FRET}{CFP} \tag{6.1}$$

where $FRET$ is the FRET channel image, CFP is the CFP channel image.
Computing this ratio amplifies small changes in FRET: as the acceptor (FRET) emission at numerator increases, the donor (mECFP) emission at denominator correspondingly decreases, and thus their ratio will FRET/mECFP will be further amplified.
In our analysis, the final FRET ratio image intensity is adopted as indicator of Rab activity.

6.2 Experiments

6.2.1 Visualizing Rab5 and Rab11 on the endosomal membrane

These experiments are aimed at visualizing Rab5 and Rab11 on the endosomal membrane in two conditions of effector level - a control (wild-type) condition and a treated (Zfyve26 overexpression) condition.
Experimental steps can be summarized into four main stages:

- Plasmids encoding fluorescent Rab5 and Rab11 versions and plasmids encoding Rab5 and Rab11 FRET activity biosensors were either provided by collaborators or synthesized by cloning.
- Mammalian cells were grown in culture and transfected with plasmids encoding Rab fluorescent versions and FRET activity biosensors. To vary the amount of effector, cell samples were partly kept as control and partly additionally transfected with plasmid DNA encoding Zfyve26 (effector overexpression).
- The recycling pathway was initiated by feeding cells with transferrin - a molecule recognized as cargo.
- Rab5 and Rab11 were imaged on endosomes with a fluorescence microscope in time-lapse mode.

Each of these steps will be illustrated in greater detail in the next subsections.

Plasmid DNA cloning

The following plasmid DNAs encoding fluorescent versions of the two Rab proteins were used:

- Plasmid encoding red fluorescent Rab5 - mCherry-Rab5 - was provided by Emilio Hirsch (Department of Molecular Biotechnology and Health Sciences, University of Turin, Italy);
- Plasmid encoding red fluorescent Rab11 - mCherry-Rab11a, was generated by PCR and restriction enzyme cloning from plasmid DNA provided by Michael Davidson.

The following plasmid DNAs encoding the two Rabs with their fluorescent activity biosensors (FRET biosensors) were used:

- Plasmid encoding Rab5 with FRET biosensor - Raichu-Rab5 - was provided by Michiyuki Matsuda (Department of Pathology and Biology of Diseases, Graduate School of Medicine, Kyoto University, Japan).
- Plasmid encoding Rab11 with FRET biosensor - AS-Rab11 - was provided by Prof. Emilio Hirsch (Department of Molecular Biotechnology and Health Sciences, University of Turin, Italy).

In addition, the following plasmid coding for Zfyve26 was used to overexpress the effector protein:

- Plasmid encoding Zfyve26 - FLAG-Zfyve26 - was generated by PCR and restriction enzyme cloning from plasmid DNA provided by Christian Hübner (Institute of Human Genetics, University Hospital, Jena, Germany).

Cell culture and transfection

5x10^4 COS-7 cells (a cell line derived from monkey kidney tissue) were plated in 8 well μ-Slide imaging dishes of 35 mm diameter (Ibidi).
Control cells were transfected with plasmid DNA encoding fluorescent Rab versions and FRET activity biosensors using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Cells intended for Zfyve26 overexpression were transfected with the same plasmids and additionally with plasmid DNA encoding Zfyve26.

Specifically, two sets of transfection experiments were performed:

- Set 1: Cells were transfected with the red fluorescent Rab11 version mCherry-Rab11a and with the Rab5 activity biosensor Raichu-Rab5. Cells for effector overexpression were additionally transfected with FLAG-Zfyve26.
- Set 2: Cells were transfected with the the Rab11 activity biosensor AS-Rab11. Cells for effector overexpression were additionally transfected with FLAG-Zfyve26.

Following transfection, cells were incubated at 37$^\circ C$.

Initiation of recycling

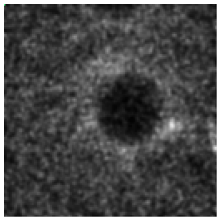
48 hours after transfection, cells were starved for two hours in serum-free medium DMEM 0.1% BSA at 37°*C* in 5% CO2.
Human transferrin (Invitrogen) was added to the cell culture medium at a concentration of 20 ug/ml, to allow continuous cargo uptake and recycling during the subsequent image acquisition.

Time-lapse microscopy

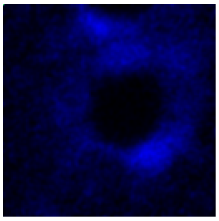
Endosomes were imaged on a two-dimensional focal plane to obtain images containing a single endosome and the surrounding region. Imaging was performed in a CO2 independent medium, Dulbecco's modified Eagle's medium without FBS (GIBCO). Time-lapse series were acquired at 37°*C* on an inverted confocal Leica SP8 microscope with AOBS equipped with 63× O2/Oil immersion objective, NA 1.40. The temperature was controlled by a climate box covering the set up. Fluorescent Rab proteins were imaged sequentially in line-interlace mode to eliminate cross-talk between the channels. mECFP was excited with a 458-*nm* laser line and imaged at 470–500-nm bandpass emission filters. YFP was excited with the 514-*nm* argon laser line and imaged through a 525–550-*nm* bandpass emission filter. mCherry/mRFP was excited with the 568-nm helium neon laser line and imaged through a 580–650-*nm* bandpass emission filter. Serial sections were acquired satisfying the Nyquist criterion, sampling submicrometric-sized pixels (50 nm). Acquisition was performed at zoom (×40) in a region of 7.3 μm in side. Exposure times and readout were fixed to obtain time-lapse sequences of roughly one frame per 158 *ms*, coherently with the rapid temporal scale of GTPase signalling variations. Images obtained were stacked as time-lapse and exported as a single TIFF file.

6.2.2 Photon counting

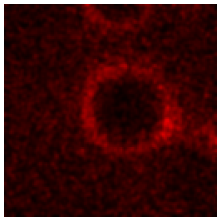
5x10^4 COS-7 cells were plated in 8 well μ-Slide imaging dishes (Ibidi) and transfected with the indicated combination of fluorescent tagged version of biomolecular switches: EGFP-Rab5/mCherry-Rab11 or EGFP-Rab11/mCherry-Rab5. After 48 hours, cells were starved for two hours in serum-free DMEM 0.1% BSA at 37°*C*



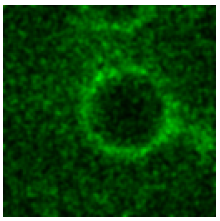
A frame example from set 1 in the mCherry channel (Rab5)



A frame example from set 1 in the eYFP channel (Rab11)



A frame example from set 1 in the eCFP channel (Rab5)



A frame example from set 1 in the FRET sensitized channel (Rab5)

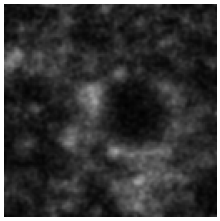
Fig. 6.5 Multiple frame examples from set 1 in different channels

in 5% CO2. Human Transferrin (Invitrogen) was added at a concentration of 20 ug/ml for the time of the acquisition. Imaging was performed in a CO2 independent medium, Dulbecco's modified Eagle's medium without FBS (GIBCO). Cells were imaged using an inverted confocal Leica SP8 microscope with AOBS equipped with 63× O2/Oil immersion objective, NA 1.40. Acquisition was performed at zoom (×11) in a region of 26 μm in side. To obtain photon counting, the Hybrid Detector of LEICA SP8 microscope was set to "Photon counting mode" with 10-line accumulation. The number of measured photons was subsequently used to quantify the amount of each Rab GTPase on the endocytic structure.

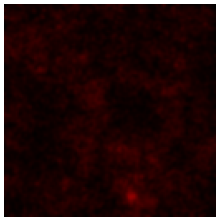
6.3 Data analysis and results

6.3.1 Analysis of time-lapse images

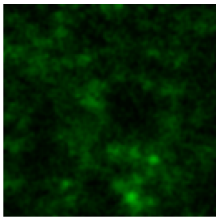
Each timelapse (provided in $.tiff$ format) is a series of approximately 300 frames. The first set of timelapses (set 1) is constituted of four channels, i.e. CFP (Rab5), FRET sensitized (Rab5), mCherry (Rab11), YFP (Rab5), whereas the second set (set 2) is constituted of three channels, i.e. CFP (Rab11), FRET sensitized (Rab11), mCherry (Rab11).



A frame example from set 2 in the mCherry channel (Rab11)



A frame example from set 2 in the eCFP channel (Rab11)



A frame example from set 2 in the FRET sensitized channel (Rab11)

Fig. 6.6 Multiple frame examples from set 2 in different channels

We will now illustrate the main steps of the analysis of fluorescence videos, that can be assembled into two main stages. First is the automatic segmentation of images to identify the region corresponding to the endosomal membrane. Second is the quantification of Rab5 and Rab11 abundances, activities and correlations on the identified region.

Image segmentation for the identification of the endosomal membrane

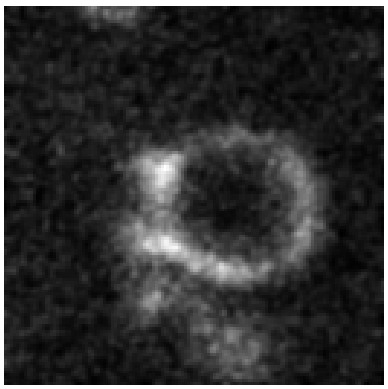


Fig. 6.7 A frame example from set 1 in the mCherry (Rab5) channel.

Each frame of each timelapse was segmented in order to correctly identify the region of interest, i.e. the endosomal membrane. The final purpose of this segmentation is to generate a ring-shaped region enclosing the endosome.
The segmenation code is implemented within a *MATLAB* environment using the *Bioformats* package, which allows to read data from a variety of microscopy image

formats, including $.tiff$ frames.

Image segmentation steps are the following:

Applying the in-built Bioformats function "bfopen()" to a timelapse file returns a cell structure in which each channel in each frame can be accessed separately. Each frame is represented by four (set 1) or three (set 2) 128x128 matrices containing the fluorescence intensity values of channels. Each matrix can be visualized in the form of a 128x128 pixel grayscale image.
For each frame of a timelapse we performed the following manipulation steps (an example is shown in Figure 6.8):

- The Rab5 fluorescence intensity channel in set 1 (mCherry) and the Rab11 fluorescence intensity channel in set 2 (mCherry) are employed as references to localize the endosomes in the two sets of timelapses.
- Fluorescence intensity values are transformed by a maximum intensity projection and then adding linear contrast.
- Noise is reduced by means of a Gaussian smoothing.
- The image is binarized according to a custom threshold.
- Since the expected ring-shaped region can still be poorly defined or unclosed, the connected components present in the image are first dilated and then eroded.
- After erosion, since the endosome might still be unclosed or unwanted objects might still be present in the image, we localize the main connected component and delete objects that show up closer to the image boundaries.
- To further ensure that only the region corresponding to the endosome is kept, we apply to the image a criterion that keeps only objects that satisfy a circularity index.
- Eventually, we delete any other object except the largest one present in the image.

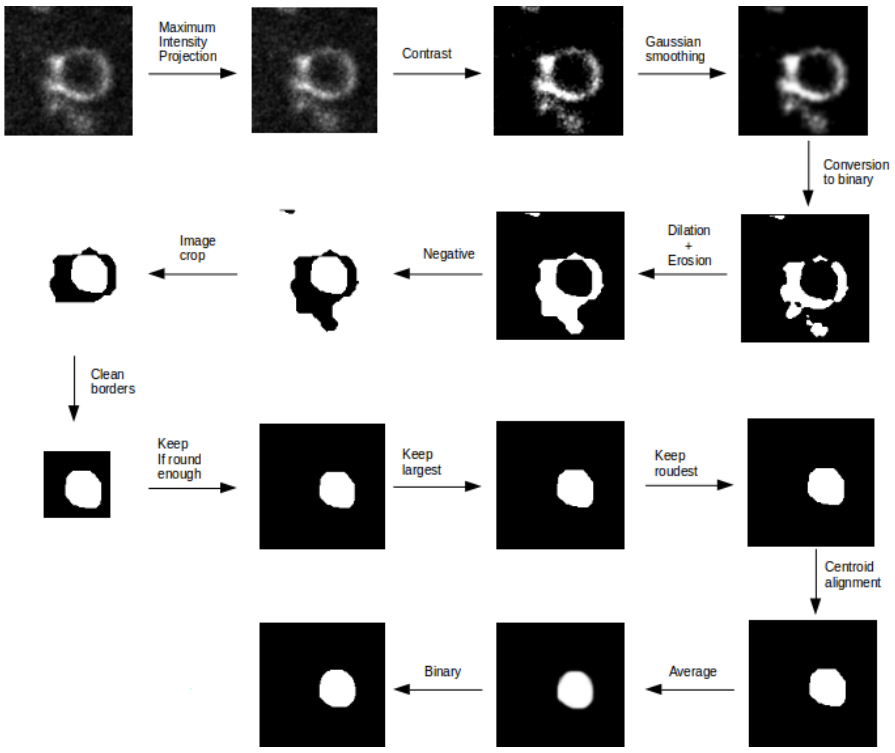
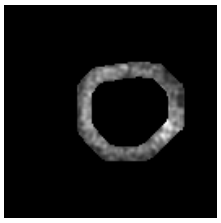


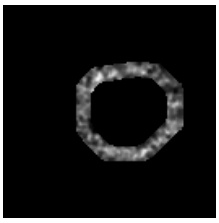
Fig. 6.8 Schematization of image manipulation steps performed with *MATLAB*.

Every image manipulation step employs a different MATLAB function that requires specific parameters. Therefore, parameters for correct recognition were optimized for each distinct timelapse to obtain correct closing of the endosomal area in as many frames as possible.
Frames that still display incorrect segmentation at the end of these steps are not considered for endosome localization.
The identified region is aligned to the same centroid point in every frame to eliminate potential displacements of the endosome during the timelapse.
Eventually, frames that presented correct segmentation are employed to estimate an average region by taking the mean of frames. The average image is binarized using a custom threshold to obtain the final area which represents the internal region of the endosomal membrane throughout the timelapse.
At this stage we were able to build a ring-shaped mask surrounding the obtained area. The membrane of endosomes has a thickness of approximately 5*nm*, but since our resolution is consistently lower, the size of the ring is chosen accordingly and set to 150*nm*, which corresponds roughly to a 10-pixel size on our 128x128 pixel images. The rest of the image is converted to black and fluorescence signals are considered

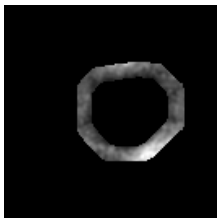
only on the ring-shaped endosome membrane region.



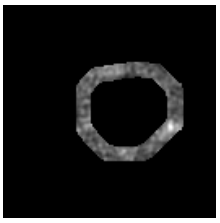
A frame example of CFP (Rab5) channel filtered through the identified ring-like region (set 1)



A frame example of FRET (Rab5) channel filtered through the identified ring-like region (set 1)

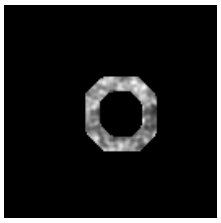


A frame example of mCherry (Rab11) filtered through the identified ring-like region (set 1)

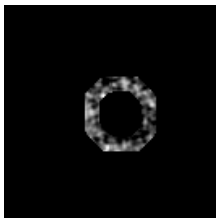


A frame example of YFP (Rab5) channel filtered through the identified ring-like region (set 1)

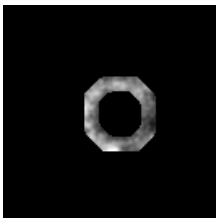
Fig. 6.9 Frame examples from set 1 in different channels



A frame example of CFP (Rab11) channel filtered through the identified ring-like region (set 2)



A frame example of FRET (Rab11) channel filtered through the identified ring-like region (set 2)



A frame example of mCherry (Rab11) filtered through the identified ring-like region (set 2)

Fig. 6.10 Frame examples from set 2 in different channels

Fluorescence data cleaning and FRET ratio computation

We cleaned fluorescence data by eliminating negative/NaN pixel values and we computed the FRET ratio image - our proxy of Rab activity - by taking the ratio of the FRET sensitized channel image and the CFP channel image for each frame. An example of FRET ratio computation is shown in Figure 6.11.

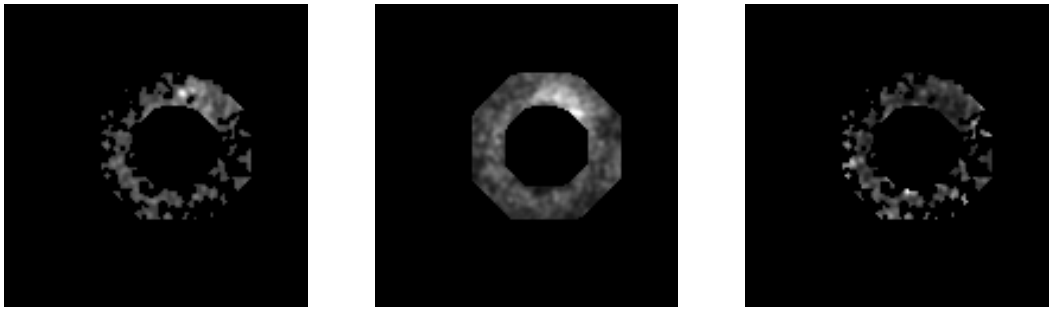


Fig. 6.11 Example of FRET ratio computation performed with *MATLAB*. The FRET ratio image represented (right panel) is the ratio between the FRET sensitized image (left panel) and the CFP image (middle panel).

Exploratory data analysis

For each timelapse, we computed the fluorescence intensity distribution of Rab fluorescence intensity channels and FRET ratio relative to each Rab.
By setting a threshold on fluorescence intensity and FRET ratio distribution values, we binarized signals and visualized them on the identified ring-shaped region to view their localization thoughout a timelapse. In the frame sequence reported in Figure 6.12 we show an example of a 6-frame evolution taken from a timelapse of set 1.

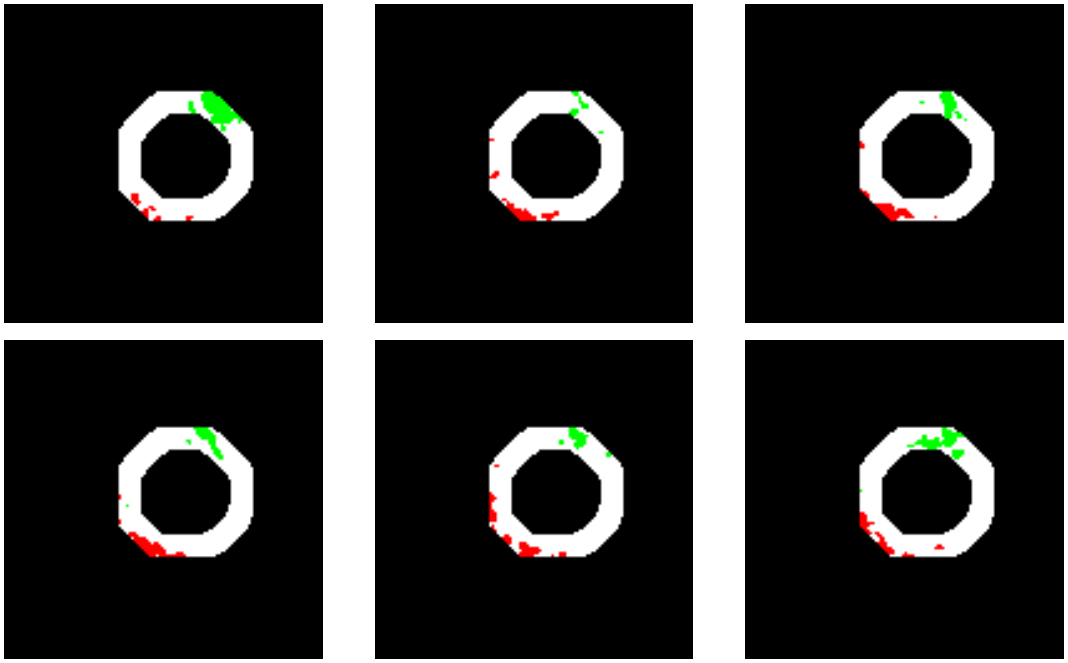


Fig. 6.12 Example of six temporally consecutive frames with binary Rab11 (mCherry, depicted in red) and Rab5 FRET ratio (depicted in green) signals.

These images suggested that Rab11 (mCherry, red patches) and active Rab5 (FRET ratio, green patches) tend to occupy separate areas of the membrane. This observation is in line with previous literature reporting that distinct Rab species mark specific regions of intracellular membranes with mutual exclusivity [90].

6.3.2 Experimental quantification of mean values

Mean fluorescence values of Rab5 mCherry (i.e. Rab5 membrane abundance), Rab5 FRET ratio (i.e. Rab5 activity), Rab11 mCherry (i.e. Rab11 membrane abundance) and Rab11 FRET ratio (i.e. Rab11 activity) were quantified by estimating the mean fluorescence of the endosomal membrane for each frame and then averaging values relative to frames. Once a mean value was obtained for each experiment, we took the mean value of experiments belonging to the control and the Zfyve26 overexpression sets. Figures 6.13, 6.14, 6.15 and 6.16 represent the obtained mean abundance and activity values, each one normalized by the mean value of the control condition:

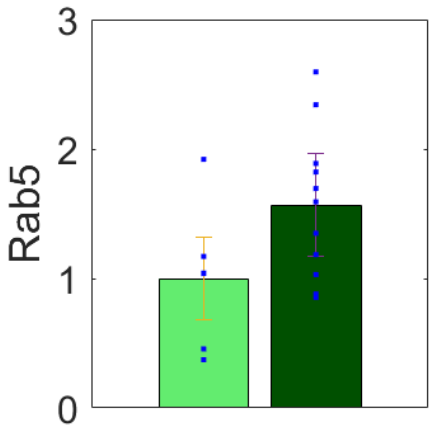


Fig. 6.13 Normalized mean Rab5 abundance (YFP fluorescence intensity) on the endocytic membrane in control (left bar) and Zfyve26 overexpression (right bar) conditions. Points are relative to experiments (i.e. timelapses). Error bars represent standard errors of the mean.

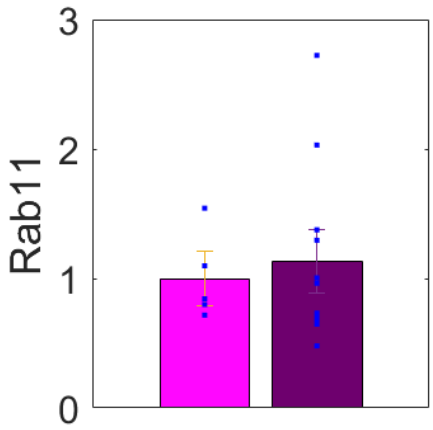


Fig. 6.14 Normalized mean Rab11 abundance (mCherry fluorescence intensity) on the endocytic membrane in control (left bar) and Zfyve26 overexpression (right bar) conditions. Points are relative to experiments (i.e. timelapses). Error bars represent standard errors of the mean.

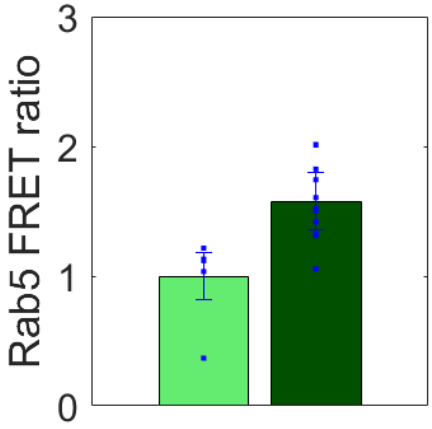


Fig. 6.15 Normalized mean Rab5 activity (Rab5 FRET ratio) on the endocytic membrane in control (left bar) and Zfyve26 overexpression (right bar) conditions. Points are relative to experiments (i.e. timelapses). Error bars represent standard errors of the mean.

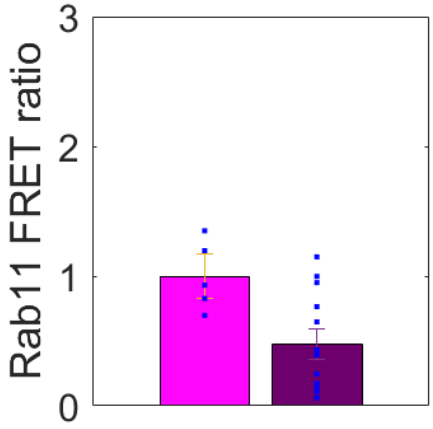


Fig. 6.16 Normalized mean Rab11 activity (Rab11 FRET ratio) on the endocytic membrane in control (left bar) and Zfyve26 overexpression (right bar) conditions. Points are relative to experiments (i.e. timelapses). Error bars represent standard errors of the mean.

Results show that Zfyve26 overexpression leads to an increase of Rab5 activity, i.e. the Rab5 FRET ratio, and to a decrease of Rab11 activity, i.e. the Rab11 FRET ratio (respective p-values obtained by two-sample t-tests: $p<0.05$, $p<0.05$). Similar results were obtained respectively for Rab5 abundance, i.e. fluorescence intensity mCherry, and Rab11 abundance, i.e. fluorescence intensity mCherry on the endosomal membrane (respective p-values obtained by two-sample t-tests: $p<0.1$, $p<1.0$).
These experimental results match qualitative predictions of models ComIII, ComIV and Cas.

6.3.3 Experimental quantification of correlations

To verify whether GTPases might be correlated as predicted by our competitive models, we computed correlations between the following pairs of signals:

- Rab5 abundance (mCherry fluorescence intensity) and Rab5 activity (Rab5 FRET ratio).
- Rab11 abundance (mCherry fluorescence intensity) and Rab11 activity (Rab5 FRET ratio).
- Rab11 abundance (mCherry fluorescence intensity) and Rab5 activity (Rab5 FRET ratio).
- Rab5 abundance (mCherry fluorescence intensity) and Rab11 abundance (YFP fluorescence intensity).

Framewise correlation values were obtained by computing the spatial correlation between molecular intensities on the membrane boundary at each timestamp as:

$$\rho(a,b) = \frac{\sum_{i=1}^{n}(a_i-\bar{a})(b_i-\bar{b})}{\sqrt{\sum_{i=1}^{n}(a_i-\bar{a})^2}\sqrt{\sum_{i=1}^{n}(b_i-\bar{b})^2}} \tag{6.2}$$

where a and b represent their pixelwise intensities, i identifies the pixel, and n indicates the total number of pixels contained in the ring-shaped region of interest.

Next, for each correlation we averaged its temporal trajectories to obtain its mean value for each single time-lapse. Note that the temporal stationarity of correlation trajectories allowed us to average correlation time series values without loss of information. Thus, by averaging such mean correlations relative to single endosomes we obtained a mean correlation value for timelapses of control condition and similarly a mean correlation value for timelapses of Zfyve26 overexpression condition. Using this approach, we obtained a mean value for the four correlations above in the two effector conditions, i.e. control and Zfyve26 overexpression, reported in Figures 6.17, 6.18, 6.19, and 6.20.

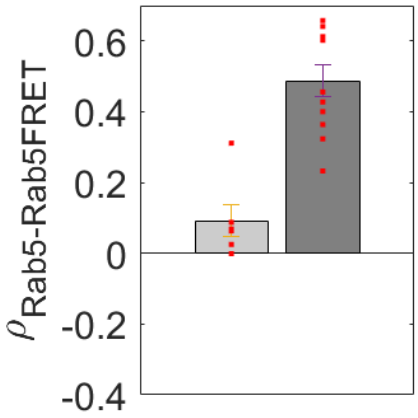


Fig. 6.17 Mean correlation between Rab5 membrane abundance (YFP) and Rab5 activity (FRET ratio) in control (left) and Zfyve26 overexpression (right) conditions. Points are means relative to experiments (i.e., timelapses). Error bars represent standard errors of the mean.

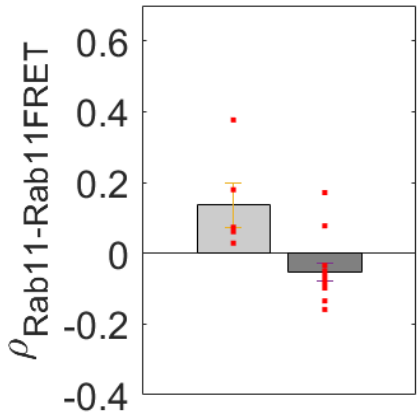


Fig. 6.18 Mean correlation between Rab11 membrane abundance (YFP) and Rab11 activity (FRET ratio) in control (left) and Zfyve26 overexpression (right) conditions. Points are means relative to experiments (i.e., timelapses). Error bars represent standard errors of the mean.

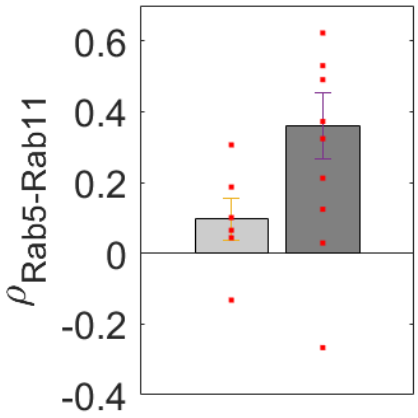


Fig. 6.19 Mean correlation between Rab5 and Rab11 membrane abundances (YFP and mCherry) in control (left) and Zfyve26 overexpression (right) conditions. Points are means relative to experiments (i.e., timelapses). Error bars represent standard errors of the mean.

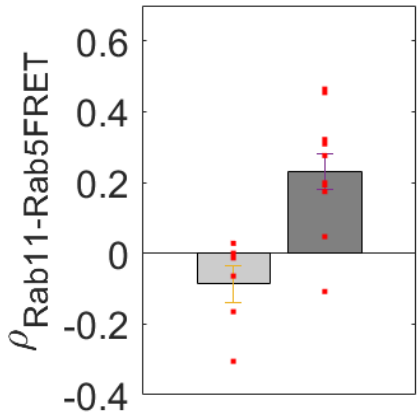


Fig. 6.20 Mean correlation between Rab11 membrane abundance (mCherry) and Rab5 activity (FRET ratio) in control (left) and Zfyve26 overexpression (right) conditions. Points are means relative to experiments (i.e., timelapses). Error bars represent standard errors of the mean.

In agreement with Cas, ComIII and ComIV model predictions on correlations, we found that expression of Zfyve26 induces an increase in the correlation between Rab5 fluorescence intensity (i.e. Rab5 abundance) and FRET ratio (i.e. Rab5 activity), as shown in Figure 6.17 (Two-sample t-test p-value p<0.001). In contrast, measured $Rab11_{Mem}$-$Rab11^{GTP}$ correlation values were found to switch from positive to negative values with Zfyve26 overexpression (Figure 6.18), in line with models ComIII and ComIV (Two-sample t-test p-value p<0.005).
Eventually, we addressed reciprocal fluctuations of the two GTPase abundances and activities, i.e. Rab11-Rab5FRET and Rab5-Rab11 correlations. We thus computed the correlation between fluorescence intensities of Rab5 and Rab11 and found that it increases with Zfyve26 overexpression (Figure 6.19, two-sample t-test p-value p<0.1). In parallel, the correlation between Rab11 fluorescence intensity and Rab5 FRET ratio was found switching from negative values in control cells to positive values in cells where Zfyve26 was overexpressed (Figure 6.20, two-sample t-test p-value p<0.001), thereby confirming predictions of both models ComIII and ComIV. These data thus suggest that Rab5 and Rab11 are connected to their shared effector Zfyve26 through a competition mechanism, and points towards models ComIII and ComIV. In these models, Rab5 binds to the effector in its active form $Rab5^{GTP}$ to trigger self-activation, whereas Rab11 interacts with Zfyve26 in its inactive form $Rab11^{GDP}$.
These results however did not allow to discriminate between the two models ComIII and ComIV, which differ in the kind of interaction - activatory or inhibitory - that the $Rab11^{GDP} : Zfyve26$ complex exerts on the GTPase.

6.3.4 Quantification of Rab5 and Rab11 from photon counting data

Photon counting experiments provided a relative Rab5/Rab11 amount ratio on the endocytic membrane of 2.93 (two-sample t-test p-value $p < 0.001$), thus resulting consistent with the range of $Rab5_{TOT}/Rab11_{TOT}$ ratio values explored in our theoretical model analysis.

6.4 Discrimination between models ComIII and ComIV

Since models ComIII and ComIV provide qualitatively similar Zfyve26-dependent predictions, we could not discriminate between the two hypotheses by relying solely on their response to different effector amounts. Due to the competition mechanism embedded in the two models, each GTPase is not only affected by the level of Zfyve26 but also by the level of the cognate protein. Since models ComIII and ComIV differ in the hypothesized mechanism acting on Rab11, we addressed the relationship between the membrane amount of Rab11, $Rab11_{Mem}$, and its activity, $Rab11^{GTP}$ resulting from varying the total amount of Rab5 fed to the system in the two competitive models.

While maintaining the effector level fixed to either a low (50 molecules) or a high (400 molecules) level, the system was fed with progressively increasing amounts of $Rab5_{TOT}$ ranging from 0 to 50 molecules. For each $Rab5_{TOT}$ value, we quantified both the membrane abundance and the activity of Rab11, corresponding to $Rab11_{Mem}$ and $Rab11^{GTP}$ in our models.

Predictions obtained from the two models show that model ComIII generates an increasing activity, $Rab11^{GTP}$, as a function of the Rab11 amount on the membrane $Rab11_{Mem}$ in the low Zfyve26 condition (50 molecules), whereas a decreasing trend if Zfyve26 is higher (400 molecules) (see Figure 6.21, left panel). Conversely, model ComIV predicts a decreasing $Rab11^{GTP}$ as a function of $Rab11_{Mem}$ for both low and high Zfyve26 conditions (see Figure 6.22).

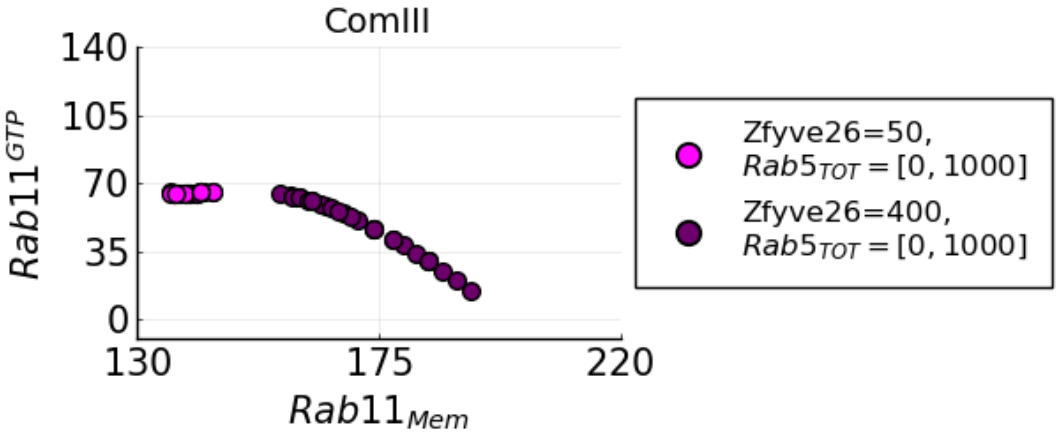


Fig. 6.21 Scatter plot of $Rab11_{Mem}$ and $Rab11^{GTP}$ obtained from model ComIII by varying the total Rab5 amount present in the system, $Rab5_{TOT}$, from 0 to 1000 molecules. Light pink dots refer to a low-effector condition (Zfyve26 = 50 molecules) and dark dots to a high-effector condition (Zfyve26 = 400 molecules).

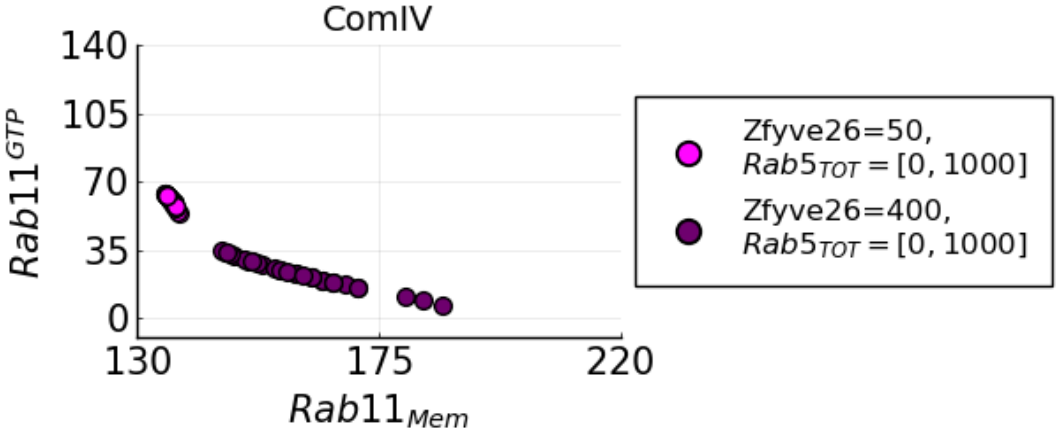


Fig. 6.22 Scatter plot of $Rab11_{Mem}$ and $Rab11^{GTP}$ obtained from model ComIV by varying the total Rab5 amount present in the system, $Rab5_{TOT}$, from 0 to 1000 molecules. Light pink dots refer to a low-effector condition (Zfyve26 = 50 molecules) and dark dots to a high-effector condition (Zfyve26 = 400 molecules).

We subsequently verified these predictions by analysing our experimental data. Although Rab5 fluorescence intensity is not available in our timelapses belonging to set 2, we assume that these endosomes present variable mean amounts of Rab5 on their membrane. Therefore, we assume that these endosomes sample a range of Rab5 values mimicking our theoretical modulation of the $Rab5_{TOT}$ amount fed to the system. We thus plotted their mean Rab11 activity (Rab11 FRET ratio value)

against their mean Rab11 abundance (fluorescence intensity) (see Figure 6.23). Interestingly, we found that the sign of the slope of this relationship matches predictions of model ComIII: positive for control (low Zfyve26) and negative for effector overexpression (high Zfyve26) condition.

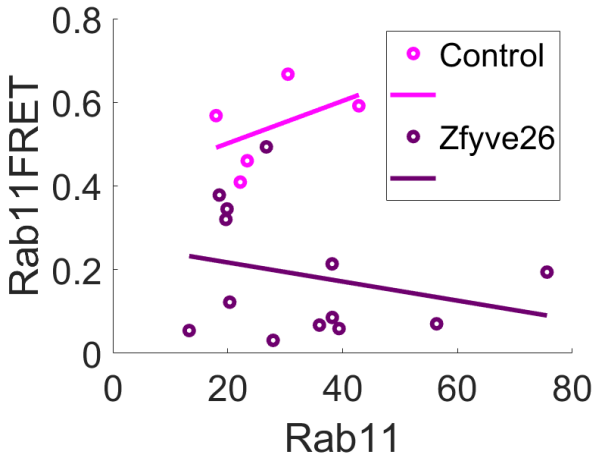


Fig. 6.23 Scatter plot of experimental mean Rab11 membrane abundance (YFP intensity) and mean Rab11 FRET ratio. Each scattered circle represents an experiment (i.e. a timelapse), either of the control condition (light pink) or the Zfyve26 overexpression condition (dark purple). Straight lines represent linear fits of points relative to the control condition (light pink) and the Zfyve26 overexpression condition (dark purple).

We thus selected model ComIII as the best model in predicting our experimental data.

In conclusion, we could discriminate between two hypotheses, which are both based on the binding between inactive Rab11 and the effector but differ in the action exerted by the Rab11-effector complex: model ComIII, where the resulting $Rab11^{GDP} : Zfyve26$ complex triggers Rab11 activation, and model ComIV, where the complex leads to Rab11 deactivation. By addressing the relationship between the amounts of active and total Rab11 present on the membrane, we selected model ComIII as the best among our hypotheses in predicting experimental data. Therefore, our findings point to a competition mechanism between Rab5 and Rab11 for binding to Zfyve26, where the effector triggers activation of both GTPases by interacting with $Rab5^{GTP}$ and $Rab11^{GDP}$.

Chapter 7

Correlations predict system saturation

7.1 Prediction of saturation effects using correlation

Altogether, our theoretical predictions and experimental data show that the competitive interaction between Rab5 and Rab11 governed by Zfyve26 generates different correlations that depend on the level of shared effector. However, our theoretical results concern the endosome as a whole, without accounting for the statistical variability that occurs across its surface even if the global mean effector abundance remains unchanged, both due to the inherent stochasticity of biochemical reactions and to the extrinsic Zfyve26 fluctuations. Indeed, the predicted GTPase abundances, activities, and correlations behaviours should vary not only from the standpoint of mean values between lower and higher Zfyve26 conditions, but should fluctuate if statistically sampled even in a fixed effector condition. In our theoretical framework, these fluctuations can be assessed by stochastically simulating our mathematical models and subsampling steady state abundances, activities and correlations from multiple subpopulations of trajectories.
We thus performed 10^5 simulations of the selected model, i.e. ComIII, with the Gillespie algorithm using fixed baseline parameter values, equal amounts of $Rab5_{TOT}$ and $Rab11_{TOT}$ ($Rab5_{TOT}$=200 molecules and $Rab11_{TOT}$=200 molecules) and fixed mean Zfyve26 amount (0, 50, 200, 400 molecules). To account for variability in the effector amount stemming from extrinsic factors, we added noise by drawing the

number of Zfyve26 molecules from a Gaussian distribution centered on its mean value (Coefficient of Variation = 0.1) at each simulation.
Then, batches of 50 trajectories were drawn from the the 10^5 simulations. For each batch, the steady-state mean abundance and activity was computed for both Rab5 and Rab11. For each GTPase, correlations between the membrane amount and the active amount ($\rho(Rab5_{Mem}, Rab5^{GTP})$ and $\rho(Rab11_{Mem}, Rab11^{GTP})$) were computed as reported in the previous chapter (see Section 5.5.1).
In this way we subsampled Rab_{Mem}, Rab^{GTP} and $\rho(Rab_{Mem}, Rab^{GTP})$ for both Rabs. With this approach we obtained their sampling distributions for the four explored Zfyve26 conditions, reported in Figures 7.1, 7.4, 7.2 and 7.5.

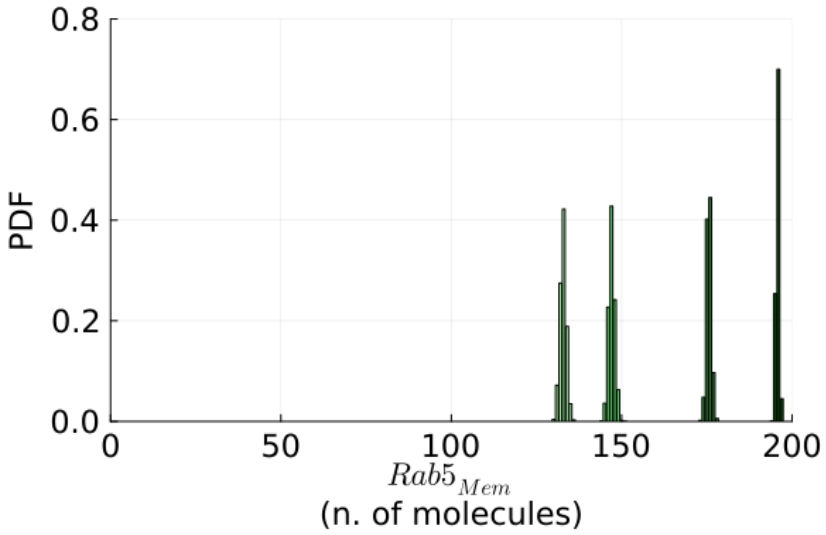


Fig. 7.1 Probability density distribution of subsampled mean $Rab5_{Mem}$ for 4 effector levels. Shades of green from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

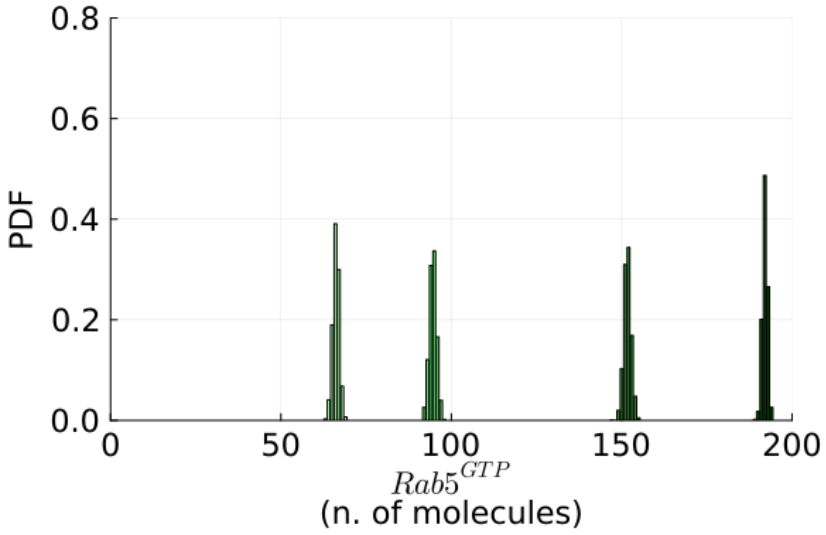


Fig. 7.2 Probability density distribution of subsampled mean $Rab5^{GTP}$ for 4 effector levels. Shades of green from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

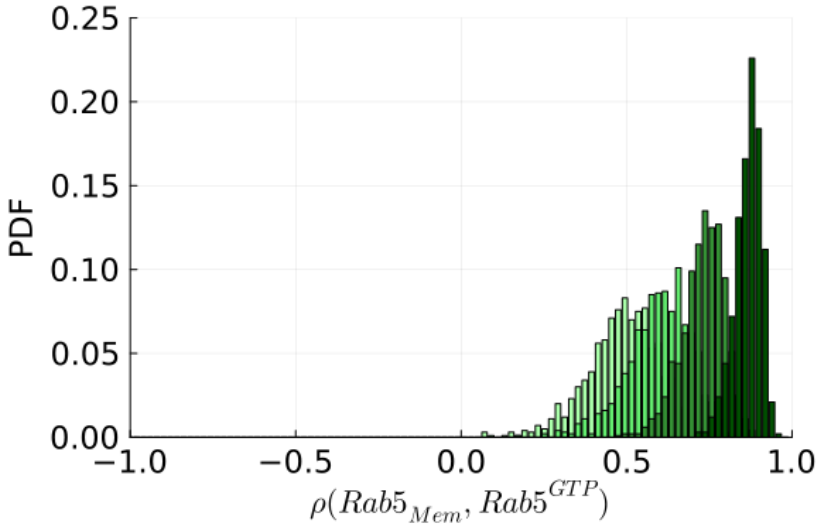


Fig. 7.3 Probability density distribution of subsampled correlation $\rho(Rab5_{Mem}, Rab5^{GTP})$ for 4 effector levels. Shades of green from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

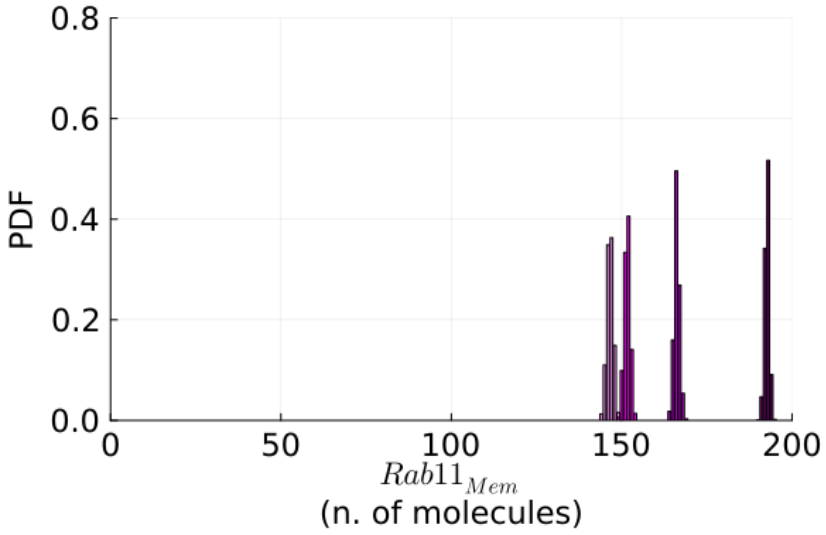


Fig. 7.4 Probability density distribution of subsampled mean $Rab11_{Mem}$ for 4 effector levels. Shades of pink from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

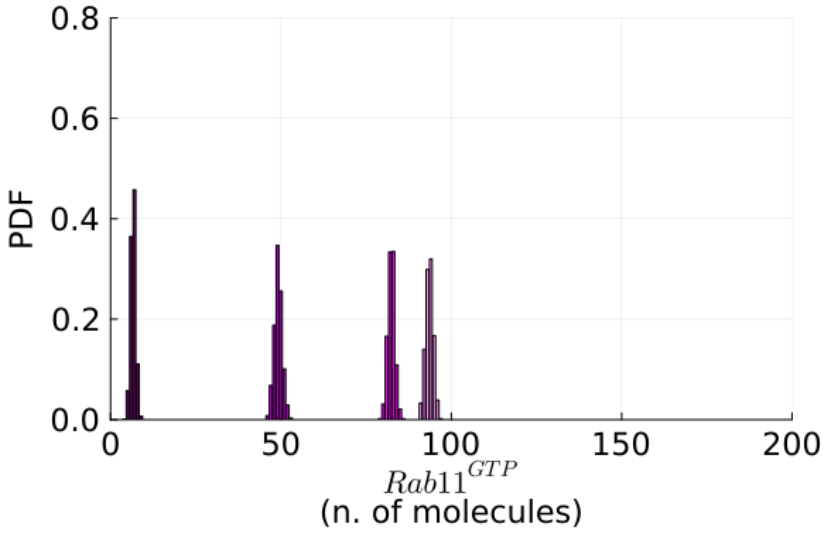


Fig. 7.5 Probability density distribution of subsampled mean $Rab11^{GTP}$ for 4 effector levels. Shades of pink from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

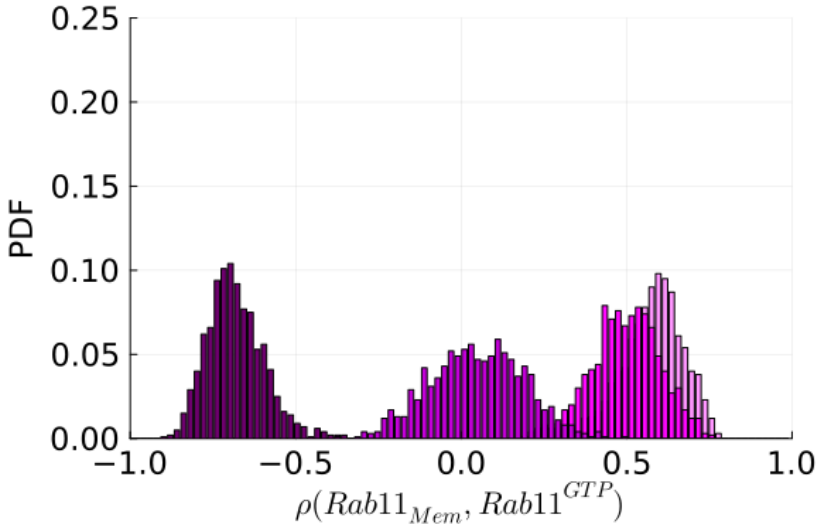


Fig. 7.6 Probability density distribution of subsampled correlation $\rho(Rab11_{Mem}, Rab11^{GTP})$ for 4 effector levels. Shades of pink from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

By relating subsampled Rab activity (Rab^{GTP}) and correlation values ($\rho(Rab_{Mem}, Rab^{GTP})$) on a scatter plot, we noticed that the two are negatively coupled for high Zfyve26 (400 molecules), while being uncorrelated in lower Zfyve26 conditions (50, 100, 200 molecules), as shown in Figures 7.7 and 7.8.

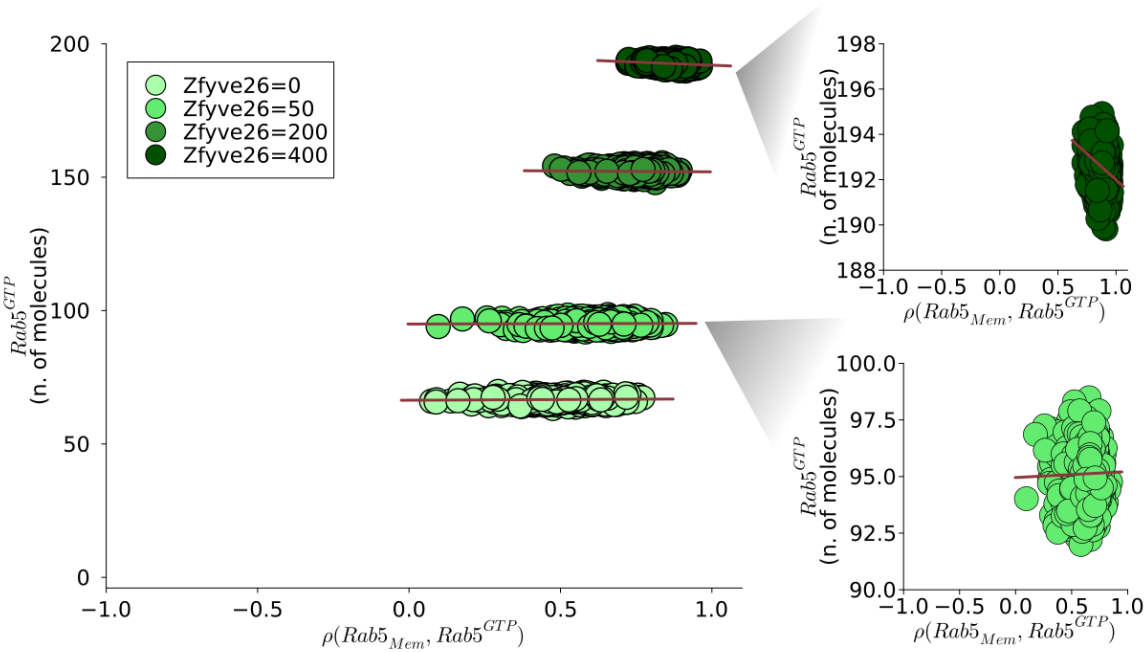


Fig. 7.7 Scatter plot of subsampled mean $Rab5^{GTP}$ and $\rho(Rab5_{Mem}, Rab5^{GTP})$ obtained for 4 effector levels (left panel). Shades of green from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules. Straight lines represent linear fits of scattered points. Magnifications (right panels) refer to Zfyve26 = 400 molecules (top) and Zfyve26 = 50 molecules (bottom).

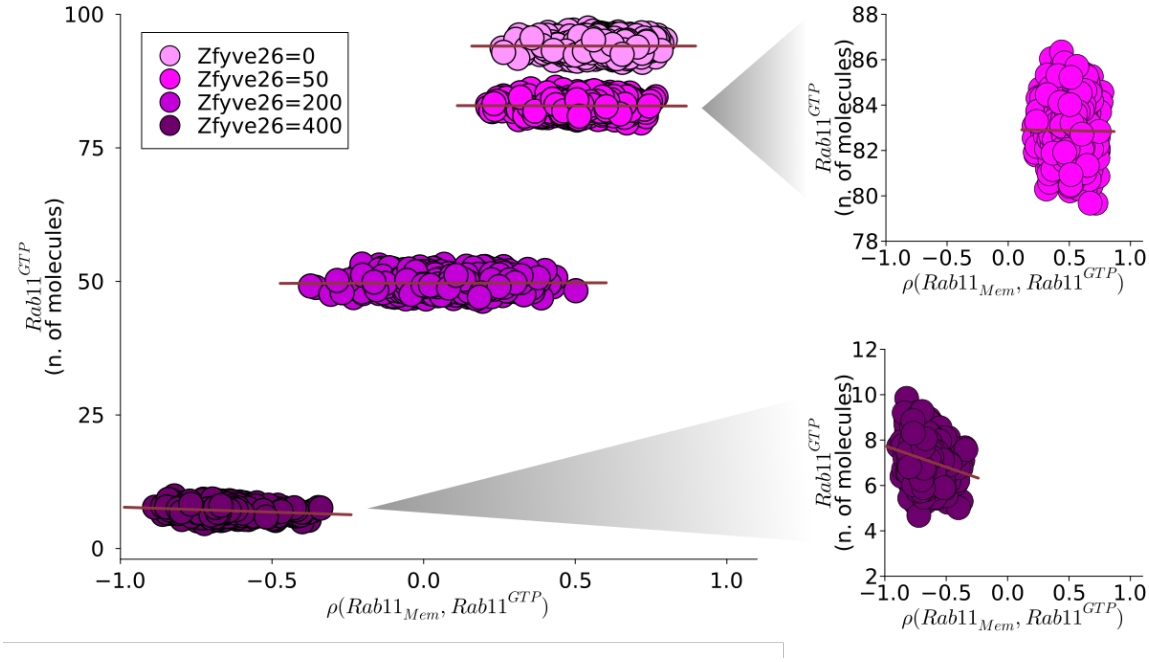


Fig. 7.8 Scatter plot of subsampled mean $Rab11^{GTP}$ and $\rho(Rab11_{Mem}, Rab11^{GTP})$ obtained for 4 effector levels (left panel). Shades of pink from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules. Straight lines represent linear fits of scattered points. Magnifications (right panels) refer to Zfyve26 = 50 molecules (top) and Zfyve26 = 400 molecules (bottom).

The coupling between Rab activity and correlation was computed for each Zfyve26 condition as:

$$c(Rab^{\mathrm{GTP}}, \rho) = \frac{\sum_{i=1}^{n}(Rab_i^{\mathrm{GTP}} - (\overline{Rab^{\mathrm{GTP}}}))(\rho_i - (\overline{\rho}))}{\sqrt{\sum_{i=1}^{n}(Rab_i^{\mathrm{GTP}} - (\overline{Rab^{\mathrm{GTP}}}))^2 \cdot \sum_{i=1}^{n}(\rho_i - (\overline{\rho}))^2}} \tag{7.1}$$

Where ρ abbreviates $\rho(Rab_{\mathrm{Mem}}, Rab^{\mathrm{GTP}})$, i identifies the batch and n represents the total number of batches.

Specifically, for Rab5 we obtained coefficient values c = 0.06, -0.02, -0.04, -0.25, and for Rab11 c = 0.00, -0.01, -0.01, -0.22 respectively for Zfyve26 = 0, 50, 200, 400 molecules, thus confirming that for high effector levels subsampled activity and correlation become negatively coupled.

This effect is a consequence of modelling the total GTPase pool as finite: when the effector saturates the entire GTPase pool, all Rabs are bound to the effector and Rab5

and Rab11 reach their maximal/minimal activity. In this situation, the magnitude of fluctuations is dampened by the presence of this superior/inferior limit given by the total amount of Rabs present in the system (see Figures 7.9, 7.10), which results in a reduced absolute correlation value.

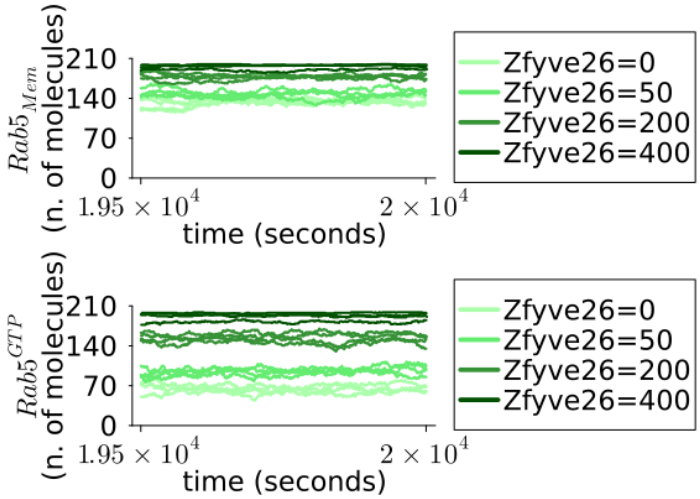


Fig. 7.9 Example of simulated temporal trajectories of $Rab5_{Mem}$ (top panel) and $Rab5^{GTP}$ (bottom panel) for 4 different levels of effector: shades of green from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

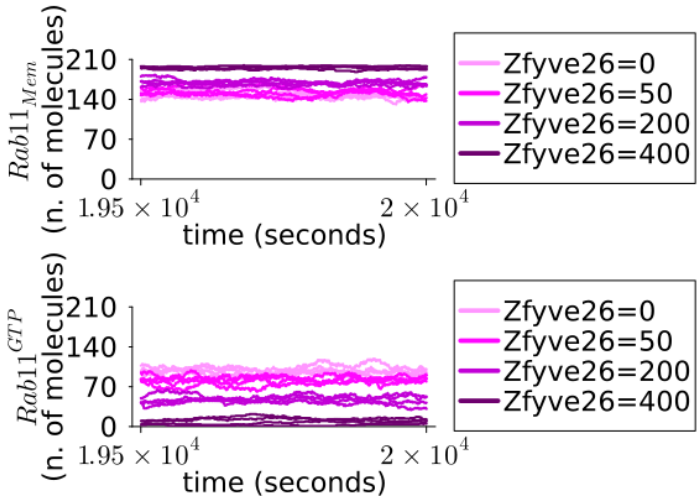


Fig. 7.10 Example of simulated temporal trajectories of $Rab11_{Mem}$ (top panel) and $Rab11^{GTP}$ (bottom panel) for 4 different levels of effector: shades of pink from the lightest to the darkest refer respectively to 0, 50, 200, 400 Zfyve26 molecules.

Thus, these predictions identify the coupling between correlation and activity as an indicator of saturation in our system.
In the next section we will illustrate the experimental validation of this saturation effect using our timelapse data.

7.2 Experimental investigation of saturation effects

To assess whether the coupling between correlation and activity is recovered in our data, we sampled Rab fluorescence intensity (Rab_{Mem}), Rab FRET ratio (Rab^{GTP}) and their correlation ($\rho(Rab_{Mem}, Rab^{GTP})$) from subregions of the endocytic membrane. With this aim, we split the ring-shaped surface into slices at each frame of our timelapses (as schematized in Figure 7.11) and we computed mean fluorescence intensities and mean FRET ratio values relative to each slice for both Rab5 and Rab11. In this way, we obtained a sampling of GTPase abundances and activities from different locations on the endocytic membrane, that we subsequently represented on bidimensional heatmaps (Figures 7.12, 7.13, 7.14, 7.15). Concomitantly, we computed slice-wise ($\rho(Rab5_{Mem}, Rab5^{GTP})$) and ($\rho(Rab11_{Mem}, Rab11^{GTP})$) to subsample correlations.
Then, to mimic our theoretical procedure, we addressed for each Rab the coupling between subsampled activity and correlation values. We computed their Pearson coefficient as:

$$c(\mathrm{FRET}, \rho(I, FRET)) = \frac{\sum_{i=1}^{n}(\mathrm{FRET}_i - (\overline{\mathrm{FRET}}))(\rho(I, FRET)_i - (\overline{\rho(I, FRET)}))}{\sqrt{\sum_{i=1}^{n}(\mathrm{FRET}_i - (\overline{\mathrm{FRET}}))^2}\sqrt{\sum_{i=1}^{n}(\rho(I, FRET)_i - (\overline{\rho(I, FRET)}))^2}} \tag{7.2}$$

where I is the Rab fluorescence intensity (corresponding to Rab_{Mem}) $FRET$ is the Rab FRET ratio (i.e. the activity corresponding to Rab^{GTP}), i identifies the slice of the endosomal membrane and n is the total number of slices.

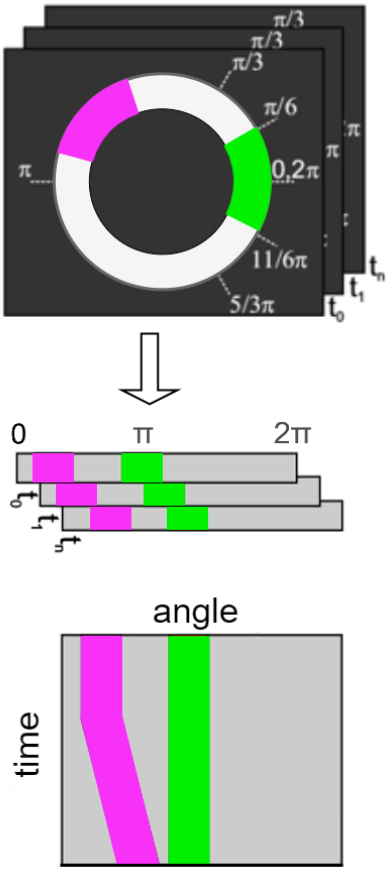


Fig. 7.11 Schematic of the procedure used to split the endosome membrane into slices and reconstruct the temporal dynamics of signals on a bidimensional heatmap.

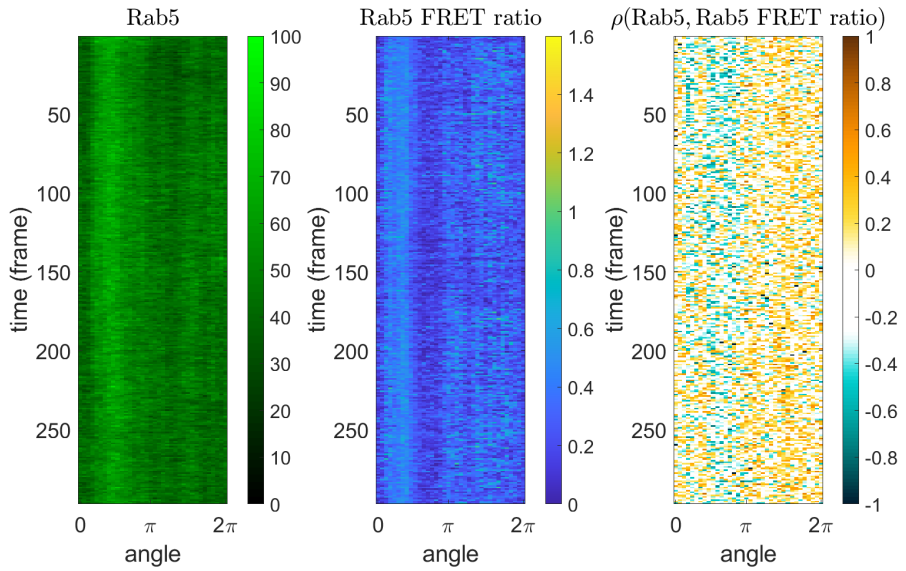


Fig. 7.12 Examples of bidimensional heatmaps referring to a single control experiment: Rab5 YFP fluorescence intensity (left panel), Rab5 FRET ratio (middle panel) and Rab5 abundance-activity correlation (right panel) as functions of angle and time (frame).

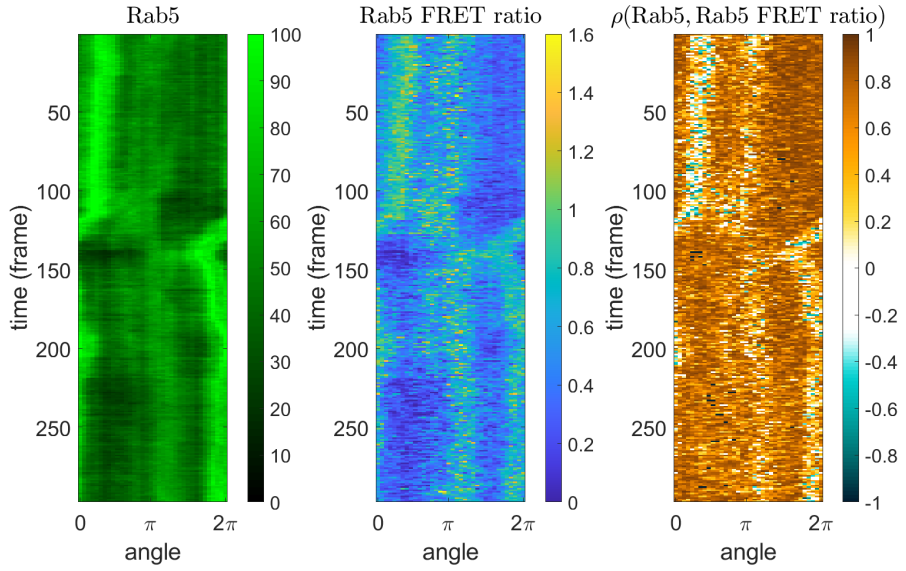


Fig. 7.13 Examples of bidimensional heatmaps referring to a single Zfyve26 overexpression experiment: Rab5 YFP fluorescence intensity (left panel), Rab5 FRET ratio (middle panel) and Rab5 abundance-activity correlation (right panel) as functions of angle and time (frame).

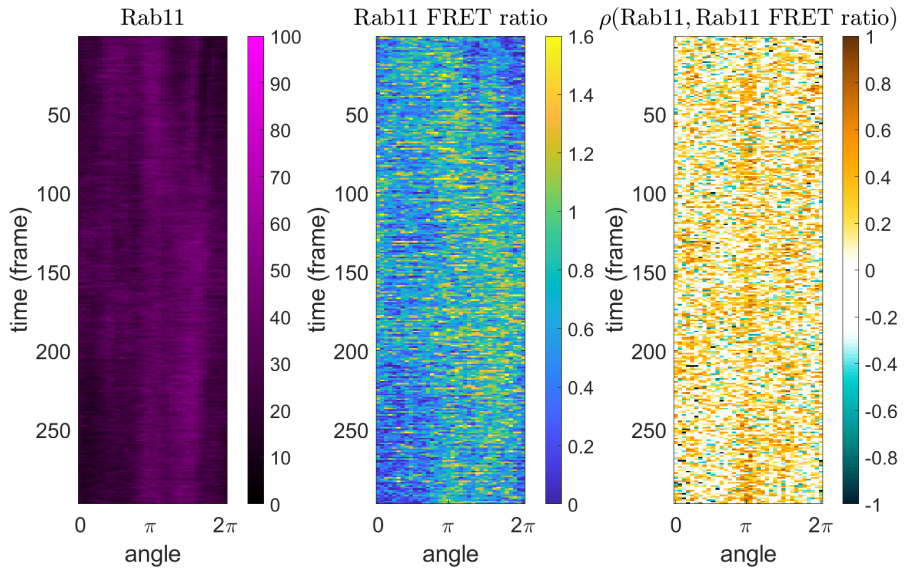


Fig. 7.14 Examples of bidimensional heatmaps referring to a single control experiment: Rab11 YFP fluorescence intensity (left panel), Rab11 FRET ratio (middle panel) and Rab11 abundance-activity correlation (right panel) as functions of angle and time (frame).

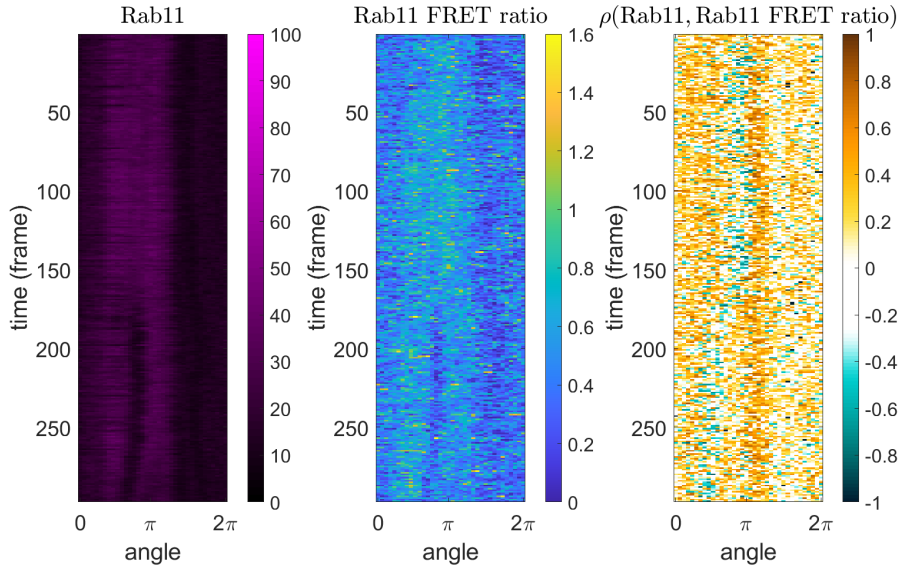


Fig. 7.15 Examples of bidimensional heatmaps referring to a single Zfyve26 overexpression experiment: Rab11 YFP fluorescence intensity (left panel), Rab11 FRET ratio (middle panel) and Rab11 abundance-activity correlation (right panel) as functions of angle and time (frame).

In agreement with our model predictions, for both Rab5 and Rab11 we found that the subsampled Rab FRET ratio (Rab^{GTP}) and subsampled correlation $\rho(I, FRETratio)$ (corresponding to $\rho(Rab_{Mem}, Rab^{GTP})$ in our models) are uncorrelated in control cells whereas anticorrelated in Zfyve26 overexpression cells (Figures 7.16 and 7.17). Specifically, we obtained mean experimental values for the coefficient c associated to Rab5: -0.05 and -0.20, respectively in control and Zfyve26 overexpression conditions (two-sided t-test p-value = 0.14). For Rab11, we obtained mean c values: -0.04 and -0.16 respectively in control and Zfyve26 overexpression condition (two-sided t-test p-value = 0.10).

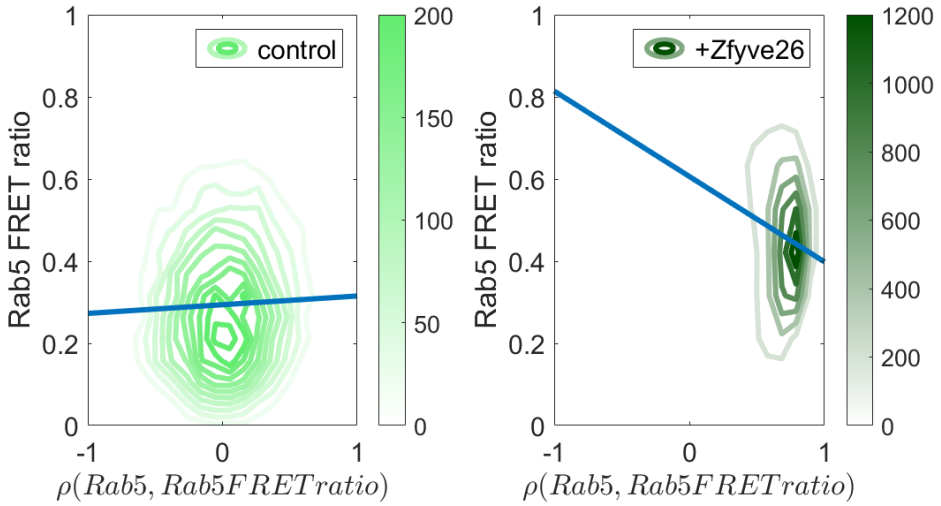


Fig. 7.16 Contour density plot of Rab5 FRET ratio and $\rho(Rab5, Rab5$ FRET ratio) density referring to a single control (light green) and a single Zfyve26 overexpression (dark green) experiment. Straight lines represent linear fits of data.

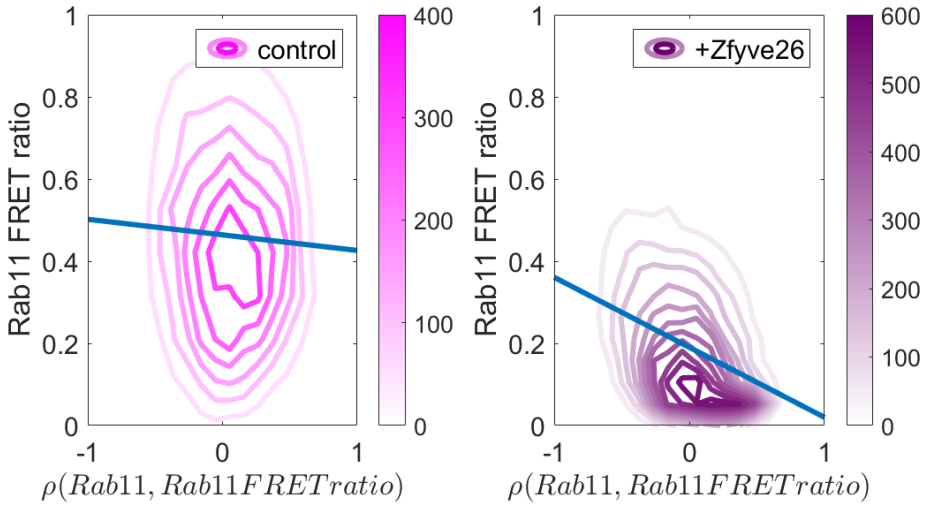


Fig. 7.17 Contour density plot of Rab11 FRET ratio and $\rho(Rab11, Rab11$ FRET ratio$)$ density referring to a single control (light pink) and a single Zfyve26 overexpression (dark purple) experiment. Straight lines represent linear fits of data.

Since in our theoretical description this effect is explained in terms of effector saturation (see previous section, 7.1), these findings indicate that the two competing GTPases are saturated on endosomes where Zfyve26 was overexpressed.
In conclusion, our results suggest the coupling between Rab activity and correlation as an indicator of saturation in our system.

Part II

Band-pass filtering in post-transcriptional gene regulation

Chapter 8

Frequency preference

The concept of frequency preference indicates the selective response of a system to particular ranges of frequencies of a signal. While some physical systems - such as those studied in electronics and acoustics - have long been associated with behaviours of frequency filtering, this phenomenon has only recently been proposed to affect biological processes. In this chapter we will introduce the concept of frequency preference borrowing some examples from both physics and biology.

8.1 Frequency preference in physical systems

The study of frequency preference and band-pass filtering behaviors plays an important role in understanding a wide range of physical phenomena where periodicity is involved. From the fundamental principles governing atomic and molecular interactions to the propagation of waves in different mediums, frequency-dependent behaviors are spread throughout the realm of physics. This section aims at giving some examples of these phenomena across the different domains of physics.
In many physical phenomena, frequency preference is tightly connected to the concept of resonance: an oscillating system responds selectively to periodic forcing with frequencies that somehow match their own. For instance, resonance can manifest as the absorption or emission of electromagnetic radiation at discrete frequencies: each atom has a set of energy levels (i.e. electron orbitals), and transitions between these levels occur when the atom absorbs or emits photons with energies matching the difference between levels. These transitions are mathematically described by the

Schrodinger equation, which yields wavefunctions describing the quantum states of electrons in an atom. The amplitude of the associated wavefunction determines the probability of finding an electron in a specific energy level: when the frequency of incident light matches the energy gap between two levels, the transition becomes dramatically more likely, leading to strong absorption or emission lines in the atom's spectrum. This is the foundation of atomic spectroscopy, a fundamental tool for indentifying chemical elements and studying the properties of atoms.

Frequency preference behaviors are present also at a molecular level, although at a greater complexity given the interactions between atoms composing the molecule. Here, resonance occurs when the frequency corresponds to a particular vibrational mode of the molecule. Consider for instance the diatomic hydrogen molecule, H_2. As the two hydrogen atoms move closer together or farther apart, the molecule vibrates at specific frequencies given by vibratonal modes that have their own quantized energy levels, and transitions between these levels occurs only at specific frequencies: when the molecule is exposed to radiation at the resonant frequencies corresponding to its vibrational modes, the probability of transitioning between these states increases significantly.

In the realm of wave propagation, the way electromagnetic waves interact with matter can be subject to frequency preference. An example is optical absorption in semiconductors, where electrons distributed among energy bands are separated by an energy gap that keeps them blocked in their band. When the semiconductor is illuminated with light, each photon carries a specific amount of energy determined by its frequency (f) according to the Planck-Einstein relation:

$$E = hf$$

where h is Planck's constant. In this case absorption exhibits frequency preference because the semiconductor primarily absorbs photons whose energy matches or exceeds the bandgap energy. Photons with energies below the bandgap are not absorbed significantly, while photons with energies at or above the bandgap energy are absorbed. This electron transition creates an electron-hole pair, contributing to electrical conductivity. The selective absorption of photons in this manner enables many devices to operate effectively within specific frequency ranges.

Frequency preference can also present itself in the form of band-pass filtering in an electric circuit. A classic example of circuit behaving as a band-pass filter is the the RLC (Resistor-Inductor-Capacitor), which is widely used in electronics for selecting

specific frequency components from an input signal. The differential response of these circuits to distinct input frequencies finds wide application for instance in the design of communication systems, fundamental in our modern world.

8.2 Frequency preference in biological systems

In biology, frequency preference was initially explored in systems that respond to physical input signals such as sound or light waves. For instance, the human auditory system is able to perceive pitch in sound thanks to a selection in frequency: our cochlea, a spiral-shaped organ in the inner ear, contains frequency-sensitive cells that vibrate preferentially in response to specific frequencies, allowing us to perceive pitch. Thus, this organ behaves as a biological frequency analyzer.
Similarly, color perception in the human visual system is sensitive to frequency: photoreceptor cells - known as cones - can be classified into different types that primarily respond to specific wavelenght ranges. Subsequently, inputs from cones are processed by the brain to create the perception of color.
Interestingly, the brain itself can behave as a band-pass filter. Neuronal frequency preference - also known as neuronal resonance - describes the tendency of a neuron to display higher responsiveness following specific input frequencies: neurons are more likely to fire in a range of preferred frequencies, and thus an input signal that matches their preferred frequency is more likely to exhibit a strong response. This preference phenomenon arises from many factors such as the neuron's membrane properties, synaptic inputs and ion channels, and has a functional significance: it allows neurons to respond to specific sensory inputs within a complex and noisy environment. Moreover, it can confer an adaptive advantage to organisms able to efficiently process specific types of stimuli.
However, frequency-dependent biological phenomena go beyond those driven by physical stimuli, and include systems where input signals themselves are biological. For instance, the ligand-receptor interaction has been predicted to exhibit maximal response when the ligand is put into the system at specific frequencies: the preferred frequencies are those that resonate with the rythmic pace of ligand binding and unbinding from the receptor [117, 118]. This simple network might provide a way to selectively amplify or attenuate certain signals, leading to precise cellular responses. To our particular interest, frequency preference behaviors have been extensively

observed in genetic networks. Genetic networks are complex systems of interaction among different genes, where a gene can affect the activity of another. In general, the activity of a gene can be oscillatory, making the level of certain proteins fluctuate rythmically over time. For instance, genes related to the circadian rythms can switch on an off in a periodic fashion, with a 24 hour periodicity. Disruptions of this frequency are known to have profound effects on an organism's well being [119]. Moreover, genetic networks can act as frequency filters, allowing specific signal frequencies to be propagated along the pathway while attenuating others. This has also been shown as an efficient way to filter noise in biological signals: cells often face sources of molecular noise, and these networks can filter out noise at specific frequencies, enabling the genetic information to be correctly transmitted.

One of the best known examples of endogenous genetic circuits exhibiting frequency preference is the p53-Mdm2 oscillator. p53 is a tumor suppressor protein - known as "the guardian of the genome" - that plays crucial roles in DNA repair and cell cycle regulation, whereas Mdm2 is a protein that targets p53 accelerating its degradation. When DNA damage occurs, p53 levels increase in the cell. Its rise triggers the transcription of Mdm2, which in turn starts to promote p53 degradation. As p53 levels turn low, also Mdm2 transcription is switched off together with its repressive action of p53, leading again to high p53 production. This interaction network generates cyclic oscillations of p53 and Mdm2.

This simple circuit is able to select input signals by their frequency: low-frequency pulses of p53 are not sufficient to lead to sustained oscillations, because they are soon dampened by Mdm2-mediated degradation. Conversely, p53 pulses with higher frequency - such as those triggered by prolonged DNA damage - lead to sustained oscillations that trigger cellular apoptosis. In other words, the high-pass filtering properties of this circuit ensure that cells with extensive DNA damage are destined to death, and are thus crucial for preventing tumor development. For this reason, p53 is considered a key tumor-suppressor.

Beyond the realm of genetic regulatory interactions, frequency preference is nowadays becoming a matter of investigation in the context of epigenetics. To our particular interest, a study published in 2014 by Kim and colleagues demonstrated that a periodic microRNA concentration with precise frequency is required for normal development in *C. Elegans* [120].

Chapter 9

Post-transcriptional gene regulation by microRNAs

9.1 MicroRNAs and their mode of action

MicroRNAs (miRNAs) are approximatley 22 nt long RNA molecules that act as inhibitors of gene activity. They are encoded in loci dispersed all over the genome, even if more than half of them reside in regions near to their target genes, called introns [121]. MiRNAs are known as "post-transcriptional" gene regulators because their inhibition of gene activity does not occur by directly preventing the target gene from being transcribed - as carried out by other regulators such as transcription factors - but after a gene has been transcribed into its messenger RNA (mRNA). Indeed, they exert their repression by binding to a gene's transcript mRNA by Watson-Crick base pairing, either at the 5' terminus (i.e. the starting region of the mRNA) or at the 3'-UTR (i.e. the final region of the mRNA) [122]. The strenght of binding depends on the chemical bonds given by the degree of complementarity between the miRNA and the mRNA. Depending on this strength, the mRNA will be subject to different kinds of repressive processes, even if the reason behind these differing actions is not clear at present: if the strength of binding is lower, the mRNA will be subject to translational repression - meaning that the miRNA will prevent the translation of the mRNA into protein, thus resulting in a lower protein level; conversely, if the strength of binding is high, the mRNA molecule will be directly degraded, thus analogously resulting in a lower protein level [123, 124]. Both ways of repression thus result in

an inhibited gene expression.
However, a miRNA and its target mRNA are required to be complementary for at least six consecutive nucleotides in order to interact [125]. This minimal sequence is often called "miRNA seed". Given the softness of this constraint, mRNAs of many different genes result as potential binding targets for a single miRNA species: indeed, a miRNA typically targets many different genes [123]. Similarly, a single gene is often targeted by multiple miRNA species [123].
Humans express over 2500 different miRNA species [124], and the last decades saw a rapidly increasing discovery of regulatory interactions involving them. Given their predictable mode of action, many of them have been uncovered by theoretical methods before being experimentally validated [126].
MiRNAs regulate fundamental biological processes ranging from enbryonic development to synaptic plasticity, and are crucially connected to diseases such as diabetes and cancer [127–129]. However, their pervasiveness and their high conservation across evolution are still partially unexplained: only a few miRNAs are demonstrated to determine phenotypes alone [130]. Indeed, the repressive action of a single miRNA is typically weak - reaching around 3- or 4-fold inhibition of protein levels (i.e. gene expression) [130].
Still, there exist exceptions such as miR-233, which was shown by M'Baya-Moutoula and collaborators to significantly determine cell metabolism and possibly affect proliferation and programmed death [131]. Some miRNAs, such as miR-34a, have been also identified to act as tumor suppressors through the regulation of immune functions [132]. Some of them also have identified as signatures of metastasis and tumor progression in combination of two, revealing that the higher metastatic risk results from joint miRNA actions [133]. Therefore, despite their individual effects are typically weak, their overall impact in gene regulatory networks appears critical, with a particular role in pathological outcomes.
MiRNA-mediated regulation has been increasingly investigated under a quantitative framework, revealing some important properties of their action. Fundamental quantitative properties of miRNA-target systems will be summarized in the next section.

9.2 MiRNAs in quantitative biology

From a mathematical standpoint, the action of miRNAs presents some interesting properties that were theoretically predicted and experimentally demonstrated in the last decades. These findings can be summarized into three main properties that will be described in the next sections: titration effect, ceRNA effect, and noise processing. Eventually, we will illustrate the novel emerging role of periodicity in miRNA-mediated regulation.

9.2.1 Titration effect

From a mathematical standpoint, the way a miRNA interacts with a target is interesting because it can introduce a nonlinear behavior in the target's response - called "titration" [134]. Titration mechanisms manifest in conditions where the interacting biochemical species are equimolar - i.e. they are present in a similar copy number - by making the system hypersensitive to small variations in their amounts: in this regime, the system is likely to fall either into a state where the target gene is completely inhibited or in a state where its activity remains unaffected by the inhibition. In the case of miRNA-mediated regulation, it has been shown both theoretically and experimentally that in a regime where a miRNA and a target mRNA are present in a similar copy number (i.e. when they are equimolar), small variations in their amounts can quickly lead the system to fall into either an unrepressed or a repressed state - meaning that the target gene's activity is either unaffected or completely silenced by the inhibitory miRNA action [135, 57, 136]. In other words, the system has been shown hypersensitive around the threshold.
Titration effects driven by miRNA-mediated regulation were originally predicted by Van Oudernaarden and colleagues [134] by means of ODE modelling. Then, the authors investigated their predictions experimentally: they exploited a miRNA - miR-20a - that is endogenously present in human embryonic kidney cells (*Hek*293 cells), and introduced in the same cells variable copy numbers of a target gene that produces a fluorescent protein. In this way they were able to measure the target gene's activity in different conditions of relative target-miRNA amounts, and they indeed confirmed that the target - measured as fluorescence intensity - presents a threshold-like behavior.
Later, Bosia, Del Giudice and coworkers [112, 113, 137] demonstrated by ODE

modelling that this hypersensitivity around the threshold can lead to bimodal distributions in the target amount. Indeed, the presence of noise drives the system either to the repressed or the unrepressed state: the prevalence of these two states leads to the spontaneous emergence of bimodality in the target activity's distribution.

9.2.2 CeRNA effect

Since many miRNA species target the same mRNA and vice versa - making the interaction combinatorial - their mode of regulation is also tighly connected to competition: different mRNA species can compete for the same regulator miRNA and similarly different miRNAs compete for binding to the same mRNA species (i.e. for repressing the same gene). As pointed out in the introduction on competitive mechanisms (chapter III), competition introduces an indirect dependence between competitors, even without their direct interaction. This coupling has been predicted to lead to correlation among targets of the same miRNA, with competitors maximally correlated in conditions of equimolarity between species [57]. Indeed, in conditions where the system is hypersensitive to small changes in molecule amounts, fluctuations in the activity of one target induce competitor targets to cross the threshold and thus cause their concomitant fluctuation.
The latter phenomenon is also identified as "competing endogenous RNA (ceRNA) effect" [138, 139].
The ceRNA effect thus marks indirect interaction among targets competing for a common pool of miRNA molecules, such as mRNAs, long non-coding RNAs [140, 141], pseudogenes [142, 143] and circular RNAs [144], which act as natural miRNA sponges, and can therefore co-regulate each other through miRNA mediation. In this competitive context, experimentally observed scenarios with interacting trios composed of one miRNA, one target mRNA and one non-coding RNA acting as a miRNA sponge have received increasing attention [145].

9.2.3 Noise processing

A recent idea is that miRNA-mediated regulation is also able to tune noise in the activity of genes [146]. Many theoretical and experimental studies suggested that miRNAs can act as fine-tuners of gene expression thanks to noise-buffering proper-

ties resulting from their way of interaction [130, 147]. In particular, the involvement of many miRNAs in double-negative feedback loops (DNFLs) - i.e. interactions of reciprocal inhibition between two molecules known to confer robustness to gene expression - supports the idea that miRNAs provide stability to phenotypes in the presence of noise. In line with this idea, Lai and coworkers showed that introducing noise into an endogenous miRNA-mediated DNFL, miRNA action functions as a noise buffer, thus conferring robustness to the network [124]. Moreover, Osella and colleagues investigated the noise buffering properties of miRNAs involved in incoherent feedforward loops (iFFLs) [148], and subsequently many studies analyzed noise processing by miRNAs in this kind of regulatory network [149–153].
However, some studies suggested that miRNAs may process noise differently depending on the expression level of genes their interact with [154], as shown for instance by Schmiedel and coworkers [130]. Moreover, other studies have highlighted that miRNA-mediated circuits can take advantage of stochasticity, rather than reducing its extent. In particular, it is thought that miRNAs can exploit noise in biological contexts where variability of gene expression is required, such as differentiation processes [155, 137].
Eventually, Garg and Sharp proposed that miRNAs may even play a role in generating cell-to-cell heterogeneity [155]. They suggested that miRNAs could enhance the variability of differentiating cells through noise in the miRNA pool - an idea that would also be consistent with the observation that miRNA expression profiles identify different phenotypes.

9.2.4 The role of periodicity in miRNA-mediated regulation

To date, periodicity has been associated to miRNA-mediated regulation mainly through the ability of miRNAs to generate gene expression oscillations. It appears that in particular feedback loops involving miRNAs - i.e. circuits of reciprocal interaction between a miRNA and a gene - can generate oscillatory gene expression [156]. For instance, sustained oscillations emerge when a transcription factor (TF) slowly activates the expression of a miRNA, that in turn rapidly inhibits the TF by repressing its translation or enhancing its degradation. Xue and colleagues [157] examined a case of interaction between the TF NF-κ B and the miRNA miR-21, and showed that this circuit can produce oscillations in the TF levels.
In general, oscillatory expression plays a crucial role in differentiation and develop-

ment, where the timing of oscillations often marks different stages of the process. An example sees miR-9 controlling the timing of oscillations of Hes1 - a fundamental TF in neuron differentiation [158]. Here, the presence of miR-9 determines whether cells commit to the mesodermal state or either differentiate as neuron stem cells [159].

Moreover, periodic gene expression associated to the circadian clock has been linked to miRNA-mediated regulation, which contributes to the robustness of oscillations. Also in this case, circadian rythms have been shown as controlled by feedback loops mediated by miRNAs [160–163].

Interestingly, miRNAs have even been found to oscillate in phase with their target genes to silence their expression oscillations when needed. For instance, Kim and collaborators [120] showed this silencing in a miRNA-mediated iFFL in *C. Elegans*: the pulsatile expression of the miRNA lin-4 dampens the expression oscullations of its target lin-14, ensuring its constant negative expression gradient necessary for a proper developmental process. This might represent a general mechanism that allows to temporarily isolate targets from upstream signals.

Despite miRNAs have displayed key roles in the generation, regulation and silencing of periodic expression patterns of target genes, little effort has been devoted to investigate the periodicity of miRNA expression itself and its implications. Our project seeks to understand if periodic miRNA expression might confer a repression advantage to miRNA-mediated regulation, and whether this kind of regulation might exhibit band-pass filtering properties.

Chapter 10

Band-pass filtering in microRNA-mediated regulation: theoretical predictions

In this chapter we investigate frequency preference behaviours driven by periodic miRNA synthesis. To this purpose, we feed a miRNA-RNA interaction model with different frequencies of periodic miRNA synthesis. For each frequency we measure the system's response in terms of RNA Fold Repression - a biologically relevant observable that quantifies the inhibitory effect of a miRNA on its target RNA.
In this way we find that periodic miRNA synthesis can outperform constant miRNA synthesis in terms of average Fold Repression. Moreover, we show that the system can exhibit frequency preference by generating higher Fold Repression for a narrow range of miRNA synthesis frequencies.
After sampling the model's parameter space in biologically relevant ranges, we analyze the role of distinct parameters in determining the appearance of such features. This analysis shows that both species' degradation parameters greatly affect the emergence of a Fold Repression advantage given by periodic miRNA synthesis, as well as the sharpness of the system's frequency preference. With these results, we provide indications for tuning experimental parameters in an artificial miRNA-target network aimed at validating these properties, which will be illustrated in the next chapter.
Next, to investigate how competition between multiple targets shapes these properties, we use a miRNA-target interaction model with two target mRNAs. With an analo-

gous parameter space sampling, we show that competition can shift the maximal Fold Repression of a target in frequency in a way that depends on the competitor's kinetics. This result highlights a potential mechanism of frequency-dependent regulation that might confer an adaptive trait to periodic miRNA-mediated regulation.

10.1 MiRNA-RNA interaction model M1 with periodic miRNA synthesis

In this model we consider one miRNA species (unbound concentration $\hat{r}$) and one RNA species (unbound concentration $\hat{R}$) with a single miRNA-binding site. Relying on a model proposed by Nordick and colleagues [164], we assume that the two species can bind to form a complex ($\hat{C}$) and both can be degraded either independently or when in complex (Figure 10.1). Thus, the eight chemical reactions included in the model are the following:

$$\varnothing \xrightarrow{s_R} \hat{R} \quad \hat{R} \xrightarrow{k_R^0} \varnothing \quad \varnothing \xrightarrow{s_r} \hat{r} \quad \hat{r} \xrightarrow{k_r^0} \varnothing \quad \hat{R}+\hat{r} \xrightarrow{k_{on}} \hat{C} \quad \hat{C} \xrightarrow{k_{off}} \hat{R}+\hat{r} \quad \hat{C} \xrightarrow{k_r^1} \hat{R} \quad \hat{C} \xrightarrow{k_R^1} \hat{r} \tag{10.1}$$

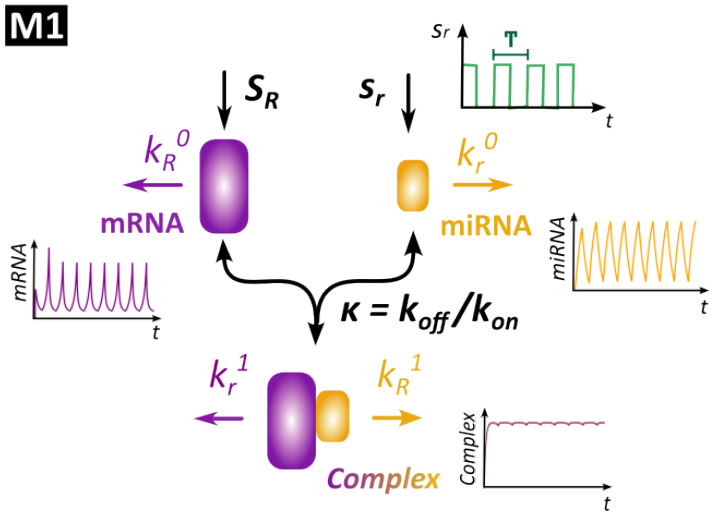


Fig. 10.1 Schematic of the miRNA-target interaction model M1

where $\hat{R}$, $\hat{r}$ and $\hat{C}$ represent respectively the RNA, the miRNA the RNA-miRNA complex.
The first two reactions represent respectively RNA synthesis and degradation, which occur with rate constants S_R and k_R^0. The third and fourth reactions describe miRNA synthesis and degradation along with their rates S_r and k_0^r. The fifth and sixth reactions represent miRNA-RNA binding and dissociation, whose rate constants are k_{on} and k_{off}. The last two reactions describe respectively miRNA and RNA degradation for the two species in complex, occurring at rates k_r^1 and k_R^1; note that these imply that the species that does not undergo degradation - either the miRNA or the RNA - is recycled back into the system, and are thus commonly called "recycling reactions".
We thereby describe the dynamics of the three molecule species by the following Ordinary Differential Equations (ODEs) according to the law of Mass Action:

$$\frac{d\hat{R}}{d\hat{t}} = s_R - k_{on}\hat{R}\hat{r} + k_{off}\hat{C} - k_R^0\hat{R} + k_r^1\hat{C} \tag{10.2}$$

$$\frac{d\hat{r}}{d\hat{t}} = s_r - k_{on}\hat{R}\hat{r} + k_{off}\hat{C} - k_r^0\hat{r} + k_R^1\hat{C} \tag{10.3}$$

$$\frac{d\hat{C}}{d\hat{t}} = k_{on}\hat{R}\hat{r} - k_{off}\hat{C} - k_R^1\hat{C} - k_r^1\hat{C} \tag{10.4}$$

To nondimensionalize the ODEs, we apply the following changes of variables:

$$\hat{t} = t/k_R^0, \quad \hat{R} = RS_R/k_R^0, \quad \hat{r} = rS_R/k_R^0, \quad \hat{C} = CS_R/k_R^0 \tag{10.5}$$

Given these changes of variables, the scaled time t results proportional to the RNA half-life: one time unit in the nondimensional equations corresponds to $\hat{t} = t_{1/2}\ln(2)$.
The nondimensionalized ODEs read as follows:

$$\frac{dR}{dt} = 1 - \kappa_{on}Rr + \kappa_{off}C - R + \beta\gamma C \tag{10.6}$$

$$\frac{dr}{dt} = \sigma - \kappa_{on}Rr + \kappa_{off}C - \gamma r + \alpha C \tag{10.7}$$

$$\frac{dC}{dt} = \kappa_{on}Rr - \kappa_{off}C - \alpha C - \beta\gamma C \tag{10.8}$$

where:

$$\sigma = s_r/s_R, \gamma = k_r^0/k_R^0, \kappa_{on} = \frac{k_{on}s_R}{{k_R^0}^2}, \kappa_{off} = \frac{k_{off}}{k_R^0}, \alpha = k_R^1/k_R^0, \beta = k_r^1/k_r^0$$

.

Here, σ represents the synthesis rate constant of miRNA relative to that of the RNA. γ is the degradation rate constant of unbound miRNA relative to that of unbound RNA. α represents the degradation rate of RNA in complex relative to degradation in its undound form. Analogously, β represents the degradation rate of the miRNA in complex relative to its unbound form.

10.1.1 Estimation of model M1 parameters

Prior to model analysis, we estimated biologically meaningful ranges of parameters. Starting from the estimation of dimensional rate constants, we then derived ranges for nondimensional model parameters as illustrated hereafter.

The median mammalian mRNA half-life in absence of post-transcriptional regulation, which corresponds to k_R^0 in our model, is estimated to be 4 hours [165]. The diameter of a typical human cell ($Hek293$) is $13\mu m$, and thus the corresponding cell volume - considering the cell as a sphere - is $1.15 \times 10^{-10}L$ [101]. Since the number of mRNA molecules per gene in a single cell can range from a few copies to tens of thousands of copies, we adopt 200 molecules as mean value, in agreement with [112] and [166]. We can therefore calculate the mean mRNA molar concentration as $\bar{R} = n_R/N_A/V = 2.9 \times 10^{-10}M$, and estimate the mRNA transcription rate as $s_R = k_R^0\bar{R} = 1.4 \times 10^{-14}Ms^{-1}$.

MiRNA half-lives are observed to be approximately four times the mRNA half-lives [167]. Therefore, we estimate the median of the scaled degradation rate $\gamma = \frac{K_r^0}{k_R^0}$ to be $1/4$. However, as miRNA half lives can vary from about 4 hours up to 48 hours [168], it is possible that some miRNAs may have shorter half lives than their target mRNAs, and we thus sample γ log-uniformly in the range $[10^{-1}/4 - 10/4]$ as suggested by [164].

For the estimation of the miRNA-mRNA association constant k_{on} we adopt ranges spanning the orders of magnitude reported in [169]: $[10^6 - 10^9]$. Since the dissociation constant $K = k_{off}/k_{on}$ was estimated to be $3.7pM$ [170], we derive a suitable range for the miRNA-mRNA dissociation rate constant as k_{off} as $[3.7 \times 10^{-6}, 3.7 \times$

$10^{-3}]s^{-1}$. Then using the estimated values of s_R and k_R^0 for scaling, we obtain ranges for nondimensional association and dissociation parameters: $\kappa_{on} = [6, 6 \times 10^3]$ and $\kappa_{off} = [7.7 \times 10^{-2}, 7.7 \times 1]$. In this way, using ranges for nondimensional parameters, we also explore scenarios where the number of mRNA molecules per cell spans smaller and greater orders of magnitude with respect to the estimated mean value of 200 molecules.

Parameters α and β were both sampled in the interval $[1/8, 16]$, estimated based on previous experimental data reported in [171, 172].

Eventually, to explore scenarios where miRNA synthesis is either faster, comparable or slower than synthesis of its target, we adopted σ values ranging from 10^-1 to 10^1.

Parameter	**Biological meaning**	**Estimated median value / range**
k_R^0	RNA degradation rate	$4.8 \times 10^{-5} s^{-1}$
s_R	mRNA transcription rate	$1.4 \times 10^{-14} M s^{-1}$
K	Dissociation constant	3.7 pM
k_{on}	Binding rate	$[10^6 - 10^9]\ M^{-1} s^{-1}$
k_{off}	Unbinding rate	$[3.7 \times 10^{-6} - 3.7 \times 10^{-3}]\ s^{-1}$

Table 10.1 Dimensional parameter ranges

Non-dim. parameter	**Biological meaning**	**Estimated range**
σ	Scaled miRNA transcription rate	$[10^{-1}, 10^1]$
κ_{on}	Scaled binding rate	$[6, 6 \times 10^3]$
κ_{off}	Scaled unbinding rate	$[7.7 \times 10^{-2}, 7.7 \times 10^1]$
α	Bound relative to unbound RNA degradation rate	[1/8, 16]
β	Bound relative to unbound miRNA degradation rate	[1/8, 16]
γ	Scaled miRNA degradation rate	$[10^{-1}/4, 10/4]$

Table 10.2 Nondimensional parameter ranges

Tables 10.1 and 10.2 report respectively estimated values of dimensional and nondimensional model parameters.

10.1.2 Modelling periodic miRNA synthesis

Periodic miRNA expression is fed to the ODE systems by modelling the relative miRNA-RNA synthesis rate as a square pulse:

$$\sigma_{pulse}(t) = \begin{cases} \sigma & \text{if } nT < t < (n+d)T, \quad n = 0,1,2,\ldots \\ 0 & \text{if } (n+d)T < t < (n+1)T, \quad n = 0,1,2,\ldots \end{cases} \tag{10.9}$$

where T represents the pulse period and d represents its duty cycle.
Note that due to the scaling of time resulting from model nondimensionalization (see section 10.1), the pulse period T and thus its frequency $f = 1/T$ are also scaled: a frequency $f = 1$ corresponds to pulses occurring every 1,44 RNA half-lives - our nondimensional time unit. Considering that the median mRNA half life corresponds to approximately 4 hours (see 10.1), a frequency $f = 1$ could typically correspond to pulses occurring roughly every 6 hours.

10.1.3 Dose conservation in model M1

Considering model M1, since the RNA is synthesized at a constant rate, its amount produced in a period T of the pulsatile input σ_{pulse} would be:

$$\int_0^T S_R \, dt = S_R \cdot T \tag{10.10}$$

where S_R is the dimensional RNA synthesis rate. Conversely, the amount of miRNA produced in one period of σ_{pulse} would depend on the duty cycle d as:

$$\int_0^T S_r \, dt = S_r \cdot T \cdot d \tag{10.11}$$

where S_r is the dimensional miRNA synthesis rate.
Therefore, considering that the scaled synthesis rate resulting from model nondimensionalization is $\sigma = \frac{S_r}{S_R}$ (see supplementary), the dose of repressor synthesized relative to that of its target in a period will be:

$$\frac{S_r \cdot T \cdot d}{S_R \cdot T} = \sigma \cdot d \tag{10.12}$$

which results independent from the period size and thus from the input's frequency. Therefore, by conserving both the nondimensional synthesis rate σ and the duty cycle d, and feeding model M1 with an integer number of input pulses, we ensure that the amount of miRNA synthesized with respect to the RNA is maintained constant as we vary the input frequency. We call the quantity $\sigma \cdot d$ "relative miRNA-RNA dose". Despite alternative strategies of dose conservation rely on varying the duty cycle d and appropriately compensating the amplitude σ, we keep both values fixed as we vary the frequency. Indeed, we adopt a fixed duty cycle $d = 0.5$ relying on the observation that endogenous gene expression pulses - such as those associated to circadian rhythms - commonly exhibit comparable ON and OFF pulse durations [173]. However, we still explore the role of pulse amplitude σ by parameter sensitivity analysis in the next section.

10.2 M1 Model analysis for frequency preference

10.2.1 Output observable: Fold Repression

RNA Fold Repression (FR) is typically employed in the experimental quantification of miRNA-mediated RNA repression; it is defined as the fold change between the basal unrepressed level of RNA - i.e. the RNA level in the absence of miRNA - and its corresponding repressed level - i.e. the RNA level reduced by miRNA action:

$$FR = \frac{\text{basal RNA level}}{\text{repressed RNA level}} \tag{10.13}$$

Since we aim at comparing the fold repression induced by periodic and constant miRNA syntheses, we compute the average fold repression achieved by both kinds of stimuli in an identical timespan. As such timespan we adopt the time needed by the constant input to induce steady state RNA level, which we call τ. We indeed consider τ - whose value is dependent on the system's kinetic parameters - as a

characteristic timescale required by the constantly stimulated system to reach its steady repression extent. In this way we address whether the average fold repression achieved by pulsatile stimuli of certain frequencies in the same timespan $[0,\tau]$ might overcome the one obtained by a constant stimulus with identical relative dose. We indicate as FR_{const} the fold repression value achieved by constant miRNA synthesis σ_{const}:

$$FR_{const} = \frac{\langle R_{basal}(t)\rangle_{[0,\tau]}}{\langle R_{const}(t)\rangle_{[0,\tau]}} \tag{10.14}$$

where $R_{basal}(t)$ represents the time-course of RNA achieved in the absence of miRNA and $R_{const}(t)$ represents the actual repressed RNA time-course induced by constant input synthesis σ_{const}.
Similarly, we indicate as FR_{pulse} the fold repression value achieved by periodic miRNA synthesis σ_{pulse}:

$$FR_{pulse} = \frac{\langle R_{basal}(t)\rangle_{[0,\tau]}}{\langle R_{pulse}(t)\rangle_{[0,\tau]}} \tag{10.15}$$

where $R_{basal}(t)$ represents again the RNA time-course achieved in the absence of miRNA and $R_{pulse}(t)$ represents the actual repressed RNA time-course induced by periodic input synthesis σ_{pulse}. Note that both fold repression measurements - FR_{const} and FR_{pulse} - are ratios between variable values and are thus independent from the scaling of variables used for model nondimensionalization.
While the value of FR_{const} depends only on the system's kinetic parameters, FR_{pulse} will depend also on the frequency of pulses and will be thus indicated as $FR_{pulse}(f)$. To compute $FR_{pulse}(f)$ as a function of the frequency f, we solve the ODE system for different frequencies of the input σ_{pulse}. Since the relative miRNA-to-RNA dose depends only on the duty cycle d and the pulse amplitude σ as long as we consider only complete pulses, the considered time interval $[0,\tau]$ must contain an integer number of periods to ensure dose conservation, thus posing a constraint on the minimum frequency value that can be explored for a fixed value of τ: the lowest possible frequency will be $f_{min} = 1/\tau$, whereas we fix the maximum frequency to $f_{max} = 10^2/\tau$.

10.2.2 Periodic miRNA synthesis can achieve higher fold repression than constant synthesis in characteristic timespan

To check whether any pulsing miRNA synthesis σ_{pulse} with a certain frequency could lead to stronger fold repression than a constant synthesis σ_{const} with identical relative dose $\sigma/2$, we chose a fixed parameter set with realistic kinetic rate values and we computed FR in the timespan $[0, \tau]$ both for the constant (FR_{const}) and the periodic ($FR_{pulse}(f)$) cases. In the periodic case, we solved the system for all frequencies ranging from $f_{min} = 1/\tau$ to an arbitrary maximum frequency $f_{max} = 10^2/\tau$.
The fixed parameter set was chosen by picking values corresponding to the median of the log-uniform distribution of ranges estimated in 10.1.1.

Interestingly, we found that some frequencies of periodic synthesis of the repressor lead to higher FR than constant synthesis (Figure 10.2). As the frequency increases, the $FR_{pulse}(f)$ curve approaches FR_{const}, showing that fold repression becomes progressively more similar to the one resulting from a constantly stimulated system.

10.2.3 Frequency response metrics: advantage, selectivity, preferred frequency

Since we are interested in comparing outcomes of constant and pulsing inputs, we define a metric A that quantifies the advantage of periodic over constant inputs as:

$$A = \frac{\int_{f_{min}}^{f_max} (FR_{pulse}(f) - FR_{const})df}{\int_{f_{min}}^{f_{max}} FR_{pulse}(f)df} \tag{10.16}$$

where f_i and f_f are respectively the lowest and the greatest frequency for which $FR_{pulse}(f)$ overcomes FR_{const}. A is bounded in the range [0, 1]. The broader the range of frequencies whose fold repression response $FR_{pulse}(f)$ overcomes that of the constant stimulus FR_{const}, and the greater the difference between $FR_{pulse}(f)$ and FR_{const}, the higher is the resulting value of A.

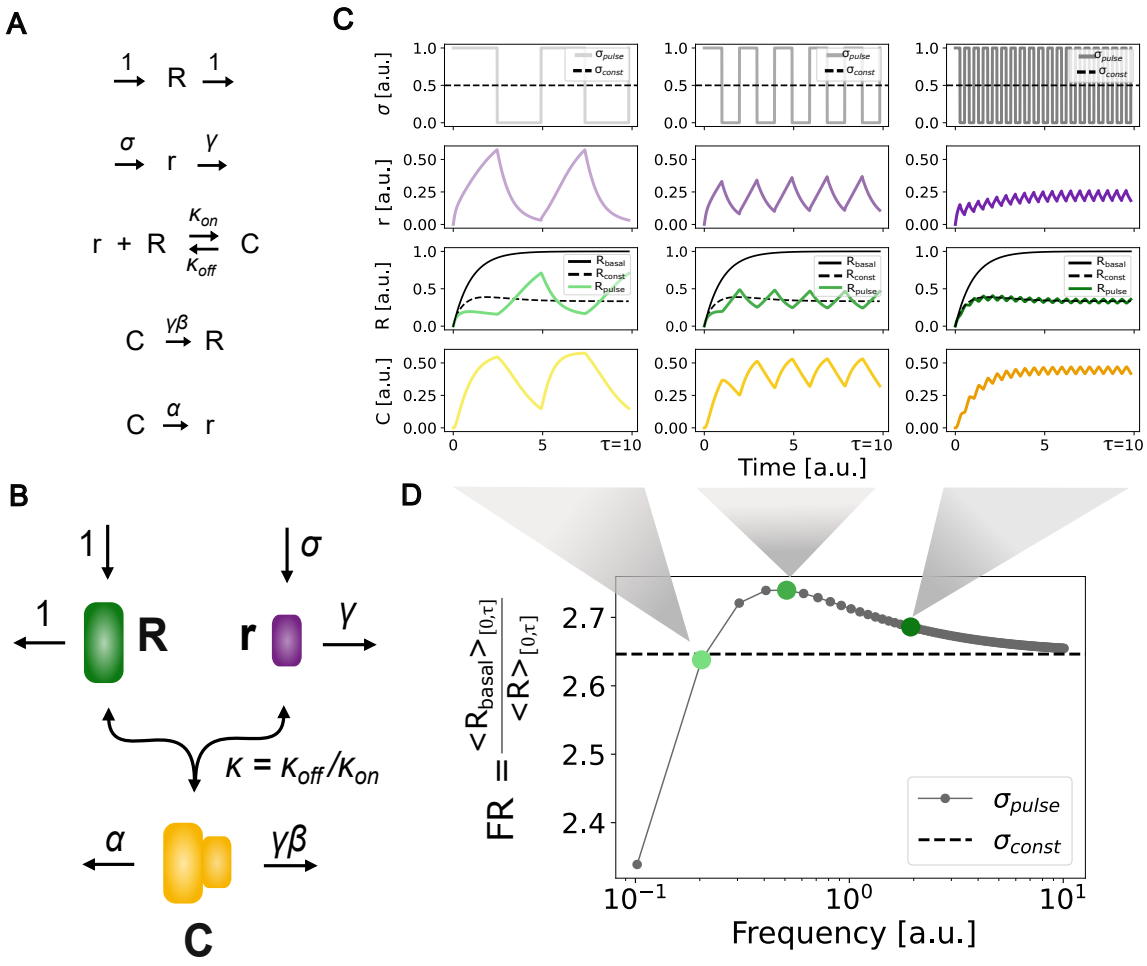


Fig. 10.2 a) Schematic of biochemical reactions employed to model microRNA-RNA interaction. b) Schematic of microRNA-RNA interaction model. c) Example of molecular species temporal trajectories. Timecourses in the interval $[0, \tau]$ are reported for three different input miRNA synthesis frequencies. d) Example of fold repression (FR) computed in the $[0, \tau]$ time interval. Green points represent FR values as a function of input miRNA synthesis pulse frequency ($FR_{pulse}(f)$), whereas the black dotted line represents the FR value reached with a constant input of identical miRNA-to-RNA relative dose (FR_{const}).

Since both the height and the width of the $FR_{pulse}(f)$ peak may change depending on the model's kinetic parameters, we use a metric proposed by Romano and colleagues [174], that quantifies an observable's selectivity S for a range of frequencies, given by:

$$S = 1 - \frac{\int_{f_{min}}^{f_{max}} FR_{pulse}(f)df}{(\log_{10}(f_{min}) - \log_{10}(f_{max}))max(FR_{pulse}(f)} \tag{10.17}$$

where f_min and f_max are the frequency range bounds.
S is also bounded in the range $[0,1]$. It considers the area and the height of the fold repression curve FR_{pulse} in relation to the frequency range covered, and thus measures the system's frequency preference regardless of comparisons with the constant synthesis case.
Also the single frequency that generates the maximal value of $FR_{pulse}(f)$ - i.e. the system's preferred frequency - may change depending on model parameters, and we compute it as:

$$f^* = argmax(FR_{pulse}(f)) \tag{10.18}$$

10.2.4 Parameter sensitivity analysis of frequency response in model M1

To investigate how parameter variations impact the frequency preference of fold repression, 10^4 parameter sets were randomly drawn from log-uniform distributions in ranges estimated in 10.1.1 using Latin Hypercube Sampling [175–177]. For each parameter set, we numerically solved the ODE system with constant miRNA synthesis rate σ_{const} and estimated the time τ needed to reach steady state using a relative tolerance threshold of 10^-6. We thus calculated the average fold repression resulting from the constantly stimulated system - FR_{const} - as defined in eq. 10.14.

Next, we defined the frequency values yielding an integer number of pulses in the range $[0,\tau]$ - a prerequisite for relative dose conservation in our framework (see section 10.1.3). Thus, for each permitted frequency we numerically solved the ODE system with periodic miRNA synthesis rate σ_{pulse} with identical relative dose, and

we subsequently computed the average fold repression as a function of frequency $FR_{pulse}(f)$ as indicated in equation 10.15.

Having the values of τ, the range of permitted frequencies $[1/\tau,\ 10^2/\tau]$, and the values of FR_{const} and $FR_{pulse}(f)$ associated to each sampled parameter set, we calculated the three metric quantities, i.e. the maximally preferred frequency f^*, the selectivity metric S and the advantage metric A, as described respectively in equations 10.18, 10.17, and 10.16.
The following grid shows examples of fold repression as a function of frequency for randomly selected parameter sets:

Then, to determine which parameters most greatly impact each metric, we calculated for each measurement - f^*, S and A - partial rank correlation coefficients (PRCC) with each nondimensional model parameter as:

$$\rho_{M,p_i\cdot(p_1,p_2,\ldots,p_{i-1},p_{i+1},\ldots,p_n)} = \frac{\rho_{M,p_i} - \sum_{j\neq i}\rho_{M,p_j}\rho_{p_i,p_j}}{\sqrt{(1-\sum_{j\neq i}\rho^2_{M,p_j})(1-\sum_{j\neq i}\rho^2_{p_i,p_j})}}$$

where:

- M indicates any of the three metric quantities, i.e., f^*, S, or A.
- ρ_{M,p_i} is the Pearson correlation coefficient between M and a specific parameter p_i.
- ρ_{M,p_j} is the Pearson correlation coefficient between M and p_j, for all $j \neq i$.
- ρ_{p_i,p_j} is the Pearson correlation coefficient between p_i and p_j, for all $j \neq i$.

Parameter	σ	κ_{on}	κ_{off}	α	β	γ
A	0.1	0.28	−0.14	0.6	−0.24	−0.57
S	0.13	0.25	−0.10	0.38	−0.26	−0.27
f^*	0.23	0.04	0.04	−0.20	0.09	0.79

Table 10.3 Partial Rank Correlation Coefficients (PRCCs) between nondimensional model parameters and the observables: fold repression advantage (A), fold repression selectivity (S), and frequency of maximum fold repression (f^*).

The analysis showed that the relative RNA degradation rate in complex, i.e. α, has the highest impact in determining the frequency selectivity of fold repression

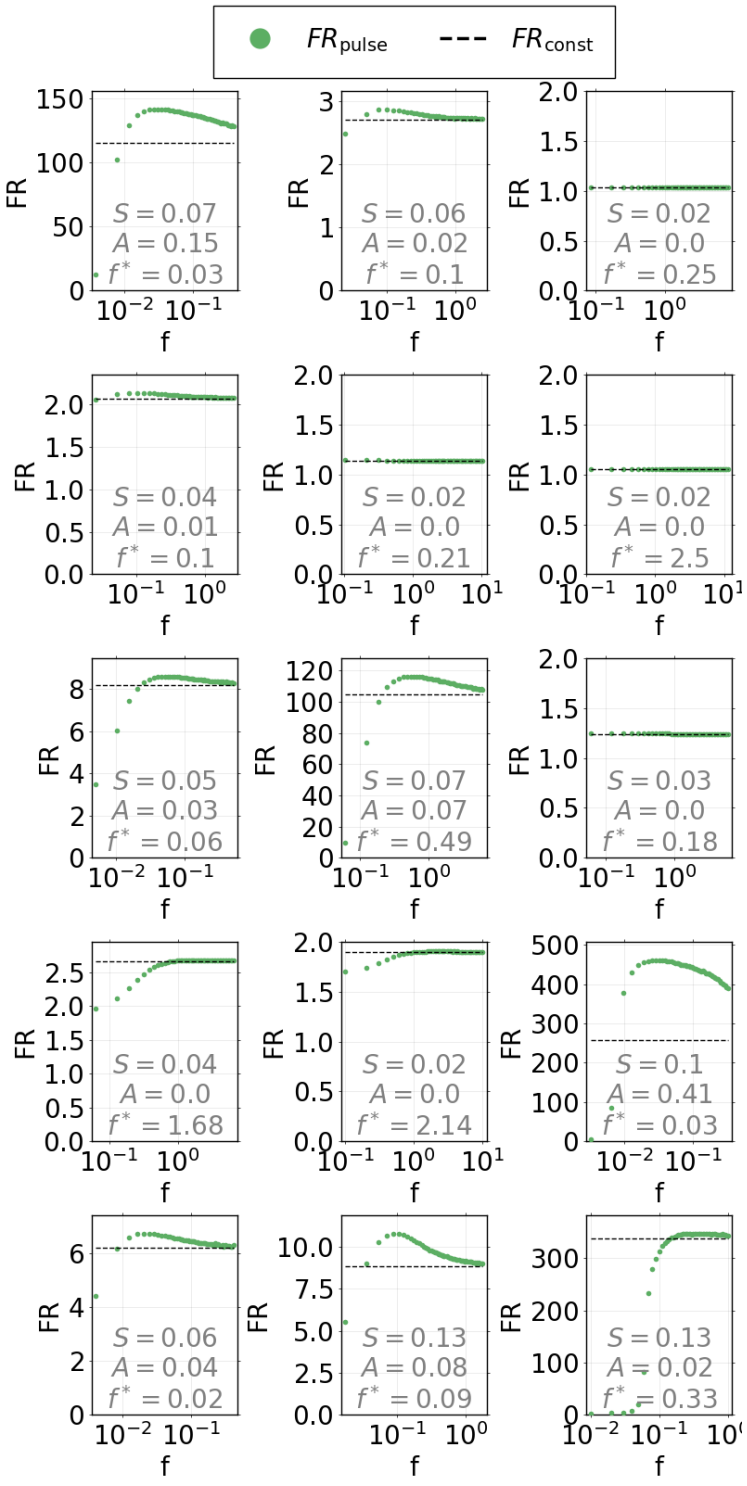


Fig. 10.3 Random sampling of fold repression computed in the $[0, \tau]$ interval for either constant or periodic miRNA input synthesis. Green dots represent $FR_{pulse}(f)$ values, whereas black dashed lines represent FR_{const} values. Each subplot refers to a parameter set sampled randomly by Latin Hypercube Sampling. For each case, advantage (A), selectivity (S) and preferred frequency (f^*) values are reported in the plot.

S, whereas degradation rates α and β both exert comparable effects to the binding scaled kinetic rate k_{on} (see table 10.3). Specifically, S resulted to be positively correlated with α and κ_{on}, whereas negatively correlated with β and γ.
We can interpret the effects dependent on the binding rate constant and on the RNA degradation in complex α as resulting from the fact that a stronger miRNA-target interaction leads to higher absolute values of fold repression, thus increasing the selectivity S. Oppositely, the reduced selectivity effect driven by γ and β can be explained by the fact that an increased decay of the miRNA leads to lower absolute fold repression values.
The metric A - that quantifies the repression advantage of periodicity - resulted correlated with each parameter in a similar way, although displaying a stronger correlation with α and γ. One can observe that A is greatly impacted by α because the fold repression difference between constant and periodic inputs depends strongly on the system's ability to respond quickly by readily exploiting the input pulses of miRNA synthesis. Conversely, a high miRNA degradation rate γ leads to lower advantage values because it reduces the absolute fold repression magnitude.
Eventually, which input frequency is the most advantageous for repression, i.e. f^*, resulted maximally correlated with the relative miRNA-RNA degradation rate γ, showing that the reciprocal timescales of repressor and target decay are crucial in determining the preferred frequency band (see 10.3). This results suggests that miRNAs with shorter lifetimes than their targets need to be synthesized with higher frequencies in order to achieve their maximal RNA fold repression, and thus confers an advantage to scenarios where the miRNA accumulates in time rather than oscillating.
The remaining parameters displayed a considerably milder impact on f^*.

10.2.5 Impact of model parameters on frequency response metrics

To further explore the dependence of advantage A on model parameters, we also computed A as a function of all possible pairs of parameters, by keeping the remaining parameters fixed. Each parameter was fixed to the median value of the log uniform distribution in the range estimated in 10.1.1, corresponding to $\sqrt{(}p_1 \cdot p_2)$ where p_1 and p_2 represent range bounds for a generic parameter p.

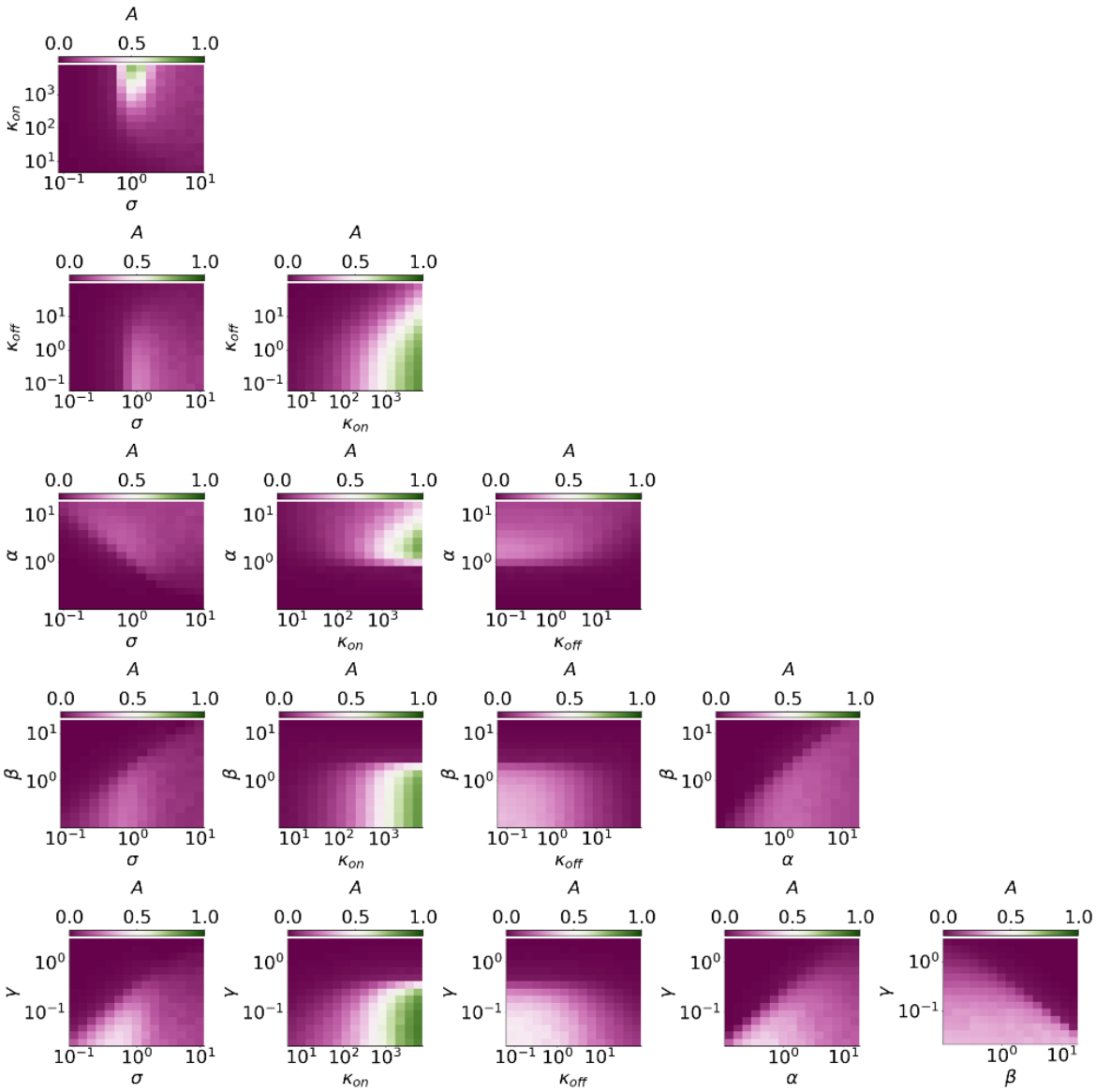


Fig. 10.4 Advantage computed as a function of parameter pairs.

Notice that parameter pairs constituted by either synthesis or degradation rates (i.e. σ, α, β, γ) concurrently generate a tradeoff behavior of A: maximal advantage values are found when the two associated parameters are reciprocally balanced. Conversely, binding and unbinding rates drive advantage effects mostly independent from the influence of remaining parameters. However, the concomitant variation of κ_{on} and κ_{off} leads to the highest advantage values when a low binding rate is coupled with a fast unbinding.

To explore more in detail the role of model parameters in determining the frequency

preference of the system, we computed the selectivity S as a function of all possible parameter pairs, keeping the rest of parameters fixed. Fixed parameter values were set to the median value of the log-uniform distribution in ranges estimated in 10.1.1. Similar to the case of advantage, maximal selectivity values are either found for concomitantly low or concomitantly high σ and γ values, thus showing that a tradeoff between any synthesis or degradation rate of the system controls frequency selectivity.

Again, effects driven by binding and unbinding rates appear independent from those generated by other parameters. Analogous to the case of advantage, the association of low binding rates to high unbinding rates generates the highest selectivity values.

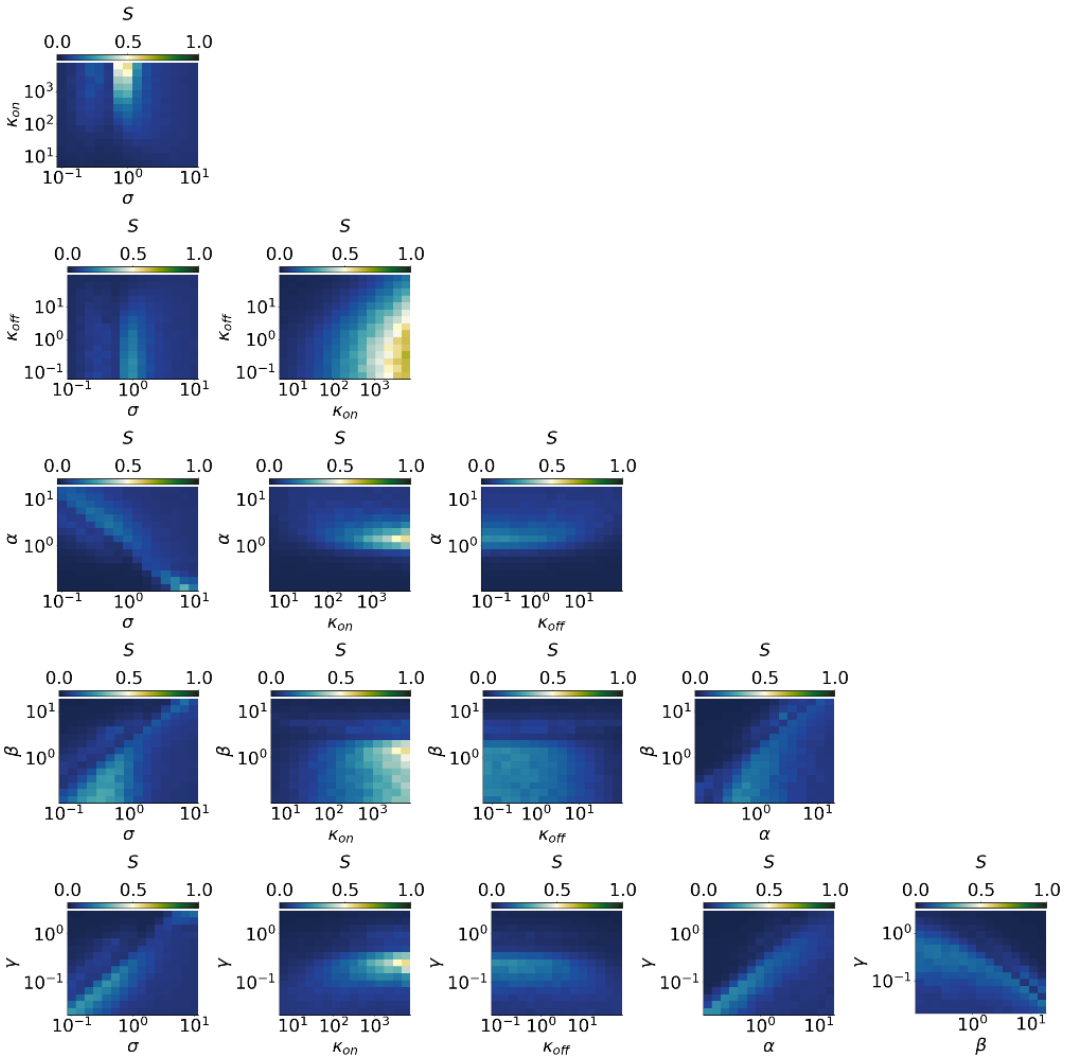


Fig. 10.5 Selectivity computed as a function of parameter pairs.

To visualize how the maximally preferred frequency depends on model parameters, we quantified also f^* as a function of parameter pairs. The remaining parameters were fixed to median values of log-uniform distributions in each range estimated in 10.1.1.

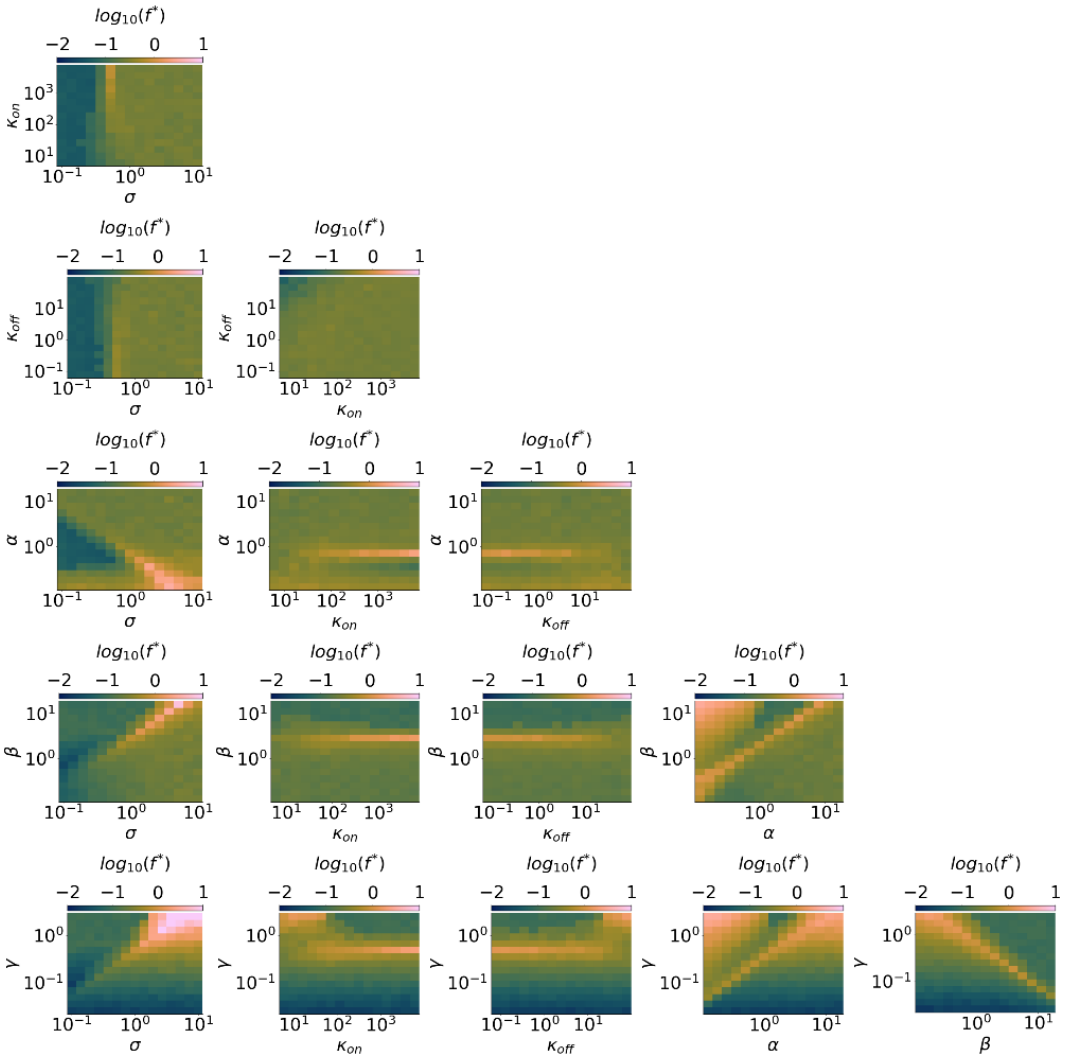


Fig. 10.6 Preferred frequency computed as a function of parameter pairs.

We notice that, similar to the first two metric quantities, the concomitant variation of synthesis and degradation parameters σ, α, β and γ generates tradeoff behaviours in the preferred frequency f^*. In particular, the association of high miRNA synthesis σ and degradation γ generates the highest preferred frequencies. Note also that a fast RNA degradation in complex α, i.e. an enhanced miRNA recycling, shifts

the preferred frequency to higher values because it favours the accumulation of the repressor.

10.2.6 Impact of initial miRNA and RNA conditions on advantage, selectivity and preferred frequency

Since the fold repression response might be affected also by initial species concentrations, we fixed model parameters to median values of log-uniform distributions in the estimated ranges and we evaluated both FR_{const} and $FR_{pulse}(f)$ starting from different miRNA and RNA concentrations. With this approach, we aim to consider different situations where either the repressor or the target or both start being expressed after a specific cellular signal and interacting with the cognate species.
In particular, we focus on four different initial conditions where: (i) both the miRNA and the target RNA start being synthesized from zero: $R_0 = 0$ and $r_0 = 0$; (ii) miRNA synthesis starts from zero and acts on a previously unrepressed RNA target, i.e. the initial RNA concentration corresponds to the non-interacting RNA steady state which is always 1 due to nondimensionalization: $R_0 = 1$ and $r_0 = 0$; (iii) conversely, the initial miRNA concentration corresponds to its maximal steady state level in the absence of target interaction, i.e. σ/γ: $R_0 = 0$ and $r_0 = \sigma/\gamma$; iv) both the miRNA and the RNA are initially expressed to their non-interacting steady state levels, corresponding respectively to σ/γ and 1, and start interacting afterwards: $R_0 = 1$ and $r_0 = \sigma/\gamma$.
By varying initial miRNA and RNA conditions in the respective ranges $r_0 = [0, \sigma/\gamma]$ and $R_0 = [0, 1]$, we explore all three situations together with intermediate conditions.

We found that both selectivity and advantage are higher if the initial miRNA and target RNA concentrations are respectively higher and lower, consistent with a stronger miRNA-mediated repression (Figures 10.7, 10.8). Similarly, the preferred frequency shifts to higher values for high initial miRNA and low initial RNA concentrations, thus privileging scenarios where the miRNA accumulates in time rather than oscillating.

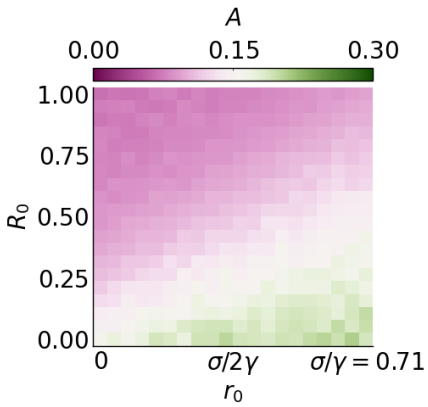


Fig. 10.7 Advantage A computed as a function of initial miRNA and RNA conditions.

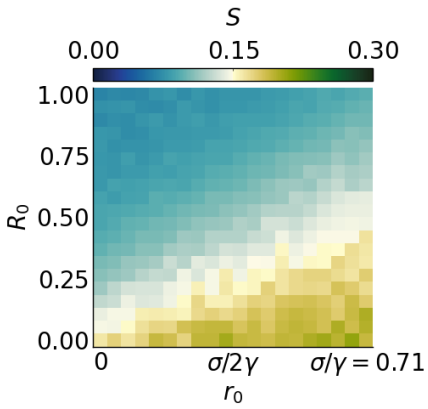


Fig. 10.8 Selectivity S computed as a function of initial miRNA and RNA conditions.

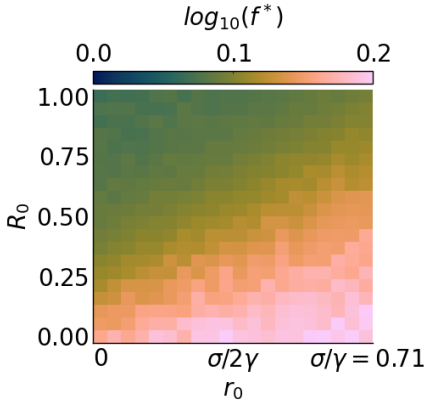


Fig. 10.9 Preferred frequency f^* computed as a function of initial miRNA and RNA conditions.

10.3 Competitive miRNA-RNA interaction model M2 with periodic miRNA synthesis

The M2 model is based on assumptions identical to model M1, but considers two target RNA species with one miRNA binding site each. Similarly to the first model, both the miRNA and the RNA species are synthesized through transcription and each RNA can form a complex with the miRNA. All species can be degraded both when unbound and when in complex, with independent miRNA and RNA degradation similar to model M1.
In this modified model we add a second RNA species able to bind to the same miRNA. We thus consider one miRNA species and two RNA species that form distinct molecular complexes with the miRNA. The chemical reactions involved in this model are the following:

$$\begin{aligned}
&\varnothing \overset{s_{R_1}}{\to} \hat{R}_1 \quad \hat{R}_1 \overset{k^0_{R_1}}{\to} \varnothing \quad \varnothing \overset{s_{R_2}}{\to} \hat{R}_2 \quad \hat{R}_2 \overset{k^0_{R_2}}{\to} \varnothing \quad \varnothing \overset{s_r}{\to} \hat{r} \quad \hat{r} \overset{k^0_r}{\to} \varnothing \\
&\hat{R}_1 + \hat{r} \overset{k_{on_1}}{\to} \hat{C} \quad \hat{R}_2 + \hat{r} \overset{k_{on_2}}{\to} \hat{C} \quad \hat{C} \overset{k_{off_1}}{\to} \hat{R}_1 + \hat{r} \quad \hat{C} \overset{k_{off_2}}{\to} \hat{R}_2 + \hat{r} \\
&\hat{C} \overset{k^1_{r_1}}{\to} \hat{R}_1 \quad \hat{C} \overset{k^1_{r_2}}{\to} \hat{R}_2 \quad \hat{C} \overset{k^1_{R_1}}{\to} \hat{r} \quad \hat{C} \overset{k^1_{R_2}}{\to} \hat{r}
\end{aligned} \tag{10.19}$$

where $\hat{R}_1$, $\hat{R}_2$, $\hat{r}$, $\hat{C}_1$ and $\hat{C}_2$ represent respectively the two RNA species, the miRNA and the two RNA-miRNA complex species.
The first two reactions represent the synthesis and degradation of the first RNA species, which occur respectively at rates s_{R_1} and $k^0_{R_1}$. Similarly, the next two reactions describe the synthesis and degradation of the second target, associated with rates s_{R_2} and $k^0_{R_2}$. The fifth and sixth reactions represent miRNA synthesis and degradation along with rates s_r and k^0_r. Next, we report reactions of binding and unbinding for both RNA targets, occurring respectively with rates k_{on_1} and k_{off_1}, and k_{on_2} and k_{off_2}. Eventually, reactions of miRNA degradation in complex (i.e. recycling of one of the two targets) occur with rates $k^1_{r_1}$ and $k^1_{r_2}$ for the two complex species, whereas those of target degradation in complex (i.e. miRNA recycling) follow respective rates $k^1_{R_1}$ and $k^1_{R_2}$.
Considering the law of mass action, we describe the dynamics of the five molecule species using the following Ordinary Differential Equations (ODEs):

$$
\begin{aligned}
\frac{d\hat{R}_1}{d\hat{t}} &= s_{R_1} - k_{on_1}\hat{R}\hat{r} + k_{off_1}\hat{C} - k^0_{R_1}\hat{R}_1 + k^1_{r_1}\hat{C}_1 \\
\frac{d\hat{R}_2}{d\hat{t}} &= s_{R_2} - k_{on_2}\hat{R}\hat{r} + k_{off_2}\hat{C} - k^0_{R_2}\hat{R}_2 + k^1_{r_2}\hat{C}_2 \\
\frac{d\hat{r}}{d\hat{t}} &= s_r - k_{on_1}\hat{R}\hat{r} - k_{on_2}\hat{R}\hat{r} + k_{off_1}\hat{C}_1 + k_{off_2}\hat{C}_2 - k^0_r\hat{r} + k^1_{R_1}\hat{C}_1 + k^1_{R_2}\hat{C}_2 \\
\frac{d\hat{C}_1}{d\hat{t}} &= k_{on_1}\hat{R}_1\hat{r} - k_{off_1}\hat{C}_1 - k^1_{R_1}\hat{C} - k^1_{r_1}\hat{C}_1 \\
\frac{d\hat{C}_2}{d\hat{t}} &= k_{on_2}\hat{R}_2\hat{r} - k_{off_2}\hat{C}_2 - k^1_{R_2}\hat{C} - k^1_{r_2}\hat{C}_2
\end{aligned}
\tag{10.20}
$$

where $\hat{R}_1$ and $\hat{R}_2$ represent concentrations of the two unbound RNA species, $\hat{r}$ describes miRNA concentration, whereas $\hat{C}_1$ and $\hat{C}_2$ describe concentrations of the two RNA-miRNA complex species.
Adopting the same approach as for the M1 model, we make the following changes of variables using the synthesis and degradation rates of the first RNA, s_{R_1} and $k^0_{R_1}$:

$$
\begin{aligned}
\hat{t} &= \frac{t}{k^0_R}, \quad \hat{R}_1 = \frac{R_1 s_{R_1}}{k^0_{R_1}}, \quad \hat{R}_2 = \frac{R_2 s_{R_1}}{k^0_{R_1}}, \\
\hat{r} &= \frac{r s_{R_1}}{k^0_{R_1}}, \quad \hat{C}_1 = \frac{C_1 s_{R_1}}{k^0_{R_1}}, \quad \hat{C}_2 = \frac{C_2 s_{R_1}}{k^0_{R_1}}
\end{aligned}
\tag{10.21}
$$

which yield the nondimensional ODE system:

$$
\begin{aligned}
\frac{dR_1}{dt} &= 1 - \kappa_{on_1}R_1 r + \kappa_{off_1}C_1 - R_1 - \gamma\beta_1 C_1 \\
\frac{dR_2}{dt} &= \delta R_2 - \kappa_{on_2}R_2 r + \kappa_{off_2}C_2 - \varepsilon R_2 - \gamma\beta_2 C_2 \\
\frac{dr}{dt} &= \sigma - \kappa_{on_1}R_1 r + \kappa_{off_1}C_1 - \gamma r + \alpha_1 C_1 - \kappa_{on_2}R_2 r + \kappa_{off_2}C_2 + \alpha_2 C_2 \\
\frac{dC_1}{dt} &= \kappa_{on_1}R_1 r - \kappa_{off_1}C_1 - \gamma\beta_1 C_1 - \alpha_1 C_1 \\
\frac{dC_2}{dt} &= \kappa_{on_2}R_2 r - \kappa_{off_2}C_2 - \gamma\beta_2 C_2 - \alpha_2 C_2
\end{aligned}
\tag{10.22}
$$

where:

$$\sigma = s_r/s_{R_1}, \kappa_{on_i} = \frac{k_{on_i} s_{R_1}}{k_{R_1}^{0\ 2}}, \kappa_{off_i} = \frac{k_{off_i}}{k_{R_1}^0}, \alpha_i = k_{R_i}^1/k_{R_1}^0, \beta_i = k_{r_i}^1/k_r^0, \gamma = k_r^0/k_{R_1}^0, \delta = \frac{s_{R_2}}{s_{R_1}}, \varepsilon = \frac{k_{R_2}^0}{k_{R_1}^0}$$

with $i = 1, 2$.

Here, σ represents the synthesis rate constant of miRNA relative to that of the first RNA target. κ_{on_i} and κ_{off_i} represent miRNA binding and unbinding rates of the i-th RNA species. α_i describes the degradation rate of the i-th RNA in complex relative to that of its unbound form. β_i is the degradation rate of miRNA in the i-th complex relative to its unbound form. γ is the degradation rate of unbound miRNA relative to that of the first RNA species. δ and ε describe respectively the synthesis and the degradation rate of the second target with respect to the first.

R_1 and R_2 represent dimensionless concentrations of the two unbound RNA species, r is the dimensionless concentration of unbound miRNA, and C_1 and C_2 represent the two molecular complexes formed with the miRNA by the two distinct RNA species.

10.3.1 Estimation of relative parameters of the two targets

For parameters σ, κ_{on_i}, κ_{off_i}, α_i, β_i and γ we adopted ranges estimated for the M1 model. For the remaining parameters δ and ε we chose the range $[10^{-1}, 10^1]$, to explore scenarios where one of the two targets is characterized by slower or faster synthesis and/or degradation kinetics with respect to the other. Parameter ranges for the nondimensional M2 model are reported in the following table, where $i = 1, 2$:

Non-dim. parameter	**Biological meaning**	**Range**
σ	Scaled miRNA synthesis rate	$[10^{-1}, 10^1]$
κ_{on_i}	Scaled binding rate	$[6, 6 \times 10^3]$
κ_{off_i}	Scaled unbinding rate	$[7.7 \times 10^{-2}, 7.7 \times 10^1]$
α_i	Bound relative to unbound RNA degradation rate	[1/8, 16]
β_i	Bound relative to unbound miRNA degradation rate	[1/8, 16]
γ	Scaled miRNA degradation rate	$[10^{-1}/4, 10/4]$
δ	Relative target synthesis rate	$[10^{-1}, 10^1]$
ε	Relative target degradation rate	$[10^{-1}, 10^1]$

Table 10.4 Nondimensional parameter ranges for model M2.

10.3.2 Dose conservation in model M2

If we now consider the modified model M2 - where the miRNA targets two distinct RNA species - σ represents the synthesis rate of miRNA relative to the first RNA, and δ represents the synthesis rate of the second RNA species relative to the first. In this case, adopting a a similar approach as for model M1, the dimensional amount of target 1 synthesized in one period of the periodic input would be:

$$\int_0^T S_{R_1}\, dt = S_{R_1} \cdot T \tag{10.23}$$

and similarly, the amount of target 2 synthesized in one period of the periodic input would be:

$$\int_0^T S_{R_2}\, dt = S_{R_2} \cdot T \tag{10.24}$$

whereas the miRNA dose synthesized in one period would be again given by eq. (10.11). Therefore, the amount of repressor synthesized in one period relative to the first RNA species would result, similar to the M1 model case:

$$\frac{S_r \cdot T \cdot d}{S_{R_1} \cdot T} = \sigma \cdot d \tag{10.25}$$

whereas its amount synthesized in one period relative to the second RNA species is:

$$\frac{S_r \cdot T \cdot d}{S_{R_2} \cdot T} = \frac{\sigma}{\delta} \cdot d \tag{10.26}$$

Thus, model M2 requires conserving both quantities $\sigma \cdot d$ and $\frac{\sigma}{\delta} \cdot d$ in order to maintain constant the dose of repressor synthesized relative to each target in a complete pulse. Therefore, since we assume a fixed duty cycle value of 0.5, we fix parameters σ and δ while varying the frequency of pulses. In the context of model M2 we will refer to the quantities $\sigma \cdot d$ and $\frac{\sigma}{\delta} \cdot d$ respectively as "relative miRNA-mRNA1 dose" and "relative miRNA-mRNA2 dose".

10.4 M2 model analysis for frequency preference

10.4.1 Output responses: Fold Repression of competing mRNAs

To analyze the competitive model's response, we use analogous fold repression measurements specific to each target, i.e. FR^i_{const} with $i = 1,2$ indicates responses generated by constant miRNA synthesis while $FR^i_{pulse}(f)$ to responses generated by periodic miRNA synthesis relative to the first and second target RNA species.
We are interested in how the frequency-dependent behavior of a target is affected by the presence of a competitor and how this influence depends on the relative kinetics of the two targets. Moreover, we are interested in kinetic features that control the mutual behavior of the two targets in terms of fold repression: we want to investigate the occurrence of cases where the two competitors are tuned to respond either to similar or to distinct frequencies of periodic miRNA synthesis, and how this synchronization/desynchronization depends on the targets' relative kinetics.
With the aim of studying the effect of adding a competitor RNA to the system, we adopt as characteristic timespan the time window $[0,\tau]$ calculated in the absence of the second target, i.e. for the single target of model M1. Thus τ corresponds to the time needed by the constant stimulus to induce steady state of the target in the absence of competition. Fold repression is then computed for both targets in such timespan as:

$$FR^i_{const} = \frac{\langle R^i_{basal}(t)\rangle_{[0,\tau]}}{\langle R^i_{const}(t)\rangle_{[0,\tau]}} \tag{10.27}$$

with the index $i = 1,2$ referring either to the first or the second target. $R^i_{basal}(t)$ represents the RNA timecourses achieved in the absence of miRNA, $R^i_{const}(t)$ represents the actual repressed RNA timecourse induced by constant input synthesis σ_{const}.
Analogously:

$$FR^i_{pulse} = \frac{\langle R^i_{basal}(t)\rangle_{[0,\tau]}}{\langle R^i_{pulse}(t)\rangle_{[0,\tau]}} \tag{10.28}$$

$R^i_{basal}(t)$ represents the RNA timecourses achieved in the absence of miRNA, and $R^i_{const}(t)$ represents the actual repressed RNA timecourse induced by periodic

input synthesis σ_{pulse}.
Since relative dose conservation is ensured as long as the timespan $[0,\tau]$ contains an integer number of periods, we use $f_{min} = 1/\tau$ and $f_{max} = 10^2/\tau$ as lower and upper bounds for the explored frequency range.

10.4.2 Frequency response metrics

We analogously compute the preferred frequencies for both RNA targets, i.e. frequencies that generate the highest values of $FR^i_{pulse}(f)$ with $i = 1,2$ as:

$$f^{*i} = argmax(FR^i_{pulse}(f)) \tag{10.29}$$

To evaluate the extent of frequency preference displayed by each target, we compute selectivities S_1 and S_2, which are always bounded in $[0,1]$ as:

$$S^i = 1 - \frac{\int_{f_{\min}}^{f_{\max}} FR^i_{\text{pulse}}(f)\,df}{(\log_{10}(f_{\min}) - \log_{10}(f_{\max}))\max(FR^i_{\text{pulse}}(f))} \tag{10.30}$$

Analogously, we compute the fold repression advantage, bounded in $[0,1]$, displayed by each target upon pulsatile stimulation with respect to the constant input:

$$A^i = \frac{\int_{f_{min}}^{f_max} (FR^i_{pulse}(f) - FR^i_{const})df}{\int_{f_{min}}^{f_{max}} FR^i_{pulse}(f)df} \tag{10.31}$$

10.4.3 Parameter sensitivity analysis of frequency response in model M2

To investigate how the kinetics of an additional target influences the fold repression of the first target as a function of frequency, we randomly sampled parameters of the competitor using Latin Hypercube Sampling [175–177]. We fixed parameters relative to the first target, i.e. R_1, and we randomly sampled the remaining parameters,

i.e. κ_{on_2}, κ_{off_2}, α_2, β_2, δ and ε. Fixed parameters correspond to median values of the loguniform distribution in ranges estimated for model M1 (section 10.1.1), whereas for the remaining parameters we drew 10^4 sets from log-uniform distributions in ranges estimated in 10.3.1. For each parameter set, we numerically solved the ODE system with constant miRNA synthesis rate σ_{const} and estimated the time τ needed by the first target to reach steady state using a relative tolerance threshold of 10^{-6}. We thus defined the frequency values yielding an integer number of pulses in the range $[0, \tau]$ to ensure relative dose conservation (see section 10.3.2). For each permitted frequency, we numerically solved the ODE system with periodic miRNA synthesis rate σ_{pulse} with the same relative dose, and we computed the average fold repression of the first target as a function of frequency, i.e. $FR^1_{pulse}(f)$ as indicated in equation 10.28. Then we fed the system with a constant miRNA synthesis rate with identical relative dose (see section 10.3.2) and computed $FR^1_{const}(f)$, as indicated in equation 10.27.
Eventually, having the values of τ, the range of permitted frequencies $[1/\tau, 10^2/\tau]$, and the values of FR_{const} and $FR_{pulse}(f)$ associated to each sampled parameter set, we calculated the three metric quantities for each target, i.e. the maximally preferred frequency f^{*1}, the selectivity metric S^1 and the advantage metric A^1, as described respectively in equations 10.29, 10.30, and 10.31.

The grid of plots in figure 10.10 shows examples of fold repression as a function of frequency for randomly selected parameter sets of the competitor target.
Then, to determine how the parameters of the competitor impact the frequency-dependent behavior of the first target, we calculated partial rank correlation coefficients (PRCC) of f^{*1}, S^1 and A^1 with each sampled parameter. Correlation results are summarized in table 10.5.

Parameter	κ_{on_2}	κ_{off_2}	α_2	β_2	δ	ε
A^1	−0.01	0.11	0.53	−0.61	−0.08	−0.51
S^1	0.01	0.09	0.51	−0.62	−0.13	−0.48
f^{*1}	−0.27	0.04	0.10	0.48	0.22	−0.21

Table 10.5 Partial Rank Correlation Coefficients (PRCCs) between nondimensional model parameters and the observables relative to the first target R_1: fold repression advantage (A^1), fold repression selectivity (S^1), and frequency of maximum fold repression (f^{*1}).

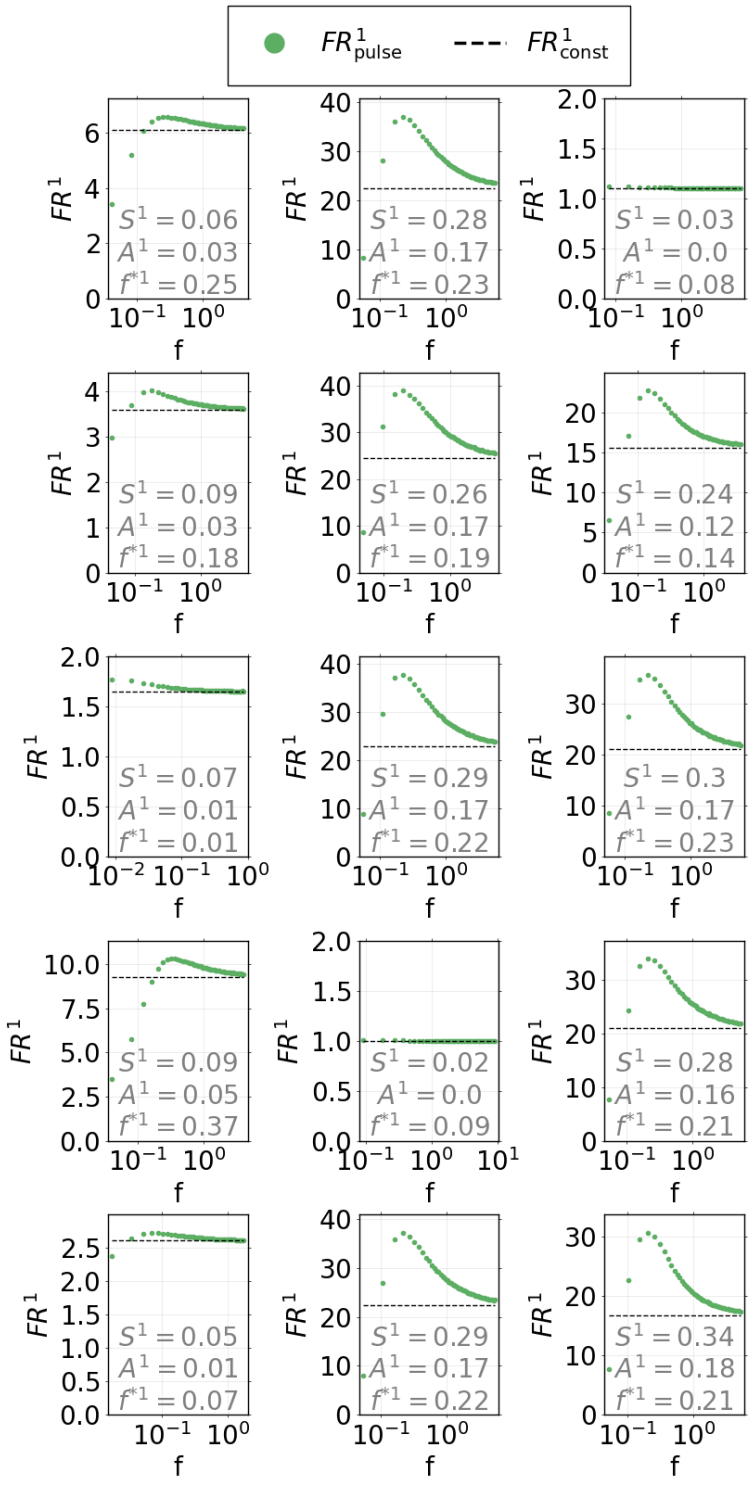


Fig. 10.10 Random sampling of fold repression computed in the $[0, \tau]$ interval for either constant or periodic miRNA input synthesis. Green dots represent $FR^1_{pulse}(f)$ values, whereas black dashed lines represent respectively FR^1_{const} values. Each subplot refers to a set of parameters of the additional target, i.e. R_2, sampled randomly by Latin Hypercube Sampling. For each case, values of the first target's advantage (A^1), selectivity (S^1) and preferred frequency (f^{*1}) are reported in the plot.

In particular, we found that both S_1 and A_1 are negatively correlated with β_2 and positively with α_2: a faster recycling of the competitor makes it dominant in the system, thus reducing the first target's possibility to be repressed by the miRNA; whereas a faster recycling of the miRNA causes the opposite effect. Interestingly, both metrics are also anticorrelated with the parameter ε, showing that selectivity and advantage of R_1 are enhanced by a longer half-life of the competitor. On the other hand, we found that the preferred frequency value of target R_1 is strongly correlated with β_2, δ and ε. Specifically, a faster degradation of the miRNA in complex (i.e. a higher value of β_2) leads to a preference of the first target's fold repression for higher frequencies. We can interpret this effect as due to the fact that a reduced availability of miRNA requires higher synthesis frequencies - and thus accumulation of the repressor - to achieve higher fold repression values. A similar result was obtained for δ, which controls the relative synthesis rate of the two competing targets: a faster synthesis of R_2 causes a reduced availability of the miRNA to the first target, thus favoring again higher synthesis frequencies that lead to the repressor's accumulation. Conversely, a faster turnover of target R_2 shifts the preference of R_1 to lower frequencies of miRNA synthesis, as displayed by the negative correlation between ε and f^{*1}.
In general, it seems that whenever the presence of a competitor becomes dominant - leaving a restricted amount of repressor available - advantage and band-pass selectivity of the first target are shrinked and its preferred frequency is concomitantly shifted to higher values. The fact that the addition of a competitor with proper kinetics can modulate the preference of a target for a specific frequency might represent an interesting regulatory property of miRNA-mediated regulation.

10.4.4 Impact of competitor RNA kinetics on frequency response metrics

To further explore the role of competitor RNA kinetic parameters in shaping a target's frequency preference, we computed the selectivity S^1, the advantage A^1 and the preferred frequency f^{*1} as a function of pairs of parameters of the competitor, keeping the rest fixed to mean range values (Figures 10.11, 10.12, and 10.13.
We found that the distinct degradation timescales of the competitor RNA tend to generate trade-off behaviours, where the first target switches from low to high advantage

and/or selectivity when varying parameters of its competitor. For instance, a high recycling of the miRNA in complex with the competitor (α_2) associated to a low enough recycling of the competitor itself (β_2) maintains A^1 and S^1 in a high-valued region. Similarly, if both β_2 and the competitor's decay relative to the first target, ε, are low enough, advantage and selectivity are maintained high. The first target's preferred frequency, i.e. f^{*1}, presents a similar trade-off behaviour: the highest frequency values are found outside the high advantage and selectivity region, where the balance between degradation channels of the competitor target is lost.
On the other hand, binding/unbinding parameters of the competitor target can generate optimality in frequency preference metrics: if κ_{on2} is high enough, the fold repression of the first target becomes more selective at critical values of ε. Thus, if the competitor RNA is slightly more unstable than the first, we achieve the highest S^1. Note that in this optimal selectivity region the system prefers lower frequency values. A similar optimality of S^1 is displayed at low κ_{off2} values (Figure 10.11). Note also that the highest advantage values do not necessarily appear in the same region, as the fold repression given by periodic miRNA synthesis can greatly outperform the one given by constant synthesis despite not presenting a sharp preference for a range of frequencies.
In general, these results thus suggest that a competitor RNA's half-life and its affinity to the miRNA can be tuned to achieve an optimal frequency-driven repression on a first target RNA in its characteristic timespan.

10.4.5 Differential target RNA regulation by desynchronization in frequency

Eventually, we are interested in addressing the conditions where two competitors might show a preference for different input frequencies in the characteristic timespan of one target. We thus sampled the complete parameter space using Latin Hypercube sampling. We drew 10^4 sets from log-uniform distributions in the estimated ranges. For each parameter set, we first numerically solved the ODE system with constant miRNA synthesis rate σ_{const} to estimate the time needed by the first RNA to reach the steady state, τ, using a relative tolerance threshold of 10^{-6}. We then defined the frequency values that yield an integer number of pulses in the range $[0, \tau]$ to ensure relative dose conservation (see section 10.3.2). For each frequency, we numerically solved the ODE system with periodic miRNA synthesis rate σ_{pulse}, and we computed

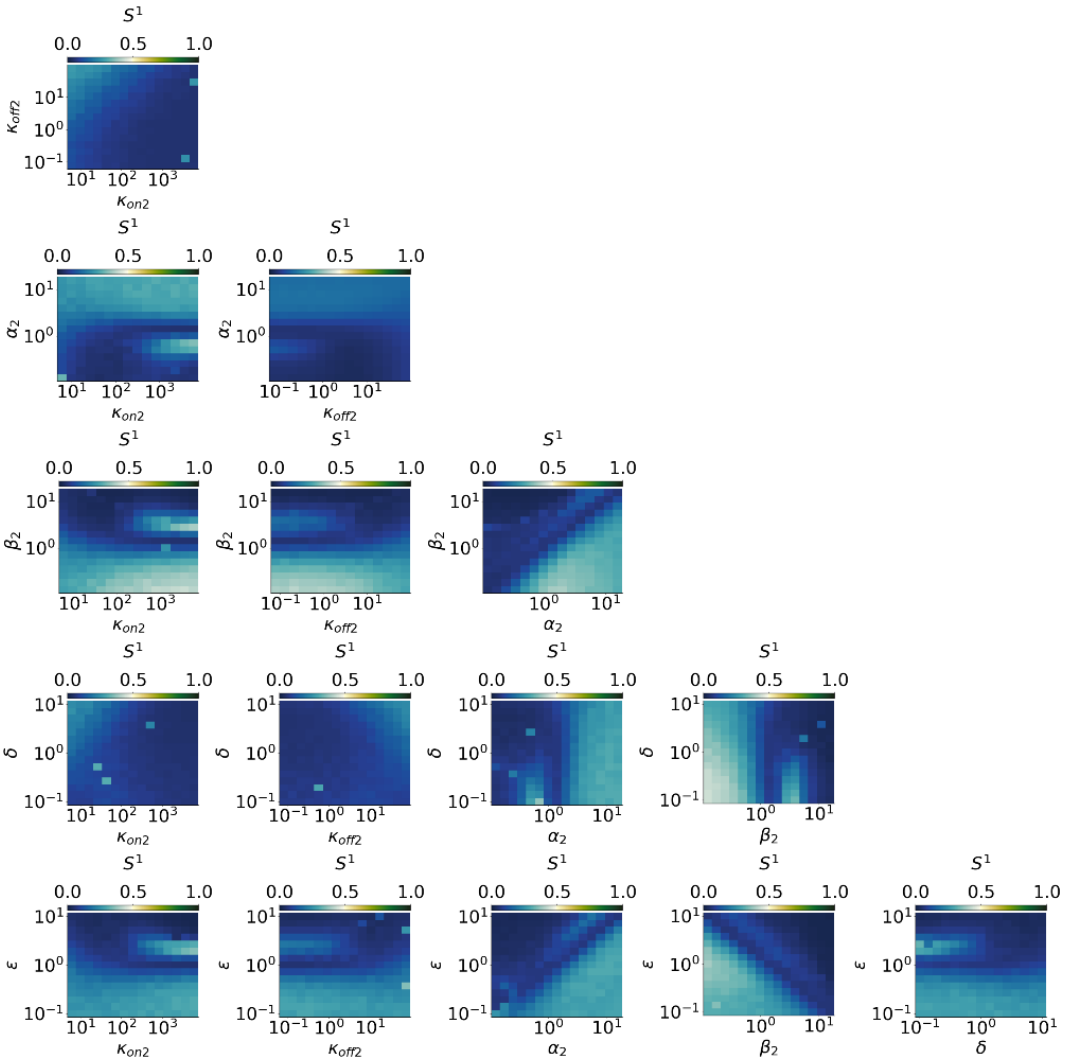


Fig. 10.11 Selectivity of target 1 computed as a function of parameter pairs of the competitor - target 2.

the average fold repression curves of both targets as a function of frequency, i.e. $FR^1_{pulse}(f)$ and $FR^2_{pulse}(f)$, as well as the corresponding fold repression given by constant synthesis, FR^1_{const} and FR^2_{const} (see grid of plots in Figure 10.14).
Thus, to capture situations where the fold repression of the two competing targets is maximal at distinct frequencies, we defined a metric quantity to capture cases where their selectivities S^1 and S^2 are simultaneously high and their preferred frequencies f^{*1} and f^{*2} are distant, as:

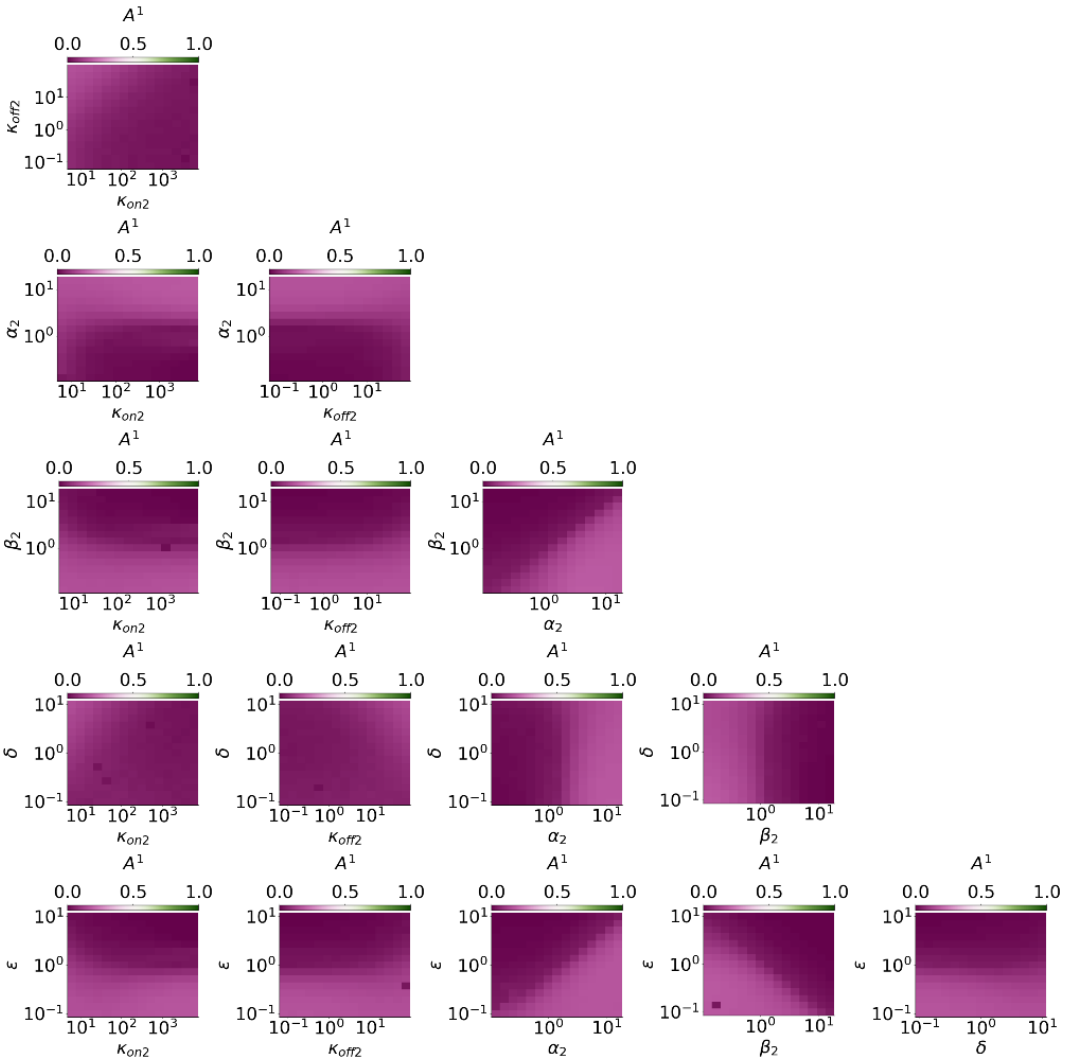


Fig. 10.12 Advantage of target 1 computed as a function of parameter pairs of the competitor - target 2.

$$DS = \frac{2}{\frac{1}{S^1} + \frac{1}{S^2}} \cdot [\log_{10}(f^{*2}) - \log_{10}(f^{*1})] \tag{10.32}$$

This metric involves the harmonic mean of selectivities, given by $\frac{2}{\frac{1}{S^1}+\frac{1}{S^2}}$, which privileges cases where both S^1 and S^2 are similarly high, penalizing cases where the two selectivities differ greatly. Given the factor $\log_{10}(f^{*2}) - \log_{10}(f^{*1})$, negative or positive DS values will result depending on the relative position in frequency of the maximal fold repression of the two target RNAs.

To investigate how the differential regulation of the two targets in frequency is

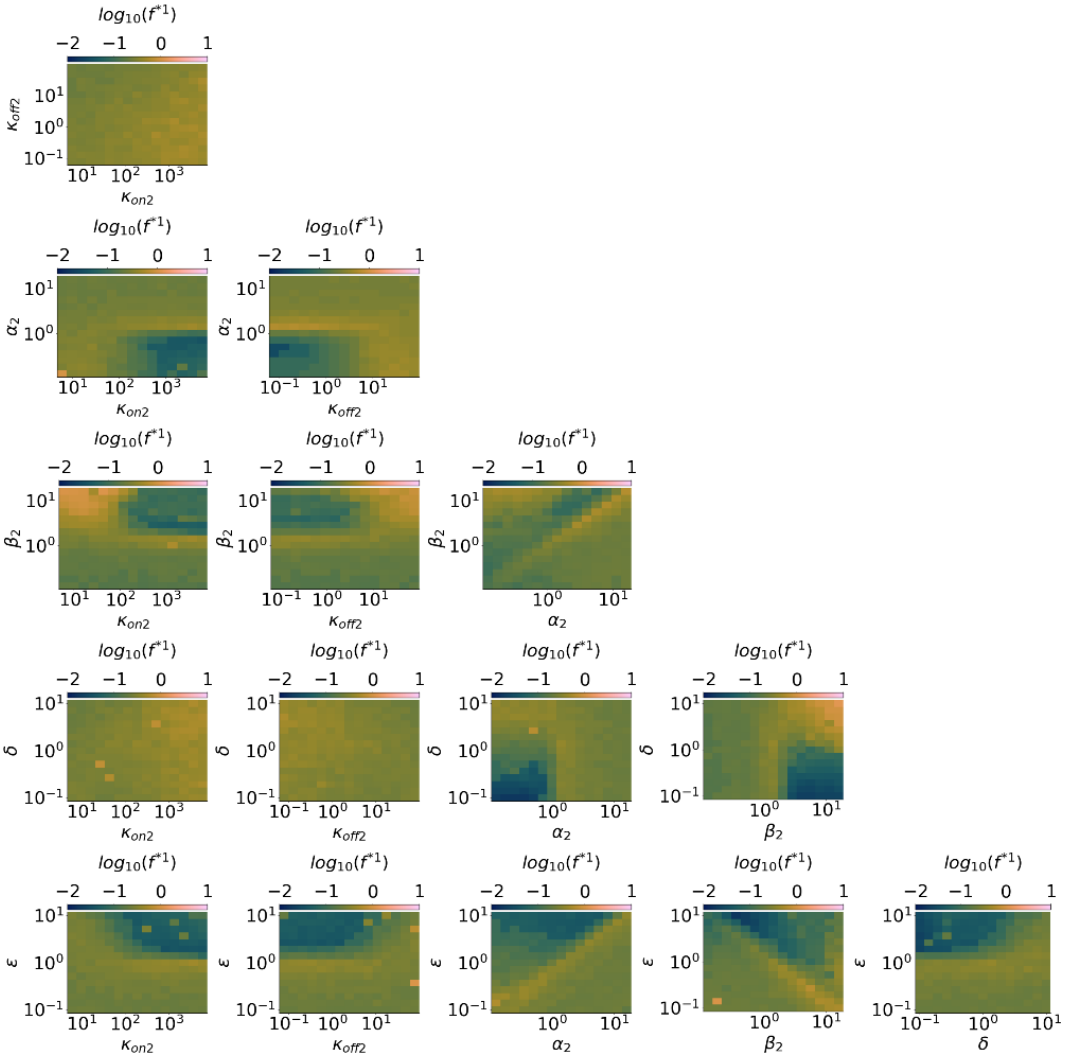


Fig. 10.13 Preferred frequency of target 1 computed as a function of parameter pairs of the competitor - target 2.

affected by their relative kinetics, we calculated PRCCs (see results in Table 10.6) between the metric quantity *DS* and the following parameter ratios: $\frac{K_1}{K_2} = \frac{\kappa_{on2}}{\kappa_{on1}} \cdot \frac{\kappa_{off1}}{\kappa_{off2}}$, $\frac{\alpha_2}{\alpha_1}$, $\frac{\beta_2}{\beta_1}$, δ and ε.

Parameter	K_2/K_1	α_2/α_1	β_2/β_1	δ	ε
DS	0.40	−0.19	−0.04	0.06	−0.12

Table 10.6 Partial rank correlation coefficients (PRCCs) between parameter ratios of the two target RNAs and the differential selectivity DS.

Interestingly, we found that *DS* presents the strongest correlation with the relative dissociation constant of the two targets, K_2/K_1, showing that the separation of miRNA binding timescales of the two targets is important for achieving differential target repression driven by frequency. We also found both the ratio between α_2 and α_1 as well as the relative degradation of the two targets, ε anticorrelated with *DS*.
These findings hint at the possibility of exclusive target regulation driven by periodic miRNA synthesis, suggesting that the experimental modulation of kinetic rates - such as the miRNA-target affinity of the two competitors - can be tuned to achieve this exclusivity.
Given the role of the relative kinetics of the two targets in determining the differential selectivity, we also explore the behaviour of such metric, *DS*, varying in pairs the parameters that presented the highest correlation values. We thus compute *DS* as a function of κ_{on2} and α_2, keeping the rest of parameters fixed, including the binding and unbinding rates of the first RNA target (Figure 10.15). In this way we find that for either low or high miRNA affinity to the second target R_2, κ_{on2}, a narrow range of α_2 values generates the highest absolute values of *DS* - either positive or negative. Similarly, by varying κ_{on2} and the relative decay rate of the two targets ε, we observe that in a narrow range of ε values the differential selectivity *DS* takes the highest absolute values, either negative or positive (Figure 10.16).

We notice that this behaviour appears in a range of ε and α_2 values where the kinetic parameters of the two targets do not differ greatly. This suggests that slight differences in the half-lives or in the degradation induced by miRNA binding of the two competitors can lead to their differential regulation in frequency, provided that their affinities to the repressor are dissimilar enough.
These insights are aimed at providing suggestions for an experimental modulation of target kinetics that might focus on the binding affinity between a miRNA and its distinct targets and/or on their relative stabililty.

10.5 Conclusions

In conclusion, our predictions suggest that periodicity in post-transcriptional repression can represent a versatile mode of regulation of gene expression. In particular, if compared to constant repressor synthesis, it can achieve variable extents of regulation

controlled by the frequency of pulses, reaching even a stronger repression. Moreover, it can generate preferential repression for a narrow range of frequencies, exhibiting a band-pass filtering behaviour. In the presence of competitor target RNAs, this preference can be modulated and shifted in frequency, in a way that depends on their kinetics, and on particular on the system's timescales of degradation.
Eventually, we suggest that periodic miRNA synthesis might enable the exclusive regulation of a target even in the presence of competitors, thanks to the selective response to a specific frequency range.

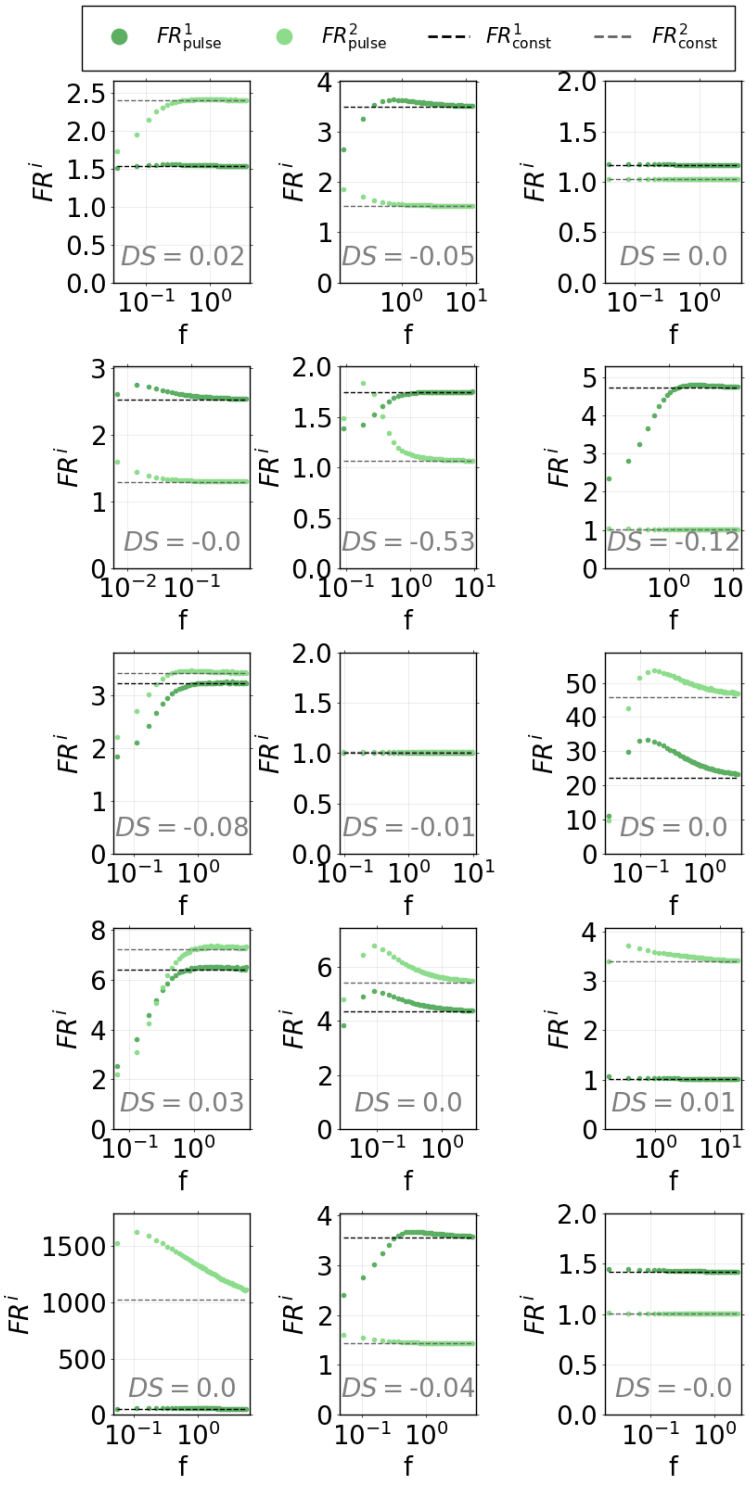


Fig. 10.14 Random sampling of fold repression computed in the $[0, \tau]$ interval for either constant or periodic miRNA input synthesis. Dark and light green dots represent respectively $FR^1_{pulse}(f)$ and $FR^2_{pulse}(f)$ values, whereas black and grey dashed lines represent respectively FR^1_{const} and FR^2_{const} values. Each subplot refers to a set of parameters sampled randomly by Latin Hypercube Sampling. For each case, the value of differential selectivity (*DS*) is reported in the plot.

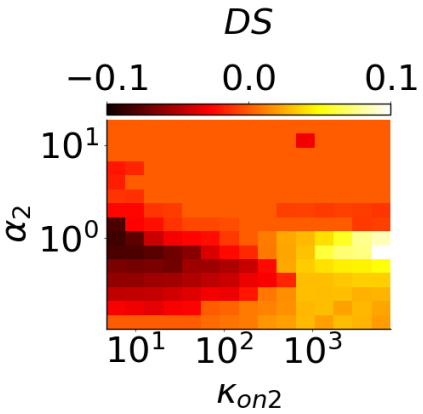


Fig. 10.15 Differential selectivity *DS* computed as a function of the R_2 degradation in complex relative to its degradation alone, α_2, and its rate of miRNA binding, κ_{on2}.

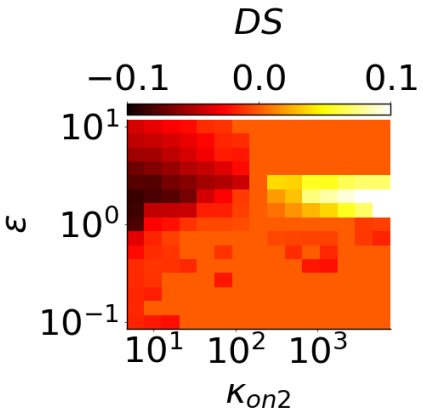


Fig. 10.16 Differential selectivity *DS* computed as a function of the relative degradation of R_2 with respect to R_1, ε, and its rate of miRNA binding, κ_{on2}.

Chapter 11

Band-pass filtering in microRNA-mediated regulation: experimental inquiry

Our experimental goal is to feed cells with periodic miRNA expression of tunable frequency, and to measure the resulting Fold Repression of a target mRNA. In this way, we would like to investigate the system's band-pass filtering properties predicted by our theoretical models.

This chapter will illustrate the building and calibration of the experimental setup aimed at generating periodic miRNA synthesis.

First, we generate a synthetic DNA plasmid that carries the gene sequence of a miRNA - miR-20a - and we verify that this plasmid correctly expresses the mature miRNA form by RNA quantification. To this plasmid, we then add a gene encoding a fluorescent protein (i.e. a reporter) that allows tracking the miRNA by fluorescence microscopy. To ensure that the fluorescence faithfully reflects the miRNA level, we design the fluorescent protein gene so that variations in protein amount mimic variations in miRNA amount: to this purpose, we modify the gene to obtain a fluorescent protein whose degradation kinetics match that of the miRNA.

Then, to make miRNA (and reporter) synthesis switchable on-and-off periodically, we further modify the plasmid in such a way that its transcription is controllable with light. This is made possible by a previously designed optogenetic system [5], which makes DNA transcription functional only upon excitation with blue light.

To test this optogenetic system, we introduce the DNA construct in human cells,

we expose them to blue light overnight, and we quantify the resulting reporter protein fluorescence. Moreover, we verify the system's ability to generate a transient expression pulse by analyzing the temporal fluorescence generated by a single blue light pulse of a few minutes.
Currently, we are working to stimulate cells with multiple ON-and-OFF pulses of blue light to induce periodic miRNA synthesis with the desired frequency. With the purpose of quantifying the resulting mRNA Fold Repression, we plan to introduce into cells an additional DNA construct encoding a gene targeted by the miRNA.
All experimental steps will be illustrated more in detail in the next sections.

11.1 Introduction to experimental techniques

11.1.1 Optogenetics

The use of light to control genetic processes in living organisms is referred to as "optogenetics", and is widely used to control gene expression in mammalian cells. Indeed, a genetic regulatory process can be engineered to depend on molecules that are naturally sensitive to light, such as chromophores, making the expression of a gene of interest controllable in time and space by illumination.
For instance, many optogenetic systems rely on engineered transcription factors whose activity is inducible with light [178]. A typical engineering approach is based on split transcription factors: two distinct domains of the molecule are fused to a chromophore whose conformation is sensitive to a specific wavelength; conformational changes caused by illumination bring the two domains together and lead to the recovery of the complete transcription factor molecule, able to initiate transcription; in this way, the expression of a gene of interest is switchable ON and OFF using light.
The optogenetic approach employed in our project, developed by Yamada and collaborators [5], relies on genetically introducing into cells a split transcription factor, TetA, fused to the light-sensitive molecules CRY2 and CIB1 - which are known to dimerize upon excitation with blue light. The dimerization caused by illumination leads to the reconstitution of the full active TetA form, which becomes able to bind its compatible gene promoter and start transcription. Exploiting this system to control

the expression of a miRNA gene, we aim to make miRNA synthesis periodically switchable ON and OFF in live cells.

11.1.2 Reverse Transcription Quantitative Polymerase Chain Reaction (RT-qPCR)

RT-qPCR is a molecular biology technique used to quantify an RNA molecule of interest. We adopt this approach in our project to quantify miRNAs and their target mRNAs.
RT-qPCR consists of three main steps : (i) The RNA sample is converted into complementary cDNA by reverse transcription. (ii) The cDNA of interest is amplified by quantitative polymerase chain reaction (qPCR). (iii) Amplification curves are analyzed to estimate the original RNA concentration in the sample.
We hereby describe each step in detail:

- Reverse transcription (RT):
  the total RNA in the sample is reverse-transcribed into complementary DNA (cDNA) using an enzyme called reverse transcriptase. Normally the reverse transcription reaction is unspecific, i.e. it converts the total RNA present in the sample without specificity for a precise RNA molecule of interest. This step is fundamental because it allows converting the RNA into a form suitable for amplification by PCR.

- Amplification by qPCR:
  DNA primers specific to the target (i.e. corresponding to the cDNA of interest) react with the previously generated cDNA, amplifying the target sequence. This is allowed by repeated cycles of heating and cooling that denature the DNA, anneal primers to the target cDNA sequence and synthesize new copies of the target using DNA polymerase. Note that during each cycle the amount of DNA doubles, leading to exponential amplification. Thus, the cDNA of interest is amplified - if present in the sample - proportionally to the initial target concentration. Fluorescent probes or DNA-binding dyes are used to monitor the accumulation of synthesized DNA in real-time, obtaining an amplification curve that is subsequently analyzed to quantify the target.

- Relative quantification:

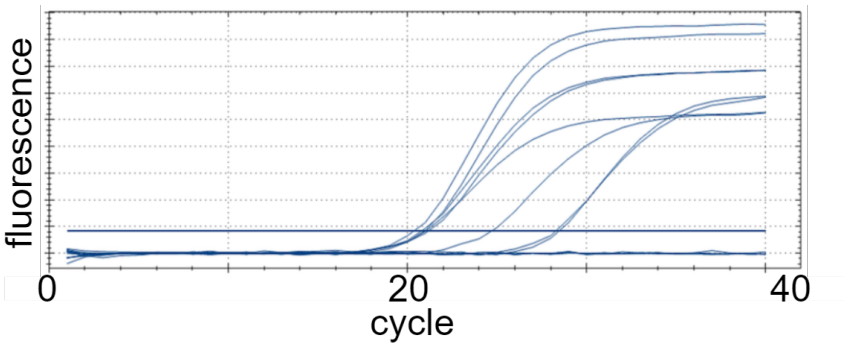


Fig. 11.1 Example of amplification curves obtained by RT-qPCR. The horizontal line represents the threshold employed to obtain Ct values for each curve.

The most widely used approach to compare RNA levels between different samples, e.g different groups or experimental conditions, is the so-called "$\Delta\Delta Ct$" method [179]. This method exploits the amplification curves obtained by qPCR (e.g. Figure 11.1) to estimate the original concentration of the RNA of interest in the samples.

Each amplification curve, corresponding to a sample, is associated to a "cycle threshold" (Ct) value, i.e. the cycle number at which the fluorescence curve crosses a set threshold (Figure 11.1). Ct values are inversely related to the amount of target present in the sample: lower Ct values indicate higher original amounts of target.

Normally, a reference RNA is amplified by RT-qPCR together with the RNA of interest, to be used as a normalizer. Typical references are genes that are constantly expressed in the cell (i.e. their expressed RNA amount is constant), and thus provide a way to correct for differences in the amount of starting material and efficiency of reverse transcription. The ΔCt is calculated by subtracting the Ct value of a reference RNA from the Ct value of the RNA of interest in each sample: $\Delta Ct = Ct_{\text{gene}} - Ct_{\text{reference}}$.

The $\Delta\Delta Ct$ value is then computed by subtracting the ΔCt value of a control sample from the ΔCt value of each sample of interest: $\Delta\Delta Ct = \Delta Ct_{\text{experimental sample}} - \Delta Ct_{\text{control sample}}$. This calculation determines the difference in the RNA of interest between a control sample and an experimental sample.

Eventually, the fold change in RNA expression between control and exper-

imental samples is calculated as $FC = 2^{-\Delta\Delta Ct}$. The fold change provides a measure of RNA amount in the experimental sample relative to the control.

11.2 Generation of a light-inducible DNA construct encoding miR-20a

To make miRNA synthesis switchable on-and-off, we chose to use optogenetics. We make DNA trascription dependent on light excitation using a preexisting optogenetic system proposed by Yamada and collaborators [5], which exploits conformational changes of light-sensitive molecules to make DNA transcription active only upon blue light stimulation (465 nm wavelength).

The system consists of two light-sensitive proteins, encoded in a DNA plasmid, which are known to dimerize if stimulated with blue light: CRY2 (Cryptochrome 2) and CIB1 (Cryptochrome-interacting basic-helix-loop-helix). CRY2 is a light-sensitive protein naturally found in plants (i.e. a chromophore), and similarly CIB1 is a protein derived from Arabidopsis thaliana, a small flowering plant. When blue light is applied to these proteins, CRY2 undergoes a conformational change that makes it active and able to bind CIB1. Due to their dimerization property being selective for blue light stimulation, this pair of proteins has been extensively used to control various cellular processes: in principle, one can partially fuse each of them to a molecule involved in a specific process, so that the complete functional molecule is reconstituted only under the light stimulus, when CRY2 and CIB1 are allowed to bind.

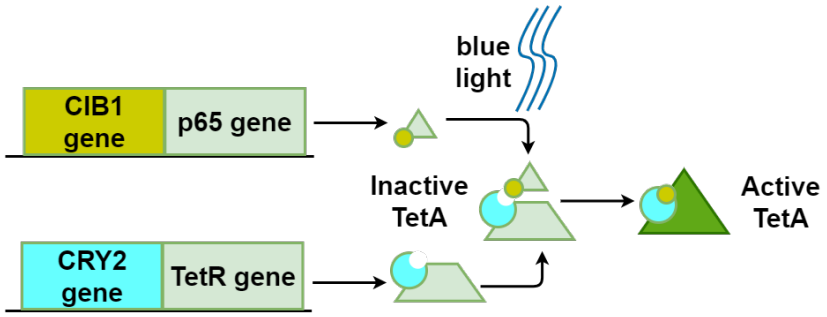


Fig. 11.2 CRY2-CIB1-TetA optogenetic plasmid system [5]. A plasmid encodes CRY2 and CIB1 fused to two domains of the TetA transcription factor: CRY2 is fused with TetR and CIB1 is fused with p65. Upon blue light excitation, CRY2 and CIB1 bind, bringing the two domains together to reconstitute TetA.

In our optogenetic system, CRY2 and CIB1 are respectively fused to two domains of the TetA transcription factor (p65 and TetR) necessary to activate DNA transcription (Figure 11.2). In the absence of light, CRY2 and CIB1 cannot dimerize and thus the two domains remain separated, leaving TetA not functional and incapable of initiating DNA transcription. Under blue light, CRY2-CIB1 dimerization leads TetA to be reconstituted: once TetA is functional, it can bind to any compatible gene promoter and initiate transcription. This system can thus in principle be used to temporally control the expression of any gene of interest, e.g. a gene encoding a miRNA.

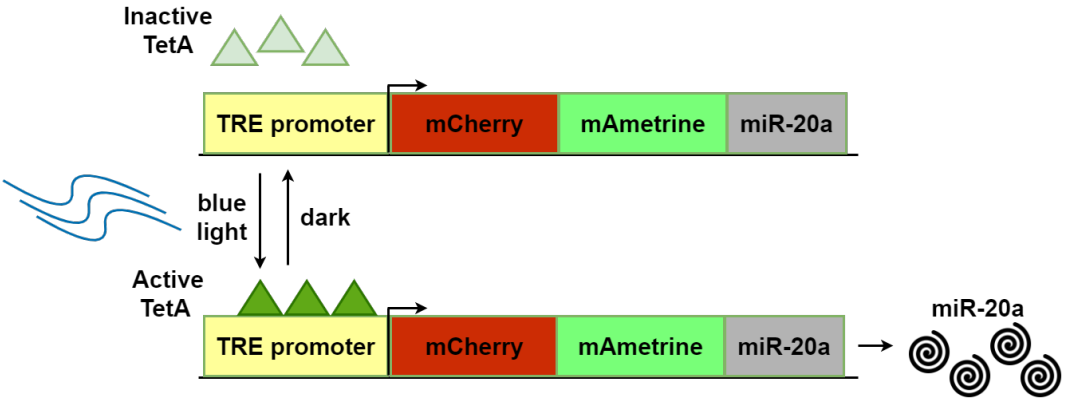


Fig. 11.3 Light-inducible plasmid carrying the miR-20a gene, a fluorescent reporter protein gene (mAmetrine) and a control fluorescent protein gene (mCherry): mCherry-mAmetrine-miR-20a.

Therefore, we cloned a gene encoding a miRNA - miR-20a - into a plasmid where transcription is controlled by the TRE promoter compatible with TetA (Figure 11.3). Moreover, to make the miRNA detectable with a fluorescent reporter, we inserted a gene encoding a green fluorescent protein - mAmetrine - in the same plasmid DNA construct. The mAmetrine gene was amplified by PCR from preexisting plasmid DNA and cloned into the plasmid by restriction DNA digestion and ligation. The miR-20a gene with compatible cloning ends was purchased from Eurofins Genomics and cloned downstream of the Ametrine protein by restriction DNA digestion and ligation. To ensure that the miRNA and the fluorescent protein are synthesized simultaneously, the miR-20a gene was inserted to the 3'UTR region of the mAmetrine gene: in this way the two genes are transcribed in a single transcript. As a further fluorescent control, we inserted an additional red fluorescent protein

gene - mCherry - upstream of the mAmetrine gene, to obtain the final mCherry-mAmetrine-miR-20a plasmid (Figure 11.3).

11.3 Testing miR-20a expression by transfection

11.3.1 Quantification of miR-20a

To verify that our DNA construct encoding miR-20a and the fluorescent reporter effectively synthesizes the mature miRNA - necessary for functional target repression - we transfected cells with the mCherry-mAmetrine-miR-20a plasmid and we quantified their expressed amount of miR-20a.
Note that this transfection was performed in the human cell line *Hek-293 Tet-off* that constitutively expresses the *TetA* transcription factor necessary for transcription; and thus in this case the blue light stimulus is not needed to activate synthesis of the miRNA and its Ametrine reporter.
The experiment consisted of the following steps:

- Cell culture:
  Hek293 Tet-off cells were plated in 24-well plates and cultured in DMEM medium with 10% FBS (Fetal Bovine Serum) for 24 hours.

- Cell transfection:
  As a control, cells were partly transfected with the plasmid lacking the miR-20a gene (mCherry-mAmetrine, 100 ng/well). The remaining cells were transfected with the mCherry-mAmetrine-miR-20a plasmid (100 ng/well).

- RNA extraction:
  After 48 hours, we extracted the total RNA from both control (mCherry-mAmetrine) and miRNA-transfected (mCherry-mAmetrine-miR-20a) cell samples.

- MiRNA quantification:
  The extracted RNA was reverse-transcribed into cDNA and the TaqMan Advanced kit was used for amplifying the mature miR-20a. As a normalization control, the TaqMan Advanced kit was also used for amplifying miR-16.

- qPCR:
  MiR-20a and the normalization control miR-16 were further amplified by RT-qPCR.

- Estimation of miR-20a fold change (FC) between control and miRNA-transfected cells:
  The miR-20a fold change between control and miRNA-transfected cells was calculated using the $\Delta\Delta CT$ method [179].

Results showed that our mCherry-mAmetrine-miR-20a plasmid effectively expresses the mature form of miR-20a:

$$\text{FC miR-20a} = \frac{\text{miR-20a expression in transfected cells}}{\text{miR-20a expression in control cells}} = 2.7 \tag{11.1}$$

11.3.2 Quantification of miR-20a endogenous target

To verify that the mature form of miR-20a expressed by our plasmid is functional, we tested its repressive action on one of its known endogenous target genes, i.e. VEGF-A. VEGF-A is a protein necessary for the formation of blood vessels, and it is constitutively expressed by Hek293 cells.
We thus transfected cells with the mCherry-mAmetrine-miR-20a plasmid and we quantified the resulting amount of VEGF-A mRNA.
Note that this transfection was performed in the human cell line *Hek-293 Tet-off* that constitutively expresses the *TetA* transcription factor necessary for transcription; and thus in this case the blue light stimulus is not needed to activate synthesis of the miRNA and its Ametrine reporter.
The experiment consisted of the following steps:

- Cell culture:
  Hek293 Tet-off cells were plated in 96-well plates and cultured in DMEM medium with 10% FBS (Fetal Bovine Serum) for 24 hours.

- Cell transfection:
  As a control, cells were partly transfected with the plasmid lacking the miR-20a gene (mCherry-mAmetrine, 100 ng/well). The remaining cells were transfected with the mCherry-mAmetrine-miR-20a plasmid (100 ng/well).

- RNA extraction:
  After 48 hours, we extracted the total RNA from both control (mCherry-mAmetrine) and miRNA-transfected (mCherry-mAmetrine-miR-20a) cells.

- VEG-FA amplification by qPCR:
  The extracted RNA was reverse-transcribed into cDNA and VEGF-A was amplified by RT-qPCR.

- Estimation of VEGF-A fold change (FC) between control and transfected cells:
  The VEGF-A fold change between control and transfected cells was calculated using the $\Delta\Delta CT$ method.

Results showed that the mature miR-20a expressed by our mCherry-mAmetrine-miR-20a plasmid effectively inhibits its target VEGF-A mRNA:

$$\text{FC VEGF-A} = \frac{\text{VEGF-A expression in transfected cells}}{\text{VEGF-A expression in control cells}} = 0.22 \tag{11.2}$$

11.4 Modification of reporter protein half-life by ubiquitination

Using a fluorescent reporter such as mAmetrine to track the temporal expression of miR-20a requires that changes in miRNA levels are mimicked by corresponding changes in reporter amount (i.e. fluorescence), and thus demands that the miRNA and the protein follow similar temporal kinetics. Thus, we had to optimize Ametrine protein kinetics to make its fluorescence intensity a reliable reporter of the amount of miRNA. To this purpose, we modified the mAmetrine gene to make the protein's degradation rate - i.e. its half-life - to match that of miR-20a.
The half-life of miR-20a was previously measured to be between 3 and 6 hours [180], making it a short-lived miRNA, whereas the typical half-lives of common fluorescent proteins are around 24 hours [181]. Therefore, adapting the degradation kinetics of a fluorescent protein such as Ametrine to that of miR-20a likely means shortening its half-life.
To shorten the half-life of the Ametrine reporter protein, we used ubiquitination: marking a protein with a chain of ubiquitin molecules - often called "ubiquitin tag" -

induces the proteasome to recognize and degrade it. We thus cloned many copies of the DNA construct containing miR-20a, Ametrine and mCherry genes, and upstream of the Ametrine we added ubiquitin tags which affect the protein's half-life. To this purpose, we used ubiquitin tags created by Chassin and colleagues, UbD, UbK, UbM, UbS, $UbVR$ [182].

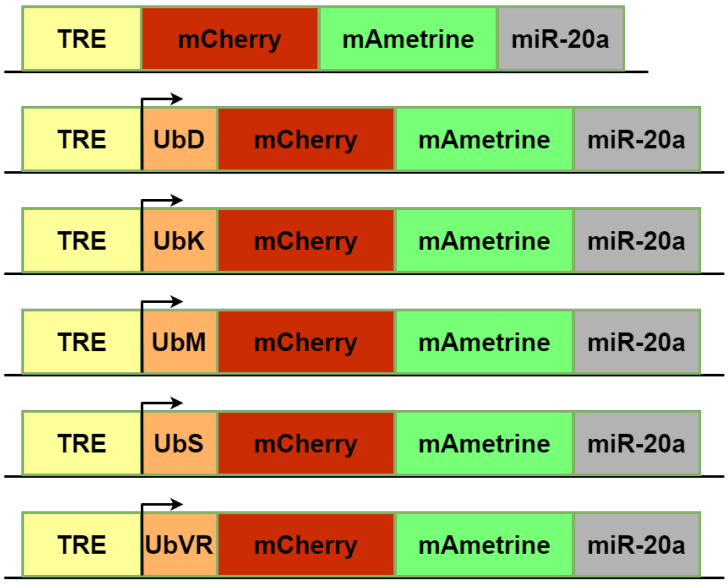


Fig. 11.4 Schematization of the six plasmid versions encoding the mAmetrine gene modified with different ubiquitin tags: no tag, UbD, UbK, UbM, UbS, $UbVR$.

In this way we obtained six copies of the mCherry-mAmetrine-miR20a plasmid - one untagged and five ubiquitin-tagged - each one expressing a destabilized Ametrine with a different half-life, schematized in Figure 11.4.
Then, we performed a series of experiments to estimate the Ametrine half-life for each ubiquitinated version.
The approach used to estimate the half-life of our reporter protein consisted of the following steps:

- Cell culture:
  Hek293 Tet-off cells were plated in 6-well plates and cultured in DMEM medium with 10% FBS (Fetal Bovine Serum) for 24 hours.
- Cell transfection with ubiquitinated plasmids:
  To induce Ametrine synthesis, we started by transfecting cells with each

plasmid version - untagged and ubiquitin-tagged. Despite these plasmids - as described in previous sections - need activation by blue light, we transfected them in a ready modified human cell line (Hek-293 Tet-off) where their expression occurs constitutively, without the need of any further stimulus. Indeed, in this human cell line the transcription factor TetA - necessary for the transcription of miR-20a and mAmetrine in our plasmids - is endogenously expressed and does not need to be generated through the described optogenetic system.

- Cell incubation:
  After transfection of cells with the ubiquitinated plasmids, we incubated cells at 37° C and we waited 48 hours for the expression of mAmetrine to reach steady state - a time window typically adopted for transient transfection experiments.

- Chemical blockage of reporter protein transcription:
  We subsequently stopped mAmetrine synthesis using a common chemical method of transcription blockage, i.e. somministrating the Doxycycline (DOX) antibiotic to cells. The Hek293 Tet-off cell line is engineered in such a way that the DOX molecule sequestrates the site needed by the transcription factor TetA for binding to the promoter and initiate transcription; thus, in the presence of DOX, our plasmid expressing the miRNA and the reporter loses functionality. The transcriptional block allows to observe the subsequent degradation timecourse of any fluorescent molecule of interest in live cells.

- Tracking the reporter fluorescence decay by time-lapse fluorescence microscopy:
  Following the transcriptional block, we tracked the progressive decay of Ametrine fluorescence by imaging cells with a fluorescence microscope (*Nikon Lipsi*) for 25 hours.

- Microscopy image segmentation for the identification of single cells:
  After acquisition, we created an automatic segmentation script that identifies single cells in our images and estimates their mean fluorescence.

- Estimation of reporter protein half-life:

  For each ubiquitinated plasmid version, we averaged mean fluorescences of thousands of cells at each time step, to obtain a single fluorescence decay curve. We then normalized each curve by its maximum.

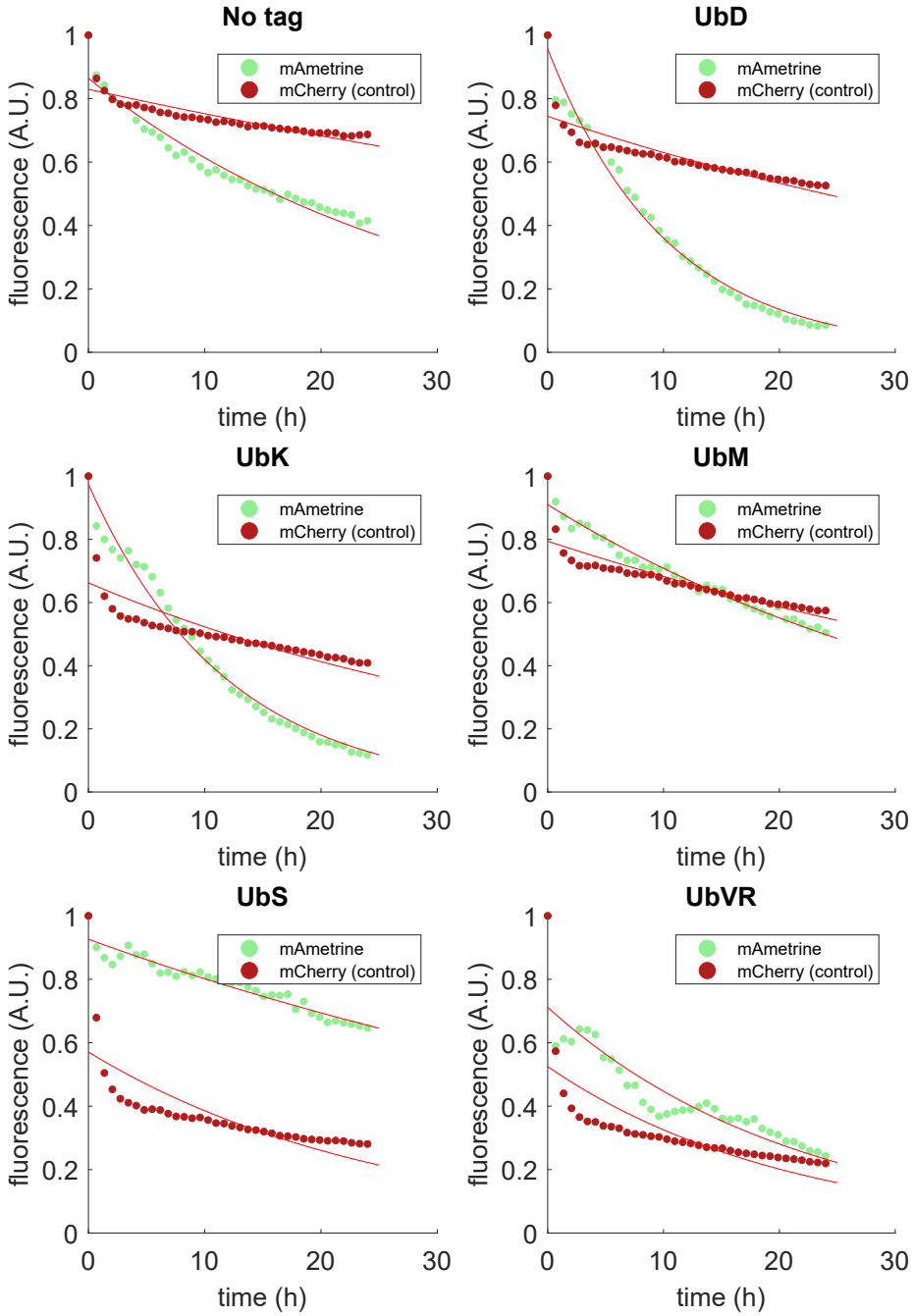


Fig. 11.5 Temporal decay of Ametrine fluorescence. Each plot represents the decay of fluorescence for a different ubiquitinated mAmetrine version, along with the fluorescence decay of its control mCherry. Points represent the normalized average of cell fluorescence mean values at each time point.

For each of the six versions, where the reporter gene is characterized by a different ubiquitin tag, the mAmetrine decay curve resulted more or less steep. Conversely, the decay curves of control mCherry resulted similar. Plots in Figure 11.5 represent the decay curves of mCherry and tagged mAmetrine in the six different versions, i.e. untagged, UbD, UbK, UbM, UbS, UbVR.
Assuming protein degradation to follow First Order kinetics (i.e. the rate of reaction depends linearly on the amount of reactant), we fitted each temporal fluorescence curve with the decreasing exponential:

$$y = y_0 \cdot e^{-kt} \tag{11.3}$$

and we subsequently estimated the corresponding half-life as:

$$\text{half-life} = \frac{ln(2)}{k} \tag{11.4}$$

We obtained the following half-lives for the different construct versions:

- Untagged mCherry-mAmetrine-MiR-20a: mAmetrine half-life = 20.27 hours
- mCherry-UbD-mAmetrine-MiR-20a: mAmetrine half-life = 8.21 hours
- mCherry-UbK-mAmetrine-MiR-20a: mAmetrine half-life = 7.09 hours
- mCherry-UbM-mAmetrine-MiR-20a: mAmetrine half-life = 27.66 hours
- mCherry-UbS-mAmetrine-MiR-20a: mAmetrine half-life = 48.04 hours
- mCherry-UbVR-mAmetrine-MiR-20a: mAmetrine half-life = 14.89 hours

Thus, to select the best miRNA fluorescent reporter for our next experiments, we chose the plasmid version with the Ametrine half-life most closely matching that of miR-20a, i.e. the one tagged with *Ubk*: mCherry-UbK-mAmetrine-MiR-20a.

11.5 Assembly of the optogenetic LED platform

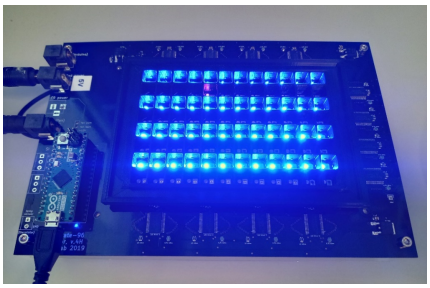


Fig. 11.6 LED light platform used to stimulate DNA transcription in cells.

To build an illumination source for our light-sensitive plasmid DNA constructs, our collaborators at ETH Zürich assembled a 12 x 8 grid of LED lights using Arduino (Figure 11.6). This LED structure fits a 96-well cell culture plate, and is thus easily employed to irradiate cell plates with a transparent bottom from below.

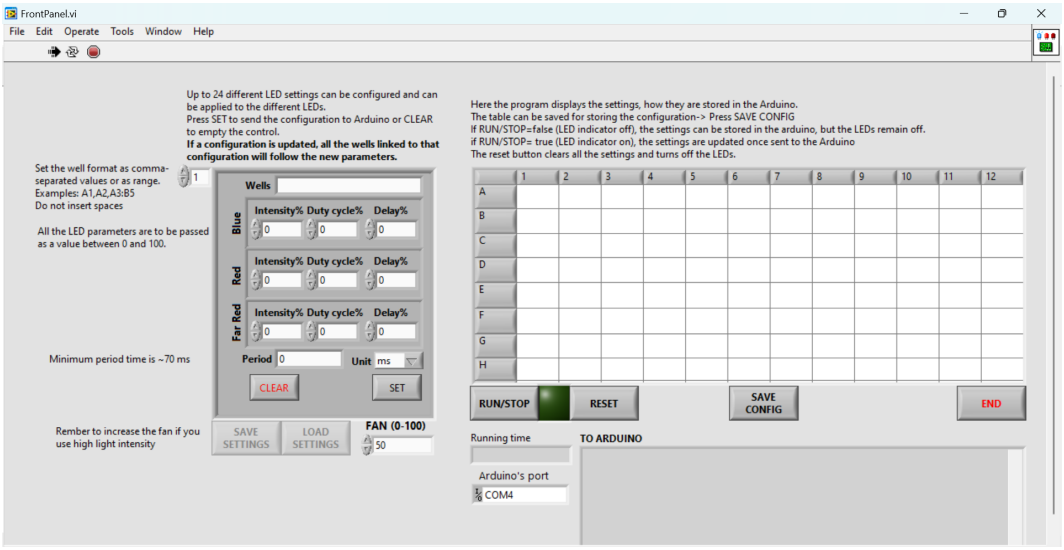


Fig. 11.7 Programming interface of the LED light platform used to stimulate plasmid DNA transcription in cells.

The platform comes along with an interface that allows to program the timing and the intensity of light pulses for each LED (Figure 11.7). In particular, it is possible to tune the period and the duty cycle of light pulses, as well as the starting delay of

the stimulus.
Available light wavelengths are the following:

- Red: 630 nm.
- Blue: 465 nm.
- Far red: 730 nm.

11.6 Testing fluorescent reporter protein expression by blue light

In order to test the responsiveness of our optogenetic system, we transfected cells with both the optogenetic plasmid CRY2-CIB1-TetA and the mCherry-UbK-mAmetrine-MiR-20a plasmid, illuminated them with blue light for 24 hours, and estimated the induced Ametrine fluorescence. With the purpose of testing the plasmid's responsiveness to light, transfection was performed in a human cell line (*Hek-293T*) that does not endogenously express the *TetA* transcription factor necessary for transcription in our plasmid; and thus in this case the blue light stimulus is needed to activate synthesis of the miRNA, of its Ametrine reporter and the control mCherry. The experiment consisted of the following steps:

- Cell culture:
  Hek-293T cells were plated in 96-well plates and cultured in DMEM medium with 10% FBS (Fetal Bovine Serum) for 24 hours.
- Cell transfection:
  Cells were co-transfected with both the mCherry-UbK-mAmetrine-miR-20a plasmid and the plasmid encoding the CRY2-CIB1 optogenetic system, with a 1:5 ratio between the two plasmids. A total of 100 ng of plasmid DNA was used for each well, i.e. 16,6 ng of miR-20a-Ametrine-UbK and 83,3 ng of CRY2-CIB1 plasmid.
- Optogenetic stimulation:
  Transfected cells were partly left in the dark (control) and partly illuminated with blue light (465 nm) using the LED optogenetic platform. Illumination

was performed in incubation at 37° for 24 hours and was somministrated in pulses of 2 seconds every 30 seconds.

- Imaging the Ametrine reporter protein by fluorescence microscopy:
  Following the 24-hour light stimulus, we imaged both control and illuminated cells with a fluorescence microscope (*Nikon Lipsi*). Figure 11.8 represents an example imaged mAmetrine field of control and illuminated cells.

- Microscopy image segmentation:
  After acquisition, we created an automatic *MATLAB* script that identifies single cells in our images and estimated their mean fluorescence.

- Estimation of mean reporter fluorescence:
  By averaging the mean fluorescence of the identified cells, we estimated the mean mCherry and mAmetrine fluorescence for both control and light-stimulated cells.

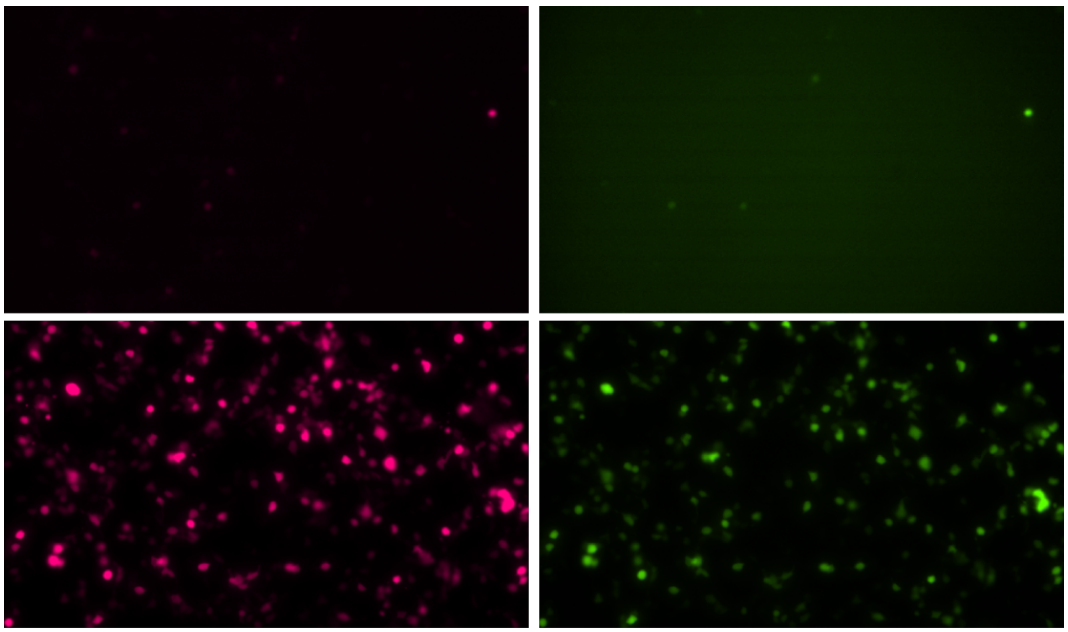


Fig. 11.8 Example of cell fields imaged in the mCherry channel and in the mAmetrine channel in control (top panels) and light-stimulated (bottom panels) cells.

Results revealed a $\sim$ 18-fold increase of both mCherry and mAmetrine fluorescence in cells that were stimulated with blue light with respect to control (Figure 11.9). This test thus confirmed the functionality of our optogenetic system.

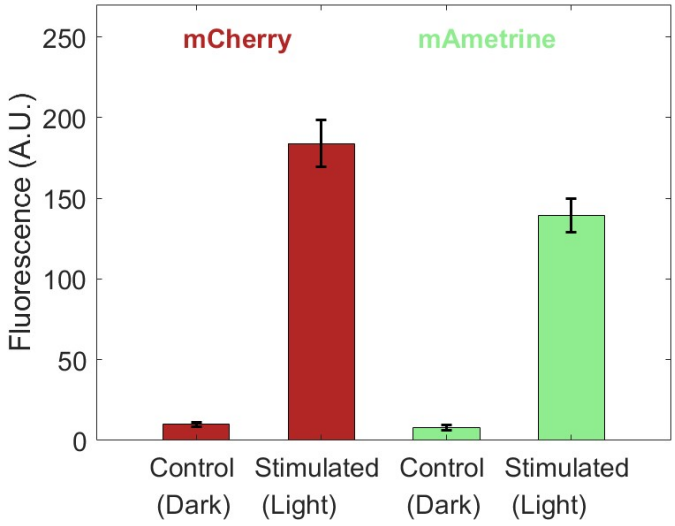


Fig. 11.9 MCherry and mAmetrine fluorescence after a 24-hour blue light stimulus. Red bars represent respectively the mean mCherry fluorescence of imaged control fields and stimulated fields. Green bars represent respectively the mean mAmetrine fluorescence of imaged control fields and stimulated fields. Errorbars represent standard errors of means.

11.7 Generation of a single reporter expression pulse

To test the generation of a single expression pulse of the mAmetrine reporter, we transfected cells with both the optogenetic plasmid CRY2-CIB1-TetA and the mCherry-UbK-mAmetrine-miR-20a plasmid, and we illuminated them with a single 15-minute long blue light pulse. However, we noticed that the light required for acquiring the mAmetrine fluorescence (406 nm excitation wavelenght) interferes with the optogenetic system leading to undesired activation of transcription. Therefore, the continuous temporal acquisition of Ametrine images would introduce a false positive protein expression that is not generated by the initial blue light stimulus. We thus overcame this issue by using multiple cell samples stimulated with progressively increasing delays, and simultaneously imaging mAmetrine in all samples at once. In this way, each sample served to reconstruct one timepoint in the temporal mAmetrine fluorescence curve following the light stimulus.
The experiment consisted of the following steps:

- Cell culture:
  Hek-293T cells were plated in 96-well plates and cultured in DMEM medium with 10% FBS (Fetal Bovine Serum) for 24 hours.

- Cell transfection:
  Cells were co-transfected with both the mCherry-UbK-mAmetrine-miR-20a plasmid and the plasmid encoding the CRY2-CIB1-TetA optogenetic system, with a 1:5 ratio between the two plasmids. A total of 100 ng of plasmid DNA was used for each well, i.e. 16,6 ng of mCherry-UbK-mAmetrine-miR-20a and 83,3 ng of CRY2-CIB1 plasmid.

- Optogenetic stimulation:
  Transfected cells were partly left in the dark (control) and partly illuminated with blue light (465 nm) using the LED optogenetic platform. Each well was exposed to the light stimulus with a delay of 1 hour with respect to the adjacent well. For each well, continuous illumination was performed in incubation at 37° for 15 minutes. Considering the timespan occurring between the stimulation of different wells, illumination of all wells lasted 24 hours.

- Imaging the reporter by fluorescence microscopy:
  Following the 24-hour experiment, we simultaneously imaged mAmetrine in both control and illuminated cells with a fluorescence microscope (*Nikon Lipsi*).

- Microscopy image segmentation:
  After acquisition, we created an automatic *MATLAB* script that identifies single cells in our images and estimates their mean fluorescence.

- Estimation of mean reporter fluorescence for each well (i.e. timepoint):
  By averaging the mean fluorescence of cells belonging to each well, we estimated the mean mCherry and mAmetrine fluorescences for both control and light-stimulated cells. In this way, we quantified the expression of the reporter and the control fluorophore at each time point ranging from 1 to 24 hours after the stimulus.

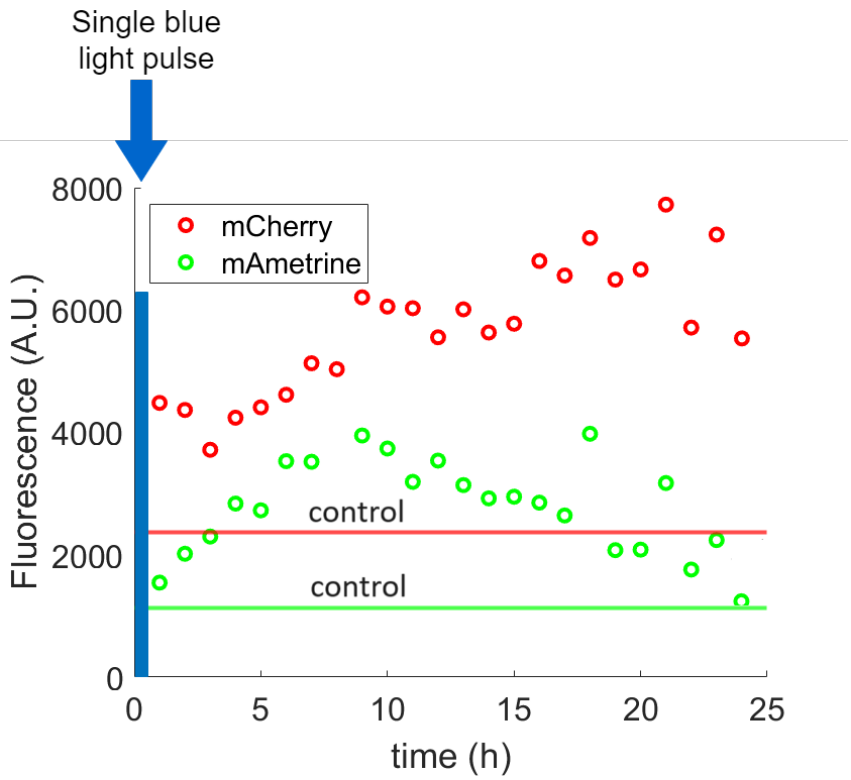


Fig. 11.10 Temporal mCherry and mAmetrine fluorescence after a single 15-minute blue light stimulus. Straight lines represent the mean fluorescence of control cells (i.e. cells left in the dark). Errorbars represent standard errors of the mean.

We obtained an mAmetrine fluorescence timecourse that reaches a peak around 10 hours after the stimulus, before starting to progressively decay to control levels in approximately 15 hours (Figure 11.10). This results is consistent with the mAmetrine half life of ~ 7 hours measured for the employed plasmid mCherry-UbK-mAmetrine-miR-20a (Section 11.4).
For mCherry - coherently with its expected slower maturation and decay compared to the ubiquitinated mAmetrine - we obtained a gradual temporal accumulation of fluorescence.

11.8 Future experiments

The forthcoming experimental efforts will be aimed at generating periodic ON-and-OFF pulses of periodic miRNA synthesis using our optogenetic system, and

quantifying the Fold Repression resulting from different frequencies.
With the purpose of quantifying Fold Repression, we are currently planning the completion of our experimental setup by adding a target of miR-20a. An artificial target could be added by transfecting cells with a further plasmid that encodes an mRNA targeted by the miRNA. One possibility would be a plasmid encoding two distinct fluorescent protein genes: an mCherry protein targeted by the miRNA (i.e. containing miR-20a binding sites) and a non-targeted YFP protein which would provide a basal unrepressed protein level.

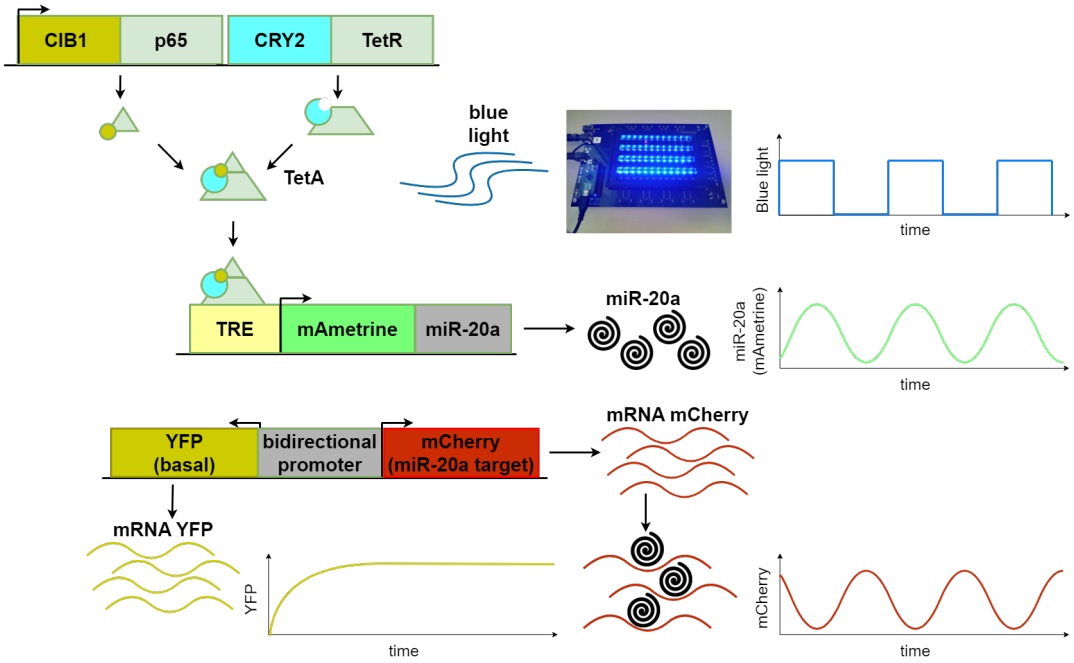


Fig. 11.11 Representation of the complete experimental setup for the generation of periodic miRNA synthesis and the quantification of target mRNA Fold Repression. Upon illumination with blue light, the TetA transcription factor is reconstituted; TetA activates miR-20a transcription; miR-20a inhibits mCherry by binding to its mRNA; YFP is expressed without inhibition.

The measurement of mCherry and YFP expression levels by fluorescence microscopy would allow to quantify the Fold Repression resulting from periodic miRNA synthesis as:

$$\text{Fold Repression} = \frac{\text{YFP expression}}{\text{mCherry expression}} \tag{11.5}$$

Figure 11.11 represents the complete setup composed by the optogenetic plasmid system, the light-inducible miR-20a plasmid, and the target plasmid encoding mCherry and YFP proteins.

References

[1] Angela Wandinger-Ness and Marino Zerial. Rab proteins and the compartmentalization of the endosomal system. *Cold Spring Harbor perspectives in biology*, 6(11):a022616, 2014.

[2] Jochen Rink, Eric Ghigo, Yannis Kalaidzidis, and Marino Zerial. Rab conversion as a mechanism of progression from early to late endosomes. *Cell*, 122(5):735–749, 2005.

[3] Perla Del Conte-Zerial, Lutz Brusch, Jochen C Rink, Claudio Collinet, Yannis Kalaidzidis, Marino Zerial, and Andreas Deutsch. Membrane identity and gtpase cascades regulated by toggle and cut-out switches. *Molecular systems biology*, 4(1):206, 2008.

[4] Carlo Cosimo Campa, Jean Piero Margaria, Abhishek Derle, Marco Del Giudice, Maria Chiara De Santis, Luca Gozzelino, Francesca Copperi, Carla Bosia, and Emilio Hirsch. Rab11 activity and ptdins (3) p turnover removes recycling cargo from endosomes. *Nature chemical biology*, 14(8):801–810, 2018.

[5] Mayumi Yamada, Yusuke Suzuki, Shinji C Nagasaki, Hiroyuki Okuno, and Itaru Imayoshi. Light control of the tet gene expression system in mammalian cells. *Cell reports*, 25(2):487–500, 2018.

[6] James D Watson and Francis HC Crick. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, 171(4356):737–738, 1953.

[7] Maurice Hugh Frederick Wilkins, WE Seeds, AR Stokes, and HR Wilson. Helical structure of crystalline deoxypentose nucleic acid. *Nature*, 172(4382):759–762, 1953.

[8] Rosalind E Franklin and Raymond G Gosling. Molecular configuration in sodium thymonucleate. *Nature*, 171(4356):740–741, 1953.

[9] Peter W Hawkes and John CH Spence. *Springer handbook of microscopy*. Springer Nature, 2019.

[10] Jacob Beutel. *Handbook of medical imaging*, volume 3. Spie Press, 2000.

[11] James Ladyman, James Lambert, and Karoline Wiesner. What is a complex system? *European Journal for Philosophy of Science*, 3:33–67, 2013.

[12] Kunihiko Kaneko. *Life: an introduction to complex systems biology*. Springer, 2006.

[13] Abhinava K Mishra, Joseph P Campanale, James A Mondo, and Denise J Montell. Cell interactions in collective cell migration. *Development*, 146(23):dev172056, 2019.

[14] Tamás Vicsek and Anna Zafeiris. Collective motion. *Physics reports*, 517(3-4):71–140, 2012.

[15] Nelson R Barton and LS Goldstein. Going mobile: microtubule motors and chromosome segregation. *Proceedings of the National Academy of Sciences*, 93(5):1735–1742, 1996.

[16] Akira Isihara. *Statistical physics*. Academic Press, 2013.

[17] Giorgio Parisi. Statistical physics and biology. *Physics World*, 6(9):42, 1993.

[18] Gašper Tkačik and Aleksandra M Walczak. Information transmission in genetic regulatory networks: a review. *Journal of Physics: Condensed Matter*, 23(15):153102, 2011.

[19] Andreas Karoly Gombert and Jens Nielsen. Mathematical modelling of metabolism. *Current opinion in biotechnology*, 11(2):180–186, 2000.

[20] Volker Grimm. Mathematical models and understanding in ecology. *Ecological modelling*, 75:641–651, 1994.

[21] William Bialek. Perspectives on theory at the interface of physics and biology. *Reports on Progress in Physics*, 81(1):012601, 2017.

[22] Johannes Kepler and Tycho Brahe. *Astronomia nova aitiologetos, seu physica coelestis, tradita commentariis de motibus stellæ Martis, ex observationibus GV Tychonis Brahe: jussu & sumptibus Rudolphi 2. Romanorum imperatoris &c: plurium annorum pertinaci studio elaborata Pragæ, a... Joanne Keplero...* 1968.

[23] AY 2 Klimenko. Complex competitive systems and competitive thermodynamics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1982):20120244, 2013.

[24] Dirk De Vos, Frank J Bruggeman, Hans V Westerhoff, and Barbara M Bakker. How molecular competition influences fluxes in gene expression networks. *PloS one*, 6(12):e28494, 2011.

[25] Ashleigh S Griffin, Stuart A West, and Angus Buckling. Cooperation and competition in pathogenic bacteria. *Nature*, 430(7003):1024–1027, 2004.

[26] Richard Svanbäck and Daniel I Bolnick. Intraspecific competition drives increased resource use diversity within a natural population. *Proceedings of the Royal Society B: Biological Sciences*, 274(1611):839–844, 2007.

[27] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[28] Jörn Bleck-Neuhaus. Mechanical resonance: 300 years from discovery to the full understanding of its importance. *arXiv preprint arXiv:1811.08353*, 2018.

[29] Fred Irons. *Active filters for integrated-circuit applications*. Artech, 2005.

[30] Isaac Edery. Circadian rhythms in a nutshell. *Physiological genomics*, 3(2):59–74, 2000.

[31] Detlev Arendt. Evolution of eyes and photoreceptor cell types. *International Journal of Developmental Biology*, 47(7-8):563–571, 2003.

[32] Uri Alon. *An introduction to systems biology: design principles of biological circuits*. CRC press, 2019.

[33] Brian P Ingalls. *Mathematical modeling in systems biology: an introduction*. MIT press, 2013.

[34] Péter Érdi and János Tóth. *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*. Manchester University Press, 1989.

[35] Kenneth A Johnson and Roger S Goody. The original michaelis constant: translation of the 1913 michaelis–menten paper. *Biochemistry*, 50(39):8264–8269, 2011.

[36] Athel Cornish-Bowden. *Fundamentals of enzyme kinetics*. John Wiley & Sons, 2013.

[37] Daniel T Gillespie. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(1-3):404–425, 1992.

[38] Paul A Gagniuc. *Markov chains: from theory to implementation and experimentation*. John Wiley & Sons, 2017.

[39] Daniel T Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58:35–55, 2007.

[40] Daniel T Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of computational physics*, 22(4):403–434, 1976.

[41] Kendall Atkinson. *An introduction to numerical analysis*. John wiley & sons, 1991.

[42] DEMOSTHENES KAZANAS and DAVID N SCHRAMM. Competition of neutrino and gravitational radiation in neutron star formation. *Nature*, 262(5570):671–672, 1976.

[43] Elena Vedmedenko. *Competing interactions and pattern formation in nanoworld*. John Wiley & Sons, 2007.

[44] Taufiq Murtadho, Sai Vinjanampathy, and Juzar Thingna. Cooperation and competition in synchronous open quantum systems. *arXiv preprint arXiv:2301.04322*, 2023.

[45] Richard R Ernst, Geoffrey Bodenhausen, Alexander Wokaun, and Alfred G Redfield. Principles of nuclear magnetic resonance in one and two dimensions, 1989.

[46] Rainer Zangerl. The methods of comparative anatomy and its contribution to the study of evolution. *Evolution*, pages 351–374, 1948.

[47] Benjamin H Good, Michael J McDonald, Jeffrey E Barrick, Richard E Lenski, and Michael M Desai. The dynamics of molecular evolution over 60,000 generations. *Nature*, 551(7678):45–50, 2017.

[48] Jeffrey K Conner. Artificial selection: a powerful tool for ecologists. *Ecology*, 84(7):1650–1660, 2003.

[49] T Ryan Gregory. Understanding natural selection: essential concepts and common misconceptions. *Evolution: Education and outreach*, 2(2):156–175, 2009.

[50] Vito Volterra. *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*, volume 2. Societá anonima tipografica" Leonardo da Vinci", 1927.

[51] Alfred J Lotka. Analytical note on certain rhythmic relations in organic systems. *Proceedings of the National Academy of Sciences*, 6(7):410–415, 1920.

[52] Debashish Chowdhury. Collective effects in intra-cellular molecular motor transport: Coordination, cooperation and competition. *Physica A: Statistical Mechanics and its Applications*, 372(1):84–95, 2006.

[53] Stuart H Orkin and Sean J Morrison. Stem-cell competition. *Nature*, 418(6893):25–26, 2002.

[54] Aida Di Gregorio, Sarah Bowling, and Tristan Argeo Rodriguez. Cell competition and its role in the regulation of cell fitness from development to cancer. *Developmental cell*, 38(6):621–634, 2016.

[55] Lei Zhang, Yuko Narita, Lin Gao, Muhammad Ali, Mamoru Oshiki, Satoshi Ishii, and Satoshi Okabe. Microbial competition among anammox bacteria in nitrite-limited bioreactors. *Water research*, 125:249–258, 2017.

[56] Ping Li, Periasamy Selvaraj, and Cheng Zhu. Analysis of competition binding between soluble and membrane-bound ligands for cell surface receptors. *Biophysical journal*, 77(6):3394–3406, 1999.

[57] Carla Bosia, Andrea Pagnani, and Riccardo Zecchina. Modelling competing endogenous rna networks. *PloS one*, 8(6):e66609, 2013.

[58] LA Sklar, J Sayre, VM McNeil, and DA Finney. Competitive binding kinetics in ligand-receptor-competitor systems. rate parameters for unlabeled ligands for the formyl peptide receptor. *Molecular Pharmacology*, 28(4):323–330, 1985.

[59] Eric R Kandel, James H Schwartz, Thomas M Jessell, Steven Siegelbaum, A James Hudspeth, Sarah Mack, et al. *Principles of neural science*, volume 4. McGraw-hill New York, 2000.

[60] Sung Ok Yoon, Patrizia Casaccia-Bonnefil, Bruce Carter, and Moses V Chao. Competitive signaling between trka and p75 nerve growth factor receptors determines cell survival. *Journal of Neuroscience*, 18(9):3273–3281, 1998.

[61] Peter W Zandstra, Douglas A Lauffenburger, and Connie J Eaves. A ligand-receptor signaling threshold model of stem cell differentiation control: a biologically conserved mechanism applicable to hematopoiesis. *Blood, The Journal of the American Society of Hematology*, 96(4):1215–1222, 2000.

[62] Pau Formosa-Jordan and Marta Ibañes. Competition in notch signaling with cis enriches cell fate decisions. *PloS one*, 9(4):e95744, 2014.

[63] Greg Lemke. Glial control of neuronal development. *Annual review of neuroscience*, 24(1):87–105, 2001.

[64] Luciano Adorini, Simone Muller, Francis Cardinaux, Paul V Lehmann, Fiorenza Falcioni, and Zoltan A Nagy. In vivo competition between self peptides and foreign antigens in t-cell activation. *Nature*, 334(6183):623–625, 1988.

[65] VI Seledtsov and GV Seledtsova. An 'antigenic ligand competition'model for antigen receptor-mediated lymphocyte selection. *Biomedicine & pharmacotherapy*, 50(3-4):170–177, 1996.

[66] Holger Hoff, Gerd-R Burmester, and Monika C Brunner-Weinzierl. Competition and cooperation: Signal transduction by cd28 and ctla-4. *Signal Transduction*, 6(4):260–267, 2006.

[67] Colin A Flaveny, Iain A Murray, Chris R Chiaro, and Gary H Perdew. Ligand selectivity and gene regulation by the human aryl hydrocarbon receptor in transgenic mice. *Molecular pharmacology*, 75(6):1412–1420, 2009.

[68] Alice Y Cheung and Sacco C de Vries. Membrane trafficking: intracellular highways and country roads. *Plant physiology*, 147(4):1451–1453, 2008.

[69] Nava Segev, Andrei A Tokarev, Aixa Alfonso, and Nava Segev. Overview of intracellular compartments and trafficking pathways. *Trafficking inside cells: pathways, mechanisms and regulation*, pages 3–14, 2009.

[70] Barth D Grant and Julie G Donaldson. Pathways and mechanisms of endocytic recycling. *Nature reviews Molecular cell biology*, 10(9):597–608, 2009.

[71] Juan S Bonifacino and Benjamin S Glick. The mechanisms of vesicle budding and fusion. *cell*, 116(2):153–166, 2004.

[72] Emanuele Cocucci, Raphaël Gaudin, and Tom Kirchhausen. Dynamin recruitment and membrane scission at the neck of a clathrin-coated pit. *Molecular biology of the cell*, 25(22):3595–3609, 2014.

[73] Heidi Hehnly and Mark Stamnes. Regulating cytoskeleton-based vesicle motility. *FEBS letters*, 581(11):2112–2118, 2007.

[74] Hans-Günther Döbereiner, J Käs, D Noppl, Irene Sprenger, and E Sackmann. Budding and fission of vesicles. *Biophysical journal*, 65(4):1396–1403, 1993.

[75] Marco Faini, Rainer Beck, Felix T Wieland, and John AG Briggs. Vesicle coats: structure, function, and general principles of assembly. *Trends in cell biology*, 23(6):279–288, 2013.

[76] David G Robinson and Peter Pimpl. Clathrin and post-golgi trafficking: a very complicated issue. *Trends in plant science*, 19(3):134–139, 2014.

[77] Bruce Alberts. *Molecular biology of the cell*. Garland science, 2017.

[78] Marko Jovic, Mahak Sharma, Juliati Rahajeng, and Steve Caplan. The early endosome: a busy sorting station for proteins at the crossroads. *Histology and histopathology*, 25(1):99, 2010.

[79] Tanmay Bhuin and Jagat Kumar Roy. Rab proteins: the key regulators of intracellular vesicle transport. *Experimental cell research*, 328(1):1–19, 2014.

[80] Miguel C Seabra, Emilie H Mules, and Alistair N Hume. Rab gtpases, intracellular traffic and disease. *Trends in molecular medicine*, 8(1):23–30, 2002.

[81] Marcellus J Banworth and Guangpu Li. Consequences of rab gtpase dysfunction in genetic or acquired human diseases. *Small GTPases*, 9(1-2):158–181, 2018.

[82] Wan Jie Chia and Bor Luen Tang. Emerging roles for rab family gtpases in human cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1795(2):110–116, 2009.

[83] Ferdi Ridvan Kiral, Friederike Elisabeth Kohrs, Eugene Jennifer Jin, and Peter Robin Hiesinger. Rab gtpases and membrane trafficking in neurodegeneration. *Current Biology*, 28(8):R471–R486, 2018.

[84] Harald Stenmark and Vesa M Olkkonen. The rab gtpase family. *Genome biology*, 2(5):1–7, 2001.

[85] Alex H Hutagalung and Peter J Novick. Role of rab gtpases in membrane traffic and cell physiology. *Physiological reviews*, 91(1):119–149, 2011.

[86] Harald Stenmark. Rab gtpases as coordinators of vesicle traffic. *Nature reviews Molecular cell biology*, 10(8):513–525, 2009.

[87] Birte Sönnichsen, Stefano De Renzis, Erik Nielsen, Jens Rietdorf, and Marino Zerial. Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of rab4, rab5, and rab11. *The Journal of cell biology*, 149(4):901–914, 2000.

[88] Philippe Chavrier, Robert G Parton, Hans Peter Hauri, Kai Simons, and Marino Zerial. Localization of low molecular weight gtp binding proteins to exocytic and endocytic compartments. *Cell*, 62(2):317–329, 1990.

[89] Frederick R Maxfield and Timothy E McGraw. Endocytic recycling. *Nature reviews Molecular cell biology*, 5(2):121–132, 2004.

[90] Suzanne R Pfeffer. Rab gtpase regulation of membrane identity. *Current opinion in cell biology*, 25(4):414–419, 2013.

[91] Sirkka Keränen and Jussi Jäntti. *Regulatory mechanisms of intracellular membrane transport*, volume 10. Springer Science & Business Media, 2004.

[92] Suzanne R Pfeffer. Rab gtpases: specifying and deciphering organelle identity and function. *Trends in cell biology*, 11(12):487–491, 2001.

[93] Francis A Barr. Rab gtpases and membrane identity: causal or inconsequential? *Journal of Cell Biology*, 202(2):191–199, 2013.

[94] Bassam R Ali, Christina Wasmeier, Lynn Lamoreux, Molly Strom, and Miguel C Seabra. Multiple regions contribute to membrane targeting of rab gtpases. *Journal of cell science*, 117(26):6401–6412, 2004.

[95] Andreas Gerondopoulos, Lars Langemeyer, Jin-Rui Liang, Andrea Linford, and Francis A Barr. Bloc-3 mutated in hermansky-pudlak syndrome is a rab32/38 guanine nucleotide exchange factor. *Current Biology*, 22(22):2135–2139, 2012.

[96] Julia Blümer, Juliana Rey, Leif Dehmelt, Tomáš Mazel, Yao-Wen Wu, Philippe Bastiaens, Roger S Goody, and Aymelt Itzen. Rabgefs are a major determinant for specific rab membrane targeting. *Journal of Cell Biology*, 200(3):287–300, 2013.

[97] Cecilia Bucci, Robert G Parton, Ian H Mather, Henk Stunnenberg, Kai Simons, Bernard Hoflack, and Marino Zerial. The small gtpase rab5 functions as a regulatory factor in the early endocytic pathway. *Cell*, 70(5):715–728, 1992.

[98] Oliver Ullrich, Sigrid Reinsch, Sylvie Urbé, Marino Zerial, and Robert G Parton. Rab11 regulates recycling through the pericentriolar recycling endosome. *The Journal of cell biology*, 135(4):913–924, 1996.

[99] Chiara Vantaggiato, Elena Panzeri, Marianna Castelli, Andrea Citterio, Alessia Arnoldi, Filippo Maria Santorelli, Rocco Liguori, Marina Scarlato, Olimpia Musumeci, Antonio Toscano, et al. Zfyve26/spastizin and spg11/spatacsin mutations in hereditary spastic paraplegia types ar-spg15 and ar-spg11 have different effects on autophagy and endocytosis. *Autophagy*, 15(1):34–57, 2019.

[100] Rafael Dominik Fritz and Olivier Pertz. The dynamics of spatio-temporal rho gtpase signaling: formation of signaling patterns. *F1000Research*, 5, 2016.

[101] Ron Milo, Paul Jorgensen, Uri Moran, Griffin Weber, and Michael Springer. BioNumbers The database of key numbers in molecular and cell biology. *Nucleic Acids Research*, 38(SUPPL.1), 10 2009.

[102] Lena Katharina Oesterlin, Olena Pylypenko, and Bruno Goud. Effectors of rab gtpases: Rab binding specificity and their role in coordination of rab function and localization. *Ras Superfamily Small G Proteins: Biology and Mechanisms 2: Transport*, pages 39–66, 2014.

[103] Alice Cezanne, Janelle Lauer, Anastasia Solomatina, Ivo F Sbalzarini, and Marino Zerial. A non-linear system patterns rab5 gtpase on the membrane. *Elife*, 9:e54434, 2020.

[104] John Chant and Lisa Stowers. Gtpase cascades choreographing cellular behavior: Movement, morphogenesis, and more. *Cell*, 81(1):1–4, 1995.

[105] Felix E Rivera-Molina and Peter Novick. A rab gap cascade defines the boundary between two rab gtpases on the secretory pathway. *Proceedings of the National Academy of Sciences*, 106:14408 – 14413, 2009.

[106] Danièle Stalder and Bruno Antonny. Arf gtpase regulation through cascade mechanisms and positive feedback loops. *FEBS Letters*, 587(13):2028–2035, 2013. St. Petersburg.

[107] Ryan S. D'Souza, Rachel Semus, Emily A. Billings, Corey B. Meyer, Kathryn Conger, and James E. Casanova. Rab4 orchestrates a small gtpase cascade for recruitment of adaptor proteins to early endosomes. *Current Biology*, 24(11):1187–1198, 2014.

[108] Heike Esters, Kirill Alexandrov, Andrei Iakovenko, Tania Ivanova, Nicolas Thomä, Vladimir Rybin, Marino Zerial, Axel J Scheidig, and Roger S Goody. Vps9, rabex-5 and dss4: proteins with weak but distinct nucleotide-exchange activities for rab proteins11edited by j. karn. *Journal of Molecular Biology*, 310(1):141–156, 2001.

[109] Vladimir Rybin, Oliver Ullrich, Mariantonietta Rubino, Kirill Alexandrov, Iris Simon, Miguel C Seabra, Roger Goody, and Marino Zerial. Gtpase activity of rab5 acts as a timer for endocytic membrane fusion. *Nature*, 383(6597):266–269, 1996.

[110] Urban Bezeljak, Hrushikesh Loya, Beata Kaczmarek, Timothy E Saunders, and Martin Loose. Stochastic activation and bistability in a rab gtpase regulatory network. *Proceedings of the National Academy of Sciences*, 117(12):6540–6549, 2020.

[111] Konstantin Gavriljuk, Emerich-Mihai Gazdag, Aymelt Itzen, Carsten Kötting, Roger S Goody, and Klaus Gerwert. Catalytic mechanism of a mammalian rab· rabgap complex in atomic detail. *Proceedings of the National Academy of Sciences*, 109(52):21348–21353, 2012.

[112] Carla Bosia, Francesco Sgrò, Laura Conti, Carlo Baldassi, Davide Brusa, Federica Cavallo, Ferdinando Di Cunto, Emilia Turco, Andrea Pagnani, and Riccardo Zecchina. Rnas competing for micrornas mutually influence their fluctuations in a highly non-linear microrna-dependent manner in single cells. *Genome biology*, 18(1):1–14, 2017.

[113] Marco Del Giudice, Carla Bosia, Silvia Grigolon, and Stefano Bo. Stochastic sequestration dynamics: a minimal model with extrinsic noise for bimodal distributions and competitors correlation. *Scientific reports*, 8(1):10387, 2018.

[114] Michael J Sanderson, Ian Smith, Ian Parker, and Martin D Bootman. Fluorescence microscopy. *Cold Spring Harbor Protocols*, 2014(10):pdb–top071795, 2014.

[115] Tae Kyung Kim and James H Eberwine. Mammalian cell transfection: the present and the future. *Analytical and bioanalytical chemistry*, 397:3173–3178, 2010.

[116] Joshua A Broussard, Benjamin Rappaz, Donna J Webb, and Claire M Brown. Fluorescence resonance energy transfer microscopy as demonstrated by measuring the activation of the serine/threonine kinase akt. *Nature protocols*, 8(2):265, 2013.

[117] P. Lánský, V. Křivan, and J. P. Rospars. Ligand-receptor interaction under periodic stimulation: A modeling study of concentration chemoreceptors. *European Biophysics Journal*, 30(2):110–120, 2001.

[118] Federico Sevlever, Juan Pablo Di Bella, and Alejandra C Ventura. Discriminating between negative cooperativity and ligand binding to independent sites using pre-equilibrium properties of binding curves. *PLOS Computational Biology*, 16(6):e1007929, 2020.

[119] Filipa Rijo-Ferreira and Joseph S Takahashi. Genomics of circadian rhythms in health and disease. *Genome medicine*, 11:1–16, 2019.

[120] Dong Hyun Kim, Dominic Grün, and Alexander van Oudenaarden. Dampening of expression oscillations by synchronous regulation of a microrna and its target. *Nature genetics*, 45(11):1337–1344, 2013.

[121] Vladimir V Galatenko, Alexey V Galatenko, Timur R Samatov, Andrey A Turchinovich, Maxim Yu Shkurnikov, Julia A Makarova, and Alexander G Tonevitsky. Comprehensive network of mirna-induced intergenic interactions and a biological role of its core in cancer. *Scientific reports*, 8(1):2418, 2018.

[122] Yimei Cai, Xiaomin Yu, Songnian Hu, and Jun Yu. A brief review on the mechanisms of mirna regulation. *Genomics, proteomics & bioinformatics*, 7(4):147–154, 2009.

[123] Zijun Luo, Xuping Xu, Peili Gu, David Lonard, Preethi H Gunaratne, Austin J Cooney, and Robert Azencott. mirna regulatory circuits in es cells differentiation: A chemical kinetics modeling approach. *PLoS One*, 6(10):e23263, 2011.

[124] Xin Lai, Olaf Wolkenhauer, and Julio Vera. Understanding microrna-mediated gene regulatory networks through mathematical modelling. *Nucleic acids research*, 44(13):6019–6035, 2016.

[125] Daniel C Ellwanger, Florian A Büttner, Hans-Werner Mewes, and Volker Stümpflen. The sufficient minimal set of mirna seed types. *Bioinformatics*, 27(10):1346–1350, 2011.

[126] Angela Re, Michele Caselle, Federico Bussolino, et al. Microrna-mediated regulatory circuits: outlook and perspectives. *Physical biology*, 14(4):045001, 2017.

[127] John Tsang, Jun Zhu, and Alexander van Oudenaarden. Microrna-mediated feedback and feedforward loops are recurrent network motifs in mammals. *Molecular cell*, 26(5):753–767, 2007.

[128] Baohong Zhang, Xiaoping Pan, George P Cobb, and Todd A Anderson. micrornas as oncogenes and tumor suppressors. *Developmental biology*, 302(1):1–12, 2007.

[129] Jayanth Kumar Palanichamy and Dinesh S Rao. mirna dysregulation in cancer: towards a mechanistic understanding. *Frontiers in genetics*, 5:54, 2014.

[130] Jörn M Schmiedel, Sandy L Klemm, Yannan Zheng, Apratim Sahay, Nils Blüthgen, Debora S Marks, and Alexander van Oudenaarden. Microrna control of protein expression noise. *Science*, 348(6230):128–132, 2015.

[131] Eléonore M'Baya-Moutoula, Loïc Louvet, Roland Molinié, Ida Chiara Guerrera, Catherine Cerutti, Ophélie Fourdinier, Virginie Nourry, Laurent Gutierrez, Patrice Morlière, François Mesnard, et al. A multi-omics analysis of the regulatory changes induced by mir-223 in a monocyte/macrophage cell line. *Biochimica Et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1864(8):2664–2678, 2018.

[132] Martin Hart, Barbara Walch-Rückheim, Lena Krammes, Tim Kehl, Stefanie Rheinheimer, Tanja Tänzer, Birgit Glombitza, Martina Sester, Hans-Peter Lenhof, Andreas Keller, et al. mir-34a as hub of t cell regulation networks. *Journal for immunotherapy of cancer*, 7:1–11, 2019.

[133] Luis Jaime Castro-Vega, Eric Letouzé, Nelly Burnichon, Alexandre Buffet, Pierre-Hélie Disderot, Emmanuel Khalifa, Céline Loriot, Nabila Elarouci, Aurélie Morin, Mélanie Menara, et al. Multi-omics analysis defines core genomic alterations in pheochromocytomas and paragangliomas. *Nature communications*, 6(1):6044, 2015.

[134] Shankar Mukherji, Margaret S Ebert, Grace XY Zheng, John S Tsang, Phillip A Sharp, and Alexander Van Oudenaarden. Micrornas can generate thresholds in target gene expression. *Nature genetics*, 43(9):854–859, 2011.

[135] Ugo Ala, Florian A Karreth, Carla Bosia, Andrea Pagnani, Riccardo Taulli, Valentine Léopold, Yvonne Tay, Paolo Provero, Riccardo Zecchina, and Pier Paolo Pandolfi. Integrated transcriptional and competitive endogenous rna networks are cross-regulated in permissive molecular environments. *Proceedings of the National Academy of Sciences*, 110(18):7154–7159, 2013.

[136] Matteo Figliuzzi, Enzo Marinari, and Andrea De Martino. Micrornas as a selective channel of communication between competing rnas: a steady-state theory. *Biophysical journal*, 104(5):1203–1213, 2013.

[137] Marco Del Giudice, Stefano Bo, Silvia Grigolon, and Carla Bosia. On the role of extrinsic noise in microrna-mediated bimodal gene expression. *PLoS computational biology*, 14(4):e1006063, 2018.

[138] Leonardo Salmena, Laura Poliseno, Yvonne Tay, Lev Kats, and Pier Paolo Pandolfi. A cerna hypothesis: the rosetta stone of a hidden rna language? *cell*, 146(3):353–358, 2011.

[139] Yvonne Tay, John Rinn, and Pier Paolo Pandolfi. The multilayered complexity of cerna crosstalk and competition. *Nature*, 505(7483):344–352, 2014.

[140] Sung Wook Chi, Julie B Zang, Aldo Mele, and Robert B Darnell. Argonaute hits-clip decodes microrna–mrna interaction maps. *Nature*, 460(7254):479–486, 2009.

[141] Donny D Licatalosi, Aldo Mele, John J Fak, Jernej Ule, Melis Kayikci, Sung Wook Chi, Tyson A Clark, Anthony C Schweitzer, John E Blume, Xuning Wang, et al. Hits-clip yields genome-wide insights into brain alternative rna processing. *Nature*, 456(7221):464–469, 2008.

[142] Laura Poliseno, Leonardo Salmena, Jiangwen Zhang, Brett Carver, William J Haveman, and Pier Paolo Pandolfi. A coding-independent function of gene and pseudogene mrnas regulates tumour biology. *Nature*, 465(7301):1033–1038, 2010.

[143] Yang An, Kendra L Furber, and Shaoping Ji. Pseudogenes regulate parental gene expression via ce rna network. *Journal of cellular and molecular medicine*, 21(1):185–192, 2017.

[144] Thomas B Hansen, Trine I Jensen, Bettina H Clausen, Jesper B Bramsen, Bente Finsen, Christian K Damgaard, and Jørgen Kjems. Natural rna circles function as efficient microrna sponges. *Nature*, 495(7441):384–388, 2013.

[145] Gerald W Dorn and Scot J Matkovich. Menage a trois: intimate relationship among a microrna, long noncoding rna, and mrna, 2014.

[146] Héctor Herranz and Stephen M Cohen. Micrornas and gene regulatory networks: managing the impact of noise in biological systems. *Genes & development*, 24(13):1339–1344, 2010.

[147] Tao Hu, Lei Wei, Shuailin Li, Tianrun Cheng, Xuegong Zhang, and Xiaowo Wang. Single-cell transcriptomes reveal characteristics of micrornas in gene expression noise reduction. *Genomics, Proteomics & Bioinformatics*, 19(3):394–407, 2021.

[148] Matteo Osella, Carla Bosia, Davide Corá, and Michele Caselle. The role of incoherent microrna-mediated feedforward loops in noise buffering. *PLoS computational biology*, 7(3):e1001101, 2011.

[149] Maria A Duk, Maria G Samsonova, and Alexander M Samsonov. Dynamics of mirna driven feed-forward loop depends upon mirna action mechanisms. *BMC genomics*, 15:1–18, 2014.

[150] MA Duk, AM Samsonov, and MG Samsonova. The dynamics of a feed-forward loop depends on the regulator type in its indirect pathway. *Biophysics*, 60:163–172, 2015.

[151] MA Duk and SA Rukolaine. The ability of feed-forward regulatory loops to adapt: Dependence on model parameters and conditions of absolute adaptation. *Biophysics*, 63:837–846, 2018.

[152] Alberto Carignano, Sumit Mukherjee, Abhyudai Singh, and Georg Seelig. Extrinsic noise suppression in micro rna mediated incoherent feedforward loops. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 4353–4359. IEEE, 2018.

[153] Pavol Bokes, Michal Hojcka, and Abhyudai Singh. Buffering gene expression noise by microrna based feedforward regulation. In *Computational Methods in Systems Biology: 16th International Conference, CMSB 2018, Brno, Czech Republic, September 12-14, 2018, Proceedings 16*, pages 129–145. Springer, 2018.

[154] Stefan Semrau and Alexander van Oudenaarden. Studying lineage decision-making in vitro: emerging concepts and novel tools. *Annual review of cell and developmental biology*, 31:317–345, 2015.

[155] Salil Garg and Phillip A Sharp. Single-cell variability guided by micrornas. *Science*, 352(6292):1390–1391, 2016.

[156] Peipei Zhou, Shuiming Cai, Zengrong Liu, and Ruiqi Wang. Mechanisms generating bistability and oscillations in microrna-mediated motifs. *Physical Review E*, 85(4):041916, 2012.

[157] Xu Xue, Wang Xia, and Hu Wenzhong. A modeled dynamic regulatory network of nf-κb and il-6 mediated by mirna. *Biosystems*, 114(3):214–218, 2013.

[158] Marc Goodfellow, Nicholas E Phillips, Cerys Manning, Tobias Galla, and Nancy Papalopulu. microrna input into a neural ultradian oscillator controls emergence and timing of alternative cell states. *Nature communications*, 5(1):3399, 2014.

[159] Taeko Kobayashi and Ryoichiro Kageyama. Hes1 regulates embryonic stem cell differentiation by suppressing notch signaling. *Genes to Cells*, 15(7):689–698, 2010.

[160] Rongmin Chen, Matthew D'Alessandro, and Choogon Lee. mirnas are required for generating a time delay critical for the circadian oscillator. *Current Biology*, 23(20):1959–1968, 2013.

[161] Jacek Krol, Volker Busskamp, Ilona Markiewicz, Michael B Stadler, Sebastian Ribi, Jens Richter, Jens Duebel, Silvia Bicker, Hans Jörg Fehling, Dirk Schübeler, et al. Characterizing light-regulated retinal microrna reveals rapid turnover as a common property of neuronal microrna. *Cell*, 141(4):618–631, 2010.

[162] Xia Wang, Guihua Tian, Zhongfeng Li, and Lei Zheng. The crosstalk between mirna and mammalian circadian clock. *Current Medicinal Chemistry*, 22(13):1582–1588, 2015.

[163] Qian Gao, Lan Zhou, Su-Yu Yang, and Ji-Min Cao. A novel role of microrna 17-5p in the modulation of circadian rhythm. *Scientific Reports*, 6(1):30070, 2016.

[164] Benjamin Nordick, Polly Y. Yu, Guangyuan Liao, and Tian Hong. Nonmodular oscillator and switch based on RNA decay drive regeneration of multimodal gene expression. *Nucleic Acids Research*, 50(7):3693–3708, 2022.

[165] Lioudmila V. Sharova, Alexei A. Sharov, Timur Nedorezov, Yulan Piao, Nabeebi Shaik, and Minoru S.H. Ko. Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Research*, 16(1):45–58, 2 2009.

[166] Cem Albayrak, Christian A. Jordi, Christoph Zechner, Jing Lin, Colette A. Bichsel, Mustafa Khammash, and Savaş Tay. Digital Quantification of Proteins and mRNA in Single Mammalian Cells. *Molecular Cell*, 61(6):914–924, 3 2016.

[167] Brian Reichholf, Veronika A. Herzog, Nina Fasching, Raphael A. Manzenreither, Ivica Sowemimo, and Stefan L. Ameres. Time-Resolved Small RNA Sequencing Unravels the Molecular Principles of MicroRNA Homeostasis. *Molecular Cell*, 75(4):756–768, 8 2019.

[168] Matteo J. Marzi, Francesco Ghini, Benedetta Cerruti, Stefano De Pretis, Paola Bonetti, Chiara Giacomelli, Marcin M. Gorski, Theresia Kress, Mattia Pelizzola, Heiko Muller, Bruno Amati, and Francesco Nicassio. Degradation dynamics of microrna revealed by a novel pulse-chase approach. *Genome Research*, 26(4):554–565, 4 2016.

[169] William E. Salomon, Samson M. Jolly, Melissa J. Moore, Phillip D. Zamore, and Victor Serebrov. Single-Molecule Imaging Reveals that Argonaute Reshapes the Binding Properties of Its Nucleic Acid Guides. *Cell*, 162(1):84–95, 7 2015.

[170] Liang Meng Wee, C. Fabián Flores-Jasso, William E. Salomon, and Phillip D. Zamore. Argonaute divides Its RNA guide into domains with distinct functions and RNA-binding properties. *Cell*, 151(5):1055–1067, 11 2012.

[171] Stephen W. Eichhorn, Huili Guo, Sean E. McGeary, Ricard A. Rodriguez-Mias, Chanseok Shin, Daehyun Baek, Shu hao Hsu, Kalpana Ghoshal, Judit Villén, and David P. Bartel. MRNA Destabilization Is the dominant effect of mammalian microRNAs by the time substantial repression ensues. *Molecular Cell*, 56(1):104–115, 2014.

[172] Manuel de la Mata, Dimos Gaidatzis, Mirela Vitanescu, Michael B Stadler, Corinna Wentzel, Peter Scheiffele, Witold Filipowicz, and Helge Großhans. Potent degradation of neuronal mi RNA s induced by highly complementary targets . *EMBO reports*, 16(4):500–511, 4 2015.

[173] Joe H Levine, Yihan Lin, and Michael B Elowitz. Functional roles of pulsing in genetic circuits. *Science*, 342(6163):1193–1200, 2013.

[174] Daniel R. Romano, Matthew C. Pharris, Neal M. Patel, and Tamara L. Kinzer-Ursem. Competitive tuning: Competition's role in setting the frequency-dependence of Ca2+-dependent proteins. *PLoS Computational Biology*, 13(11):1–26, 2017.

[175] Michael D McKay, Richard J Beckman, and William J Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1):55–61, 2000.

[176] V Eglajs and P Audze. New approach to the design of multifactor experiments. *Problems of Dynamics and Strengths*, 35(1):104–107, 1977.

[177] Ronald L Iman, Jon C Helton, and James E Campbell. An approach to sensitivity analysis of computer models: Part i—introduction, input variable selection and preliminary variable assessment. *Journal of quality technology*, 13(3):174–183, 1981.

[178] Jens Duebel, Katia Marazova, and José-Alain Sahel. Optogenetics. *Current opinion in ophthalmology*, 26(3):226, 2015.

[179] Kenneth J Livak and Thomas D Schmittgen. Analysis of relative gene expression data using real-time quantitative pcr and the 2- $\delta\delta$ct method. *methods*, 25(4):402–408, 2001.

[180] An-Jou Chen, Ji-Hye Paik, Hailei Zhang, Sachet A Shukla, Richard Mortensen, Jian Hu, Haoqiang Ying, Baoli Hu, Jessica Hurt, Natalie Farny, et al. Star rna-binding protein quaking suppresses cancer via stabilization of specific mirna. *Genes & development*, 26(13):1459–1472, 2012.

[181] Erik Lee Snapp. Fluorescent proteins: a cell biologist's user guide. *Trends in cell biology*, 19(11):649–655, 2009.

[182] Hélène Chassin, Marius Müller, Marcel Tigges, Leo Scheller, Moritz Lang, and Martin Fussenegger. A modular degron library for synthetic circuits in mammalian cells. *Nature communications*, 10(1):2013, 2019.