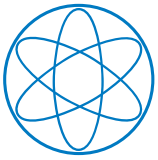
DEPARTMENT OF PHYSICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Dissertation

Spatio-Temporal Organization of Enzymatic Reactions

Giovanni Giunta



Fakultät für Physik
Technische Universität München

Spatio-Temporal Organization of Enzymatic Reactions

Giovanni Giunta

Vollständiger Abdruck der von der Fakultät für Physik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:
Prof. Dr. Friedrich C. Simmel
Prüfende der Dissertation:
1. Prof. Dr. Ulrich Gerland
2. Prof. Dr. Martin Zacharias

Die Dissertation wurde am 16.10.2020 bei der Technischen Universität München eingereicht und durch die Fakultät für Physik am 17.11.2020 angenommen.

Abstract

All living organisms rely on the catalytic activity of enzymes to carry out most of their biological functions. Enzymes catalyze, i.e. accelerate, the conversion of the so-called substrates into products. The efficiency of enzymatic reactions depends on the encounter rate between enzyme and substrate. This, in turn, is affected by the enzyme and substrate motion and also by their spatial arrangement.

Typically, in cells, substrate molecules diffuse while being exposed to several loss mechanisms. To intercept them and shorten their diffusive trajectories, enzymes can be placed in proximity of substrate sources. For example, in cases in which the substrate is imported into the cytoplasm from the external environment, enzymes are sometimes localized at the external membrane. In contrast, when the substrate is produced inside organelles and diffuses into the cytoplasm through the organelle membrane, enzymes are found to be bound to the organelles. In this thesis, we derive an optimal allocation principle for the static spatial arrangement of enzymes so as to maximize the reaction flux of an enzymatic reaction. The allocation principle applies to systems where substrate enters either through the external boundary or the boundary of internal compartments. By drawing an analogy between the biophysical enzyme spatial allocation problem and an economical portfolio optimization, we are able to solve for the optimal strategy. The solution allows us to characterize a transition in the optimal enzyme arrangement from a fully bound configuration into a more extended profile, where enzymes are also present in the interior of the system. Moreover, the optimal strategy permits us to derive a deterministic construction algorithm to systematically build up the optimal enzyme arrangement as the amount of enzymes in the system gradually increases. The construction algorithm can help in the design of optimal synthetic bio-reactors. In fact, the amount of enzymes bound and their localization can be controlled with the use of several techniques.

While substrate molecules undergo a diffusive motion, recent experiments suggest that the motion of enzymes can be affected by the substrate concentration profile. Different models have been proposed over the last decade to describe the enzyme motion and try to explain the experimental results, which are at times contradictory. These models study how enzymes move in response to pre-imposed substrate profiles. They neglect the feedback from enzyme transport onto the substrate caused by the reaction. In this thesis, we show that such a feedback leads to spontaneous spatial

pattern formation from initial homogeneous concentrations for one of the proposed models. For this model, enzymes can move in response to substrate and product gradients due to nonspecific interactions of the enzyme with the substrate and the product. These interactions generate cross-diffusion terms in the enzyme motion that can drive the accumulation of enzymes in regions of low substrate and high product concentrations. In these regions, the reaction then amplifies local substrate and product fluctuations, steepening the substrate and product gradients, hence further driving the accumulation of enzymes. Experimental analysis of spontaneous pattern formation can in principle determine whether or not nonspecific interactions play a role in the enzyme motion.

Zusammenfassung

Alle lebenden Organismen sind für die Erfüllung des Großteils ihrer biologischen Funktionen auf die katalytische Aktivität von Enzymen angewiesen. Enzyme katalysieren, d.h. sie beschleunigen die Umwandlung der so genannten Substrate in Produkte. Die Effizienz enzymatischer Reaktionen hängt von der Begegnungsrate zwischen Enzym und Substrat ab. Diese wiederum wird von der Bewegung der Enzyme und Substrate sowie von deren räumlicher Anordnung beeinflusst.

Typischerweise diffundieren Substratmoleküle in Zellen, während sie verschiedenen Verlustmechanismen ausgesetzt sind. Um sie abzufangen und ihre diffusiven Wege zu verkürzen, können Enzyme in der Nähe von Substratquellen platziert werden. In Fällen, in denen das Substrat aus der äußeren Umgebung in das Zytoplasma transportiert wird, werden Enzyme beispielsweise manchmal an der äußeren Membran lokalisiert. Wenn das Substrat dagegen innerhalb von Organellen produziert wird und durch die Organellenmembran in das Zytoplasma diffundiert, werden Enzyme an den Organellen gebunden. In dieser Arbeit leiten wir ein optimales Allokationsprinzip für die statische räumliche Anordnung der Enzyme ab, um den Reaktionsfluss einer enzymatischen Reaktion zu maximieren. Das Allokationsprinzip gilt für Systeme, bei denen das Substrat entweder durch eine äußere Begrenzung oder durch die Begrenzungen innerer Kompartimente eintritt. Indem wir eine Analogie zwischen dem biophysikalischen räumlichen Allokationsproblem der Enzyme und einer ökonomischen Portfoliooptimierung ziehen, sind wir in der Lage, die optimale Strategie zu finden. Diese Lösung erlaubt es uns, einen Übergang in der optimalen Enzymanordnung von einer vollständig gebundenen Konfiguration in ein erweitertes Profil, in dem auch Enzyme im Inneren des Systems vorhanden sind, zu charakterisieren. Darüber hinaus können wir von der optimalen Strategie einen deterministischen Konstruktionsalgorithmus ableiten, um bei allmählicher Zunahme der Menge der Enzyme im System systematisch die optimale Enzymanordnung aufzubauen. Der Konstruktionsalgorithmus kann beim Entwurf optimaler synthetischer Bioreaktoren helfen. Tatsächlich kann die Menge der gebundenen Enzyme und ihre Lokalisierung mit Hilfe verschiedener Techniken kontrolliert werden.

Während Substratmoleküle einer diffusiven Bewegung unterliegen, deuten neuere Experimente darauf hin, dass die Bewegung der Enzyme durch das Substratkonzentrationsprofil beeinflusst werden kann. Im letzten Jahrzehnt wurden verschiedene

Modelle vorgeschlagen, um die Enzymbewegung zu beschreiben und zu versuchen, die manchmal widersprüchlichen experimentellen Ergebnisse zu erklären. Diese Modelle untersuchen, wie sich Enzyme als Reaktion auf vorgegebene Substratprofile bewegen, vernachlässigen dabei aber die Rückkopplung des durch die Reaktion verursachten Enzymtransports auf das Substrat. In dieser Arbeit zeigen wir, dass eine solche Rückkopplung bei einem der vorgeschlagenen Modelle zur spontanen räumlichen Musterbildung aus anfänglich homogenen Konzentrationen führt. Bei diesem Modell können sich die Enzyme als Reaktion auf Substrat- und Produktgradienten aufgrund unspezifischer Wechselwirkungen des Enzyms mit dem Substrat und dem Produkt bewegen. Diese Interaktionen erzeugen Kreuzdiffusionsterme in der Enzymbewegung, die die Akkumulation von Enzymen in Regionen mit niedrigen Substrat- und hohen Produktkonzentrationen antreiben können. In diesen Regionen verstärkt die Reaktion dann lokale Substrat- und Produktfluktuationen, wodurch die Substrat- und Produktgradienten steiler werden und somit die Anreicherung von Enzymen weiter vorangetrieben wird. Im Prinzip kann man über die experimentelle Untersuchung der spontanen Musterbildung bestimmen, ob unspezifische Wechselwirkungen bei der Enzymbewegung eine Rolle spielen oder nicht.

Contents

1. What is an enzyme and how does it work?

Enzymes are catalysts that accelerate biochemical reactions at fast enough rates to sustain life. As catalysts, they have the role of lowering the energy barrier for chemical reactions to proceed. They are mostly proteins and are used by all living cells to catalyze the conversion of food to energy, the formation of building blocks and the disposal of waste chemicals. The enzyme catalysis begins with the binding of the enzyme to a certain molecule called substrate. In this bound state the enzyme and the substrate form a temporary complex, the structure of which favors the conversion of the substrate into a product. Once the reaction occurs, the product detaches from the enzyme and the enzyme is again ready for another catalytic cycle. The enzymatic activity of an ensemble of many enzymes can depend on the amount of substrate and product available in the surroundings, or it can be inhibited by other chemicals. In this chapter, we provide a more detailed description of enzyme catalysis from an energetic and structural point of view. Moreover, we present different enzyme kinetic models describing how the rate of product formation is affected by changes in the amount of substrate, product and how it is affected in the presence of inhibitors.

1.1. Discovery of enzymes

The word ‘enzyme’ was first coined in 1876 by the German physiologist Wilhelm Kühne, from Ancient Greek *ἐν* (en,"in") + *ζύμε* (zúmē, "leavened" or "in yeast") [1]. It referred to the substances used by alive yeast cells to carry out the fermentation of sugars to alcohol. It was only in 1897 that Eduard Buchner discovered that fermentation could take place with cell-free extracts [2]. From this moment on, the word enzyme started to be used to refer to extracts and substances able to catalyze biochemical reactions. However, the nature and the mechanism behind the enzyme catalysis were largely unknown until mid 1900s. It was in 1926 that James B. Sumner crystallized the enzyme urease (responsible for breaking down urine into ammonia and carbon dioxide) and showed that it is a pure protein [3]. Similarly, John H. Northrop in 1929 showed that the enzymes pepsin, trypsin and chymotrypsin (active in the digestive process) are pure proteins as well [4, 5]. Since these discoveries, enzymes have been long thought

to be exclusively proteins, i.e. composed of long chains of amino acids. However in 1978 and 1982, respectively, Thomas R. Cech and Sidney Altman showed that folded RNA molecules can also catalyze biochemical reactions [6, 7]. Such RNA molecules have been named 'ribozymes' (ribonucleic acid enzymes) [7] and to date they are the simplest form of enzymes. The discovery of ribozymes is important in the context of the origin of life on Earth. A long standing question in evolutionary biology is: What came first, the proteins needed for DNA replication or the DNA strands containing information about the proteins? In the 'RNA world hypothesis' we have RNA strands that are both capable of containing information in form of sequences of nucleic acids, just like DNA, and also show catalytic activity [8]. Hence a likely scenario for early life on Earth is an environment where RNA strands have the opportunity of self-replicate and store information encoded as nucleic acids sequences.

Although ribozymes might have played a crucial role in the early life, most of the known enzymes in current living organisms are proteins or a mixture of proteinaceous structures and RNA sequences with catalytic activity. Hence the main focus of this and the following chapters will be on protein enzymes.

1.2. Enzyme catalysis

1.2.1. Energetics of enzyme catalysis

Enzyme catalysis initiates when a certain molecule S called substrate binds to the enzyme E, forming the enzyme-substrate complex ES. In the complexed form, the enzyme-substrate interaction favors the formation of the biochemical product P. Once the reaction is completed the product detaches from the enzyme, the enzyme reverts to its original form and is able to bind another substrate molecule to repeat the catalysis [9]. An overview of the thermodynamics of enzyme catalysis can be seen in Fig. 1.1. The figure shows how the Gibbs free energy changes as the reaction proceeds in the absence (purple dashed line) and in the presence (orange line) of the enzyme. In the absence of the enzyme the reaction can occur spontaneously if the difference in Gibbs free energy between the substrate and the product is negative [10]. This implies that the product is in a favorable, more stable energetic state than the substrate. However a high energy barrier, known as activation energy, needs to be overcome for the reaction to occur. The activation energy depends on the stability of the substrate molecule. The more stable the substrate molecule, the higher the barrier and the longer it takes for thermal fluctuations to push the reaction forward. The rate of the reaction strongly depends on the rate at which the state corresponding to the top of the energy barrier is reached. Such state is called transition state and once it is passed, the substrate quickly transforms into the product. The presence of the enzyme has the overall effect

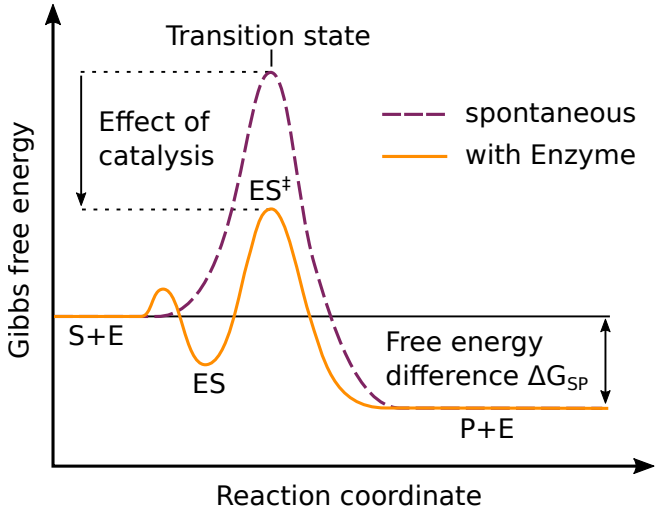


Figure 1.1.: Thermodynamics of enzyme catalysis. The Gibbs free energy is plotted as the reaction of a substrate S into a product P proceeds from left to right. The purple dashed line and the orange solid line represent the change in free energy in case of a spontaneous reaction and a reaction in presence of the enzyme E respectively. The effect of catalysis is to lower the activation energy of the transition state, which corresponds to the highest energetic state. First S and E form the enzyme-substrate complex ES. In the complexed form the transition state $ES^{\ddagger}$ is reached more easily. Once the transition state is passed, the product is quickly formed. Hence in the presence of E, the rate of reaction is increased. The most right hand side of the reaction coordinate represents the state where P is detached from E.

of lowering the activation energy needed for the reaction to occur. According to the typical coarse-grained picture, the substrate and the enzyme first bind, forming the ES complex, and then the real step of catalysis takes place. The enzyme-substrate interaction favors the reaction and the new transition state $ES^{\ddagger}$ has a lower activation energy as compared to the transition state of the spontaneous reaction. The transition state is reached more easily and the rate of product formation is increased [11]. Once the catalysis happens and the product is formed, the product detaches from the enzyme leading us to the end of the catalytic cycle. In Fig. 1.1, this corresponds to the most right hand side of the reaction coordinate, where enzyme and product are present separately. The enzyme is now available for the binding of a new substrate molecule and the catalytic cycle can start over.

How much faster can reactions be in presence of enzymes? One of the most proficient enzymes is the arginine decarboxylase that catalyzes the conversion of L-arginine into agmatine and carbon dioxide. The acceleration of such reaction is of 19 order of magnitudes [12]. In this case, the spontaneous reaction would occur in a time almost as long as the age of the Earth. Instead, the presence of the enzyme speeds up the reaction rate to $\approx 10^3$ reactions per second. Another example is the carbonic anhydrase which is considered to be one of the fastest enzymes and it catalyzes the interconversion of carbon dioxide and water to hydrogen protons and bicarbonate ions [13]. Carbon

anhydrase speeds up the reaction rate by 7 order of magnitudes taking the rate from $\approx 10^{-1}s^{-1}$ to $\approx 10^{6}s^{-1}$ [12]. These examples show the key role of enzymes in living systems: although they do not alter the chemical equilibrium of a reaction [10], they help biochemical reactions to be on relevant timescales.

1.2.2. Structural view of enzyme catalysis

Enzymes are usually much larger than their corresponding substrate. For example, the to date smallest protein enzyme, the monomer of 4-Oxalocrotonate tautomerase (4-OT), is a chain of 62 amino acids residues [14] catalyzing the isomerization of β,γ-unsaturated enones to their α,β-isomers, which are organic compounds of $\approx$ 156Da (hence comparable to the molecular weight of a single amino acid). Just 4 out of the 62 amino acids of the monomeric 4-OT enzyme are involved in the catalysis [15]. For a generic enzyme this is often the case: only a small part of the enzyme is involved in the catalysis, the so called 'catalytic site'. The catalytic site is often adjacent to the binding site, which is the site at which the substrate binds the enzyme. Together the binding site and the catalytic site form the 'active site', which is the site where all the steps of catalysis as depicted in Fig. 1.1 take place. In Fig. 1.2 the structure of the enzyme

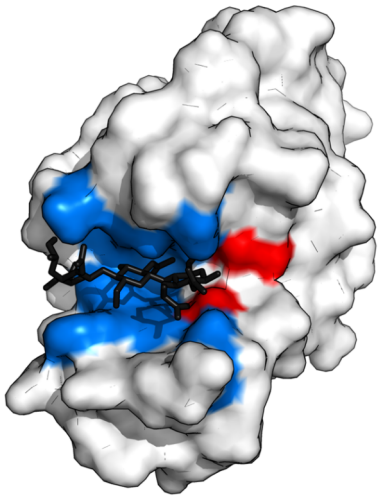


Figure 1.2.: An example of enzymatic structure. In the figure we see the structure of the enzyme lysozyme (white molecule), responsible for the breaking down of the cell wall of gram positive bacteria. The substrate (black) binds the enzyme at the binding site (blue area). Adjacent to the binding site is the catalytic site (red area), where the catalitic step of the enzymatic reaction takes place. The binding site and the catalitic site constitute together the active site. Figure adapted from Wikimedia Commons: 'Enzyme structure' by Thomas Shafee under CC BY 4.0 License. The protein structure original sources are PDB:9LYZ and [16].

lysozyme is shown (white molecule). The lysozyme is an enzyme composed of 129 amino acids that plays an important role in the innate immune systems of animals. It is responsible for breaking the bonds between residues of the cell wall of gram positive bacteria, ultimately causing their lysis. More specifically, it catalyzes the hydrolysis of 1,4-beta-linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues in the peptidoglycan [16]. The peptidoglycan is the polymer consisting of sugars

and amino acids that forms the cell wall. In Fig. 1.2, we can see how the active site, composed of the binding site (blue) and the catalytic site (red) is only a small part of the enzyme. Moreover, it is immediately visible that the substrate molecule (black) is much smaller than the enzyme.

Now that we know how a typical enzyme looks like, we can analyze the molecular mechanisms underlying the catalysis.

Substrate binding

The binding of substrate to the active site is highly specific [17]. To characterize this binding, Emil Fisher proposed the so called 'lock and key' model in 1894. He suggested that enzyme and substrate have complementary geometrical structures that fit just like a key into a lock [18]. Although the model well explains the high specificity of binding, it describes the enzyme and the substrate as rigid objects which is not the case. Moreover, the model does not explain how the energy of the transition state can be lowered. A more realistic view of substrate binding and catalysis in general is provided by the 'induced fit' model proposed by Daniel Koshman in 1958 [19]. According to the model, after the substrate binds, the enzyme and especially its active site are continuously reshaped by the interaction with the substrate in a way that the catalytic site of the enzyme is properly aligned with the substrate. In this way the substrate gets tightly bound to the enzyme and the interaction with the catalytic site favors the substrate into its transition state, hence lowering the activation energy as illustrated in Fig. 1.1. Moreover, the conformational changes of the induced fit model can contribute to a conformational proofreading mechanism that helps the enzyme in selecting the right substrate in noisy biochemical environments with many similar molecules [20].

In Fig. 1.3 we see an example of induced fit for the enzyme hexokinase. Hexokinase catalyzes the addition of a phosphate group (phosphorylation) to six-carbon sugar molecules (hexoses). The most important substrate of hexokinase is glucose. Its phosphorylation is the first step of glycolysis, which is a very important metabolic pathway that takes place in cells of all living organisms, and leads to the conversion of sugar molecules into energy. More specifically, glycolysis converts glucose into pyruvate acid and an hydrogen ion. The energy released during the different steps of the pathway is stored in high-energy molecules, such as adenosine triphosphate (ATP) and reduced nicotinamide adenine dinucleotide (NADH), which are the main sources of energy in cells. In Fig. 1.3, we can see how the hexokinase changes structure upon binding of adenosine diphosphate (ADP) and the sugar xylose. Note that upon binding of glucose the structural change of hexokinase is very similar to the one induced by the binding of xylose, see [21] for details. The substrates are shown in black and the enzyme binding site is evidenced in blue. The yellow element represents the co-factor

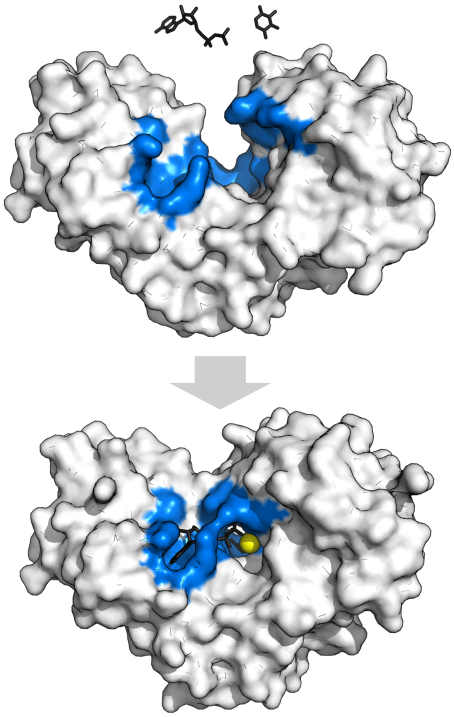


Figure 1.3.: An example of enzyme induced fit structural change. In the figure we see the conformational changes induced by subsrate binding onto the enzyme hexokinase. The substrates adenosine diphosphate and xylose are shown in black, the binding site is shown in blue. On the upper panel the enzyme is free and it is in an open conformation. In the lower panel the substrate is bound and the enzyme assumes a closed conformation. The yellow particle represents the co-factor Mg^{2+}, which binds to the enzyme to favor the reaction. The figure was adapted from Wikimedia Commons: 'Hexokinase induced fit' by Thomas Shafee under CC BY 4.0 License. The protein structures original sources are PDB:2E2N and PDB:2E2Q, both published in [21].

Mg^{2+}, which binds to the enzyme to help the reaction. Upon substrate binding the enzyme passes from an open state (upper panel) into a closed configuration (lower panel). The conformational change followed by the enzyme consists in a rotation of about 25° of the small domain (right domain in the figure) relative to the large domain (left domain in the figure) [21]. The closed conformation stabilizes the binding of substrate and ensures efficient catalysis. The enzyme adenylate kinase (ADK), which catalyzes the reversible conversion of ATP and AMP to two ADPs, follows a similar mechanism. ADK has two substrate binding domains and once the substrates bind, the enzyme appears to be preferably in a closed state with two flexible lid domains closing the binding sites [22]. ADK presents a third energetically stable state in which the two flexible lids are half-closed. This third state is the result of a trade-off between rapid substrate exchange and tight closing that favor efficient catalysis [23].

Catalytic step

How does catalysis happen after the catalytic site of the enzyme is oriented towards the substrate? One of the mechanism is via electrostatic forces [24]. The catalytic site can have complementary charges to that of the substrate transition state, hence favoring the substrate into the transition state, i.e. lowering the energy barrier for the

reaction to occur. Alternatively the enzyme could participate in the reaction forming covalent intermediates for which the energy barrier is lower. A typical example of this is the digestion of peptide chains by the serine proteases. In this case, at the active site of the enzyme, the protein polypeptide binds either via hydrophobic or ionic interactions. Then a proton transfer, mediated by favorable electrostatic interactions, causes the cleavage of the peptide bond and a part of the polypeptide is detached from the enzyme. However the other part is still covalently attached to the enzyme, forming an intermediate complex. A second proton transfer, favored by the presence of water, frees the binding site from the remaining polypeptide [9]. The opposite process, where a peptide bond is formed can also be favored by the presence of an enzyme. In this case the enzyme just serves as a template where the two substrates are brought to contact and oriented in a way to favor their reaction [9].

Generally, different enzymes follow different mechanisms, sometimes involving extra cofactors [17] as for example shown in Fig. 1.3. However the underlying physical forces in play are similar to the ones illustrated above. Non-covalent interactions, such as hydrogen bonds, electrostatic interactions, steric, ionic and hydrophobic forces can all contribute to the catalysis. Sometimes these can also trigger the formation of temporary intermediates that have lower activation energies. Note that in this case the Gibbs free energy of the reaction does not behave as in Fig. 1.1, but it presents more than two peaks [25]. However the overall effect of the enzyme would still be the one of lowering the activation energy barrier.

1.3. Enzyme kinetics

In the current section we are going to discuss different kinetics models used in the characterization of enzymatic reactions. The different models describe the rate of reaction, i.e. the rate at which products are formed, as the concentrations of substrate, enzyme and product are varied.

1.3.1. Michaelis-Menten equation

In 1913 Leonor Michaelis and Maud Menten presented a model describing the kinetics of an enzyme following the same steps of the catalysis illustrated in Fig. 1.1 [26]. First the substrate and the enzyme form a complex, then the catalytic step of the reaction takes place and the product is formed. They assumed the complex formation to be a reversible process with forward and reverse rates k_1 and k_{-1} respectively. However they considered the catalytic step to be an irreversible process. This assumption is valid if the dissociation of the enzyme-product complex is fast and if the amount of product at the beginning of the reaction is low [25]. Hence, the reverse reaction of product and

enzyme to the enzymatic complex can be neglected and the reaction scheme takes the form

$$\mathrm{E} + \mathrm{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{k_{\mathrm{cat}}} \mathrm{E} + \mathrm{P}\,. \tag{1.1}$$

Starting from this reaction scheme, Michaelis and Menten derived a formula characterizing the velocity of reaction v as a function of substrate concentration s. The reaction velocity corresponds to the rate of product formation, which according to the law of mass action [27] takes the form

$$v = \frac{\mathrm{d}p}{\mathrm{d}t} = k_{\mathrm{cat}} c\,, \tag{1.2}$$

where p and c are the concentrations of product and enzyme-substrate complex respectively and t is time. Here we rederive the Michaelis-Menten formula by following the derivation of Briggs and Haldane [28], i.e. by assuming that the concentration of the enzyme-substrate complex does not change on the timescale of product formation. This assumption, known as quasi-steady-state assumption, implies that the rate of complex formation equals the rate of dissociation

$$k_1 e_0 s = k_{-1} c + k_{\mathrm{cat}} c\,. \tag{1.3}$$

Here e_0 is the concentration of free enzymes which is equal to $e_0 = e - c$, where e is the total concentration of enzymes as given by the sum of free and complexed enzymes. Note that since the enzyme is a catalyst, its total amount does not change over the course of the reaction. From relation (1.3), we obtain

$$c = \frac{e\,s}{\frac{k_{-1}+k_{\mathrm{cat}}}{k_1} + s} := \frac{e\,s}{K_{\mathrm{M}} + s}\,, \tag{1.4}$$

where we defined the Michaelis-Menten constant $K_{\mathrm{M}} = (k_{-1} + k_{\mathrm{cat}})/k_1$. Finally, by inserting the above expression for c into Eq. (1.2), we get the famous Michaelis-Menten expression, describing the velocity of an enzymatic reaction as a function of the substrate and enzyme concentration

$$v = k_{\mathrm{cat}} \frac{e\,s}{K_{\mathrm{M}} + s}\,. \tag{1.5}$$

In Fig. 1.4 is shown the nonlinear behavior of the Michaelis-Menten relation as a function of substrate concentration s, for a constant amount of enzymes e. We see that we have a linear regime for $s \ll K_{\mathrm{M}}$ and a saturated regime for $s \gg K_{\mathrm{M}}$, where the reaction rate tends to its maximal value $v_{\max} = k_{\mathrm{cat}} e$. The value of K_{M} represents the value of substrate concentration for which the reaction rate is half of the maximal value. In the linear regime (see Fig. 1.4) most of the enzymes are free. The reaction

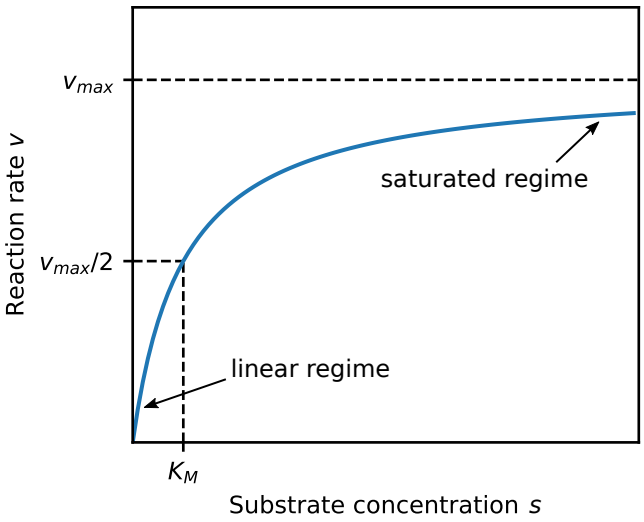


Figure 1.4.: Michaelis-Menten kinetics. The reaction rate v is plotted as a function of the substrate concentration s. For values of $s \ll K_\mathrm{M}$, most of the enzymes are free and the reaction rate v depends linearly on s (linear regime). For values of $s \gg K_\mathrm{M}$ most of the enzymes are complexed with the substrate; the reaction rate v tends to its maximal value $v_\mathrm{max} = k_\mathrm{cat} e$ (saturated regime). The point at which $s = K_\mathrm{M}$ corresponds to the half-maximum of the reaction rate with $v = v_\mathrm{max}/2$.

proceeds as a second order chemical reaction with rate $v = (k_\mathrm{cat}/K_\mathrm{M}) e\, s$. The ratio $k_\mathrm{cat}/K_\mathrm{M}$ corresponds to the rate constant of the reaction and it is known as catalytic efficiency. For example, one can use values of the catalytic efficiencies to analyze the preferences of an enzyme for different substrates [29]. For a very efficient enzyme for which $k_\mathrm{cat} \gg k_{-1}$, the highest possible value of $k_\mathrm{cat}/K_\mathrm{M}$ is given by the rate of formation of the ES complex k_1. In Sec. 2.1 we will see that for the so-called diffusion-limited enzymes, the upper bound of k_1 is $k_1 \approx 10^9 \mathrm{s}^{-1}\mathrm{M}^{-1}$. The limit is set by the rate of diffusive encounters between the enzymes and the substrates, assuming that each encounter leads to the formation of an ES complex. Catalytic efficiencies are in the range of 10^2-$10^9 \mathrm{s}^{-1}\mathrm{M}^{-1}$ with a median at $10^5 \mathrm{s}^{-1}\mathrm{M}^{-1}$, meaning that many enzymes operate far from the most efficient regime [30]. However, enzymes like fumarase and carbonic anhydrase have efficiencies close to the upper limit, in the range of 10^8-$10^9 \mathrm{s}^{-1}\mathrm{M}^{-1}$. In the saturated regime, most of the enzymes are in their complexed form, i.e. occupied by the substrate. In this regime the reaction proceeds as a zeroth order reaction with constant rate $v_\mathrm{max} = k_\mathrm{cat} e$. The Michaelis-Menten constant K_M is an approximate measure of the dissociation constant K_d of the enzyme-substrate complex. For low values of K_M the ES complex is very stable and enzymes saturate at relatively low concentrations of substrate. For high values of K_M, the ES complex is less stable, hence substrate molecules detach more easily and higher concentrations of substrate are needed to saturate the enzymes. Note that K_M becomes exactly the dissociation constant in the limit $k_{-1} \gg k_\mathrm{cat}$, $K_\mathrm{M} \to K_\mathrm{d} = k_{-1}/k_1$.

Surprisingly, the Michaelis-Menten Eq. (1.5) also holds at a single molecule level. By looking at long time traces of a single enzyme, it has been shown that subsequent enzyme turnovers tend to cluster and are not independent from each other. This

is due to a 'memory effect' caused by slow conformational fluctuations of the ES complex [31]. Nevertheless, the inverse of the waiting time for an enzymatic reaction to occur still depends on s, as given by Eq. (1.5). However, the microscopic interpretations of k_{cat} and consequently K_{M} are different [32]. During catalysis the enzyme-substrate complex ES undergoes many conformations and instead of having a single value for the catalytic rate k_{cat}, there is a distribution of values. The prefactor of Eq. (1.5) becomes the weighted harmonic mean of the k_{cat} values corresponding to the catalytic rates of the different enzyme-substrate complex conformations.

Overall, the Michaelis-Menten model provides a phenomenological law that describes the kinetics of many different enzymes. However, there are enzymatic reactions for which the model assumptions do not hold. For the reaction scheme followed by the Michaelis-Menten model, Eq. (1.1), the catalytic step of the reaction is assumed to be irreversible. This assumption holds in case product molecules are quickly released by the enzymatic complex after their formation and we are at the beginning of the reaction, when the product has not been yet accumulating in the system. Moreover enzymatic reactions are often affected by the binding of other molecules. These usually alter the enzyme activity by either inhibiting or increasing it. In the following sections, we will discuss kinetic models that include some of these effects.

1.3.2. Reversible Michaelis-Menten equation

As mentioned in the previous section the Michaelis-Menten model assumes an irreversible catalytic step for the reaction. In case product molecules have accumulated in the system, this assumption no longer holds and the reverse reaction needs to be considered [25]. We can therefore consider a reaction scheme that is similar to Eq. (1.1) but considers a fully reversible catalysis:

$$\text{E} + \text{S} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{ES} \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} \text{E} + \text{P}\,, \tag{1.6}$$

where k_2 and k_{-2} represent the forward and the reverse rate of the catalytic step of the enzymatic reaction, respectively. The reaction rate v can be derived by proceeding in a similar manner as done for the derivation of the Michaelis-Menten equation, i.e. by assuming that the complex ES is at steady state. Note that since now the process is totally reversible, it is possible to define a Michaelis-Menten constant for product formation, $K_{\text{M}}^{\text{S}} = (k_{-1} + k_2)/k_1$, but also one for substrate formation, $K_{\text{M}}^{\text{P}} = (k_{-1} + k_2)/k_{-2}$. By defining the catalytic rates for the forward and backward reaction

as $k_S = k_1 k_2/(k_{-1} + k_2)$ and $k_P = k_{-1} k_{-2}/(k_{-1} + k_2)$ respectively, we obtain

$$v = \frac{k_S\, s\, e - k_P\, p\, e}{1 + \frac{s}{K_M^S} + \frac{p}{K_M^P}}\,. \tag{1.7}$$

Note that when $p = 0$, we get back the Michaelis-Menten equation: when no product is present in the system, the reverse reaction can be neglected. However when product accumulates, the reverse reaction becomes relevant and causes the reaction rate to decrease.

For the reversible reaction scheme considered here, it is possible to relate the kinetic parameters to the equilibrium constant. At the equilibrium $v = 0$, and

$$k_S s_{eq} e = k_P p_{eq} e\,, \tag{1.8}$$

where with s_{eq} and p_{eq} we refer to the equilibrium concentrations of substrate and product respectively. From Eq. (1.8), we can express the equilibrium constant K_{eq} as a function of the kinetic parameters,

$$K_{eq} = \frac{k_S}{k_P} = \frac{p_{eq}}{s_{eq}}\,. \tag{1.9}$$

This expression is known as the Haldane relationship. More complicated reaction schemes, e.g. multiple substrates reacting on the same enzyme, lead to a set of several more complex Haldane relationships; however there is at least one relationship between the kinetic parameters and the equilibrium constant as given by Eq. (1.9).

1.3.3. A model for cooperativity: the Hill's function

There are enzymes that do not follow the Michaelis-Menten kinetics Eq. (1.5), nor the reversible Michaelis-Menten one Eq. (1.7). These enzymes show a more complicated kinetics where the reaction rate v is a sigmoidal function of the substrate concentration s, see Fig. 1.5. Typically these enzymes have sharper increases in the reaction rate as the concentration s is varied as compared to the simple Michaelis-Menten kinetics. One could say that such enzymes are more sensitive to perturbation in substrate concentrations. The sharp increase in reaction rate upon small variations in s helps in defining two different kinetic states, a saturated or active state and an inactive or low activity state. This is used as a regulation mechanism for many enzymatic reactions and metabolic processes. For these reasons such enzymes are called 'regulatory enzymes' [25].

But how do regulatory enzymes work on a structural level? Many enzymes are

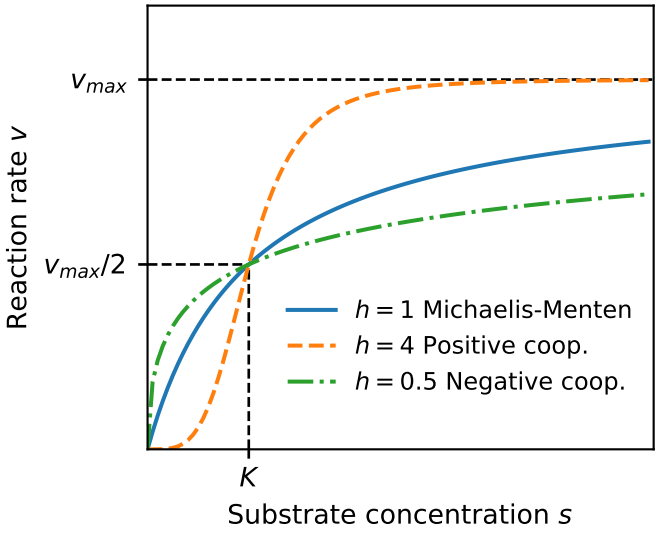


Figure 1.5.: Effects of cooperative binding on enzyme kinetics. We plot the reaction rate v as a function of the substrate concentration s for different Hill functions. We can see that for a Hill coefficient $h = 1$ the Hill function coincides with a Michaelis-Menten kinetics. In case of positive cooperative binding, $h = 4$, the reacion rate rapidly shifts from a low activity state where $v \approx 0$ to a state of high activity where v reaches its maximal value $v_{\max}$. For a negative cooperative effect, in a regime where $s > K$, higher values of s are needed to reach similar v values as compared to the Michaelis-Menten kinetics. The point at which $s = K$ corresponds to the half-maximum of the reaction rate with $v = v_{\max}/2$.

present in solution as oligomers, i.e. enzymes of the same type bind to form larger structures composed of few units, each called a monomer. An example of this is the mammalian hemoglobin, protein responsible for carrying oxygen in blood vessels, that presents four monomers with four binding sites. The hemoglobin is therefore able to bind up to four oxygen molecules. Once an oxygen molecule binds, it induces structural changes in the oligomeric structure that further favor the biding of an other substrate molecule [33]. This type of interactions, where the structure of an enzyme is affected by the binding of a molecule to a site different than the substrate binding site, is called 'allosteric interaction' and the enzyme is called 'allosteric enzyme' [34]. The binding kinetics shown by the hemoglobin also goes under the name of cooperative binding, where the binding of one molecule favors the binding of a molecule of the same type. Many allosteric enzymes are also cooperative, and vice versa. However the two terms are not interchangeable as there can be enzymes that show allosteric interactions but these do not necessarily lead to a cooperative effect, nor to a regulatory function [25].

Archibald Vivian Hill proposed a phenomenological model to describe the degree of cooperativity of regulatory enzymes, now known as Hill equation [35]. The Hill equation takes the form

$$v = \frac{v_{\max} s^h}{K^h + s^h}, \tag{1.10}$$

where K is the Hill constant and corresponds to the value of s at which $v = v_{\max}/2$

and h is called Hill coefficient, which is a measure for the degree of cooperativity of the enzyme. For $h = 1$ the enzyme follows a Michaelis-Menten kinetics, see Fig. 1.5. If $h > 1$ we have a positive cooperative effect. This is the case of hemoglobin that has a Hill coefficient in the range of $h = 1.7$-3.2. In Fig. 1.5 we see how for $h = 4$ we have a sigmoidal kinetics that steeply increases for values of s around K. The reaction rate goes rapidly from a state of low activity where $v \approx 0$ to a state of high activity where v approaches its maximal value $v_{\max}$. In general, it is also possible to have a negative cooperative effect for $h < 1$. In this case the allosteric interactions are such to disfavor the binding of further molecules. Note that the Hill equation could correspond to the reaction scheme $\mathrm{hS + E} \rightleftharpoons \mathrm{ES_h}$, however such a reaction scheme is usually not followed by enzymatic reactions. Moreover h is often a non-integer number, hence one should interpret the Hill equation simply as a phenomenological law, rather than a descriptive model for the reaction.

1.3.4. Enzyme inhibiton

The enzyme activity can often be reduced by the binding of specific molecules, which are known as inhibitors. In this section we briefly describe different types of product inhibition and we analyze how the enzyme kinetics is affected in the different cases. All types of inhibition follow a Michaelis-Menten kinetics with apparent catalytic activity $v_{\max}^{\mathrm{app}}$ and apparent Michaelis-Menten constant $K_{\mathrm{M}}^{\mathrm{app}}$ [25], as given by the following expression

$$v = \frac{v_{\max}^{\mathrm{app}} s}{K_{\mathrm{M}}^{\mathrm{app}} + s}\,. \tag{1.11}$$

With I we will refer to the inhibitor species and with i to its concentration, as similarly done so far for the other chemical species S, E and P.

- **Competitive inhibition:** A competitive inhibitor binds to the active site of the enzyme, forming the inhibitor-enzyme complex EI and it therefore impedes the binding of substrate molecules. Often in this case the inhibitor I is a molecule that closely resembles the substrate S. However, I is not able to undergo catalysis. This leads to an increase in apparent Michaelis-Menten constant $K_{\mathrm{M}}^{\mathrm{app}}$ but no change in $v_{\max}^{\mathrm{app}}$. In fact, as compared to the case with no inhibitor, more substrate is required to reach the same amount of substrate-enzyme complexes ES. The new Michaelis-Menten constant takes the apparent value of $K_{\mathrm{M}}^{\mathrm{app}} = K_{\mathrm{M}}(1 + i/K_{\mathrm{i}}^{\mathrm{c}})$, where $K_{\mathrm{i}}^{\mathrm{c}}$ is called competitive inhibition constant. An example of competitive inhibition can be seen in the reaction catalyzed by the enzyme succinate dehydrogenase, which oxidizes succinate to fumarate. A competitive inhibitor for this reaction is malonate, a very similar molecule as compared to succinate. The two differ

simply by the fact that malonate presents just a single methylene group instead of two.

- **Pure non-competitive:** A non-competitive inhibitor binds to a different site than the substrate binding site. Hence the binding-unbinding of substrate is unaffected and we have no change in K_M. However, through allosteric interactions, the binding of I reduces the catalytic activity to $v_{max}^{app} = v_{max}/(1 + i/K_i^u)$, where K_i^u is known as uncompetitive inhibition constant. Typical non-competitive inhibitors are heavy-metal ions and protons.

- **Uncompetitive inhibition:** An uncompetitive inhibitor can only bind to the ES complex but not to the free enzyme. With its binding, the inhibitor inactivates the ES complex. This mechanism affects both the apparent catalytic activity and the apparent Michaelis-Menten constant. We have that $v_{max}^{app} = v_{max}/(1 + i/K_i^u)$ and $K_M^{app} = K_M/(1 + i/K_i^u)$. An example of uncompetitive inhibitor can be found in the lithium used to treat manic depression, which inhibits the myo-inositol monophosphatase, an enzyme that plays a key role in cell signaling [36].

- **Mixed inhibition:** A mixed inhibitor binds to a site on the enzyme that is not the substrate binding site. However, allosteric interactions caused by the inhibitor binding affect both the binding-unbinding of substrate and the catalytic activity of the enzyme. This inhibition mechanism results in a mix of the competitive and uncompetitive inhibitions and we have that $v_{max}^{app} = v_{max}/(1 + i/K_i^u)$ and $K_M^{app} = K_M(1 + i/K_i^c)/(1 + i/K_i^u)$. Note that it depends on both inhibition constants K_i^c and K_i^u.

- **Irreversible inhibition:** An irreversible inhibitor permanently binds and inactivates the enzyme, usually by forming a covalent bond with the enzyme. An important example of irreversible inhibitor is the penicillin, the first discovered bacterial antibiotic, which acts as an inhibitor for an enzyme needed to keep the balance between cell wall production and degradation in bacteria. Hence the use of penicillin causes growing bacteria to rapidly die [37].

As we have already mentioned in Sec. 1.3.2, enzymes can also be inhibited by their product [25]. By looking at the reaction rate of a reversible Michaelis-Menten kinetics, Eq. (1.7), we can see how by increasing the amount of product in the system the reaction rate decreases. This decrease is significant only if the reaction is reversible. However, product inhibition is observed even in cases for which the reaction is basically irreversible. In these cases, the product can bind to the enzyme without being further processed back to substrate. In this way the product sequesters the enzyme from the substrate, behaving like a competitive inhibitor.

2. Spatial organization and dynamics of enzymatic reactions

Enzymes speed up most of the biochemical reactions that take place in the cells of living organisms. Often the product of an enzymatic reaction is further processed by a different enzyme and, in turn, the product of this second reaction has to be further processed to perform its biological function. This leads to chains of enzymatic reactions, which are known as 'metabolic pathways'. Different metabolic pathways can be topologically linked to one another in case enzymes from the different pathways target the same substrate. This leads to an intricate network of enzymatic reactions, which is known as 'metabolism', responsible for the conversion of food to energy and building blocks, and the elimination of metabolic waste. In each cell it is therefore important that a substrate, which could also be the product of an enzymatic reaction, finds the corresponding enzyme in an efficient way. The encounter rate between the enzyme and the substrate depends on their motion and on their spatial arrangement. In this chapter, we first analyze the role that a purely diffusive motion of the enzyme and the substrate has in defining the kinetic rates of a single-step enzymatic reaction. Then we discuss several examples of natural and synthetic strategies for the spatial organization of metabolic pathways that facilitate the encounter rate. We also discuss recent experimental findings showing that the enzyme motion can be coupled to the substrate concentration profile, leading to effective equations describing the enzyme dynamics that are more complex than a simple diffusion equation.

2.1. The role of diffusion on enzyme kinetics

The rate of an enzymatic reaction depends on the rate of encounter of the free enzyme E with substrate molecules S. The higher is the rate of encounter, the higher are the chances of forming the enzyme-substrate complex ES. The Michaelis-Menten kinetics presented in Sec. 1.3.1 does not present a direct dependence on such encounter rate. However, the binding rate k_1 and the unbinding rate k_{-1} depend on it. In the following we derive the dependencies of k_1 and k_{-1} on the encounter rate, provided that the substrate and the enzyme perform Brownian motion [38]. The Brownian motion is the random motion of particles suspended in a fluid, driven by the collisions with the fluid

molecules. Such a motion is described by the diffusion equation,

$$\frac{\partial s(\mathbf{r},t)}{\partial t} = D\nabla^2 s(\mathbf{r},t)\,, \tag{2.1}$$

where $s(\mathbf{r},t)$ is the concentration of the suspended molecule at position $\mathbf{r}$, at time t and D is the diffusion constant of the suspended species. The kinetics of an association process between diffusive molecules was first modeled by M. v. Smoluchowski [39] and then extened by Collins and Kimball [40]. These models consider the association process to be irreversible. However, to see how diffusion affects the kinetic rates of the Michaelis-Menten equation, we need to consider a reversible association-dissociation process as considered by M. Schurr [41] or D. Shoup and A. Szabo [42]. Here we follow the derivation carried out by M. Schurr [41].

We start by considering the diffusion equation Eq. (2.1) to describe the dynamics of substrate molecules around a free enzyme molecule, which we assume to be at the origin of our reference frame. This implies that the substrate has an effective diffusion coefficient D as given by the sum of the enzyme and substrate diffusion coefficients $D = D_\mathrm{e} + D_\mathrm{s}$. The substrate and the enzyme can form a complex when they are in direct contact, i.e. at a distance $R = R_\mathrm{e} + R_\mathrm{s}$. The net reaction flux J of complex formation is given by the rate of association between E and S into ES minus the rate of dissociation from ES into E and S. Moreover, the net reaction flux J is also assumed to be equal to the diffusive flux of substrate molecules entering in contact with the enzymes, hence

$$J = 4\pi r^2 D e_0 \left.\frac{\partial s(r)}{\partial r}\right|_{r=R} = \tilde{k}_1 e_0 s(r) - \tilde{k}_{-1} c_0\,, \tag{2.2}$$

where e_0 is the initial concentration of free enzymes, c_0 the one of enzyme-substrate complex and $\tilde{k}_1$, $\tilde{k}_{-1}$ are the so-called intrinsic rate constants of the forward and reverse step respectively of the association-dissociation reaction upon contact of substrates with enzymes. Eq. (2.2) provides a boundary condition to the second order diffusion equation for the substrate. A second boundary condition is given by the assumption that far away from the enzyme the concentration of substrate relaxes to the concentration s_0, i.e. $s(r \to \infty) = s_0$. By solving for the steady state of the diffusion equation Eq. (2.1) with the above boundary conditions, we find that the net reaction flux of complex formation is

$$J = \underbrace{\frac{k_\mathrm{D}\tilde{k}_1}{\tilde{k}_1 + k_\mathrm{D}}}_{k_1} e_0 s - \underbrace{\frac{k_\mathrm{D}\tilde{k}_{-1}}{\tilde{k}_1 + k_\mathrm{D}}}_{k_{-1}} c_0\,, \tag{2.3}$$

where k_D is the Smoluchowski rate constant of diffusion-limited reactions $k_\mathrm{D} = 4\pi R D = 4\pi(R_\mathrm{s} + R_\mathrm{e})(D_\mathrm{s} + D_\mathrm{e})$ and it corresponds to the rate at which substrate

and enzyme molecules of a certain size diffuse towards each other. The rates k_1 and k_{-1} correspond to the effective binding and unbinding rates, as given in the scheme of a Michaelis-Menten enzymatic reaction Eq. (1.1). By substituting the above values of k_1 and k_{-1} into the Michaelis-Menten equation (Eq. (1.5)), we obtain

$$v = k_{\text{cat}} \frac{e\, s}{\tilde{K}_{\text{M}} + k_{\text{cat}}/k_{\text{D}} + s}\,, \tag{2.4}$$

where e and s are the enzyme and substrate concentrations respectively and $\tilde{K}_{\text{M}} = (\tilde{k}_{-1} + k_{\text{cat}})/\tilde{k}_1$ is the intrinsic Michaelis-Menten constant. The catalytic efficiency, defined as $k_{\text{cat}}/K_{\text{M}}$ (see Sec. 1.3.1), fulfills the equality $(k_{\text{cat}}/K_{\text{M}})^{-1} = (k_{\text{cat}}/\tilde{K}_{\text{M}})^{-1} + (k_{\text{D}})^{-1}$.

In the so-called 'reaction-limited' regime where $k_{\text{D}} \gg (k_{\text{cat}}/\tilde{K}_{\text{M}})$, the limiting step of the enzymatic reaction is the binding of the substrate to the enzyme. The diffusion is in fact very fast and the transport of molecules is not a limiting factor. In this scenario, the fast diffusion smooths out any concentration gradients that might form, i.e. we are in a 'well-mixed' system. The assumption to be in a well-mixed system is implicit in the Michaelis-Menten description. For this reason, in the reaction-limited regime the kinetic rates of the Michaelis-Menten model equal the intrinsic rates, $k_1 = \tilde{k}_1$ and $k_{-1} = \tilde{k}_{-1}$.

In the regime where $k_{\text{D}} \ll k_{\text{cat}}/\tilde{K}_{\text{M}}$, known as 'diffusion-limited' regime, the reaction is limited by the rate of diffusive transport of substrates to the enzyme. We are not anymore in a well-mixed system and concentration gradients emerge due to the fast reaction around the enzymes. In the diffusion-limited regime the forward rate of the Michaelis-Menten equation is solely determined by the rate of diffusive transport $k_1 = k_{\text{D}}$ and the reverse rate $k_{-1} = k_{\text{D}}\tilde{k}_{-1}/\tilde{k}_1$. The catalytic efficiency $k_{\text{cat}}/K_{\text{M}}$, in the extreme case of $k_{\text{cat}}/\tilde{K}_{\text{M}} \to \infty$, becomes equal to k_{D}, which sets an upper limit for the enzymatic efficiency. This limit corresponds to the case of very fast enzymes immediately converting any encountered substrate molecule, provided that the enzyme and the substrate motion is described by the diffusion equation Eq. (2.1). By considering $D = 100\mu\text{m}^2\text{s}^{-1}$ and $R = R_{\text{e}} + R_{\text{s}} = 1\text{nm}$, we obtain the theoretical upper limit of the catalytic efficiency to be $k_{\text{cat}}/K_{\text{M}} \approx 10^9\text{M}^{-1}\text{s}^{-1}$. As already anticipated in Sec. 1.3.1, catalytic efficiencies are in the range of 10^2-$10^9\text{s}^{-1}\text{M}^{-1}$, with a median at $10^5\text{s}^{-1}\text{M}^{-1}$ [30]. This means that, although most of the enzymes operate far from the most efficient regime, there are enzymes like fumarase and carbonic anhydrase that have enzymatic efficiencies in the same order of magnitude as the theoretical upper limit of $10^9\text{M}^{-1}\text{s}^{-1}$.

The model described here neglects some of the details of enzymatic reaction. For instance, it assumes spherical symmetry and a uniform reaction on the enzyme surface, cf. Eq. (2.2), whereas in real enzymes the reaction takes place only at the active site.

Different extensions have been made to include the effects of rotational diffusion together with the presence of an active site on the surface of the enzyme [43], or to consider interaction potentials between the enzyme and the substrate [44]. Such extensions are important to understand how the efficiency of some enzymatic reactions can reach values that are up to one order of magnitude higher than the limit set by diffusion. For example, electrostatic interactions cause the association rate between the *lac*-repressor and the *lac*-operon to be $7 \cdot 10^9 \mathrm{M}^{-1}\mathrm{s}^{-1}$ [45]. The *lac*-repressor is a DNA binding protein that is positively charged and binds to the phosphate groups (negatively charged) of the *lac*-operon, which is a portion of DNA responsible for the expression of proteins needed for the digestion and transport of lactose in *E.coli*. The attractive electrostatic interaction generates a drift in the motion of the *lac*-repressor towards its binding site, hence favoring their encounter rate. Other examples are given again by the enzymes fumarase and carbonic anhydrase that, in case electrostatic interactions are not screened, can show catalytic efficiencies of up to $3 \cdot 10^{10}\mathrm{M}^{-1}\mathrm{s}^{-1}$ and $10^{10}\mathrm{M}^{-1}\mathrm{s}^{-1}$, respectively. In the case of fumarase both electrostatic interactions and the presence of an active site justify the incredibly high rate [46]. In the case of carbonic anhydrase the substrate is assumed to first perform a 3-dimensional diffusion motion to reach the enzyme, followed by a faster 2-dimensional diffusion on the surface of the enzyme to reach the active site [47].

2.2. Metabolic channeling

Living organisms have evolved different strategies to make enzymatic reactions more efficient. Some of them rely on an efficient transport of substrate molecules to the corresponding enzymes. These strategies go under the name of 'metabolic channeling'. Historically, metabolic channeling was initially referring to the intermediate products of metabolic pathways that are directly channeled from an enzyme of the pathway to the next via structural features [48, 49], e.g. tunnels, swinging arms or electrostatic guidance. However, other strategies that guarantee an efficient transport of substrates and at the same time prevent its diffusion into the cytoplasm are nowadays also referred to as metabolic channeling [50, 51]. These strategies include the colocalization of multiple consecutive enzymes into large clusters known as 'metabolons' [52], or simply having two consecutive enzymes in close proximity without being directly linked. In this section, we consider the advantages of metabolic channeling and we present some well-known examples.

2.2.1. Advantages of metabolic channeling

Different mechanisms of metabolic channeling can lead to the same physiological advantages [53]. Here we present a list of some of these advantages.

- **Reduction of transient time:** If a substrate is directly channeled to the active site of the corresponding enzyme, or its diffusive motion is constrained, the overall reaction velocity is increased. This implies that the catalytic efficiency of an enzyme can be higher than the expected limit k_D, as already discussed in Sec. 2.1.
- **Processing of unstable or volatile substrates:** Substrate molecules can leak through the cell membrane [54, 55], or can be destabilized by entering in contact with the solvent [56]. Metabolic channeling can protect the structure of such molecules and prevent the loss of substrate.
- **Protection from toxic substrates:** Within a metabolic pathway, some of the intermediate products could be toxic for the cells if they were freely diffusing in the cytoplasm. Hence, the compartmentalization of entire pathways [57] or the direct transport of the intermediates to the enzyme [58] can protect the cells from intoxication.
- **Avoiding competing side reactions:** Some substrates can be targeted by multiple enzymes from different metabolic pathways. However, cells might want to direct the metabolism in a preferred direction. By using metabolic channeling, the substrate can be sequestered, i.e. isolated, from the unwanted enzymes [59].
- **Making otherwise unfavorable reactions possible:** Some reactions would not spontaneously proceed if the enzyme and the substrate were well-mixed in the cytoplasm. However via metabolic channeling the intermediate product can be immediately transported to the active site of the next enzyme. This keeps the concentration of the intermediate low at the active site of the first enzyme, hence impeding the reverse reaction from occurring. In this way the reaction can proceed, although this would be thermodynamically unfavored in a well-mixed system [60].
- **Enhancement of product flux:** Metabolic channeling enhances the formation of products for fixed amounts of enzymes and substrates as compared to a well-mixed scenario. In fact, a channeled substrate is protected from the different loss mechanisms it can undergo (e.g. side reactions, diffusive loss via the membrane, spontaneous decay). Hence, a high product flux can be generated by keeping the amount of substrate low. The enhancement of product flux strongly depends on

the channeling strategy used. For example, it is under debate whether simply having two consecutive enzymes in close proximity without being directly linked (proximity channeling) can provide a measurable enhancement of product flux or not [61, 62, 63, 64]. Proximity channeling corresponds to the case where an intermediate simply diffuses from the active site of an enzyme to the active site of the next enzyme. According to a purely diffusive model such channeling strategy does not guarantee measurable enhancements in product flux unless the two enzymes are at distances which are comparable to their sizes, i.e. < 10nm [65]. However this seem to depend on whether the enzymes are reaction or diffusion-limited. For diffusion-limited enzymes, the advantages of proximity channeling are more evident [66].

In the following we present more specific examples of natural and synthetic metabolic channeling strategies and we will see how such strategies provide some of the above listed advantages.

2.2.2. Metabolic channeling via structural features

Enzymes catalyzing the sequential reactions of a metabolic pathway are sometimes found to assemble into multi-enzyme complexes [67]. Moreover, these complexes evolved structural features that favor the channeling of the intermediates. In Fig. 2.1 we present an overview of such features.

In Fig. 2.1a we show the structure of the tryptophan synthase complex. This is a bi-enzyme complex catalyzing the last two steps of the synthesis of L-tryptophan, which is an amino acid and an important building block for protein synthesis. We can see how an intramolecular tunnel of 25Å directly connects the two active sites of the complex [68]. At the first active site indole-3-glycerol phosphate is converted into indole and glyceraldehyde-3-phosphate. The intermediate indole is then transferred through the molecular tunnel to the second active site where it reacts with L-serine to form L-tryptophan and water. The intramolecular tunnel favors the rapid transport of indole to the second active site [70] and plays a crucial role in isolating the highly volatile indole molecules from the cytoplasm [54]. Another example of channeling via an intramolecular tunnel is found in the carbamoyl phosphate synthetase (CPS) complex. The CPS complex is responsible for catalyzing the synthesis of carbamoyl phosphate from bicarbonate, glutamine, and two molecules of adenosine triphosphate (ATP) [71]. Carbamoyl phosphate is an intermediate metabolite of the urea cycle, which is responsible for nitrogen disposal. The CPS complex has a 96Å long tunnel connecting three active sites [72]. The tunnel again provides a good sequestration mechanism for the unstable intermediates produced.

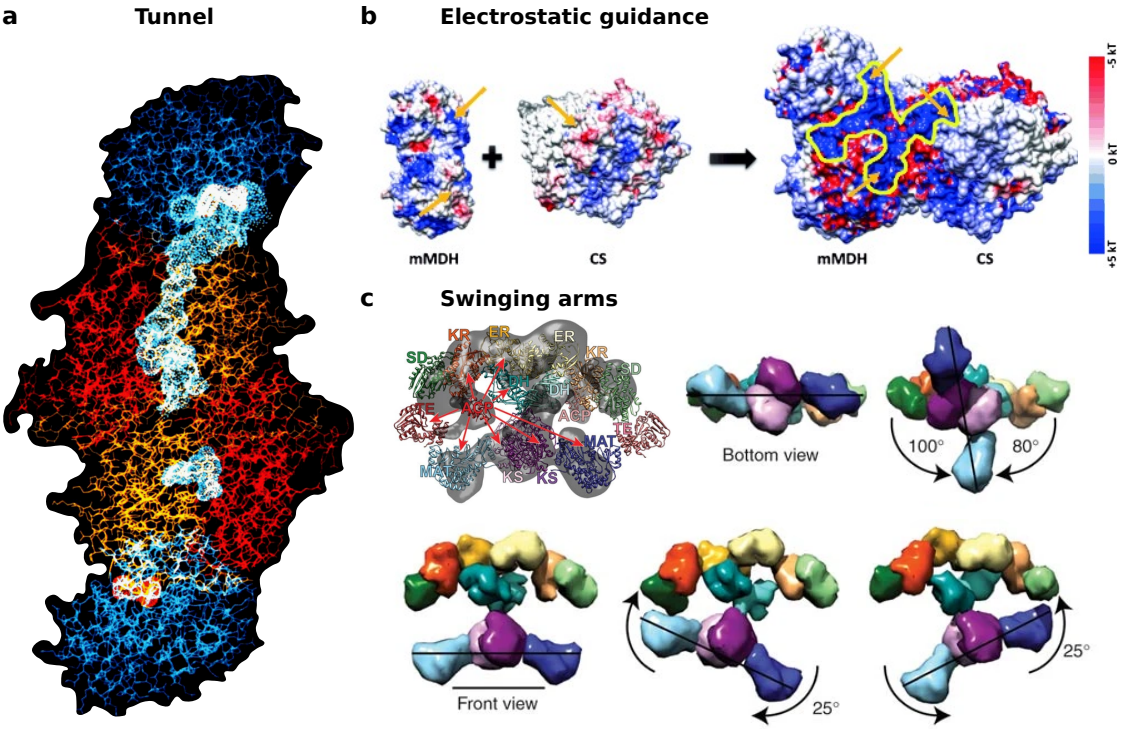


Figure 2.1.: Substrate channeling via structural features. In this figure we illustrate different channeling mechanisms that make use of structural features. **a** Tunnel. The structure of the tryptophan synthase complex is shown. A tunnel of 25Å, highlighted in light blue, connects the two active sites present in the complex. The figure is adapted with permission from [68]. **b** Electrostatic guidance. The enzymes malate dehydrogenase (MDH) and citric synthase (CS) form the MDH-CS complex. the two enzymes catalyze two consecutive steps of the citric acid cycle (TCA cycle). A negatively charged intermediate, oxaloacetate, is produced at the active site of MDH (location indicated by the orange arrow on the left). Once produced the intermediate's diffusion is confined onto the blue positively charged surface patch of the MDH-CS complex. The intermediates diffuse on this surface patch until it reaches the second active site (CS active site). The active site of CS is also present on the positively charged surface patch as indicated by the right orange arrow. The intermediate is therefore channeled with the aid of an electrostatic guidance. The figure is adapted with permission from [60]. **c** Swinging arms. The structure of the fatty acid synthase (FAS) is shown. FAS is responsible for the synthesis of fatty acids. The growing chain of a fatty acid covalently binds to the acyl carrier protein (ACP) of the complex. Via a series of structural changes the ACP enters in contact with the different active sites of the complex as indicated by the red arrows. We show a bottom and front view of such structural changes. The figure is adapted with permission from [69].

In addition to molecular tunnels, channeling can be favored by an electrostatic guidance mechanism (see Fig. 2.1b). In this case, the intermediate is guided to the next

active site by electrostatic forces that confine its diffusion to a patch of the complex surface presenting both active sites. The intermediate and the surface patch have opposite electrostatic charges. An example of this can be seen in the citric acid cycle (CAC or TCA cycle), where the enzyme malate dehydrogenase (MDH) and the enzyme citric synthase (CS) form a complex [60], see Fig. 2.1b. The diffusion of the negatively charged intermediate, oxaloacetate, is bound to a positively charged area of the complex surface (blue) connecting the two active sites [73]. The formation of the MDH-CS complex is remarkable because it pushes the reaction in the opposite direction as compared to a well-mixed system. In fact, the equilibrium constant for the reaction catalyzed by MDH is unfavourable in the forward direction of the TCA cycle ($K_{\text{eq}} = (2.86 \pm 0.12) \cdot 10^{-5}$ [74]). However the intermediate oxaloacetate is efficiently transported to the CS active site and its concentration is therefore kept low at the MDH active site. This guidance mechanism, thus, favors the direction of the reaction to be the one needed for the TCA cycle [50]. Another example of electrostatic guidance is given by the bifunctional enzyme dihydrofolate reductase- thymidylate synthase (DHFR-TS), catalyzing two steps of the thymidylate pathway, which supplies 2-deoxythymidylate for DNA synthesis. Similarly to the MDH-CS case, a negatively charged intermediate is bound to a positively charged patch on the surface of the complex that bridges the two active sites [75].

As shown in Fig. 2.1c, the use of swinging arms can also be used for channeling substrates [76]. An example of this is given by the fatty acid synthase (FAS), which is one of the largest known multi-enzyme complex and is responsible for the synthesis of fatty acids [69]. The growing chain of fatty acids is covalently bound to an acyl carrier protein (ACP) that moves the intermediate between the active sites of ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) in an assembly line fashion. After the chain has reached the length of 16 units the fatty acid is released by the action of thioesterase (TE). In Fig. 2.1c we can see how the carrier protein ACP enters in close contact with the different active sites. Another well characterized example of channeling via swinging arms is given by the pyruvate dehydrogenase complex (PDC) that converts pyruvate to acetyl coenzyme A, linking glycolysis to the TCA cycle. The PDC is formed by three different enzymes (called E1, E2 and E3 respectively) that assemble in a remarkably symmetric complex [77, 78]. Here E2 enzymes form a core to which E1 and E3 binds. After the first reaction is catalyzed by E1, the intermediate is transferred to the lipoamide swinging arm of E2 which brings the intermediate in contact with the active site of E2 where a second reaction takes place. E3 then has the role of reactivating the swinging arm for the binding of another intermediate molecule [50].

2.2.3. Metabolon formation

Metabolons are clusters of consecutive metabolic enzymes held together by protein-protein interactions (i.e. non-covalent forces such as electrostatic, hydrophobic or van der Waals interactions) and interactions with structural elements of cells, e.g. the cytoskeleton or integral membrane proteins [52]. Although the enzymes composing a metabolon are not necessarily structurally linked, the fact that they are present in large copy numbers and held in close proximity is enough to constrain the diffusion and favor the channeling of intermediates. Metabolons are often dynamically assembled and disassembled upon changes in the concentration of one of the metabolites that they process. Typically, when cells are starved of a certain nutrient, metabolons are assembled to process the low amount of nutrient in a more efficient way. Hence, they often play a key upregulatory function in metabolic pathways [79].

A well characterized example is the case of the purinosome, a metabolon assembled in human tumor cells (HeLa cells) to upregulate the de novo purine biosynthetic pathway [80]. Six enzymes catalyze the 10-steps reactions of the pathway, converting the phospho-ribosyl pyrophosphate (PRPP) into inosine monophos-phate (IMP), which is the final product of the pathway. IMP is a precursor of adenosine and guanosine nucleotides, which are building blocks of DNA and RNA. The six enzymes assemble into the purinosome via protein-protein interactions [81] with the aid of heat shock proteins [82]. G protein–coupled receptors (GPCRs), responsible for the transmission of signals to the nucleus, might trigger the formation of the purinosomes [83]. Once formed, the purinosomes can synthesize up to 3-fold higher amounts of IMP as compared to cells not having the metabolons [84].

Recently, also enzymes catalyzing the glycolysis have been shown to assemble into metabolons in different organisms. As already mentioned in Sec. 1.2.2, glycolisis is an important pathway for extracting energy in form of ATP and NADH from the sugar glucose. In yeast cells, glycolytic enzymes form aggregates to upregulate the conversion of glucose under hypoxic stress [85, 86] and lack of glucose [87]. In *C. Elegans* neurons, under energetic stress, glycolytic enzymes form metabolons in proximity of the synapses to maintain high enough local levels of ATP. This is necessary for a normal synaptic functioning [88]. Moreover, glycolytic enzymes are found to assemble on the membrane of human red blood cells [89, 90]. Despite glycolytic enzymes form metabolons in many different organisms, it is not yet clear whether there are common principles behind the assembly process. An interesting observation has been recently presented by Zhao et al. [91], who showed that glycolytic enzymes drift upstream gradients of their own substrate. This can favor the contact among the sequential glycolytic enzymes. An enzyme would drift towards high concentration of its substrate that is produced by a previous enzyme of the pathway. By following the gradient, the enzyme would then

swim towards the previous enzyme of the pathway. In Ch. 4, we will compare the model for the enzyme motion suggested by Zhao et al. with other models and we will see how spatial patterns can spontaneously form if the enzyme motion is driven by nonspecific interactions with the substrate. Understanding how enzymes respond to the presence of their own substrate can be key in better understanding the formation of metabolons.

We conclude this subsection by noting that there exist many other cases of metabolon formation. For example for the regulation of the TCA cycle [92], the formation of defense and other types of compounds in plants [93, 94]; or the purine biosynthetic pathway in yeast [87].

2.3. Organelles and microcompartments

The word organelles usually refers to compartments enclosed in lipid membranes that are typical of eukaryotic cells. Eukaryotic cells are in fact well known for the high degree of spatial organization of their enzymatic reactions. Confining reactions and entire metabolic pathways into organelles can have the same benefits as the ones obtained by metabolic channeling, cf. Sec. 2.2.1. Cells can concentrate enzymes and substrates in a smaller volume, hence reducing the transient times of the reactions and enhancing product fluxes, while keeping the amount of molecules relatively low as compared to a well-mixed scenario. An example of this can be seen in the mitochondria, which are subcellular compartments highly specialized in the production of energy in form of ATP via the TCA cycle [95]. Moreover, compartmentalization allows the isolation of substrates from competing reactions, or of toxic substrates. The enzymes contained in the lysosome are a typical example of such a sequestration mechanism. The lysosome is an organelle responsible for the digestion of proteins, lipids and nucleic acids and if the enzymes contained in the lysosome were free to diffuse in the cytoplasm, they would impair many functional activities [96]. Another example is the peroxisome, an organelle that encapsulates oxidative enzymes involved in the production and consumption of the toxic intermediate hydrogen peroxide [97]. The peroxisome also provides a different microenvironment as compared to the cytoplasm, presenting an unphysiological alkaline pH to favor the activity of the enzymes oxidase and catalase by speeding up their reactions [98].

Organelles encapsulated by lipid membranes are a characteristic of eukaryotic cells. However, also prokaryotes have been recently shown to present microcompartments enclosed by proteinaceous shells. The best studied example of such microcompartments is the carboxysome [99]. The carboxysome is used in cyanobacteria for the carbon concentrating mechanism (CCM), which is crucial for efficient carbon fixation in the

Calvin cycle of photosynthesis, i.e. the conversion of carbon dioxide (CO_2) into glucose. On the proteinaceous external shell of the carboxysome there are positively charged pores that favor the access of the anion bicarbonate (HCO_3^-). Inside the microcompartment, the enzyme carbonic anhydrase converts the bicarbonate into the uncharged CO_2. The CO_2 is then trapped inside the compartment because of its neutral charge. This is very important for the carbon fixation process, because the cellular membranes are very permeable to small uncharged molecules such as CO_2 and in this way loss of CO_2 is prevented [100, 101]. In the carboxysome, CO_2 is then fixed by the relatively inefficient enzyme ribulose-1,5-biphosphate carboxylase oxygenase (RuBisCO). Given the low efficiency of RuBisCO, the high concentration of CO_2 in the compartment helps the fixation process. Moreover, since RuBisCO also catalyzes the futile fixation of oxygen, the sequestration of RuBisCO inside the carboxysome directs the fixation process in the desired direction [99]. Overall, it has been recently shown that the compartmentalization of the carbon fixation process in the carboxysome is optimal given the properties of the protein shell of the compartment [102]. Other examples of microcompartments in prokaryotes can be found in so-called Pdu and Eut microcompartments in enteric and gut bacteria respectively. The Pdu compartment are used for the effective digestion of 1,2-propanediol, which is a product of the degradation process of plant cell wall sugars and it is used as a carbon and energy source by the enteric bacteria. The Eut microcompartments are similarly used in gut bacteria to digest phosphatidyl ethanolamine, which serves as a carbon, energy and nitrogen source for the bacteria. As already seen for the carboxysome, the function of the Pdu and Eut microcompartments is to sequester volatile and toxic intermediates that are produced by the enzymatic reactions therewithin [57, 103].

2.4. Membrane bound enzymes

Often, substrates are imported into the cytoplasm through the external membrane or the membrane of internal organelles. The enzymes that catalyze reactions involving these substrates can be found to be membrane bound, either directly or via transmembrane proteins, i.e. proteins that span the entire membrane. As a consequence, the substrates are directly intercepted by the enzymes and their diffusion into the cytoplasm is hindered. This intercepting mechanism can present the same physiological advantages of metabolic channeling presented in Sec. 2.2.1. An additional advantage can be provided if enzymes colocalize with transmembrane transporters importing their substrates. In this case, the substrate is immediately depleted from the enzyme and this guarantees a steady gradient of substrate concentration across the membrane. This steady gradient results in a reduction of the back-flux of substrates through the

transporters and in an increased import rate [104].

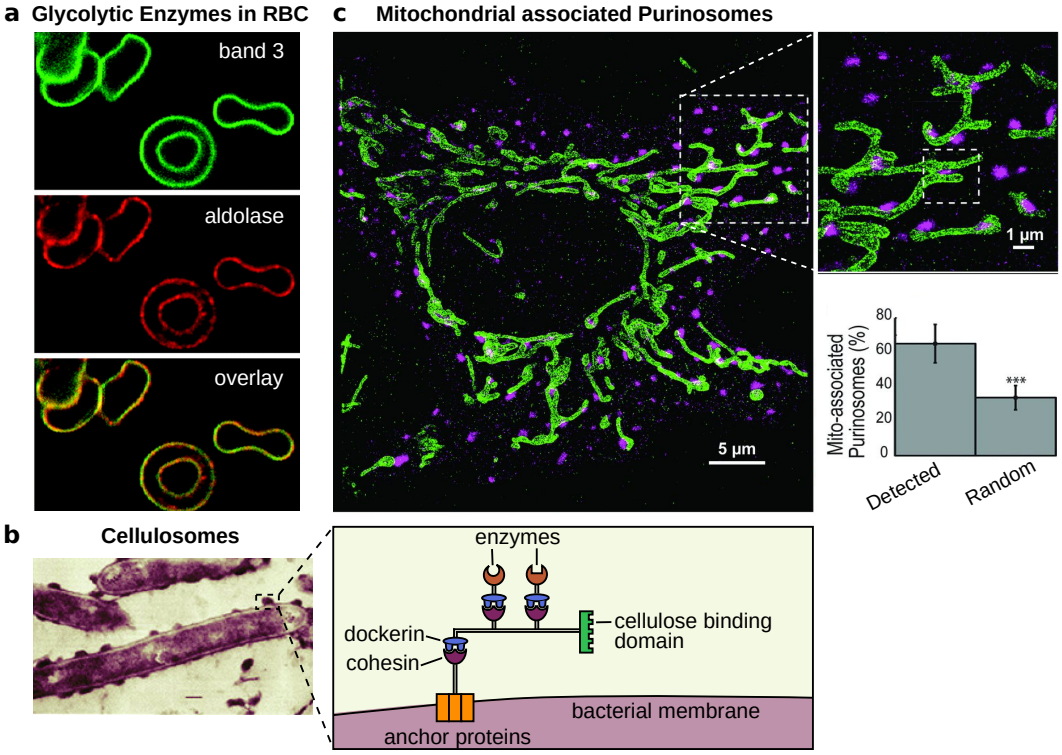


Figure 2.2.: Examples of spatial organization of enzymes on membranes. **a** In oxygenated red blood cells (RBC), the enzymes responsible for glycolysis bind to the band 3 transport protein on the membrane. As an example, here is shown the colocalization of band 3 (top, green) with the enzyme aldolase (middle, red), which catalyzes the 4th step of glycolysis. Having glycolytic enzymes at the membrane where glucose is imported, makes glycolysis more efficient. Figure adapted with permission from [89]. **b** Cellulosomes at the surface of the bacteria *Clostridium thermocellum*. Anchor proteins anchor the cellulosome to the membrane. The binding of two proteins, dockerin and cohesin, maintains the non catalytic scaffold (scaffoldin) bound to the anchor proteins and guarantee the attachment of enzymes onto the scaffoldin. Cellulose binding domains are present on the scaffold and favor the binding of the bacteria to the cellulose present on plant cell wall. The enzymes break the cellulose into smaller compounds that bacteria can more easily import. Figure adapted with permission from [105]. **c** Mitochondrial associated purinosomes. In magenta we see one of the core enzymes of the purinosome (FABS), responsible for catalyzing the 4th step of the de novo purine biosynthesis. A protein present on the outer membrane of the mitochondria has been tagged in green. The percentage of mitochondrial associated purinosome is much higher than the percentage expected for a random distribution of purinosomes. Having purinosomes bound to the mitochondrial membrane makes the synthesis of purines more efficient as some of the intermediates needed for the pathway are directly released at the mitochondrial membrane. Figure adapted with permission from [106].

An example of this can be seen in the organization of the glycolytic enzymes of red blood cells (RBC) [89]. In red blood cells, glycolysis is the only pathway responsible for the storage of energy in form of ATP molecules. The role of red blood cells is in fact to carry oxygen in blood vessels. The oxygen load cannot be used for the cellular respiration that, in other type of cells, takes place in the mitochondria and uses oxygen to form ATP molecules. When the red blood cell is oxygenated, glycolytic enzymes are observed to bind at the cellular membrane, see Fig. 2.2a. Here, glucose is imported and directly converted by the glycolytic enzymes into pyruvate and ATP molecules. This causes the depletion of glucose in the vicinity of the glucose transporters, hence impeding glucose backflux through the transporters themselves [104]. Moreover, the ATP molecules are subsequently used by ion pumps, which are membrane bound proteins that make use of the energy released by the ATP to pump ions in and out of the cells [107]. Another example of enzymes active at the external cell membrane is the multienzymatic complex 'cellulosome', used by several anaerobic bacteria to digest the cellulose contained in the cell wall of plants [108, 109]. In this case, the enzymes are secreted outside of the cell and a noncatalytic subunit called scaffoldin secures the anchoring of the complex to the cell wall and the incorporation of the various enzymatic subunits via the so-called cohesin-dockerin interaction (see Fig. 2.2b). The cohesin-dockerin interaction consists of pairs of complementary protein domains present on the scaffolds and on the different enzymes. Moreover, a cellulose binding domain (CBD) guarantees the binding to cellulose [109]. The different enzymes then breaks the cellulose into elemental subunits that can be more easily imported by the bacteria [105].

Enzymes can also be localized at the membranes of internal organelles. They can either favor the channeling of substrates inside the organelles or can intercept substrates that are being released from the organelles into the cytoplasm. An example of the first scenario is the binding of glycolytic enzymes to the mitochondrial membrane. The increased level of membrane bound glycolitic enzymes is associated with an increase in respiratory demand by cells of different species [110, 111]. The product of the glycolysis is the acetyl-CoA, which is the starting point of the TCA cycle taking place in the mitochondria. Hence, having glycolytic enzymes directly bound at the mitochondrial membrane favors the formation and channeling of the TCA precursor into the organelles. An example of the second scenario, where substrates are intercepted while being secreted into the cytoplasm, can be seen in the purinosome, see Fig 2.2c. As already mentioned in Sec. 2.2.3, the purinosome is responsible for boosting the de-novo purine synthesis in human tumor cells under purine depleted conditions. Purinosomes seem to form at nucleation sites that are located at the largest filaments of proteins that compose the cytoskeleton, i.e. the microtubules [112]. Once purinosomes are assembled on the microtubules, they seem to undergo a directed

motion towards the mitochondrial membrane [113]. This causes the purinosomes to be largely associated to the mitochondrial membrane [106]. This is thought to be beneficial for the cells for two reasons. First, the de-novo purine biosynthesis provides the nucleotide precursors necessary for the formation of mitochondrial ATP. Second, the ATP produced in the mitochondria, together with other metabolites like glycine, aspartate and formate, are used by the purinosome to carry out its reactions. This corresponds to a case where a substrate secreted by an organelle, the mitochondrion, is directly intercepted by enzymes located at the organelle's membrane, the purinosomes. In this way the turnover efficiency of the reactions carried out by the purinosome increases as it has been recently shown with the use of gas cluster ion beam secondary ion mass spectrometry (GCIB-SIMS) [114]. Other examples of active enzymes bound to the membrane of organelles can be found in many pathways of the secondary metabolism of plants, e.g. phenylpropanoid and flavonoid metabolism. The enzymes catalyzing reactions for these pathways form large complexes on the membrane of the endoplasmic reticulum (ER) [115, 116]. These complexes often form at cytochrome P450 enzymes, which are localized at the ER membrane and provide an anchor for the complexes [117, 118]. Their assembly is often reversible and enzymes switch from a soluble cytoplasmatic state to a membrane bound state [94, 93].

The main focus of the next chapter of the thesis, Ch. 3, will be on models of systems where a substrate is imported into the cytoplasm through the membrane of the cell or of internal organelles. We will consider the enzyme catalyzing the reaction to be both membrane bound or spatially arranged in the cytoplasm and we will study the enzyme arrangement that maximizes the reaction flux. Surprisingly, we will see that having all enzymes directly located at membranes is not always optimal. We will characterize the strategy behind the optimal profile and the transition in the optimal enzyme arrangement from a fully bound configuration into a more extended profile, where enzymes are also present in the cytoplasm.

2.5. Synthetic strategies

Inspired by natural strategies for the spatial arrangement of metabolic pathways, synthetic biologists developed several techniques to finely tune the spatial allocation of enzymes in cells as well as in completely synthetic systems like protein or DNA scaffolds [119].

To regulate the amount of enzymes bound to cell or organelles membranes, enzymes could be fused to transmembrane proteins, either directly [120] or via a protein scaffolds [121]. In such systems, transgenic genes are inserted into the cells to encode for modified version of the enzymes and/or transmembrane proteins. The genes encode

extra information for the synthesis of peptide linkers and/or the formation of scaffolds to which then the enzyme and the proteins can bind. Thomik et. al [121], for example, showed that by fusing an endogeneous (i.e. native) sugar transporter to the heterologous (i.e. not native) enzyme xylose isomerase in yeast cells, can lead to a substrate channeling mechanism, enhancing the conversion of the imported sugar xylose to ethanol, see Fig. 2.3a. The channeling mechanism drives the conversion of xylose into a precursor of ethanol, the xylulose. The reaction is carried out by the otherwise inefficient enzyme xylose isomerase. Moreover, the presence of xylose isomerase in proximity of the xylose transporter maintains a steep concentration gradient of xylose across the membrane, causing an increase in the xylose import rate. Another technique to enhance the presence of enzymes onto cell membranes is to exploit the orthogonal dockerin-cohesin pairs of the cellulosomes of different species, cf. Sec. 2.4. This allows for a rational design of mini-cellulosomes that can be expressed in bacteria [122] and in yeast cells [123, 124, 125] to favor the anchoring of several enzymes to the cellular membrane. In all of these cases, a reaction flux increased by 1.5 to 5-fold has been observed as compared to the free wild-type enzymes [122, 125]. The synthetic design of enzymatic assemblies onto intracellular lipid droplet membranes can also be achieved. In yeast cells such assemblies have been achieved by recruiting upstream enzymes of the ester biosynthetic pathway to the native location of the last enzymatic reaction that takes place onto the membrane of intracellular lipid droplets [126], see Fig. 2.3b. The recruitment has been obtained by expressing programmed protein scaffolds composed of oleosin (a plant based lipid droplet protein) and cohesin-dockerins pairs. The assembled enzymes show an increased reaction flux for the pathway of about 2-fold as compared to the case of unstructured enzymes.

Enzymes localization can also be regulated in the cytoplasm or any membrane enclosed space by folding programmed RNA assemblies into scaffolds. These predesigned RNA sequences fold into programmed structures and exhibit binding sites for functionalized enzymes presenting complementary RNA sequences to those of the binding sites [127]. By assembling a hydrogen producing pathway in bacteria onto discrete, one-dimensional and two-dimensional RNA scaffolds, the production of hydrogen has been shown to increase by 4, 11 and 48-fold respectively. Instead of using RNA scaffolds, protein scaffolds can also be used. It has been shown that an heterologous three enzymes pathway used to convert acetyl-CoA into mevalonate, can be optimized in bacteria to have a 77-fold increase in mevalonate production [128], see Fig. 2.4a. The first of the three enzymes (AtoB) is endogenous, however the other two (HMGS and HMGR) are expressed by genes taken from yeast. The enzymes are functionalized with peptide linkers that allow the binding of the enzymes to protein scaffolds in a similar manner as seen for the dockerin-cohesin interactions in the cellulosomes. The proteins forming the scaffold (GBD, SH3 and PDZ) are expressed from heterologous

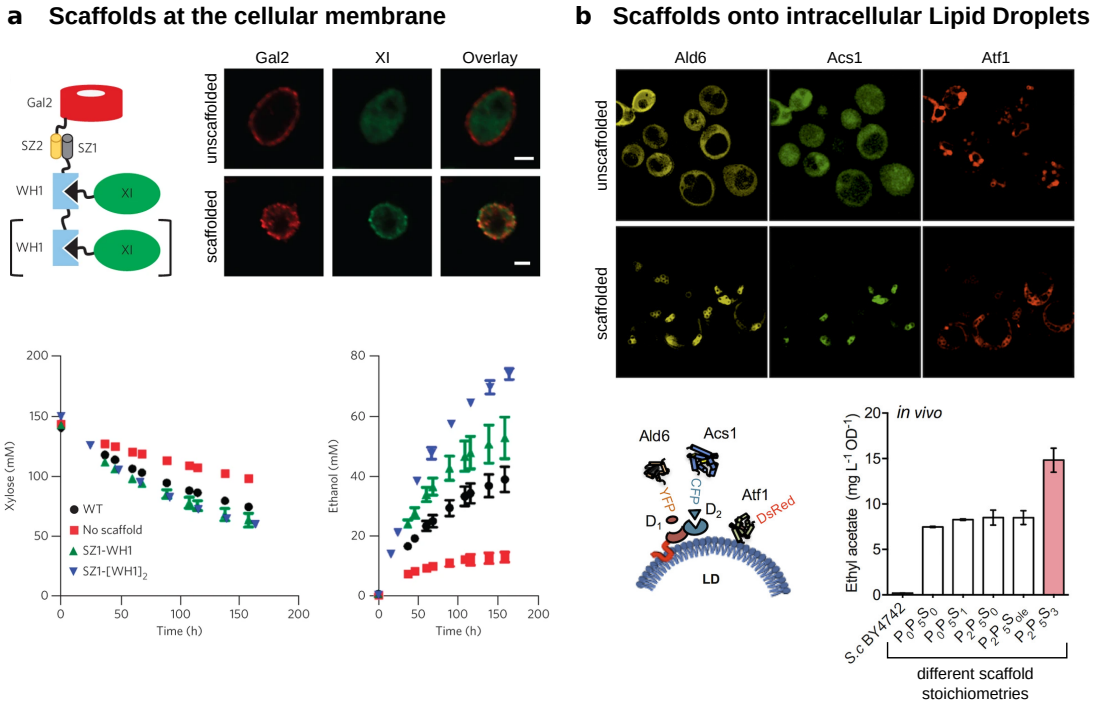


Figure 2.3.: Synthetic strategies to regulate the amount of membrane bound enzymes. **a** Scaffolds at the cellular membrane. Top-left: In genetically engineered yeast cells, the enzyme xylose isomerase (XI) is fused to the xylose importer Gal2 via a protein scaffold. In this way, the xylose is channeled directly to the xylose isomerase and it is efficiently converted into the intermediate xylulose. Xylulose is then subsequently converted to ethanol in the yeast cell. Top-right: We can see how Gal2 and XI are bound only in the presence of a scaffold. Bottom: When Gal2 and XI are bound the production of ethanol is increased. Figure adapted with permission from [121]. **b** Scaffolds onto intracellular Lipid Droplets. Example of sequential enzymes arranged onto scaffolds on the intracellular lipid droplets of synthetic yeast cells. Bottom-left: The scaffold is composed of a membrane bound protein oleisin (red) and cohesin-dockerin interactions mediate the binding of the enzymes. Top: We can see how the enzymes aldehyde dehydrogenase (Ald6), acetyl-coA synthetase (Acs1) and alcohol-O-acetyltransferase (Atf1) all localize on the membrane of lipid droplets in presence of a scaffold. These enzymes catalyze the conversion of acetaldehyde into ethyl acetate. Bottom-right: We can see how the production of ethyl acetate is enhanced for different scaffold structures. For details about the different structures we refer the reader to [126]. Figure adapted with permission from [126].

metazoan genes. The 77-fold increase in mevalonate production has been obtained by optimizing the stoichiometry of the assembled enzymes. Enzymes can also be directly fused via small peptide linkers. If the linkers do not impair the enzymatic activity, it can be shown that fluxes of metabolic pathways can be increased with such a simple technique. For example, in yeast cells an increase of 15-fold has been measured for

the production of the defensive compound resveratrol [129]. Enzyme fusions can also result in the formation of large enzymatic clusters used to regulate the direction of metabolic fluxes at metabolic branch points. If two enzymes compete for an intermediate produced by an upstream enzyme, the latter can be fused with either one of the competing enzymes. In this way the intermediate is channeled only to one of the competing enzymes and the reaction flux is shunted [59]. Enzymes can also be linked to DNA-binding proteins such as zinc fingers that very specifically target unique DNA sequences with nanomolar dissociation constant. Hence, by expressing DNA scaffolds *in vivo*, target sites for the zinc fingers can be designed to obtain a specific spatial arrangement of the enzymes. Such strategy has been succesfully applied in *E. coli* cells for three different pathways, all showing an enhanced product flux as compared to the case of unassembled enzymes [130].

DNA scaffolds can also be used *in vitro* to arrange enzymes with nanometer precision. Their application can be useful in synthetic devices such as microfluidic chips. In several studies, the throughput of the reaction catalyzed by the two consecutive enzymes glucose oxidase (GOx) and horseradish peroxidase (HRP) has been analyzed as the two enzymes were tethered onto DNA scaffolds at different distances. The GOx catalyzes the conversion of glucose and oxygen into a gluconic acid and hydrogen peroxide. The HRP then uses the hydrogen peroxide to mediate the oxidation of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid), $ABTS^{2-}$, to a colored product that can be easily detected via spectrophotometry. The two enzymes were located onto two and four hexagonal DNA scaffolds at a distance of 13nm and 33nm respectively. The product flux resulting from the enzymatic cascade onto the two hexagons showed an increase of 1.2-fold as compared to the flux produced by the enzymes on the four hexagonal scaffold. This was justified by considering the different distances that the intermediate hydrogen peroxide needs to cover via diffusion to reach the next enzyme [62]. A more systematic study of the inter-enzyme distance has been carried out by tethering GOx and HRP onto two-dimensional DNA origami tiles at varying distances in the range of 10-65nm [63], see Fig. 2.4b. DNA origami tiles are different as compared to the hexagonal scaffolds discussed before as they form from a long single strand of DNA instead of the binding of different oligomers [131]. The reaction flux is enhanced by 20-fold at 10nm and just by 3-folds at distances higher than 20nm. In other studies the enzyme pair has been enclosed into a DNA origami tunnel showing a further enhancement in the enzyme activity [132, 133]. Furthermore, the introduction of swinging arms has been shown to further increase the reaction flux for the GOx, HRP pair [63] and also for two consecutive dehydrogenase enzymes [134]. Moreover, the stoichiometry of the enzymes and of the swinging arms can be optimized, leading to a further increase in the reaction flux that, in the case of the two dehydrogenase enzymes, got up to almost 300-fold.

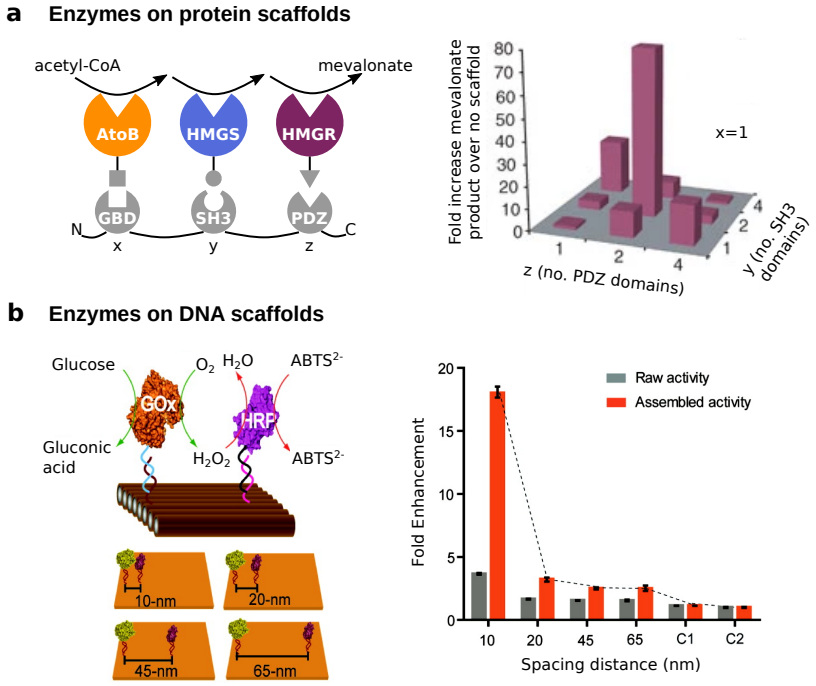


Figure 2.4.: Protein and DNA scaffolds for the arrangement of enzymes in bulk. **a** Enzymes on protein scaffolds. An heterologous three-enzymes pathway converting acetyl-CoA into mevalonate can be inserted into bacteria. The enzymes AtoB, HMGS and HMGR bind to the GBD, SH3 and PDZ scaffold proteins respectively (left). As the stoichiometry of the scaffold is varied, the mevalonate production rate changes. For scaffolds composed of 1 GBD domain, 2 PDZ domains and 2 SH3 domains it is observed a 77-fold increase in the mevalonate production as compared to the case with no scaffolds. Figure adapted with permission from [128]. **b** Enzymes on DNA scaffolds. The consecutive enzymes GOx and HRP are arranged with nanometer precision onto DNA origami tiles. On the left we can see the two-steps pathway catalyzed by the enzymes (top) and the four different distances considered (bottom). On the right the fold enhancement of the reaction compared to free enzyme in solution versus the spacing distance is shown. The raw activity (grey) is given by both assembled enzymes and enzymes present in solution. The assembled activity (orange) includes corrections accounting for the percentage of enzymes correctly assembled onto the DNA origami tiles. Only at a distance of 10nm we have a significant enhancement of $\approx$ 20-fold. At distances larger than 20nm the fold enhancement quickly drops to a 3-fold enhancement. Figure adapted with permission from [63].

All of the above techniques can be used to spatially allocate enzymes in cells and in synthetic devices to maximize the formation of metabolites. In the next chapter, Ch. 3, we will study a wide class of systems where a substrate is imported into the bulk of the system via the external boundary or the boundaries of internal compartments. Once

in the system, the substrate undergoes diffusive motion and reacts with an enzyme that is considered to be static and can be freely arranged, being also membrane bound, and that catalyze the substrate conversion into a desired product. We ask: What is the enzyme arrangement that maximizes the reaction flux? We will analyze in detail the optimal strategy leading to the optimal enzyme profiles and characterize a transition in the optimal configuration. Finally, we will derive a set of rules for the systematic construction of the optimal enzyme arrangement that, in principle, could be achieved by using the techniques presented in this section.

2.6. Enzyme motion

In Sec. 2.1, we have seen how the binding and unbinding rates of an enzymatic reaction are affected by a purely diffusive motion of the enzyme and the substrate. However, recent experiments suggest that the motion of enzymes can be influenced by the concentration profile of the corresponding substrate [135]. Although some of these experimental findings are still being debated, it seems that the motion of enzymes in presence of their substrate is not always simply diffusive. In the following, we present these experimental findings and the different effective equations proposed to describe the enzyme motion.

2.6.1. Enhanced diffusion

Recent experiments suggest that several enzymes exhibit a higher diffusion coefficient as the concentration of the corresponding substrate in solution increases [136, 137, 138, 139, 140, 141, 142, 91]. This phenomenon goes under the name of 'enhanced diffusion' of enzymes. The enhanced diffusion turns out to be directly proportional to the Michaelis-Menten reaction rate of the enzymatic reaction. Hence, the effective diffusion coefficient of the enzyme $D_\mathrm{e}(s)$ can be written in terms of the diffusion coefficient in absence of substrate D_e^0 as

$$D_\mathrm{e}(s) = D_\mathrm{e}^0 \left(1 + \alpha \frac{s}{K_\mathrm{M} + s}\right) , \tag{2.5}$$

where α corresponds to the maximal relative increase $(D_\mathrm{e}(s) - D_\mathrm{e}^0)/D_\mathrm{e}^0$, obtained for $s \gg K_\mathrm{M}$. The maximal measured relative increases are in the range of $\alpha \approx 0.15$-0.80, being usually ≈ 0.30. Most of these measurements were performed by using Fluorescence Correlation Spectroscopy (FCS). This technique is based on the measurement of the correlations of fluctuations of the fluorescence intensity produced by the motion of fluorescently labeled enzymes in a small volume. The resulting correlation function

is then fit to a model that assumes the diffusive motion of the labeled particles. Some of the above experimental results have been proven to be due to artifacts affecting the FCS measurements [143, 144]. For example the 0.80 relative increase in diffusion coefficient of the enzyme alkaline phosphatase has been shown to be due to quenching of the fluorophore caused by the substrate [143, 145]. Nevertheless, enzymes enhanced diffusion is still observed in experiments using different techniques [137, 146, 147, 148].

It is still unknown what the cause of enhanced diffusion could be. On the one hand, experiments suggest that catalysis plays a crucial role. In fact, the enhanced diffusion seems to correlate with the energy released by the enzymatic reaction [140]. Different mechanisms explaining how this energy release could speed up the enzyme diffusion have been proposed. For example, the heat dissipated by a single enzymatic reaction could generate a local temperature increase, which in turn would cause an increased diffusion as given by the Einstein relation $\Delta D = \mu k_\mathrm{b} \Delta T$, where μ is the enzyme mobility, k_b the Boltzmann constant and ΔT the temperature increase [38]. However, the heat dissipated by a typical enzymatic reaction can cause increases that are at most $\Delta T < 1\mathrm{K}$ [140]. Although these are too small to justify the experimental observations, it has been suggested that the heat released by multiple synchronous reactions could lead to global temperature increases $\Delta T \approx$ 5-10K [149]. Nevertheless, these temperature increases are not observed in experiments. Another model suggests that the energy released could cause a "chemoacoustic" effect for which the center of mass of the enzymes would be displaced by an asymmetric pressure wave propagating through the enzyme [140]. However, the energy released by the reactions is not compatible with the chemoacoustic effect [149, 150]. According to the latest measurements, it seems that the enhanced diffusion can be determined by the Gibbs free energy release rate [148]. This is consistent with super resolution experiments suggesting that the energy released from the catalysis could be fueling a ballistic component of the enzyme motion over a time interval of 1-10µs, for a distance of the order of tens of nm [142, 151]. For such a model $\alpha \sim k_\mathrm{cat} l_\mathrm{B}^2 / D_\mathrm{e}^0$, where l_B is the distance covered by the enzyme due to the energetic boost. Note that no microscopic models have been proposed to justify the ballistic motion observed.

On the other hand, experiments performed with other enzymes suggest that the catalytic step of the reaction might play a less relevant role. In fact, enhanced diffusion of these enzymes is observed in presence of inhibitors that can bind and unbind to the enzyme without reacting [136, 141, 152]. These experiments suggest that enhanced diffusion could be the result of conformational changes of the enzyme upon substrate or inhibitor binding. In Sec. 1.2.2, Fig. 1.3, we have seen how the binding of substrate can induce changes in the enzyme structure. Usually the enzyme becomes more compact and stiffer. This causes a reduction in the hydrodynamic radius and consequently an increase in diffusion coefficient [141, 153]. For such models $\alpha = (D_\mathrm{c} - D_\mathrm{e}^0)/D_\mathrm{e}^0$, where

D_c corresponds to the diffusion coefficient of the enzyme-substrate or enzyme-inhibitor complex and $D_c > D_e^0$. Having $D_c > D_e^0$ has been confirmed by Brownian dynamics simulation [154] and it has been measured for enzymes undergoing conformational changes [155]. However, whether these changes can cause a relative increase of enhanced diffusion higher than 0.15 is still argued [154, 155]. Moreover, conformational changes have also been suggested to generate hydrodynamic collective effects that could be responsible for the enhanced diffusion [156, 157].

For some enzymes, the enhanced diffusion has been related to the dissociation of enzyme oligomers into smaller subunits [158, 147]. As we have seen in Sec. 1.3.3 for the hemoglobin, oligomers are structures composed of multiple copies of the same enzyme. The oligomerization of enzyme complexes can depend on the concentration of substrate or inhibitors around the complexes and it follows a Michaelis-Menten equation of the same form of Eq. (2.5). The smaller subunits then show a higher diffusion coefficient due to their size. Although such a dissociation process has been shown to play a role for the enzymes F1-ATPase [158], hexokinase and aldolase [147], it only plays a limited role for the enzyme acetylcholinesterase and urease [147, 146]. For instance, urease shows an enhanced diffusion that can be almost 3-fold higher than in the absence of substrate in a viscous environment and not due to the dissociation of the enzyme oligomers [146]. This increase has been assessed by looking at the mean squared displacement of single enzyme molecules with total internal reflection fluorescence (TIRF) microscopy. To favor the tracking, the enzyme motion has been slowed down by the addition of methylcellulose as a viscous agent. The distribution of urease oligomers, in the presence and absence of substrate, has been controlled by single molecule photobleaching experiments.

In summary, the enhanced diffusion of enzymes can always be effectively described using Eq. (2.5). However, the exact cause of the relative increase in diffusion coefficient α depends on the specific model and the enzyme considered.

2.6.2. Directed motion in substrate gradients

Enzymes not only exhibit enhanced diffusion in the presence of their substrate but also show directed motion in substrate gradients. However, whether this motion is directed upstream or downstream depend on the experimental conditions. In some experiments enzymes have been observed to drift upstream substrate gradients, performing 'chemotaxis'. In others enzymes appear to drift downstream substrate gradients, performing 'antichemotaxis'. In the following we discuss and compare the two experimental scenarios together with the different models proposed to describe the observed enzyme motion.

Chemotaxis

The typical experimental setup used to analyze enzyme chemotaxis is shown in Fig. 2.5a. A three-inlet one-outlet microfluidic channel is employed. A fluorescently tagged enzyme (in this case urease) flows into the channel only through the middle inlet, whereas an equal concentration of buffer solution is used in all of them. The only element breaking the symmetry is the inflow of substrate (in this case urea) through the bottom inlet. This generates a lateral concentration gradient of the substrate. By using confocal microscopy is possible to analyze the enzyme profile further down the microfluidic channel. In absence of substrate the enzymatic concentration profile appears to be symmetric, see Fig. 2.5b. However, as substrate is added into the channel we see a shift on one flank of the profile. This suggests that the enzyme molecules move towards higher substrate concentrations. Moreover, the shift appears to be dependent on the substrate concentration used, see Fig. 2.5c. Similar setups have been used to characterize the chemotactic response of several enzymes. Besides urease, catalase [139], RNA polymerase [137] and several glycolytic enzymes showed a similar response [91]. Remarkably, it has been shown that the consecutive enzymes of enzymatic cascades can co-localize due to such chemotactic response [139, 91]. In fact, the intermediate produced by an enzyme of the cascade (upstream enzyme) triggers the chemotactic shift of the next enzyme in line (downstream enzyme). The downstream enzyme then drifts towards high intermediate concentrations, hence co-localizing with the upstream one.

The chemotactic response of the enzymes can be explained by considering a cross-diffusion term in the equation describing the enzyme motion. Cross-diffusion describes the response of the enzyme to forces generated by substrate gradients. Mathematically, it corresponds to an off-diagonal element in the diffusion matrix describing the motion of both the enzyme and the substrate. Such cross-diffusion has been suggested to be due to the local, attractive, short-ranged interactions between the substrate and the enzyme binding site [152, 159, 91]. Such interactions are usually referred to as specific interactions [160]. In presence of substrate gradients, enzymes that are not substrate bound would then be attracted to regions of high substrate concentration, see F. Mohajerani et al. [152] for details. The equation describing the enzyme motion, which considers both the effects of enhanced diffusion and cross-diffusion due to specific interactions, takes the form:

$$\partial_t e(x,t) = \partial_x D_{\mathrm{e}}(s)\partial_x e + \partial_x [D^{\mathrm{s}}_{\mathrm{xd}}(s,e)\partial_x s]\,, \tag{2.6}$$

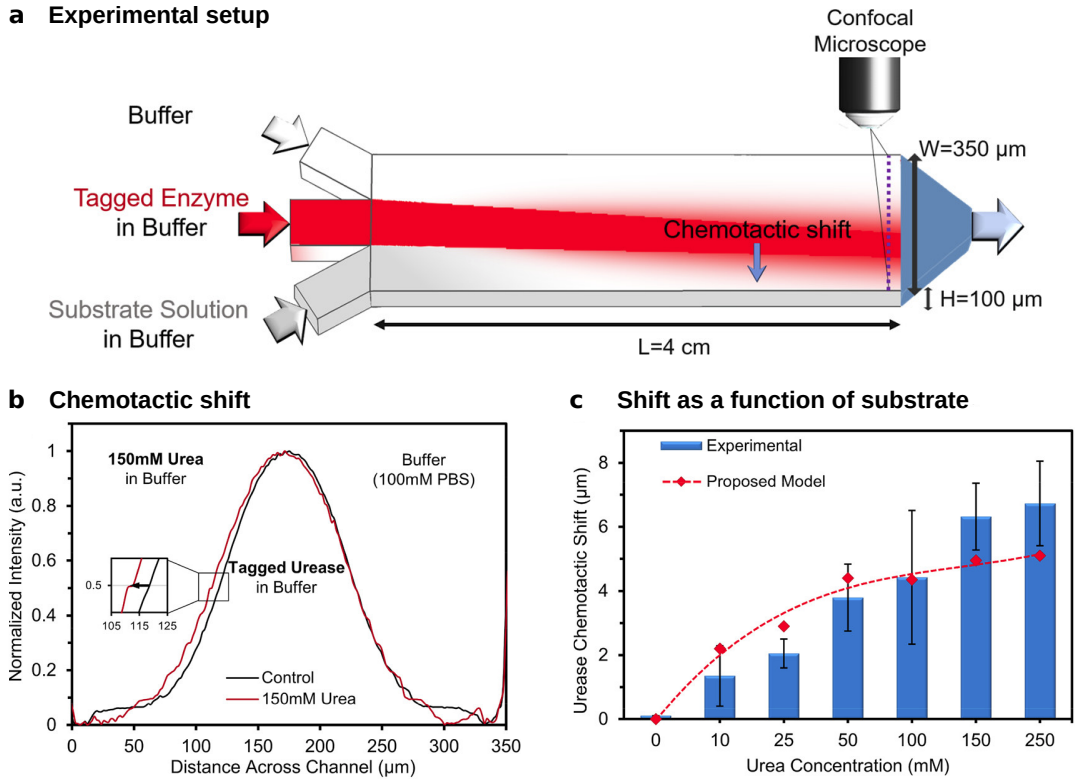


Figure 2.5.: Chemotaxis of enzymes in presence of substrate gradients. **a** Typical experimental setup used to observe enzyme chemotaxis. A three-inlet one-outlet microfluidic channel is employed. The enzyme (urease) enters the microfluidic channel through the middle inlet. The substrate (urea) enters via the bottom inlet, while equal concentrations of buffer are used in all the inlets. **b** A chemotactic shift in the enzyme profile is observed via confocal microscopy along the channel. The shift shows the tendency of enzymes to move towards the side of the channel where the concentration of substrate is the highest. **c** The chemotactic shift in the enzyme profile increases with the substrate concentration used in the bottom inlet. The blue bars are the experimental results and the red diamonds are the results of simulations that use the equation Eq. (2.7) for the enzyme motion. This equation includes the effects of enzyme enhanced diffusion and enzyme cross-diffusion generated by specific interactions with the substrate. Figures adapted with permission from [152].

where $D_e(s) = D_e^0\left(1 + \alpha \frac{s}{K_M+s}\right)$ (same as Eq. (2.5)) and the cross-diffusion is given by

$$D_{xd}^s(s,e) = -D_e^0 \frac{K_M}{K_M+s}\,\frac{e}{K_d+s}\,. \tag{2.7}$$

Eq. (2.7) shows that both the Michaelis-Menten constant K_M and the dissociation

constant $K_d = k_{-1}/k_1$ of an enzyme are determining factors for the cross-diffusion derived by F. Mohajerani et al. [152]. The first fraction in Eq. (2.7) is the fraction of free enzymes. Only free enzymes are in fact subject to forces due to specific interactions. The second fraction has to do with the magnitude of these forces. Their magnitude is higher if the binding has a high rate $k_1 \gg k_{-1}$, i.e. for values of $K_d \to 0$. The Eq. (2.7) has been used in simulations of the setup shown in Fig.2.5c to measure the shift due to chemotaxis in the enzyme profile. The results are the red diamonds shown in Fig.2.5c.

The chemotaxis of enzymes in substrate gradients would make enzymatic reactions more efficient as free enzymes would swim towards higher substrate concentrations. It could also be the mechanism behind spontaneous metabolon formation (cf. Sec. 2.2.3), by driving the co-localization of consecutive enzymes. However, these findings are still controversial and need to be further validated. In fact, another study has observed that the same enzyme used in the experiment of Fig. 2.5, urease, does not chemotax but rather antichemotax in a similar experimental setting.

Antichemotaxis

The enzyme urease that seems to chemotax in the experiment illustrated in Fig. 2.5, shows antichemotaxis in the experimental setup considered by Jee et al. [142, 151] and described in Fig. 2.6. A microfluidic channel is again employed for the analysis of the enzyme motion, see Fig. 2.6a. Using sequential dilution from two parent reservoirs, enzyme and substrate enters the channel with different profiles. The enzyme is uniform across the channel, whereas the substrate (urea) forms a linear gradient. Downstream the channel entry, urease concentration is measured by using FCS. In Fig. 2.6b, we can see how the enzyme accumulates where substrate concentration is low. This suggests that the enzyme is able to swim downstream substrate gradients, i.e. it antichemotaxes. In contrast to chemotaxis, antichemotaxis can be justified exclusively by considering enhanced diffusion. In Fig. 2.6c, we can see how the diffusion coefficient of the enzyme urease (extracted via FCS measurements) decreases across the channel together with the substrate concentration. The diffusion coefficient goes from being $\approx 20\%$ higher than the one in the absence of substrate on the left, to be approximately equal on the right.

This type of enzyme motion can be described by the following equation of motion [142]:

$$\partial_t e(x,t) = \partial_x^2 [D_e(s) e] \,, \tag{2.8}$$

where the enhanced diffusion $D_e(s)$ has the Michaelis-Menten form of Eq. (2.5). The steady state response to a linear substrate gradient given by Eq. (2.8) is given by $e \sim 1/D_e(s)$. Such a profile, fits very well the experimentally observed measurements

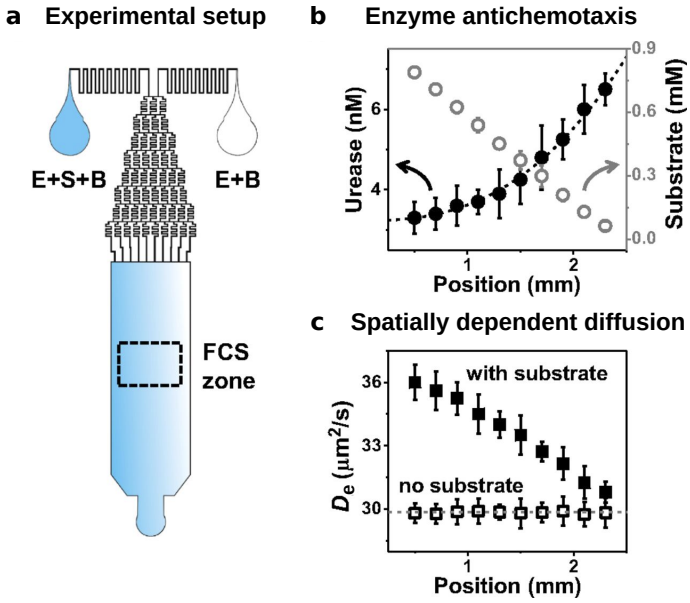


Figure 2.6.: Antichemotaxis of enzymes in presence of substrate gradients. **a** Experimental setup used in the experiments that observed enzyme antichemotaxis. A microfluidic channel is employed. The solution enters the channel via a sequential dilution from two parents reservoirs. Equal concentrations of the enzyme urease and buffer solution are employed in both reservoirs, while the substrate urea is present only in the left reservoir. This generates a linear substrate gradient at the entry of the channel as shown by the cyan shading. The enzyme and the buffer instead have initial homogeneous concentrations. **b** Enzyme antichemotaxis. The enzyme concentration is measured via FCS across the channel. The results are shown by the black solid circles. The grey open circles represent the linear substrate gradient measured in the absence of enzymes. The enzyme accumulates on the right, where the concentration of substrate is low. **c** Spatially dependent diffusion coefficient. The diffusion coefficient of the enzyme is measured via FCS across the channel with and without substrate. In presence of substrate, the diffusion coefficient of the enzyme is higher for higher concentration of substrate. However, in the absence of substrate, the diffusion coefficient is uniform across the channel. Figures adapted with permission from [142].

as can be seen by looking at the dashed black line in Fig. 2.6b. By coupling a super resolution microscopy technique (stimulated emission depletion microscopy, i.e. STED) to FCS, Jee et al. [142, 151] observed a ballistic component of the enzyme motion over distances l_B of the order of tens of nanometers, with a transient time between 5 and 10μs. They suggested that such ballistic motion could be fueled by the realease of the Gibbs free energy of the reaction (10-20k_BT) over the observed time interval. As already mentioned in Sec. 2.6.1, this would generate a maximal relative increase in diffusion coefficient of $\alpha \sim k_{cat} l_B^2 / D_e^0$, where the prefactor would depend on whether or not the

timescale of the boost is comparable to that of enzyme rotational diffusion [151].

Note that in Eq. (2.8), $D_e(s)$ is inside the second spatial derivative; this is in contrast with the model presented in the previous section where $D_e(s)$ was outside the second derivative, cf. Eq. (2.6). Both models consider a spatial dependence of the diffusion coefficient. The difference among the two models accounts for different microscopic mechanisms behind the spatial dependence [161]. Let us imagine to have discretized space, the diffusion coefficient of the enzyme molecules is then proportional to the hopping rate between adjacent positions. The description given by the enhanced diffusion term of Eq. (2.6) corresponds to a case in which the hopping rate of enzymes is symmetric with respect to the exchange of positions; whereas the model given by Eq. (2.8) describes an enhanced diffusion whose hopping rates are asymmetric and only depends on the starting position.

All possible mechanisms leading to an equation of motion for the enzymes as given by Eq. (2.8) will result in enzyme antichemotaxis. An example is given by the de-oligomerization of the enzyme urease observed for high substrate concentrations [147]. In such a regime, the enzyme urease seems to antichemotax, following a motion as given by Eq. (2.8) [151].

A model capable of both chemotaxis and antichemotaxis

We have seen how specific interactions between enzyme and substrate can cause cross-diffusion of enzymes in response to forces generated by substrate gradients, see Eq. (2.7). However, the possible effects of nonspecific interactions have been neglected. Nonspecific interactions refer to forces to which the substrate molecules (in general any molecule) are subject to once they are in proximity of the enzyme surface [160]. They are not specific to the binding site and they include interactions such as steric, electrostatic, van der Waals, etc. Unlike the specific interactions, they can be both attractive and repulsive. J. Agudo-Canalejo et al. [162] considered the equation of motion for an enzyme subject to such interactions. Moreover, while the model considered by F. Mohajerani et al. [152] leading to Eq. (2.7) is a thermodynamic model, the one considered by J. Agudo-Canalejo et al. starts from a microscopic picture where hydrodynamics effects are also taken into account.

The model is illustrated in Fig. 2.7. A single enzyme molecule is assumed to be in a bath of substrate molecules with a fixed gradient. Depending on whether the enzyme is free or complexed with the substrate, we assume it to have different diffusion coefficients, D_e^0 and D_c respectively. Moreover, we assume that the nonspecific interactions are also dependent on whether the enzyme is free or not. We denote the potentials characterizing such interactions with ϕ_{fs} and ϕ_{cs}, respectively. As already mentioned in Sec. 2.6.1, assuming two different diffusion coefficients, depending on

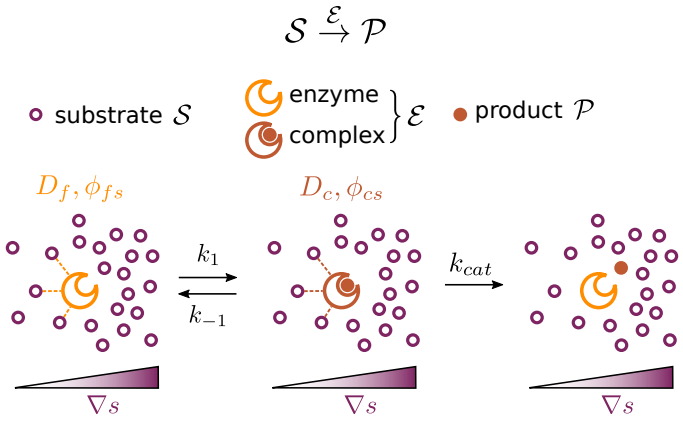


Figure 2.7.: Model for the enzyme motion driven by nonspecific interactions. The model considers a single enzyme molecule in a bath of substrate. The substrate concentration is assumed to have a fixed gradient as depicted by the purple shading. The model assumes a different diffusion coefficient for the enzyme, depending on whether the enzyme is free (orange) or it forms a complex with the substrate (brown). Moreover, the enzyme interacts with the substrate molecules via nonspecific interactions. We indicate with ϕ_{fs}, ϕ_{cs} the interaction potentials that are assumed to depend on whether the enzyme is free or not, respectively. The enzyme is assumed to catalyze a Michaelis-Menten reaction. The substrate molecules bind and unbind with rates k_1 and k_{-1} and the catalysis has a catalytic rate k_{cat}. Figure adapted with permission from [162].

the enzyme form, implies that the enhanced diffusion takes the form

$$D_{\text{e}}(s) = D_{\text{e}}^0 \left[1 + \frac{(D_{\text{c}} - D_{\text{e}}^0)}{D_{\text{e}}^0} \frac{s}{K_{\text{M}} + s} \right] , \tag{2.9}$$

which is a Michaelis-Menten interpolation between the free enzyme diffusion coefficient D_{e}^0, obtained for low substrate concentration $s \ll K_{\text{M}}$, and the diffusion coefficient of the complex D_{c} obtained in the saturated regime $s \gg K_{\text{M}}$.

The nonspecific interactions instead generate an effective cross-diffusion

$$D_{\text{xd}}^{\text{s}} = -\frac{N_{\text{A}} k_{\text{B}} T}{\eta} \left[{\lambda_{\text{f}}^{\text{s}}}^2 + ({\lambda_{\text{c}}^{\text{s}}}^2 - {\lambda_{\text{f}}^{\text{s}}}^2) \frac{s}{K_{\text{M}} + s} \right] e , \tag{2.10}$$

where η is the viscosity of the fluid, k_{B} the Boltzmann constant, N_{A} the Avogadro number, T the temperature and $\lambda_{\text{c/f}}^{\text{s}}$ the Derjaguin lengths associated to the nonspecific interactions. The Derjaguin length is a parameter capturing the lengthscale associated to the nonspecific interactions. It is typically a few angstroms [163, 164] and can at most be as large as the Debye length (screening length), which in typical buffer conditions is about 1nm [155]. It is expressed via the integral ${\lambda_{\text{c/f}}^{\text{s}}}^2 = \int_0^\infty dh h (e^{-\phi_{\text{cs/fs}}(h)(k_{\text{B}}T)^{-1}} - 1)$.

$\lambda^{s\,2}_{c/f}$ is positive for attractive interactions and negative for repulsive ones [162]. Note that the derivation of $\lambda^{s\,2}_{c/f}$ is similar to that of the second virial coefficient for a real gas [165], but it also includes hydrodynamic effects and is computed by assuming that the size of the enzyme is much larger than the interaction length. In the simple case that $\lambda^{s\,2}_{f}$ and $\lambda^{s\,2}_{c}$ have the same sign, from Eq. (2.10), we can see that the sign of $\lambda^{s\,2}_{c/f}$ determines the sign of the cross-diffusion. In presence of a substrate gradient the enzyme will be pushed to the side of higher substrate concentration if the nonspecific interactions are attractive ($\lambda^{s\,2}_{c/f} > 0$ and $D^{s}_{xd} < 0$) and it will be pushed in the opposite direction if the interactions are repulsive ($\lambda^{s\,2}_{c/f} < 0$ and $D^{s}_{xd} > 0$).

The equation of motion of an enzyme subject to nonspecific interaction is therefore given by

$$\partial_t e(x,t) = \partial_x^2[D_e(s)e] + \partial_x[D^{s}_{xd}(s,e)\partial_x s]\,, \tag{2.11}$$

where $D_e(s)$ and $D^{s}_{xd}(s)$ follow Eqs. (2.9), (2.10) respectively. The equation of motion Eq. (2.11) is a mathematical generalization of the equation presented by Jee et al. [142, 151] (Eq. (2.8)). The only difference lies in the microscopic mechanisms behind the enhanced diffusion and, therefore, the two models present different relative increases in the diffusion coefficient α. Nevertheless, the two $D_e(s)$ functions have the same functional dependence on s. The equation of motion proposed by J. Agudo-Canalejo et al. [162] generalizes also some of the models proposed in the literature to justify enzyme chemotaxis [159, 91]. However, the model proposed by J. Agudo-Canalejo et al. [162] and the one proposed by Mohajerani et al. [152] (given by Eq. (2.6)) present some differences. The enhanced diffusion of Eq. (2.11) generates a drift in the enzyme motion given by $[\partial_s D_e(s)]e = \alpha D_e^0 K_M/(K_M + s)^2$ that is very similar in form to the one suggested by Mohajerani et al. [152] and given by Eq. (2.7). Considering that often $K_M \approx K_d$, the two differs only by a prefactor α and by having opposite signs.

Overall, we have seen that different models have been proposed to describe the enzyme motion in presence of their substrate. These studies focused on the response of the enzyme to pre-imposed substrate profiles. They did not consider the effects that the enzymatic reaction has in shaping the substrate profile, which in turns has an effect on the enzyme motion. In Ch. 4, we will show that a positive feedback between enzyme accumulation due to nonspecific interactions and reaction can drive spatial pattern formation starting from initial homogeneous concentrations. We will therefore demonstrate that studying the effect that the reaction has on the substrate profile can be important in selecting the correct model of enzyme motion. In fact, patterns can only form for the model proposed by J. Agudo-Canalejo et al. [162], but not for the other models.

3. Design principles for the optimal spatial arrangement of enzymes*

Due to evolutionary pressure and limited amount of resources, cells face numerous allocation problems demanding decisions on how to distribute their resources. To regulate enzymatic reactions, cells not only decide the amounts and types of enzymes to be produced, but they also regulate their spatial arrangement (cf. Ch. 2). For example, enzymes can be positioned to maximize the product flux of the reaction that they catalyze. In this chapter, we derive an optimal allocation principle for the spatial distribution of enzymes in a wide class of reaction-diffusion systems for which the enzyme's substrate is imported into the system through the external boundary or the boundaries of internal compartments. So far, for systems presenting a single source of substrate and spherical symmetry, these distributions have been analyzed by stochastic optimization algorithms. However, these methods were unable to provide a systematic construction of the optimal profile for systems without specific symmetries. To address this issue, here we derive an optimal allocation principle demanding that the available enzymes are distributed such that the marginal flux returns at each occupied position are equal. This homogeneous marginal returns (HMR) principle generalizes the Kelly criterion [166] (widely used for betting and portfolios optimization) to a case where each bet globally feeds back onto all payoffs (see graphical abstract Fig. 3.1). The HMR criterion allows us to analytically understand and characterize a localization-delocalization transition in the optimal enzyme distribution with respect to the source of substrate that was previously observed numerically [167, 168]. In particular, our analysis reveals the generality of the transition, and produces an experimentally relevant test for the optimality of enzyme localization at the substrate source, by comparing the reaction flux to the influx of substrate. We also leverage the HMR criterion to devise a deterministic construction algorithm to build up the optimal enzyme arrangements systematically rather than by using stochastic algorithms. Taken together, our results reveal a common principle in allocation problems from biology and economics, which can also serve as a design principle for synthetic biomolecular systems.

*The chapter is adapted from the manuscript: "A generalized Kelly criterion for the optimal spatial allocation of enzymes", by G. Giunta, F. Tostevin, U. Gerland, which has been submitted for publication.

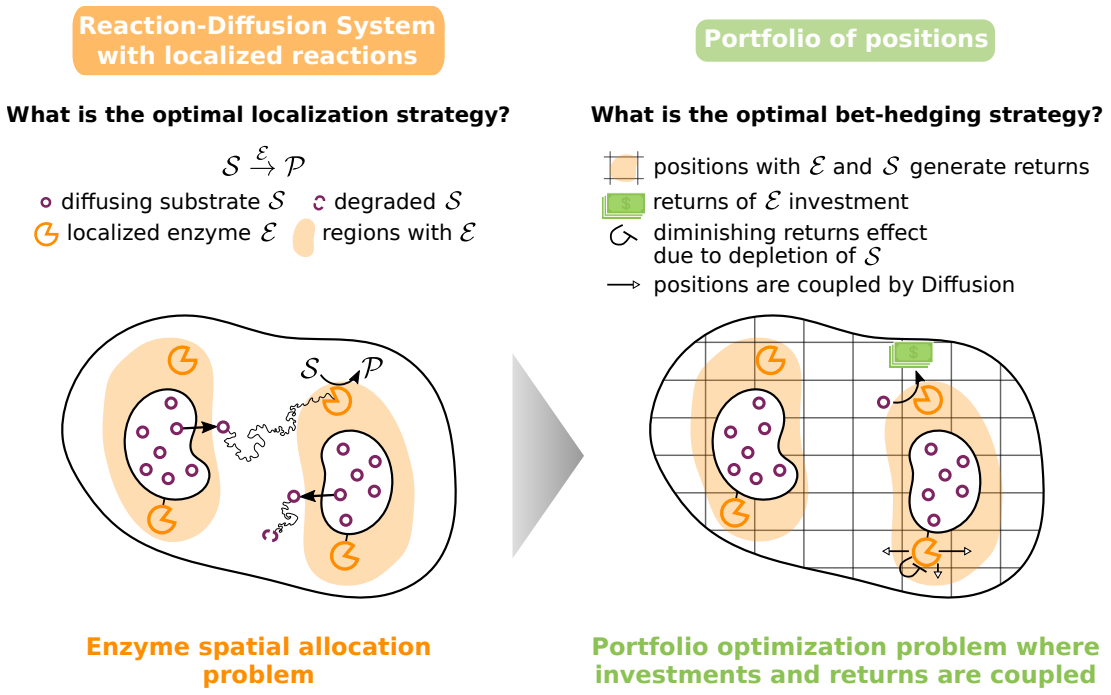


Figure 3.1.: Graphical abstract: Design principles for the optimal spatial arrangement of enzymes. We consider a class of reaction-diffusion systems where an enzyme $\mathcal{E}$ converts a substrate $\mathcal{S}$ into a product $\mathcal{P}$. The class of systems is such that $\mathcal{S}$ is imported into the system either from the external boundary or the boundary of internal compartments. Once into the system, $\mathcal{S}$ can be degraded, lost to competing reactions, or react with $\mathcal{E}$, implying that $\mathcal{S}$ is not uniformly distributed in the system at steady state. Moreover, while $\mathcal{S}$ diffuses, $\mathcal{E}$ is assumed to be localized (static). We ask: What is the enzyme spatial arrangement that maximizes the reaction flux? To answer the question we draw an analogy between the enzyme spatial allocation problem and a portfolio optimization problem. The system can be interpreted as a portfolio, where each position represents an asset generating a certain amount of returns in terms of reaction flux. Enzymes are capital that can be invested in the system at the different positions. As opposite to classical economical systems where investments and returns are assumed to be independent, here the reaction-diffusion dynamics causes the returns in reaction flux and the enzyme investments to be coupled. More specifically, since the reaction between enzyme and substrate at a given location causes the substrate to be depleted, this generates a diminishing returns effect for which the addition of extra enzymes at that position is gradually less efficient. Moreover diffusion causes the different positions to be coupled. By generalizing a widely used betting criterion (the Kelly criterion [166]) to include these couplings, we show that the enzyme spatial allocation problem can be solved in terms of a bet-hedging strategy for a portfolio optimization problem where investments and returns are coupled.

3.1. Introduction

Living organisms have a limited amount of resources at their disposal that need to be allocated among different molecular processes. To cope with such limitations, organisms present a variety of bet-hedging strategies that allows them to efficiently use their resources. Such strategies share the basic characteristics of other resource allocation problems, where a given amount of a resource must be distributed among competing alternatives to maximize the expected performance of the system [169]. On the cellular level, for example, cells of a colony of microbes can differentiate into different cell types, each specialized in a task that contributes to the growth of the colony as a whole [170]. This allows a more efficient enzymes usage as cells compartmentalize metabolic pathways, avoiding cross-talk of toxic intermediates. On the molecular level, enzymes are also efficiently allocated. For instance, a limited number of ribosomes must produce different types of proteins in different ratios to achieve balanced cell-growth [171, 172, 173]. Moreover, cells need to dynamically adjust these ratios to support their growth under changes in environmental conditions [174, 175]. Besides regulating production ratios, we have seen in Ch. 2 that cells are able to control the spatial arrangement of enzymes to optimize or regulate metabolic fluxes [79, 176]. Sec. 2.2.2 showed how consecutive enzymes evolved structural features to directly channel intermediates from one another. Sec. 2.2.3 showed how the formation of large enzymatic clusters (metabolons) enhances the flux of the final product of a pathway and reduces the transient times of reactions by constraining the diffusion of intermediates. We have also seen that synthetic strategies can be used to regulate the spatial allocation of enzymes, both *in vivo* and *in vitro*, see Sec. 2.5. Such strategies can be used to boost the yields of useful products such as drug molecules [128, 177, 178] and biofuels [122, 123].

Here, we solve the spatial allocation problem where multiple copies of a single type of enzyme $\mathcal{E}$ can be assigned to different spatial positions in a wide class of reaction-diffusion systems, see Fig. 3.2. This class of systems applies to domains of arbitrary shape, in which the substrate $\mathcal{S}$ is imported either from the external boundary and/or the boundary of internal compartments. The enzyme catalyzes the conversion of the substrate $\mathcal{S}$ into a product $\mathcal{P}$. As a measure for the performance of the system, we consider the total reaction flux, i.e. the overall rate at which product is generated. We are interested in finding the optimal enzyme profile $e^*(\mathbf{r})$. Why would the spatial distribution $e(\mathbf{r})$ of the enzymes $\mathcal{E}$ affect the total reaction flux? The rate at which an enzyme catalyzes reactions depends on the local concentration of its substrate $\mathcal{S}$. The $\mathcal{S}$ molecules, in turn, are not necessarily uniformly distributed within cells: (i) As already mentioned, they are often imported into the cytoplasm from the cell exterior or from an internal organelle (see Sec. 2.4 and Fig. 3.2b-Import), or they may be produced by

other non-uniformly distributed enzymes. (ii) They are subject to loss mechanisms, e.g. leakage through a membrane [48], conversion by a competing enzyme [59], or spontaneous decay of unstable molecules such as PRPP [179] (see Fig. 3.2b-Loss). (iii) The enzyme $\mathcal{E}$ also affects the concentration profile of its substrate $\mathcal{S}$ by depleting it locally. (iv) Substrate molecules are stochastically transported via diffusion and possibly also actively transported by cytoplasmic streaming [180] or cargo-carrying molecular motors [181] (see Fig. 3.2b-Transport). The interplay of these mechanisms can produce stable non-uniform substrate profiles, as can be seen in a variety of different contexts, including embryo-morphogenesis [182], cell mitosis [183], asymmetric division of bacteria [184], virulence strength [185] and control of cell length in yeast [186]. Having non-uniform substrate profiles implies that different spatial enzyme distributions $e(\mathbf{r})$ generate different total reaction fluxes. Note that potential co-localization of $\mathcal{E}$ with other enzymes [65, 187, 59, 66] is incorporated in our model only to the extent that the $\mathcal{S}$ influx at the boundaries may be caused by other, boundary-localized enzymes.

As we already discussed in Sec. 2.4 for cases where the substrate is produced inside organelles (e.g. the mitochondria or endoplasmic reticulum), the enzymes tend to be bound to the organelle membrane, or located in their proximity also being present in the cytoplasm [115, 111, 106, 93], see Fig. 2.2c and Fig. 3.2a. In contrast, when the substrate is imported into the cytoplasm from the external environment, some of the associated metabolic enzymes are localized to the external membrane [89, 105], see Fig. 2.2a-b. This intuitively guarantees an efficient substrate turnover as losses of substrate should be prevented, cf. Sec. 2.4. Surprisingly, a previous study of a minimal one-dimensional model within the class of systems illustrated in Fig. 3.2 observed a phase transition in the optimal enzyme distribution $e^*(\mathbf{r})$ as a function of a dimensionless reaction-diffusion parameter [167]: When reactions are slow compared to diffusion, localizing all $\mathcal{E}$ at the source boundary is optimal, while a more extended profile with enzymes also in the interior is optimal for faster reactions. The existence of this transition was subsequently found to be robust with respect to the spatial dimension and specific loss mechanism and reaction kinetics considered [168]. However, the physical principles governing the optimal configurations and the generality of the transition have remained elusive, and a systematic way of constructing the optimal enzyme arrangement is lacking.

The optimal allocation of limited resources is extensively studied in economics. Rules such as Kelly's criterion [166] are used to determine optimal betting strategies or portfolios [188]. Kelly's criterion specifies how to place bets to maximize the long-term growth rate of a gambler's capital when betting on several mutually exclusive events. We draw an analogy between the biological problem of spatial enzyme allocation and economical investment problems. Within this analogy, enzyme molecules represent capital that can be invested at different positions within the system. Each position can

be interpreted as an asset, generating a certain expected return in terms of reaction flux. An important assumption of Kelly's problem is that the gambler's bets do not directly affect the payoffs/returns of the different events. This is certainly not the case for the enzyme allocation problem, where enzymes placed at different positions have direct effects on the substrate profile and on the reaction flux. Drawing an analogy between the two problems helps in better understanding the enzyme allocation problem and it also helps in providing a generalization of the Kelly criterion for cases where payoffs and bets are directly coupled. We can therefore ask: can we use strategies of the type of Kelly's one for the determination of the optimal enzyme arrangements? Moreover, the localization-to-delocalization transition observed in the optimal enzyme arrangement is reminiscent of a transition in investment strategy from investing everything on a single asset to having a diversified portfolio where investments are partitioned among multiple assets. Therefore we can also ask: Can we explain the generality of the localization-to-delocalization transition through the analogy with a diversification investment strategy? Lastly: just as Kelly's criterion provides a set of rules for the optimal repartition of bets, can we derive an algorithm for the construction of the optimal arrangement of enzymes? To address these questions, we devise a variational approach to characterize the optimal enzyme arrangements and reveal the underlying investment diversification strategy.

3.2. Model

We consider catalytic reactions where a substrate $\mathcal{S}$ is converted into a product $\mathcal{P}$ by an enzyme $\mathcal{E}$ (Fig. 3.2a). The substrate enters the system either from the exterior or from an internal compartment (Fig. 3.2b-Import). Enzymes may be located both in the bulk of the system and at the boundaries where the substrate is imported. The former fraction is described by the concentration field $e(\mathbf{r})$, with $\mathbf{r}$ denoting positions inside the system, while $e_S(\mathbf{s})$ is the density of $\mathcal{E}$ on the boundary surface at position $\mathbf{s}$. The concentration of $\mathcal{S}$ at time t is $\rho(\mathbf{r}, t)$. We consider the distribution of $\mathcal{E}$ as stationary and explore all possible spatial arrangements of the enzymes. We assume that the rate at which $\mathcal{S}$ is converted into $\mathcal{P}$ is described by $k_{\text{cat}} e F[\rho]$, where k_{cat} is the enzyme turnover rate and $F[\cdot]$ is a monotonically-increasing reaction function that can be chosen to represent a particular reaction scheme, e.g. Michaelis-Menten kinetics, $F[\rho] = \rho/(K_{\text{M}} + \rho)$, or linear mass-action kinetics, $F[\rho] = \rho$, or a Hill function $F[\rho] = \rho^h/(K^h + \rho^h)$ (cf. Sec. 1.3.3).

Within the system, substrate molecules undergo diffusive transport with diffusion coefficient D and may also be advected by a velocity field $\mathbf{v}(\mathbf{r})$ (Fig. 3.2b-Transport). If the substrate is intrinsically unstable or subject to competing reaction pathways, it decays with rate constant σ (Fig. 3.2b-Loss). Taking into account these different

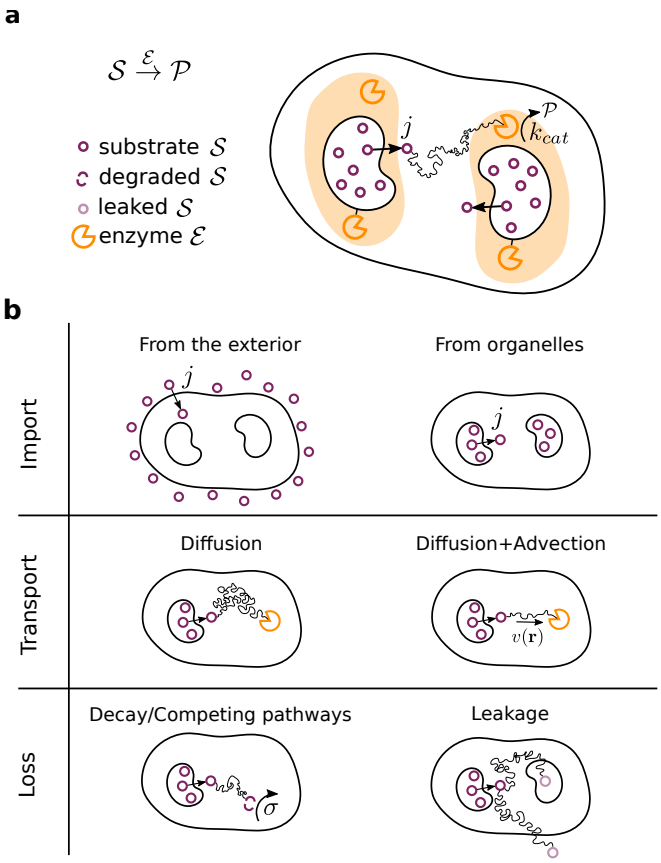


Figure 3.2.: Class of reaction-diffusion systems considered for the optimal spatial allocation of enzymes. **a** We consider a single-step catalytic reaction where a substrate $\mathcal{S}$ is converted by an enzyme $\mathcal{E}$ into product $\mathcal{P}$. The enzyme molecules can be freely placed, being either bound to the external or the internal boundaries or placed anywhere within the system. The substrate has influx j and it is processed by the enzyme with a catalytic rate k_{cat}. **b**-Import: $\mathcal{S}$ is either imported from the exterior or produced within internal compartments. **b**-Transport: Once into the cytoplasm the substrate can follow an advection-diffusion transport, where the advection can be caused by the active transport of substrate molecules or cytoplasmic streaming. **b**-Loss: During its transport the substrate can also be degraded with rate σ, leaked through the external membrane or absorbed into an other organelle.

processes, the dynamics of $\mathcal{S}$ follows

$$\partial_t \rho(\mathbf{r},t) = D\nabla^2\rho(\mathbf{r},t) - \nabla\cdot[\rho(\mathbf{r},t)\mathbf{v}(\mathbf{r})] - \sigma\rho(\mathbf{r},t) - k_{\text{cat}}e(\mathbf{r})F[\rho(\mathbf{r},t)]\,, \tag{3.1}$$

where the first two terms on the right hand side describe the transport of $\mathcal{S}$ by diffusion and advection, the third term describes decay, and the last describes the reaction

of substrate with enzymes. In the following, we focus on steady-state conditions, $\partial_t \rho(\mathbf{r}, t) = 0$, where $\rho(\mathbf{r}, t) = \rho(\mathbf{r})$.

We consider two types of boundaries for the system: boundaries at which substrate is imported and absorbing boundaries. For the first type, by including the reaction of substrate with the enzymes that may be located at such a boundary, we have a boundary condition of the form

$$j(\mathbf{s}) = \left[D \frac{\partial \rho}{\partial \mathbf{n}(\mathbf{s})} - \rho(\mathbf{s}) \mathbf{v}(\mathbf{s}) \cdot \hat{\mathbf{n}}(\mathbf{s}) \right] + k_{\text{cat}} e_S(\mathbf{s}) F[\rho(\mathbf{s})], \tag{3.2}$$

where $j(\mathbf{s})$ is the influx of $\mathcal{S}$ at position $\mathbf{s}$, $\hat{\mathbf{n}}(\mathbf{s})$ is a unit vector normal to the boundary, pointing outwards from the system, and $\frac{\partial}{\partial \mathbf{n}(\mathbf{s})}$ represents the magnitude of the normal component of the gradient. This condition enforces flux conservation at the boundary: the influx $j(\mathbf{s})$ must equal the sum of the local transport and reaction flux. We allow for $j(\mathbf{s})$ to be a discontinuous function of $\mathbf{s}$ as there could be regions on the boundary without influx, $j(\mathbf{s}) = 0$ (see Appendix A.1 for details on how such discontinuities are treated). The second kind of boundaries we consider are absorbing boundaries with $\rho(\mathbf{s}) = 0$, from which substrate is rapidly lost (see Fig. 3.2b-Loss). For convenience, we combine these two types of boundary condition

$$0 = k(\mathbf{s}) \left\{ j(\mathbf{s}) - \left[D \frac{\partial \rho}{\partial \mathbf{n}(\mathbf{s})} - \rho(\mathbf{s}) \mathbf{v}(\mathbf{s}) \cdot \hat{\mathbf{n}}(\mathbf{s}) \right] - k_{\text{cat}} e_S(\mathbf{s}) F[\rho(\mathbf{s})] \right\} - h(\mathbf{s}) \rho(\mathbf{s}) \,, \tag{3.3}$$

where $k(\mathbf{s})$ and $h(\mathbf{s})$ are set to represent absorbing boundaries ($k = 0$, $h = 1$) or boundaries with influx of substrate ($k = 1$, $h = 0$).

The total flux $J^{\mathcal{P}}$, i.e. the rate at which the whole system produces $\mathcal{P}$, has contributions from enzymes at the boundaries and within the bulk,

$$J^{\mathcal{P}} = \int_S k_{\text{cat}} e_S(\mathbf{s}) F[\rho(\mathbf{s})] \mathrm{d}\mathbf{s} + \int_V k_{\text{cat}} e(\mathbf{r}) F[\rho(\mathbf{r})] \mathrm{d}\mathbf{r} \,, \tag{3.4}$$

where the first integral is over all boundary surfaces S and the second is over the volume V of the system. The question is: Given a fixed amount of available enzymes, $E_{\text{tot}} = \int_S e_S(\mathbf{s}) \mathrm{d}\mathbf{s} + \int_V e(\mathbf{r}) \mathrm{d}\mathbf{r}$, how should these enzymes be positioned such as to maximize the rate $J^{\mathcal{P}}$ of product formation?

This optimization problem can be approached by defining a functional of the form $\mathcal{L} = J^{\mathcal{P}} - \sum_i \lambda_i C_i$, where C_i represent different constraints with associated Lagrange multipliers λ_i. For our model, these constraints are that (i) the total amount of enzymes must equal E_{tot}, (ii) $\rho(\mathbf{r})$ and $e(\mathbf{r})$ must jointly satisfy the reaction-diffusion-advection equation, Eq. (3.1), at each point $\mathbf{r}$, and (iii) $\rho(\mathbf{s})$ and $e_S(\mathbf{s})$ must jointly satisfy the boundary condition, Eq. (3.3), at each boundary point $\mathbf{s}$. The resulting Lagrangian has

the form

$$\mathcal{L} = J^{\mathcal{P}} - \lambda_e \left[\int_S e_S(\mathbf{s}) \mathrm{d}\mathbf{s} + \int_V e(\mathbf{r}) \mathrm{d}\mathbf{r} - E_{\mathrm{tot}} \right] + \int_V \lambda_V(\mathbf{r}) \mathcal{D}\left[e(\mathbf{r}), \rho(\mathbf{r}), \mathbf{r}\right] \mathrm{d}\mathbf{r} + \int_S \lambda_S(\mathbf{s}) \mathcal{B}\left[\rho(\mathbf{s}), e_S(\mathbf{s}), \mathbf{s}\right] \mathrm{d}\mathbf{s}, \tag{3.5}$$

where $\mathcal{D}\left[e(\mathbf{r}), \rho(\mathbf{r}), \mathbf{r}\right]$ and $\mathcal{B}\left[\rho(\mathbf{s}), e_S(\mathbf{s}), \mathbf{s}\right]$ represent the right-hand side of Eqs. (3.1) and (3.3) respectively, and the signs in front of the Lagrange multipliers have been chosen such that they result to be positive. Maximizing $\mathcal{L}$ with respect to $e_S(\mathbf{s})$ and $e(\mathbf{r})$ while simultaneously minimizing with respect to all Lagrange multipliers yields the optimal enzyme arrangements satisfying the constraints.

The optimal enzyme arrangement may consist of both regions with a finite density of enzymes and regions with no enzymes. Indeed, previous numerical studies have shown that optimal enzyme distributions are often discontinuous [167, 168, 59]. This is generally incompatible with the Lagrangian optimization formalism, since the optimization is performed over the space of smooth functions. We work around this restriction by splitting the volume V into different sub-spaces (and similarly for the boundaries S); within each sub-space the enzyme density $e(\mathbf{r})$ must be continuous, but at the interfaces it need not be. The number of sub-spaces as well as the shape of the interface between each pair of sub-spaces are then also variables to be optimized. Except where specifically noted, this complication is not important for understanding the following results (see Appendix A.1 for details).

We will see below that the variational approach is useful even when it is impossible to solve for the optimal enzyme arrangement analytically. In cases where it is possible to extract the exact functional form of the optimal profile, the problem can usually be reduced to a low-dimensional optimization over a small number of parameters, which is then easily performed numerically (see Appendix A.5 for an example and Fig. 3.4, magenta lines). For systems where the functional form of $e(\mathbf{r})$ cannot be found analytically, such as those with complex geometries, we provide an algorithmic procedure for constructing the optimal enzyme arrangement.

3.3. Results

Our optimization problem can be understood in terms of an optimal betting strategy by drawing an analogy with the Kelly problem [166]. Before presenting the optimal allocation principle for the enzyme arrangement, we discuss the optimal betting strategy found by Kelly and we generalize it to a case where bets and payoffs are coupled.

3.3.1. Generalization of Kelly criterion

Kelly found a criterion for the optimal allocation of bets that is widely used in betting and portfolio theory [188]. He considered the case of a gambler who wants to maximize the long-term growth rate G of his capital over a succession of repeated games. At each game, the bets are placed on mutually exclusive events s. We indicate with b the fraction of the gambler's capital that is not invested in the game, with $a(s)$ the fraction invested on the event s and with α_s the odds, i.e. the money payed back to the gambler in case s is the winning event. Given an initial capital V_0, after one game the gambler capital becomes $V_1 = V_0(b + a(s)\alpha_s)$ if s is the winning event. Note that the gambler can go bankrupt if he bets all of his capital on the losing events, i.e. for $b = 0$ and $a(s) = 0$. After a succession of N games, the gambler's capital is

$$V_{\mathrm{N}} = V_0 \prod_s (b + a(s)\alpha_s)^{W_s}\,, \tag{3.6}$$

where W_s is the number of times that s is the winning event. The long term capital growth rate is given by

$$G := \lim_{N\to\infty} \frac{1}{N} \log\left(\frac{V_{\mathrm{N}}}{V_0}\right) = \lim_{N\to\infty} \sum_s \frac{W_s}{N} \log\left[b + a(s)\alpha_s\right] = \sum_s p(s) \log\left[b + a(s)\alpha_s\right], \tag{3.7}$$

where $p(s) = \lim_{N\to\infty} W_s/N$ is the probability for s to be the winning event.

Kelly's aim was to determine the optimal way of partitioning the gambler's capital given that the gambler can neither borrow nor lend money, hence his aim was to determine the optimal set of $a(s)$ and b, under the constraint $\sum_s a(s) + b = 1$. We can therefore consider the functional

$$\mathcal{G} = \sum_s p(s) \log\left[b + a(s)\alpha_s\right] - \lambda \left[\sum_s a(s) + b - 1\right], \tag{3.8}$$

where λ is the Lagrange multiplier associated with the constraint on the gambler's capital. To find the optimal set of $a(s)$ and b we perform the following derivatives:

$$\frac{\partial \mathcal{G}}{\partial a(s)} = \frac{p(s)\alpha_s}{b + \alpha_s a(s)} - \lambda = 0 \quad \text{for } s \in K \tag{3.9}$$

$$\frac{\partial \mathcal{G}}{\partial b} = \sum_s \left[\frac{p(s)}{b + \alpha_s a(s)}\right] - \lambda = 0 \tag{3.10}$$

$$\frac{\partial \mathcal{G}}{\partial a(s)} = \frac{p(s)\alpha_s}{b} - \lambda \leq 0 \quad \text{for } s \in K'\,, \tag{3.11}$$

where we indicated with K the set of events for which it is optimal to bet and $a(s) > 0$, and with K' the set of events for which betting is suboptimal and $a(s) = 0$. The derivatives above can be interpreted as the variation of the constrained capital growth rate $\mathcal{G}$ per extra bet placed (3.9), (3.11) or per extra fraction of capital kept (3.10). The first terms of the derivatives are the net marginal increases of G generated by either placing the bets (3.9), (3.11) or by keeping an extra fraction of capital (3.10), if no constraint is considered. The $-\lambda$ terms correspond to the marginal cost of betting the extra capital. The optimal strategy is to bet on those events $s \in K$ up to the point for which the marginal gain generated equals the marginal cost. Moreover the optimal way of allocating the gambler's capital, known as Kelly criterion, can be obtained by solving for the above derivatives, together with the use of the constraint:

$$\lambda = 1 \tag{3.12}$$

$$b = \frac{1 - \sum_{s \in K} p(s)}{1 - \sum_{s \in K} 1/\alpha_s} \tag{3.13}$$

$$a(s) = \frac{p(s)\alpha_s - b}{\alpha_s} \quad \text{for } s \in K\,. \tag{3.14}$$

To understand how to determine the set K of events for which it is optimal to bet, we refer the readers to Kelly's original paper [166].

Here we ask: What would happen to Kelly's optimization problem if a feedback between the bets placed $a(s)$ and the odds α_s were to be considered? To answer the question let us imagine that the odds are now functions of the bets placed $\alpha_s = f_s(\mathbf{a})$, where $\mathbf{a}$ is the vector $(a(1), a(2), ...)$. A new functional $\mathcal{G}$ can be considered:

$$\mathcal{G} = \sum_s p(s) \log\left[b + a(s)\alpha_s\right] - \lambda \left[\sum_s a(s) + b - 1\right] - \sum_s \lambda_s \left[\alpha_s - f_s(\mathbf{a})\right]\,, \tag{3.15}$$

where λ_s are the Lagrange multipliers corresponding to the constraints on the odds α_s being functions $f_s(\cdot)$ of the bets placed $\mathbf{a}$. When we take the derivatives of $\mathcal{G}$ with respect to $a(s)$ we now have extra terms:

$$\frac{\partial \mathcal{G}}{\partial a(s)} = \frac{p(s)\alpha_s}{b + \alpha_s a(s)} + \sum_{s'} \left(\lambda_{s'} \frac{\partial f_{s'}(\mathbf{a})}{\partial a(s)}\right) - \lambda = 0 \tag{3.16}$$

The contribution of the terms in the sum $\sum_{s'}(\cdot)$ comes from the coupling between the bets placed and the payoffs/returns generated. The exact form of $f_s(\mathbf{a})$ depends on the details of the model considered for the game played by the gambler. For an optimal allocation of capital the derivative $\frac{\partial \mathcal{G}}{\partial a(s)} = 0$, meaning that the addition of the first term

with the sum $\sum_{s'}(\cdot)$ again equal the marginal cost λ. Hence, we found that even in the presence of feedback between the bets placed and the payoffs/returns generated it results optimal to invest up to the point for which the net marginal gain expected from our investment (including the contribution due to the feedback) equals the marginal cost of betting extra capital λ.

3.3.2. Optimal allocation principle

In analogy to Kelly's problem, the enzymes in our model can be considered to represent capital that can be invested in the system at different positions. Each position can be interpreted as an asset, generating a certain expected return in terms of reaction flux. The optimization problem consists in determining the optimal fraction of enzymes to be invested at the different locations such as to maximize the steady state reaction flux $J^{\mathcal{P}}$.

Analogously to Kelly's treatment we examine the variation of the Lagrangian, Eq. (3.5), with respect to the resources to be invested in the system, i.e. with respect to the enzyme density,

$$\frac{\delta \mathcal{L}}{\delta e(\mathbf{r})} = \underbrace{k_{\text{cat}} F[\rho(\mathbf{r})] - \lambda_V(\mathbf{r}) k_{\text{cat}} F[\rho(\mathbf{r})]}_{\equiv \frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(\mathbf{r})}\ \text{“marginal return”}} - \lambda_e \ . \tag{3.17}$$

This variation corresponds to the change in the constrained flux $\mathcal{L}$ as enzymes are added at position $\mathbf{r}$. Three different terms contribute to this change. The first is the increase in flux that would be observed upon adding enzymes at position $\mathbf{r}$ if everything else in the system were to remain unaffected, which corresponds mathematically to $\frac{\delta J^{\mathcal{P}}}{\delta e(\mathbf{r})}$. However, changing $e(\mathbf{r})$ also alters the substrate density profile $\rho(\mathbf{r})$, thereby affecting the rate of reactions at all positions. The reaction between enzyme and substrate at a given location causes the substrate to be depleted, this generates a diminishing returns effect for which the addition of extra enzymes at that position is gradually less efficient. Moreover, the substrate motion, due to diffusion and possibly advection, couples the enzyme and substrate profiles at the different positions of the system. The diminishing returns effect and the couplings due to substrate transport are captured by the second term in Eq. (3.17), where the Lagrange multiplier $\lambda_V(\mathbf{r})$ ensures that $\rho(\mathbf{r})$ and $e(\mathbf{r})$ satisfy the constraint of the reaction-diffusion equation. Thus the sum of the first two terms is the total rate of change of the reaction flux as extra enzymes are added at position $\mathbf{r}$. In the following, we refer to this quantity as the "marginal return" on an "investment" of enzymes, and denote it by $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(\mathbf{r})}$. Finally, the third term, $-\lambda_e = -\frac{\partial \mathcal{L}}{\partial E_{\text{tot}}}$, corresponds to a marginal cost of adding extra enzymes into the system. Note that this

term can also be interpreted as the reduction in flux contributions from other positions $\mathbf{r}' \neq \mathbf{r}$ in the optimal configuration, as enzymes are moved from these positions to $\mathbf{r}$ in order to satisfy the constraint of constant total enzyme number.

A key difference between Kelly's original problem and ours is that the payoffs for each of Kelly's events are fixed and independent of the bets placed. Thus, Kelly's original formalism lacks an equivalent of the second term in Eq. (3.17), which captures how an investment of enzymes at one position affects the expected returns at this and all other positions, cf. (3.9). We have seen in the previous section, Sec. 3.3.1 that the Kelly criterion can be generalized to include a feedback between the bets placed and the payoffs/returns, resulting in an expression for the marginal expected growth of capital that has the same form as Eq. (3.17), cf. Eq. (3.16).

As similarly done in Sec. 3.3.1 for the optimal placement of bets, in the optimal enzyme profile $e^*(\mathbf{r})$, we must distinguish between regions where adding enzymes is optimal ($e^*(\mathbf{r})$ finite), and empty regions where adding enzymes is suboptimal ($e^*(\mathbf{r}) = 0$). Wherever $e^*(\mathbf{r}) > 0$, the variation $\frac{\delta \mathcal{L}}{\delta e(\mathbf{r})}\big|_{e^*} = 0$. Importantly, the functional derivative with respect to the surface density $\frac{\delta \mathcal{L}}{\delta e_S(\mathbf{s})}$ has the same form as Eq. (3.17). Thus, the optimal profile follows a 'homogeneous marginal returns criterion' (HMR criterion): it is such that at any position with enzymes, the marginal return on added enzymes equals a constant value, namely $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(\mathbf{r})}\big|_{e^*} = \lambda_e$. By contrast, in empty regions, the marginal flux gain from adding enzymes is less than the associated marginal cost, $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(\mathbf{r})}\big|_{e^*} < \lambda_e$.

As an example of this balance between gain and cost that underlies the optimal enzyme allocation strategy, we used our analytic formalism to solve exactly the case of a linear reaction, $F[\rho] = \rho/K_\mathrm{M}$, with no advection or decay of $\mathcal{S}$ ($|\mathbf{v}| = \sigma = 0$), in a one-dimensional domain with a source of substrate at $x = 0$ ($j(0) = j_0$, $k(0) = 1$, $h(0) = 0$), and absorbing boundary at $x = L$ ($k(L) = 0$, $h(L) = 1$). The resulting optimal enzyme arrangement is (see Appendices A.2, A.3 for details)

$$e_S^*(0) = \begin{cases} E_\mathrm{tot} \\ \sqrt{\frac{E_\mathrm{tot}}{\alpha}} \end{cases} \quad e^*(x) = \begin{cases} 0 & \alpha E_\mathrm{tot} \leq 1 \\ \frac{E_\mathrm{tot}}{L}\Theta\left(\frac{x_0 - x}{L}\right) & \alpha E_\mathrm{tot} > 1 \end{cases} \tag{3.18}$$

where $\Theta(x)$ is the Heaviside step function, $\alpha = Lk_\mathrm{cat}/(K_\mathrm{M}D)$, $x_0 = L\left[1 - (\alpha E_\mathrm{tot})^{-1/2}\right]$. Thus the optimal enzyme profile undergoes a transition from having all enzymes bound at the boundary at $x = 0$ when $\alpha E_\mathrm{tot} < 1$, to an extended profile with enzymes both at the boundary and in the interior of the system when $\alpha E_\mathrm{tot} > 1$. Our analytic calculation confirms the phenomenology of the system that had previous been found numerically [167].

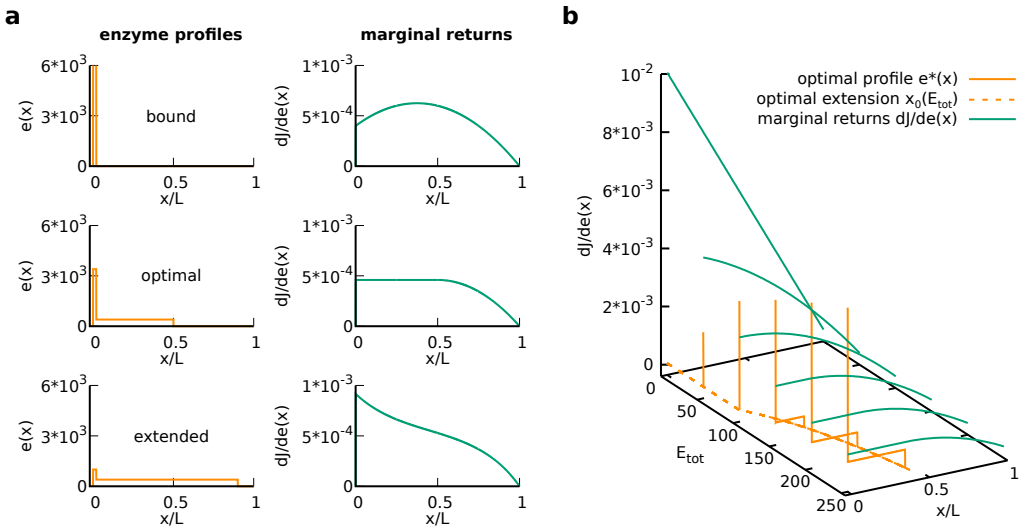


Figure 3.3.: Homogeneous Marginal Returns (HMR) criterion. **a** Different enzyme arrangements (left) and the corresponding marginal returns (right). In all cases $\alpha = 10^{-2}$ and $E_{\text{tot}} = 400$. Top: Bound configuration $e(x) = \frac{E_{\text{tot}}}{L}\delta(x/L)$. The marginal returns is peaked at $x > 0$, indicating that enzymes should be moved away from the source. Middle: Optimal configuration. The marginal return is constant over the domain where enzymes are present, i.e. the HMR criterion is satisfied. Bottom: Extended profile. The marginal return is monotonically decreasing, indicating that enzymes should be moved towards the source at $x = 0$. **b** Optimal enzyme arrangement and the corresponding marginal return landscape for different values of E_{tot}. The dashed line represents the extension of the optimal profile, x_0.

Fig. 3.3a demonstrates how the shape of the marginal returns landscape reflects the behavior of the system. When the enzyme profile is optimal (Fig. 3.3a, middle), we observe a plateau in $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(x)}$ across the entire extent of the enzyme profile, $0 \leq x \leq x_0$. Beyond this region, $x > x_0$, the marginal returns decrease. If the actual enzyme profile deviates from the optimum by having a larger proportion of enzymes located at $x = 0$ (Fig. 3.3a, top), the marginal returns landscape becomes non-monotonic with a peak at a position $x > 0$. The elevated enzyme density in this region depletes the substrate locally, reducing the turnover by each enzyme and diminishing the returns from adding additional enzymes. A higher flux could be achieved by moving some enzymes away from the boundary, thereby allowing each enzyme to operate more efficiently. In contrast, if we consider an overextended enzyme profile (Fig. 3.3a, bottom), the marginal returns landscape has a maximum at $x = 0$. The flux could be increased by moving enzymes from large-x regions, where the substrate density is low due to proximity to the absorbing boundary at $x = L$, towards $x = 0$ where the substrate concentration is high.

Figure 3.3b shows how the marginal returns landscape and the optimal enzyme profile change as E_{tot} is increased. Below the transition, $E_{\text{tot}} \leq \alpha^{-1}$, the marginal returns decrease monotonically with distance from the source. The maximal value of $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(x)}$ occurs only at $x = 0$. Above the transition, $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(x)}$ takes a constant value over the region $0 \leq x \leq x_0$ where enzymes are present. The level of this plateau, λ_e, decreases with increasing E_{tot}.

How relevant is the transition for biological systems? Given that the ratio $k_{\text{cat}}/K_{\text{M}}$ varies by several orders of magnitude: 10^2-$10^{10}(\text{Ms})^{-1}$ [30], the corresponding $\alpha = Lk_{\text{cat}}/(K_{\text{M}}D)$ for a system of $L = 1\mu\text{m}$, where $D = 100\mu\text{m}^2\,\text{s}^{-1}$, is in the range 10^{-7}-10^{1}, where K_{M} is measured in units of length μm^{-1} and where we consider an area of $1\mu\text{m}^2$ underlying the 1-D symmetry of the system. The transition then takes place for values of enzyme numbers E^t_{tot} in the range 10^{-1}-10^{7}, largely overlapping with the range of protein numbers per single cell found in *S. Cerevisiae* yeast cells, which is 1-10^6 [189], suggesting that localization is optimal for only some of the enzymes.

3.3.3. General condition for transitions in the optimal enzyme arrangement

Transitions in the optimal enzyme arrangement of the above type, from a regime in which enzymes are colocalized with the source of $\mathcal{S}$ to a regime where enzymes are distributed within the system, have been observed in a range of reaction-diffusion models [168]. We therefore asked whether there was an underlying principle determining when such transitions occur, and whether it can be generalized, e.g. to systems with more complex geometries lacking symmetry.

Fig. 3.3b suggests that the marginal returns landscape leads to a general condition for localization-delocalization transitions. For E_{tot} below the transition value, the marginal returns landscape is monotonically decreasing, with the position generating the highest returns coinciding with the source of substrate. As the transition is reached and passed, the landscape becomes flat as positions in the vicinity of the source begin to generate the same returns as the source position. This behavior generalizes to higher dimensional systems (see Appendix A.4 for details): Enzymes should be placed only at the source on the surface and not in the interior, as long as

$$\frac{\partial}{\partial \mathbf{n}(\mathbf{s})} \frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e} > 0\,, \tag{3.19}$$

corresponding to a negative slope of the marginal returns landscape from the boundary at $\mathbf{s}$ into the interior of the system, as in Fig. 3.3b for $E_{\text{tot}} < \alpha^{-1}$. By contrast, if Eq. (3.19) evaluated at $\mathbf{s}$ becomes an equality, the positions adjacent to $\mathbf{s}$ generate the same returns as $\mathbf{s}$, so that the optimal enzyme profile features enzymes both at the boundary $\mathbf{s}$ and

in the interior of the system adjacent to $\mathbf{s}$.

The theoretical condition Eq. (3.19) can be turned into an experimentally meaningful condition (see Appendix A.4) by expressing it as a comparison between the local net diffusive flux $j^D(\mathbf{s})$ of $\mathcal{S}$ away from the surface, and the local flux $j^{\mathcal{P}}(\mathbf{s})$ of product formation due to surface enzymes,

$$j^D(\mathbf{s}) \equiv D\frac{\partial \rho}{\partial \mathbf{n}(\mathbf{s})} > k_{\text{cat}} e_S(\mathbf{s}) F[\rho(\mathbf{s})] \equiv j^{\mathcal{P}}(\mathbf{s}) \,. \tag{3.20}$$

If this inequality holds, no enzymes should be placed in the interior of the system in the vicinity of $\mathbf{s}$. For low amounts of enzymes, the optimal strategy is to place enzymes at position $\mathbf{s}$ to counter the diffusive flux $j^D(\mathbf{s})$. This is optimal up to the point at which the diffusive flux $j^D(\mathbf{s})$ equals the reaction flux $j^{\mathcal{P}}(\mathbf{s})$, which defines the transition point of the localization-delocalization transition.

Further combining Eq. (3.20) with the boundary condition Eq. (3.2) yields $j^{\mathcal{P}}(\mathbf{s}) < (j(\mathbf{s}) - \rho(\mathbf{s})\mathbf{v}(\mathbf{s}) \cdot \hat{n}(\mathbf{s}))/2$, which simplifies in the case of no advection ($\mathbf{v} = 0$) to $j^{\mathcal{P}} < j(\mathbf{s})/2$. Thus, in systems with purely diffusive $\mathcal{S}$ transport, a localization-delocalization transition occurs when the reaction flux equals exactly half the influx of $\mathcal{S}$. Note that Eq. (3.20) is general, independent of $\mathbf{v}$ and σ and other details of the system in question. It is required only that the boundary with influx follows Eq. (3.2). Other aspects of the model will affect $\rho(\mathbf{r})$, but not the form of Eq. (3.20), showing that the diffusive motion of $\mathcal{S}$ and the strength of the reaction are the crucial determinants of the transition point. To demonstrate this generality we compared the fluxes $j^D(\mathbf{s})$ and $j^{\mathcal{P}}(\mathbf{s})$ as E_{tot} was varied for a number of systems, including different spatial dimensions as well as systems with and without decay and advection of $\mathcal{S}$. As shown in Fig. 3.4, in all cases both Eq. (3.20) below the transition value of E_{tot} and equality above the transition were satisfied.

The inequality condition Eq. (3.20) is a strictly local condition. In systems with multiple or spatially varying sources of substrate there will in general not be a single global transition but rather a sequence of local transitions at different positions $\mathbf{s}$ at different E_{tot} levels. We therefore investigated a system with two unequal sources of $\mathcal{S}$. In such a system the optimal enzyme profile typically undergoes three transitions (see Appendix A.6, Fig. A.1). For small E_{tot}, enzymes accumulate where the influx of $\mathcal{S}$ is the highest. As E_{tot} is increased, the marginal returns for placing enzymes in the vicinity of the two sources of $\mathcal{S}$ become more similar, until at a threshold value of E_{tot} they become equal. Above this value of E_{tot}, enzymes accumulate at both sources in a configuration with two unequal clusters. Increasing E_{tot} further, a second transition is reached where the optimal enzyme profile begins to extend from the stronger source into the system. Finally, at yet larger E_{tot}, the optimal enzyme profile begins to extend

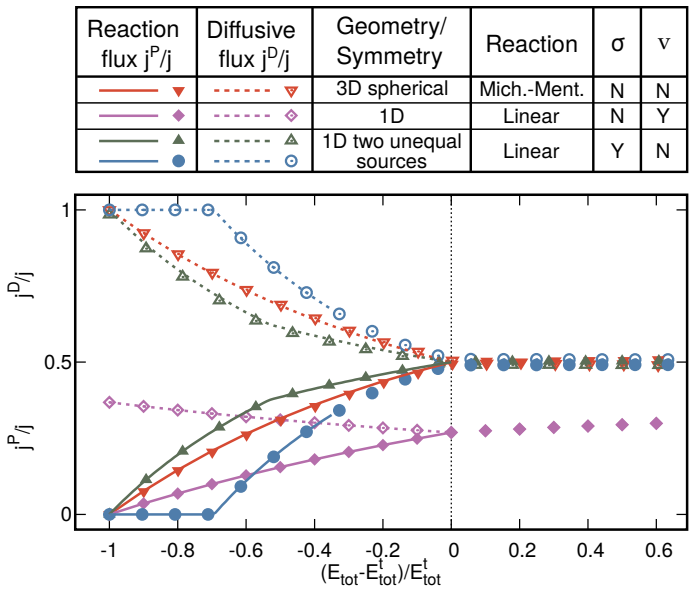


Figure 3.4.: Transition from a fully bound to an extended optimal profile in terms of local reaction and diffusion fluxes. The local reaction flux $j^{\mathcal{P}}(\mathbf{s})$ and net diffusive flux $j^{D}(\mathbf{s})$, as a fraction of $\mathcal{S}$ influx $j(\mathbf{s})$, are plotted for different models against the amount of enzymes in the system E_{tot}, rescaled by the transition amount E^t_{tot}. The various models differ in the physical dimension and/or symmetry, reaction function, loss mechanism and transport of $\mathcal{S}$. The results for each model are plotted with a different color. Lines show analytical solutions of the constrained optimization. Points show results of numerical optimizations computed either by solving numerically an analytically-derived optimization condition (see Appendix A.5), or using the construction algorithm derived in Sec. 3.3.5. The blue and the green lines refer to the same two-source model, the two curves are plotted using the different values of E^t_{tot} for the corresponding source position. See Fig. A.1 for the corresponding plot using the same rescaling of E_{tot}. In all cases, $j^D > j^{\mathcal{P}}$ until they coincide at the transition point, as predicted by Eq. (3.20). When $\mathbf{v} = 0$, $j^{\mathcal{P}} = j^D = j/2$ at and above the transition, whereas, when $\mathbf{v} \neq 0$, the pink solid and dashed lines still coincide but for values different than $j/2$.

also from the weaker source. For the latter two transitions, at which the optimal enzyme arrangement changes from membrane-localized to extended, we confirmed that the inequality Eq. (3.20) held below the transition point and equality above (Fig. 3.4 green and blue).

3.3.4. Geometrical analysis

Our optimization problem can be interpreted geometrically, analogously to a model of investment portfolio optimization introduced by Markowitz [190]. Markowitz considered the problem of minimizing the variance of the returns of a portfolio (due to

the unpredictability of the market) at a constant level of returns, finding that optimal portfolios are points where lines of constant expected returns are tangent to lines of constant variance. To see this, Markovitz plotted the lines of constant variance and constant expected returns in the space of portfolios, i.e. the space determined by the fraction of resources invested in different securities. Here we consider maximizing flux at a constant level of enzymes. We can therefore plot lines of constant reaction flux and lines of constant level of enzymes in the space of enzyme configurations, i.e. the space determined by the amount of enzymes invested at different positions.

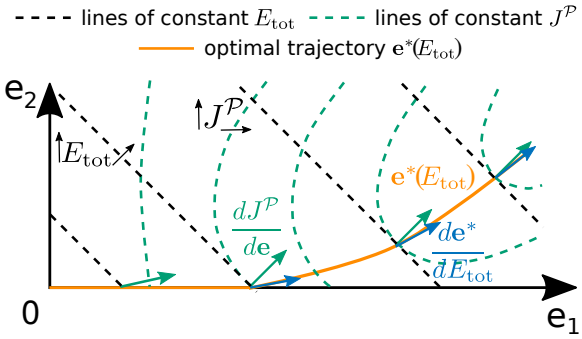


Figure 3.5.: Geometrical analysis of the optimal enzyme configuration in a system with two sites. Dashed black lines show lines of constant $E_{\mathrm{tot}} = e_1 + e_2$ and dashed green lines show lines of constant reaction flux $J^{\mathcal{P}}$. The optimal configuration $\mathbf{e}^* = (e_1^*, e_2^*)$ for each value of E_{tot} forms an optimal trajectory $\mathbf{e}^*(E_{\mathrm{tot}})$ (orange). Green arrows show the marginal returns vector $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}}$. The tangent vector $\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\mathrm{tot}}}$ (blue arrows) shows how enzymes should be added at each point.

Fig. 3.5 shows this schematically for a system where enzymes can be partitioned between two positions. The dashed black lines are lines of constant E_{tot} and the dashed green curves are lines of constant $J^{\mathcal{P}}$. In the Markowitz model, variability in returns is taken to be independent of the investment strategy, leading to lines of constant variance that are ellipses in the space of investment portfolios. For our model, lines of constant flux do not take such a simple form, since their shape is determined by the reaction-diffusion dynamics of the substrate that couples the productivity of enzyme investments at all positions. The constrained optimization problem corresponds to maximizing $J^{\mathcal{P}}$ along a given E_{tot} line, which results in the solid orange line. For the case illustrated in Fig. 3.5, above a threshold value of E_{tot} in the region where the optimal configuration has both $e_1, e_2 > 0$, the principle of equal returns means that the marginal return vectors in the optimal configurations (Fig. 3.5, green arrows) point in the $(1, 1)$ direction, which is normal to the lines of constant E_{tot} values. Since the marginal returns vector is itself the gradient of the reaction flux, it is by definition normal to lines of constant flux, implying that these lines are tangent to the lines of constant E_{tot}.

3.3.5. Construction Algorithm

The above geometrical analysis leads us to an additive construction algorithm, which generates the trajectory of optimal enzyme configurations by determining its tangent at each point. We consider a system in which space is discretized into N sites, denoting by $e_i = e(\mathbf{r}_i)$ the density of enzymes at position $\mathbf{r}_i$ and by $\rho_i = \rho(\mathbf{r}_i)$ the corresponding density of $\mathcal{S}$. For a given total enzyme level E_{tot}, there will be an optimal enzyme configuration $\mathbf{e}^* = (e_1^*, e_2^*, \ldots, e_N^*)$ that maximizes $J^{\mathcal{P}}$ on the $(N-1)$-dimensional manifold defined by $E_{\text{tot}} = \sum_{i=1}^{N} e_i$. As in Fig. 3.5, the optimal configurations at different E_{tot} values trace out a trajectory $\mathbf{e}^*(E_{\text{tot}})$ in $\mathbf{e}$ space. We seek a procedure to construct $\mathbf{e}^*(E_{\text{tot}})$. We begin with no enzymes and iteratively add enzymes until the target E_{tot} value is reached. According to the optimization principle of equal marginal returns, the largest components of $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}}$ define the sites at which enzymes should be added in the next step. Addition may initially be restricted to a single site, with further "construction sites" being introduced gradually as different E_{tot} thresholds are crossed.

Since enzymes should be added at all positions where the marginal returns take on their maximal value, we must determine how to partition new enzymes between these positions. This partitioning corresponds to the tangent vector $\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\text{tot}}}$ of the optimal trajectory, which we determine by considering how the marginal returns change as enzymes are added (see Appendix A.7),

$$\frac{\mathrm{d}}{\mathrm{d}E_{\text{tot}}}\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}} = \tilde{\mathbf{H}}\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\text{tot}}} \,. \tag{3.21}$$

Here, derivatives $\frac{\mathrm{d}}{\mathrm{d}E_{\text{tot}}}$ are taken tangent to the line of optimal enzyme configurations and $\tilde{\mathbf{H}}$ is the $N \times N$ Hessian matrix of $J^{\mathcal{P}}$ in the space of enzyme densities, $H_{ij} = \frac{d^2 J^{\mathcal{P}}}{de_i de_j}$. In the subspace of only those sites where enzymes should be added (sites at which $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i} = \lambda_e$), the left hand side of Eq. (3.21) becomes

$$\frac{\mathrm{d}}{\mathrm{d}E_{\text{tot}}}\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}_{(n)}} = \frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\text{tot}}}\mathbf{1}_{(n)} \,, \tag{3.22}$$

where with $\mathbf{1}$ we refer to a vector of 1s and the subscript (n) indicates that we are considering only the n-dimensional subspace of sites with equal returns, yielding

$$\tilde{\mathbf{H}}_{(n)}\frac{\mathrm{d}\mathbf{e}^*_{(n)}}{\mathrm{d}E_{\text{tot}}} = \frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\text{tot}}}\mathbf{1}_{(n)} \,. \tag{3.23}$$

This n-dimensional linear system can be solved for the tangent $\frac{\mathrm{d}\mathbf{e}^*_{(n)}}{\mathrm{d}E_{\text{tot}}}$ to the optimal

trajectory. In general we do not know $\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}$ *a priori*. However, this scalar prefactor affects only the length, and not the direction, of $\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\mathrm{tot}}}$. Hence it suffices to treat $\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}$ as an arbitrary constant, and subsequently to rescale the resulting $\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\mathrm{tot}}}$ to unit length once its direction is known. Taken together, we obtain the following additive construction algorithm for the optimal enzyme arrangement:

1. Begin with $\mathbf{e} = 0$.
2. For each site $\mathbf{r}_i$, evaluate the marginal returns, $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i}$, for adding enzymes at that position.
3. Identify the subset of positions at which the returns are equal (within numerical tolerance) to the maximal value.
4. For this subset of positions, evaluate the Hessian matrix of the flux with respect to the enzyme densities, $\tilde{\mathbf{H}}_{(n)}$.
5. Solve $\tilde{\mathbf{H}}_{(n)}\delta\mathbf{e}^* = \mathbf{1}$ for $\delta\mathbf{e}^*$, and normalize such that the total change in the amount of enzymes $\sum_i \delta e_i^* = \delta E_{\mathrm{tot}}$.
6. Update the optimal enzyme arrangement, $\mathbf{e}^* \to \mathbf{e}^* + \delta\mathbf{e}^*$.
7. Repeat from step 2 until E_{tot} reaches the target amount of total enzymes.

Fig. 3.6 illustrates the algorithm for a two-dimensional system featuring multiple sources of substrate (dashed rectangles) with uniform influx on their surfaces and an absorbing outer boundary. Fig. 3.6a compares two optimal enzyme configurations with different amount of enzymes. For low E_{tot}, enzymes localize non-uniformly at the boundaries of the sources. For higher E_{tot}, the optimal enzyme profile extends into the interior of the system. However, the highest densities of enzymes are found in the regions between the sources and the outer boundary of the system, where the gradient of substrate concentration is steepest. A significant fraction of enzymes are thus devoted to substrates that, by the nature of where they are produced, are likely to rapidly diffuse out of the system.

Remarkably, the sites with the highest enzyme densities at small E_{tot} do not necessarily retain the highest densities as E_{tot} is increased. Fig. 3.6b compares the enzyme densities as a function of E_{tot} at three positions (marked by colored squares in Fig. 3.6a). The blue location is the first at which enzymes are added. However, although enzymes are added at the green and red position positions only later, the densities here become higher, see Inset of Fig. 3.6b. Fig. 3.6c shows the respective ratios of the returns of the red and green position to the returns from the blue position. Enzymes are added at the green and red position only when the returns are equal and the ratios are one. At low

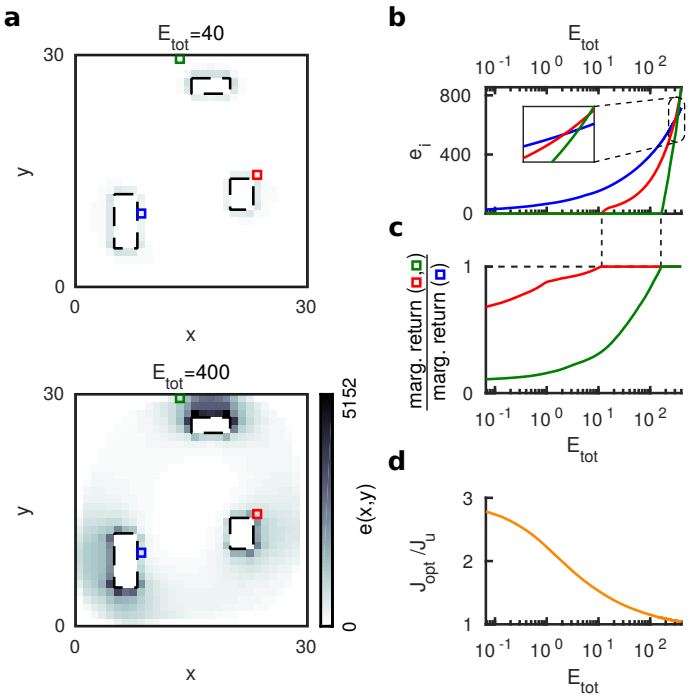


Figure 3.6.: Construction Algorithm to determine the optimal enzyme arrangement in a 2-dimensional system with multiple sources of substrate. The substrate enters the system from the sources (dashed rectangles) and it is absorbed at the external boundary. The optimal enzyme arrangement is determined by using the construction algorithm for which enzymes are iteratively added, starting from a system with no enzymes. **a** Snapshots of the optimal enzyme distribution at two E_{tot} levels. **b** Optimal enzyme densities at the positions indicated by the blue, red and green squares in **a**. Inset: Although the blue position is the first one at which enzymes are added, for high enough E_{tot} values the red and green positions get higher optimal enzyme densities. **c** Ratio of the marginal returns at the red and green position over the marginal return of the blue position, as indicated by colored squares in **a**. **d** The optimized reaction flux relative to that produced by uniformly distributed enzymes, as a function of E_{tot}.

enzyme amounts, the reaction flux of the optimal configuration is about 3-fold higher than for the uniform configuration, whereas this advantage gradually decreases for higher amounts of enzyme (Fig. 3.6d). This is consistent with previous studies showing that in the limit of infinite amount of enzymes the uniform configuration becomes the optimal one [167, 168].

3.4. Methods for the computation of the marginal returns vector and the Hessian for a discrete system

In this section we explicitly derive the expressions for the marginal returns vector and the Hessian used in the simulations of the systems shown in Fig. 3.4 and Fig. 3.6.

3.4.1. Linear reactions

For a discrete reaction-diffusion system with linear reactions, following the dynamics Eq. (3.1)

$$\partial_t \rho(\mathbf{r},t) = D\nabla^2\rho(\mathbf{r},t) - \nabla\cdot[\rho(\mathbf{r},t)\mathbf{v}(\mathbf{r})] - \sigma\rho(\mathbf{r},t) - k_{\text{cat}}e(\mathbf{r})F[\rho(\mathbf{r},t)]\,, \tag{3.24}$$

we can define a discrete diffusion operator $\tilde{\mathbf{D}}$ and a discrete advection operator $\tilde{\mathbf{V}}$. By indicating with $\boldsymbol{\rho}$ the substrate density and $\mathbf{A}$ the substrate influx vector, the steady-state equation for a linear reaction can be written in matrix form as

$$(\tilde{\mathbf{D}} + \tilde{\mathbf{V}})\boldsymbol{\rho} - \alpha\mathbf{e}\odot\boldsymbol{\rho} - \sigma\boldsymbol{\rho} = \mathbf{A}\,, \tag{3.25}$$

with $\odot$ denoting the element-wise (Hadamard) product. The reaction flux $J^{\mathcal{P}}$ is given by the scalar product $J^{\mathcal{P}} = \alpha\,\mathbf{e}\cdot\boldsymbol{\rho}$. Thus $\mathbf{e}$ enters $J^{\mathcal{P}}$ both directly and through $\boldsymbol{\rho}$, via the solution to Eq. 3.25. These multiple dependencies of $J^{\mathcal{P}}$ on $\mathbf{e}$ can be summarized by the dependency graph shown in Fig. 3.7, where the matrix $\tilde{\mathbf{M}} = \tilde{\mathbf{D}} + \tilde{\mathbf{V}} - \alpha\,\mathrm{diag}(\mathbf{e}) - \sigma\,\mathrm{diag}(\mathbf{1})$ is the operator that applied to $\boldsymbol{\rho}$ gives the substrate source vector $\tilde{\mathbf{M}}\boldsymbol{\rho} = \mathbf{A}$, $\mathrm{diag}(\mathbf{e})$ is the diagonal matrix with diagonal elements given by $\mathbf{e}$ and $\mathrm{diag}(\mathbf{1})$ is the identity matrix.

The marginal returns vector $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}}$ can be evaluated by back-propagating the derivative of the reaction flux with respect to the enzyme vector through the dependency graph. We find that

$$\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}} = \frac{\partial J^{\mathcal{P}}}{\partial\mathbf{e}} + \left(\frac{\mathrm{d}\boldsymbol{\rho}}{\mathrm{d}\mathbf{e}}\right)^t \frac{\partial J^{\mathcal{P}}}{\partial\boldsymbol{\rho}} \tag{3.26}$$

$$= \frac{\partial J^{\mathcal{P}}}{\partial\mathbf{e}} + \left(\frac{\mathrm{d}(\tilde{\mathbf{M}}^{-1}\mathbf{A})}{\mathrm{d}\tilde{\mathbf{M}}}\frac{\mathrm{d}\tilde{\mathbf{M}}}{\mathrm{d}\mathbf{e}}\right)^t \frac{\partial J^{\mathcal{P}}}{\partial\boldsymbol{\rho}}\,, \tag{3.27}$$

where $\frac{\mathrm{d}\boldsymbol{\rho}}{\mathrm{d}\mathbf{e}}$ is a matrix with $(i,j)^{\text{th}}$ element given by $\frac{\mathrm{d}\rho_i}{\mathrm{d}e_j}$, and ^{t} denotes the transpose. Each of the derivatives appearing in Eq. 3.27 has a simple form, which we can substitute to

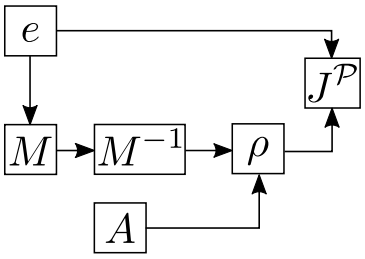


Figure 3.7.: Dependency graph for a discrete reaction-diffusion system with linear reaction kinetics. The reaction flux is given by the scalar product of the enzyme and the substrate vector $J^{\mathcal{P}} = \alpha\, \mathbf{e} \cdot \boldsymbol{\rho}$. Moreover $\boldsymbol{\rho}$ and $\mathbf{e}$ are coupled via the reaction-diffusion equation Eq. (3.25). For a given enzyme arrangement, there is an operator $\tilde{\mathbf{M}}$ that, applied to the substrate profile, gives the source vector $\tilde{\mathbf{M}}\boldsymbol{\rho} = \mathbf{A}$. The substrate profile corresponding to a certain enzyme arrangement can then be found via $\boldsymbol{\rho} = \tilde{\mathbf{M}}^{-1}\mathbf{A}$.

obtain the marginal returns vector,

$$\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}} = \alpha\boldsymbol{\rho} \odot \left[\mathbf{1} + \alpha(\tilde{\mathbf{M}}^{-1})^{t}\mathbf{e}\right] , \tag{3.28}$$

where $\mathbf{1}$ is a vector of ones.

By taking a second derivative with respect to $\mathbf{e}$, and again by back-propagating the derivatives, we can determine the Hessian,

$$\tilde{\mathbf{H}} = \alpha^2 \Big\{ \left[\tilde{\mathbf{M}}^{-1} \odot \Big(\{\mathbf{1} + \alpha(\tilde{\mathbf{M}}^{-1})^{t}\mathbf{e}\} \otimes \boldsymbol{\rho}\Big)\right] + \left[\tilde{\mathbf{M}}^{-1} \odot \Big(\{\mathbf{1} + \alpha(\tilde{\mathbf{M}}^{-1})^{t}\mathbf{e}\} \otimes \boldsymbol{\rho}\Big)\right]^{t} \Big\} , \tag{3.29}$$

where $\otimes$ indicates the outer product. The Hessian Eq. (3.29) is then used in step (4) and (5) of the construction algorithm, presented in Sec. 3.3.5, to derive the optimal way of allocating the extra enzymes added into the system.

3.4.2. Non-linear reactions

For non-linear reaction kinetics the reaction-diffusion equation becomes

$$(\tilde{\mathbf{D}} + \tilde{\mathbf{V}})\boldsymbol{\rho} - \alpha\mathbf{e} \odot \mathbf{F} - \sigma\boldsymbol{\rho} = \mathbf{A} , \tag{3.30}$$

where $\mathbf{F}$ is the vector with $F_i = F[\rho_i]$. Since $F[\rho]$ is non-linear, it is no longer possible to write the reaction-diffusion equation as a linear system that can be solved for $\boldsymbol{\rho}$. As a result the dependency graph includes a loop, making it inconvenient to back-propagate derivatives. Instead, taking derivatives of the reaction-diffusion equation with respect

to $\mathbf{e}$, we can directly solve for

$$\frac{\mathrm{d}\boldsymbol{\rho}}{\mathrm{d}\mathbf{e}} = \alpha \tilde{\mathbf{N}}^{-1} \operatorname{diag}(\mathbf{F}), \tag{3.31}$$

where $\tilde{\mathbf{N}} = \tilde{\mathbf{D}} + \tilde{\mathbf{V}} - \alpha \operatorname{diag}(\mathbf{e} \odot \mathbf{F}') - \sigma \operatorname{diag}(\mathbf{1})$ and $F_i' = F'[\rho_i]$. We then take derivatives of $J^{\mathcal{P}} = \alpha \mathbf{e} \cdot \mathbf{F}$ with respect to $\mathbf{e}$, to obtain the marginal returns vector,

$$\begin{aligned} \frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}} &= \alpha \left[\mathbf{F} + \left(\frac{\mathrm{d}\boldsymbol{\rho}}{\mathrm{d}\mathbf{e}} \right)^t \{ \mathbf{e} \odot \mathbf{F}' \} \right] \\ &= \alpha \mathbf{F} \odot \left[\mathbf{1} + \alpha (\tilde{\mathbf{N}}^{-1})^t \{ \mathbf{e} \odot \mathbf{F}' \} \right] . \end{aligned} \tag{3.32}$$

By taking a further derivative of the returns vector with respect to $\mathbf{e}$ we find the components of the Hessian,

$$\begin{aligned} \tilde{\mathbf{H}}_{j,k} = \alpha^2 \Big\{ & \tilde{\mathbf{N}}^{-1}_{j,k} \left[\mathbf{F} \otimes \left(\mathbf{F}' + \alpha (\tilde{\mathbf{N}}^{-1})^t \{ \mathbf{e} \odot \mathbf{F}'^2 \} \right) \right]_{k,j} \\ & + \tilde{\mathbf{N}}^{-1}_{k,j} \left[\mathbf{F} \otimes \left(\mathbf{F}' + \alpha (\tilde{\mathbf{N}}^{-1})^t \{ \mathbf{e} \odot \mathbf{F}'^2 \} \right) \right]_{j,k} \Big\} \\ & + \alpha^3 F_j F_k \sum_i \tilde{\mathbf{N}}^{-1}_{i,j} e_i F_i'' \tilde{\mathbf{N}}^{-1}_{i,k} \left[1 + \alpha \left[(\tilde{\mathbf{N}}^{-1})^t \{ \mathbf{e} \odot \mathbf{F}' \} \right]_i \right], \end{aligned} \tag{3.33}$$

which simplifies to Eq. 3.29 for linear reactions with $F'' = 0$.

3.5. Discussion

We provided a solution for the problem of optimally allocating enzymes in space, to maximize the reaction flux of an enzymatic reaction. Our solution is valid for a wide class of reaction diffusion systems with potentially complex geometries (Fig. 3.2). This class of systems is characterized by the fact that the substrate enters the system via internal or external boundaries, is transported via diffusion and possibly advection, and can be lost by leakage or competing reactions. The solution is based on the concept of a 'marginal returns landscape' (Fig. 3.3), defined as the derivative $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(\mathbf{r})}$ of the reaction flux with respect to the local enzyme density (Eq. (3.17)). The optimal arrangement is such that the marginal returns are spatially homogeneous over all positions with enzymes. This Homogeneous Marginal Returns (HMR) criterion can be interpreted as a generalization of the widely used Kelly criterion for optimal betting [166] to a scenario where each bet globally feeds back onto all returns (cf. Sec. 3.3.1). In the enzyme allocation problem, this feedback has two causes. First, enzymes locally deplete their

substrate, causing a local diminishing returns effect. Second, the moving substrate couples the reaction flux generated at different positions. The definition of marginal return (Eq. (3.17)) captures these effects by incorporating the flux increase caused by just adding enzymes, as well as the reaction-diffusion feedback between enzymes and substrate.

The HMR criterion leads to a general local condition for the occurrence of a transition in the optimal enzyme arrangement from a fully bound configuration into a more extended profile where enzymes are also located in the interior of the system next to the substrate source (Eqs. (3.19), (3.20)). Such transitions had previously been observed numerically for systems with simple geometries and one source of substrate [167, 168]. The condition derived here applies locally in systems with complex geometries and with multiple sources of substrate. It compares the local diffusive flux $j^D(\mathbf{s})$ to the local reaction flux $j^{\mathcal{P}}(\mathbf{s})$ at a source position $\mathbf{s}$: Binding enzymes at the source is optimal as long as $j^D(\mathbf{s}) > j^{\mathcal{P}}(\mathbf{s})$, since the reaction counters the diffusive flux. In contrast, above the transition point, the enzymes are optimally distributed also in the interior of the system and the two fluxes are kept equal (Fig. 3.4). Notably, this condition is not affected by any advective transport of the substrate. The reason ultimately is that the advective transport is deterministic, whereas the HMR principle is fundamentally a bet-hedging strategy to deal with the probabilistic nature of diffusive transport.

The same way the Kelly criterion provides rules for optimally placing bets or optimally constructing portfolios, the HMR criterion allows to derive an algorithm for constructing optimal enzyme arrangements (Figs. 3.5, 3.6). The algorithm is obtained by analyzing how the marginal returns change at every position as the amount of enzymes in the system is increased (Eqs. (3.21), (3.23)). It constitutes a deterministic construction principle, as opposed to the stochastic optimization algorithms that were previously used [167, 168].

While we have presented here the simplest general form of this algorithm, various modifications can be made to improve its efficiency in special cases. In particular, using the system geometry as a guide it is often possible to limit the sites at which the marginal returns need to be evaluated in each update step. For example, at low E_{tot} values the positions with the highest returns coincide with sources of $\mathcal{S}$. Generally it suffices to compute the marginal returns only at source positions, sites where enzymes are already present, and connected lattice sites. Additionally, larger values of δE_{tot} may be enabled by more sophisticated update schemes for $\mathbf{e}^*$, such as a Runge-Kutta-type method [191, 192]. The value of δE_{tot} could also be made adaptive to achieve a good trade-off between accuracy and speed.

We explicitly considered absorbing and reflective boundaries, allowing for non uniform and discontinuous influx profiles. However, our results can be extended to systems with permeable boundaries through which substrate molecules are allowed

to flow, corresponding to $k(\mathbf{s}) = 1$ and $h(\mathbf{s}) = p(\mathbf{s}) > 0$ in Eq. (3.3), where $p(\mathbf{s})$ is the permeability. For such a system, the HMR principle still holds. Similarly, the condition of Eq. (3.19) will still determine transitions in the optimal enzyme profile. However, expressing this condition in term of substrate fluxes will lead to a more complicated expression than Eq. (3.20). We have seen that a permeable boundary typically increases the transition threshold E^t_{tot}. For example, in the one-dimensional system in Fig. 3.3b, a non-zero permeability at the origin shifts the threshold value of E_{tot} from α^{-1} to $\alpha^{-1}(1 + p(0)L/D)$.

Our continuous reaction-diffusion model neglects the finite size of enzyme molecules that limits the attainable enzyme density. Imposing a maximal density condition $e(\mathbf{r}) \leq e_{\max}$ at each point [59] could be incorporated into our analytical framework as an additional constraint in the Lagrangian. This would introduce an effective position-dependent cost in Eq. (3.17), resulting in a marginal return landscape that is a function of position. Nevertheless, the construction algorithm presented here could still be employed if the condition that enzymes cannot be added at positions where $e_i = e_{\max}$ is imposed. Eq. (3.23) would then only hold at positions with sub-maximal enzyme densities.

The construction algorithm presented in this chapter could be used for the design of optimal synthetic bioreactors. In Sec. 2.5, we have seen how synthetic strategies can be used to spatially arrange enzymes. For instance, the amount of enzymes bound to the external membrane or to the membranes of internal organelles can be regulated by fusing enzymes to transmembrane proteins [121, 120] or by using protein scaffolds [124, 126], see Fig. 2.3. We have also discussed how the arrangement of enzymes can be regulated in the interior of membrane-bound systems and synthetic devices, see Fig. 2.4. For example, in the interior of cells, enzyme positioning can be partially controlled with RNA and DNA assemblies [127, 130], synthetic protein scaffolds [128], or fusion proteins [59]. In synthetic devices, the arrangement of enzymes can be regulated with nanometer precision via single-molecule cut-and-paste surface assembly [193, 194], or by arranging enzymes on programmed DNA scaffolds [61, 62, 63, 134]. The experimental use of our construction algorithm to design optimal synthetic bioreactors seem to be within technological reach. For example, it could be accomplished by combining techniques regulating the amount of enzymes bound to membranes and at the same time their positions in the interior of the system.

4. Cross-diffusion induced patterns for a single-step enzymatic reaction*

While in Ch. 2 we have studied how enzymes can be statically arranged in space to maximize the reaction flux of an enzymatic reaction, in this chapter we focus on how the motion of enzymes can be coupled to the substrate and the product concentrations and how these couplings can lead to spontaneous pattern formation. In Sec. 2.6.1, we have discussed the results of recent experiments showing that several different enzymes display an apparent diffusion coefficient that increases with substrate concentration. Moreover, in Sec. 2.6.2, we have seen how their motion becomes directed in the presence of substrate gradients. Whether this motion is directed upstream or downstream the substrate gradient seems to depend on the experimental conditions and several competing models have been proposed to explain the observed phenomena. The experiments and the models study the response of enzyme motion to fixed substrate gradients, they do not consider the feedback from enzyme transport onto the substrate profile. In this chapter, we show that such feedback can generate spontaneous spatial patterns in enzyme, substrate and product distributions, with just a single-step catalytic reaction (see graphical abstract Fig. 4.1). However, patterns form only for a subclass of transport models. For these models, nonspecific repulsive interactions between the enzyme and the substrate, or attractive interactions between the enzyme and the product cause the enzyme to accumulate in regions of low substrate concentration. Reactions then amplify local substrate and product fluctuations, steepening the substrate and product gradients, hence causing enzymes to further accumulate where substrate is low and product is high. Experimental analysis of this spontaneous pattern formation process could discriminate between the different transport models proposed so far.

4.1. Introduction

Recent experiments found that at least eight different enzymes present a higher diffusion coefficient when the concentration of the corresponding substrate in solution

*The chapter is adapted from the article: "Cross-diffusion induced patterns for a single-step enzymatic reaction", by G. Giunta, H. Seyed-Allaei, U. Gerland, published in *Communications Physics* 3, 167 (2020), under CC BY 4.0 License. See Ref. [195].

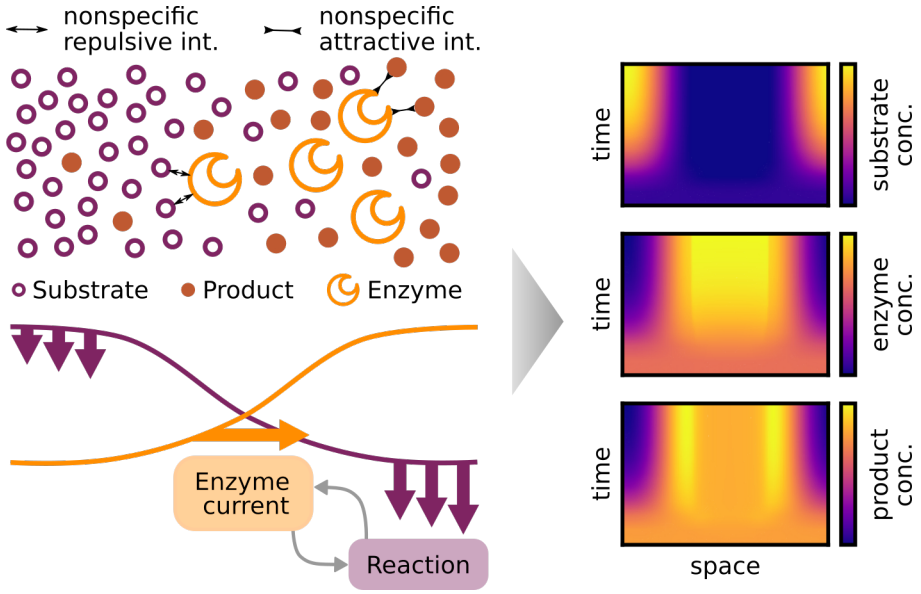


Figure 4.1.: Graphical abstract: Cross-diffusion induced patterns for a single-step enzymatic reaction. Repulsive nonspecific interactions between substrate and enzyme, or attractive nonspecific interactions between product and enzyme cause enzyme molecules to accumulate in regions of low substrate concentrations. Here the enzyme reacts with the substrate. The substrate is further depleted and more product is formed at these locations as compared to locations with higher amount of substrate. Hence the reaction causes the substrate and product gradients to get steeper. This in turn drives further accumulation of enzymes in regions of low substrate. This feedback between reaction and enzyme accumulation amplifies small fluctuations in the concentration profiles and drives spontaneous pattern formation.

is increased [136, 138, 140, 141, 142, 91, 137]. The increases were detected to be in the range of 24-80% relative to the diffusion coefficients in absence of substrate. Most of these measurements were performed by using Fluorescence Correlation Spectroscopy (FCS). Although FCS has been shown to introduce possible artifacts [143, 144], recent findings using other techniques validated the phenomenon [146, 147], which is often referred to as "enhanced diffusion" of enzymes. As we have seen in Sec. 2.6.1, the underlying mechanism is still under debate [155]. A question is whether the catalytic step of the reaction plays a role or not. On the one hand, enhanced diffusion of some enzymes has been shown to correlate with the degree of exothermicity of the enzymatic reaction [140]. Moreover, super resolution experiments suggest that the energy released from catalysis could be fueling a ballistic component of enzyme motion [142, 151]. On the other hand, some experiments indicate that enzymes present enhanced diffusivity upon binding-unbinding of inhibitors, suggesting that catalysis might play a less important role [136, 141, 152]. What contributes to enhanced diffusion could also be

dependent on the enzyme considered. For example, according to the latest experiments performed with the enzyme urease, it seems that both the binding and the catalysis step of the enzymatic reaction contribute to the phenomenon [151, 146].

In Sec. 2.6.2, we have discussed the results of recent experiments showing that enzymes also perform directed motion in presence of substrate gradients. However, we have seen that whether enzymes motion is directed upstream or downstream the gradients is still unclear. Some experiments suggest that enzymes drift downstream substrate gradients, performing "antichemotaxis" [142, 151]. Others suggest that enzymes move upstream substrate gradients, performing "chemotaxis" [91, 152, 137]. Antichemotaxis can be explained just by considering the enhanced diffusivity [142, 196]. Enzymes have a higher diffusion coefficient in regions with high substrate concentration and accumulate in regions with low substrate, where they have a lower diffusion coefficient. Chemotaxis instead cannot be generated by enhanced diffusion alone. A possible cause for enzyme chemotaxis is cross-diffusion [91]. Cross-diffusion describes the response of the enzyme to forces generated by substrate gradients. Mathematically, it corresponds to an off-diagonal element in the diffusion matrix describing the combined motion of enzyme and substrate. It has been suggested that cross-diffusion can be due to specific interactions (ligand binding) between enzyme and substrate [152, 159, 91] or due to nonspecific interactions (e.g. steric, electrostatic, van der Waals) [162]. Specific interactions refer to the local, attractive, short-ranged interactions between the substrate and the enzyme binding site, whereas nonspecific interactions refer to forces to which substrate molecules are subject to once they are in proximity of the enzyme surface [160]. Specific interactions only lead to chemotaxis, while nonspecific interactions can cause enzymes to move both up- or downstream the substrate gradient, depending on whether the interactions are attractive or repulsive, respectively. The model including nonspecific interactions [162] can be considered as a mathematical generalization of other existing models [142, 159, 91].

While the existing models and experiments study how enzymes move in pre-imposed substrate gradients, they do not consider the feedback between enzyme motion and the effect that reaction has on the substrate gradient. Here, we analyze the effects of this feedback starting from the most general transport model [162]. We show that spatial patterns can emerge in initially homogeneous systems if nonspecific interactions contribute to the accumulation of the enzyme in regions where the concentration of substrate is low. Enzymes accumulating in these regions further deplete the substrate, causing the substrate gradient to become steeper, hence further increasing the accumulation of the enzyme. We obtain a set of conditions for the parameter range in which patterns form. We see that patterns arising from initial homogeneous concentrations can emerge only for nonspecific interactions driving the enzyme away from the substrate, hence only for the model proposed by Agudo-Canalejo et al. [162], but not for the models proposed by

Zhao et al. [91], Mohajerani et al. [152] and Jee et al. [142], suggesting that the analysis of spontaneous pattern formation experiments can be used to discriminate between the different proposed models. Our findings imply that patterns can arise for a single-step enzymatic reaction even in the absence of autocatalytic activity or allosteric regulation. This is surprising given that the formation of conventional Turing patterns [197] with such simple reaction schemes requires at least a reaction network of three states, where forward and backward reactions are catalyzed by two different enzymes [198].

4.2. Model

Our starting model system is depicted in Fig. 4.2. We consider a single-step enzymatic reaction in a narrow reaction chamber of length L connected via a permeable membrane to a large substrate reservoir. We assume only substrate and product molecules can diffuse through the permeable membrane, while enzymes are confined to the reaction chamber. This is justified because enzymes are typically larger than their substrates and products. We indicate with γ_s and γ_p the permeation rates of substrate and product respectively, i.e. the rate at which substrate and product molecules diffuse through the membrane. These rates are obtained by dividing the substrate and product permeabilities by the reaction chamber thickness. The reservoir has a fixed concentration of substrate s_R and no products. For the reaction occurring in the bulk of this effectively 1D system we assume a Michaelis-Menten scheme (Fig. 4.2, inset). The substrate $\mathcal{S}$ binds to the enzyme $\mathcal{E}$ with rate k_on, forming a complex, and it can unbind with rate k_off. The catalytic step of the reaction has a rate constant k_cat and catalysis is irreversible. This leads to a turnover rate per enzyme $k_\mathrm{cat}F(s) := k_\mathrm{cat}s(K_\mathrm{M}+s)^{-1}$, where s is the substrate concentration and $K_\mathrm{M} = (k_\mathrm{off}+k_\mathrm{cat})k_\mathrm{on}^{-1}$. Following Agudo-Canalejo et al. [162] (cf. Sec 2.6.2) we assume that the diffusion coefficient of the enzyme depends on whether the enzyme is free, D_f, or it is in its complexed form, D_c. In Sec. 1.2.2, we have seen how the binding of substrate can induce changes in the enzyme conformation (see Fig. 1.3). Typically, binding induces a reduction of the hydrodynamic radius and of the conformational fluctuations of the enzyme, meaning that the enzyme becomes more compact and stiffer [141, 153, 154]. This implies that $D_\mathrm{c} > D_\mathrm{f}$, which is consistent with the experimentally observed trend. We also assume that short-range nonspecific interactions can either cause the enzyme to move towards or away from the substrate, effectively generating a phoretic drift velocity that can also be interpreted as cross-diffusion [162]. We denote these interactions via pairwise potentials ϕ_fs, ϕ_cs depending on whether the enzyme is free or in the complexed form, respectively, as depicted in Fig. 4.2 (inset).

Under the assumptions that (a) the enzyme is very dilute and (b) the system is locally

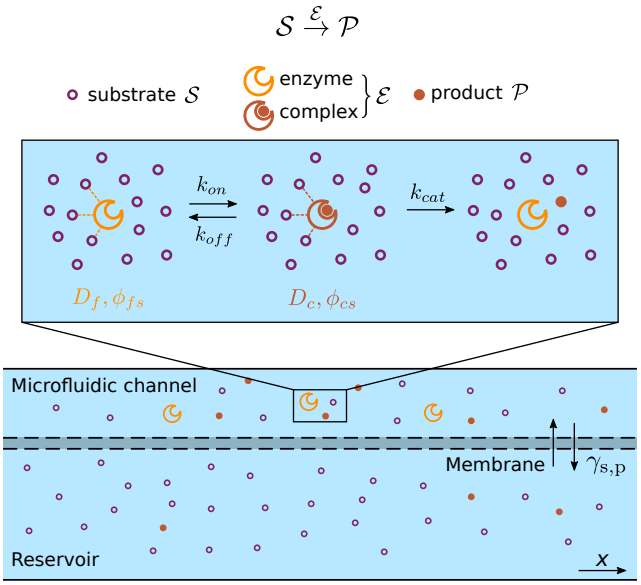


Figure 4.2.: Model scenario considered for cross-diffusion induced patterns for a single-step enzymatic reaction. A single-step enzymatic reaction takes place in a narrow reaction chamber that extends in the x-direction and is coupled to a reservoir through a permeable membrane. The membrane allows for the exchange of only substrates and products with permeation rates γ_s, γ_p respectively, but no exchange of enzymes. Inset: Reaction scheme. The substrate $\mathcal{S}$ is converted into a product $\mathcal{P}$ by an enzyme $\mathcal{E}$. The reaction follows a Michaelis-Menten scheme where substrate can bind to the free enzyme with rate k_{on}, forming a complex, and unbind with rate k_{off}. The catalytic step of the reaction has a rate k_{cat}. We assume that the enzyme has a diffusion constant D_f or D_c depending on whether the enzyme is free or in its complexed form. Nonspecific pairwise interactions between enzyme and substrate ϕ_{fs}, ϕ_{cs} can depend on the enzyme form. How nonspecific interactions between enzyme and product affect the results is discussed in Sec.4.3.4.

in chemical equilibrium (i.e. the timescales of diffusion and cross-diffusion are slower than the chemical reactions), one can derive an effective transport equation for the enzymes [162],

$$\partial_t e(x,t) = \partial_x^2[D_e(s)e] + \partial_x[D^s_{xd}(s,e)\partial_x s]\,, \tag{4.1}$$

within the quasi-1D reaction chamber (Fig. 4.2) oriented along the x axis. Here, $e(x,t)$ denotes the local enzyme concentration (regardless of whether free or complexed) and $s(x,t)$ the substrate concentration. The effective diffusion coefficient of the enzyme,

$$D_e(s) = D_f + (D_c - D_f)F(s), \tag{4.2}$$

is a function of substrate concentration, interpolating between the diffusion coefficient of the free enzyme and the complex, with $F(s)$ as defined above. For $s \ll K_{\mathrm{M}}$, most of the enzymes are free and $D_{\mathrm{e}}(s) \sim D_{\mathrm{f}}$, whereas for $s \gg K_{\mathrm{M}}$ most of the enzymes are bound to substrate and $D_{\mathrm{e}}(s) \sim D_{\mathrm{c}}$. Note that the disassembly of enzyme oligomers into monomers can also contribute to enhanced diffusion [158, 144, 147]. Eq. (4.1) would then describe the enzyme motion irrespective of its oligomeric state. The cross-diffusion term $D^{\mathrm{s}}_{\mathrm{xd}}(s,e)$ of Eq. (4.1) describes how enzymes respond to substrate gradients due to short-range nonspecific and hydrodynamic interactions,

$$D^{\mathrm{s}}_{\mathrm{xd}}(s,e) = -\left[C^{\mathrm{s}}_{\mathrm{f}} + (C^{\mathrm{s}}_{\mathrm{c}} - C^{\mathrm{s}}_{\mathrm{f}})F(s)\right] e\,. \tag{4.3}$$

Here we define what we will refer to as "cross-diffusion strength" $C^{\mathrm{s}}_{\mathrm{c/f}} = N_{\mathrm{A}} k_{\mathrm{B}} T \lambda^{\mathrm{s}\,2}_{\mathrm{c/f}} \eta^{-1}$, where η is the viscosity of the fluid, k_{B} the Boltzmann constant, N_{A} the Avogadro number, T the temperature and $\lambda^{\mathrm{s}}_{\mathrm{c/f}}$ the Derjaguin length [199, 162]. The Derjaguin length is a parameter capturing the effective short-range interaction between the complex/free enzyme and the substrate. It is typically a few angstroms [163, 164] and can at most be as large as the Debye length (screening length), which in typical buffer conditions is about 1nm [155]. It is expressed via the integral $\lambda^{\mathrm{s}\,2}_{\mathrm{c/f}} = \int_0^\infty \mathrm{d}h h(e^{-\phi_{\mathrm{cs/fs}}(h)(k_{\mathrm{B}}T)^{-1}} - 1)$. $\lambda^{\mathrm{s}\,2}_{\mathrm{c/f}}$ is positive (negative) when the interaction is attractive (repulsive) [162]. The derivation of $\lambda^{\mathrm{s}\,2}_{\mathrm{c/f}}$ is similar to that of the second virial coefficient for a real gas [165], but also includes hydrodynamic effects and is computed by assuming that the size of the enzyme is much larger than the interaction length. The sign of $\lambda^{\mathrm{s}\,2}_{\mathrm{c/f}}$ determines the sign of $C^{\mathrm{s}}_{\mathrm{c/f}}$ and it has the opposite sign of $D^{\mathrm{s}}_{\mathrm{xd}}(s,e)$. For attractive interactions $D^{\mathrm{s}}_{\mathrm{xd}}(s,e) < 0$, i.e. the enzyme drifts towards higher concentrations of substrate (chemotaxis). The enzyme performs antichemotaxis for repulsive interactions, for which $D^{\mathrm{s}}_{\mathrm{xd}}(s,e) > 0$. Note that the effect of nonspecific interactions can also be written as a phoretic drift [162] by swapping $\partial_x s$ in Eq. (4.1) with e in the definition (4.3), with a drift velocity directly proportional to the substrate gradient $v_{\mathrm{ph}}(s, \partial_x s) = \left[C^{\mathrm{s}}_{\mathrm{f}} + (C^{\mathrm{s}}_{\mathrm{c}} - C^{\mathrm{s}}_{\mathrm{f}})F(s)\right] \partial_x s$.

The enzyme dynamics as given by Eq. (4.1) is valid if short-range nonspecific interactions between product and enzyme are neglected. We first neglect such interactions, but will discuss their effects further below. This allows us to derive simple analytical conditions for the determination of the relevant parameter ranges for the formation of spontaneous patterns in the enzyme, substrate and product profiles. This approximation is valid for $\gamma_{\mathrm{p}} \gg \gamma_{\mathrm{s}}$. In such a regime products quickly permeate out of the reaction chamber, impeding the formation of large product gradients and causing product-induced cross-diffusion to be negligible. Note that the values of γ_{p} and γ_{s} can vary by several orders of magnitude. For a reaction chamber that is 1-10µm thick, the permeation rates are in the range of $\gamma_{\mathrm{s,p}} \approx 10^{-9}$-$10^{4}\mathrm{s}^{-1}$ [200]. In the regime where the enzyme is dilute and Eq. (4.1) is valid, the cross-diffusion that enzyme molecules

would induce on substrate and product molecules can be neglected. Moreover the cross-diffusions related to the substrate-product interaction can also be neglected as product and substrate molecules are typically much smaller than the enzyme and they do not affect each other as much as they affect the enzyme motion. Hence, the reaction chamber of Fig. 4.2 is described by the coupled reaction-transport equations

$$\begin{cases} \partial_t e(x,t) = \partial_x^2[D_{\mathrm{e}}(s)e] + \partial_x[D^{\mathrm{s}}_{\mathrm{xd}}(s,e)\partial_x s] \\ \partial_t s(x,t) = D_{\mathrm{s}}\partial_x^2 s - k_{\mathrm{cat}} e F(s) - \gamma_{\mathrm{s}}[s - s_{\mathrm{R}}] \\ \partial_t p(x,t) = D_{\mathrm{p}}\partial_x^2 p + k_{\mathrm{cat}} e F(s) - \gamma_{\mathrm{p}} p\,, \end{cases} \tag{4.4}$$

where $p(x,t)$ denotes the product concentration. The dynamics of the system of Eq. (4.4) is determined by the interplay of the different physical processes described above, each associated with a different timescale. The diffusion processes have timescales $\tau_{\mathrm{diff.}} = L^2 D^{-1}_{\mathrm{e,s,p}}$. The cross-diffusion process has a timescale $\tau_{\mathrm{xd}} = L^2 e_{\mathrm{h}} (D^{\mathrm{s}}_{\mathrm{xd}} s_{\mathrm{R}})^{-1}$ and it represents the time it takes an enzyme to explore a length L, driven by a constant gradient $s_{\mathrm{R}} L^{-1}$, where s_{R} is the maximal substrate concentration allowed in the system. Equivalently, we can say that the nonspecific interactions cause the enzyme to drift with velocity $v_{\mathrm{ph}}(s, \partial_x s)$, with associated timescale $\tau_{\mathrm{ph.}} = L v_{\mathrm{ph}}^{-1} = \tau_{\mathrm{xd}}$. The reaction has a timescale $\tau_{\mathrm{react.}} = k_{\mathrm{cat}}^{-1}$ and the permeations have timescales $\tau_{\mathrm{perm.}} = \gamma_{\mathrm{s,p}}^{-1}$.

Having the reaction chamber coupled to a reservoir avoids product accumulation and substrate depletion, generating a nonzero homogeneous steady-state, with concentrations e_{h}, s_{h}, and p_{h} for enzyme, substrate, and product, respectively. For simplicity, we express e_{h} and p_{h} as functions of s_{h},

$$e_{\mathrm{h}} = \gamma_{\mathrm{s}} \frac{s_{\mathrm{R}} - s_{\mathrm{h}}}{k_{\mathrm{cat}} F(s_{\mathrm{h}})}, \tag{4.5}$$

$$p_{\mathrm{h}} = \frac{k_{\mathrm{cat}} e_{\mathrm{h}} F(s_{\mathrm{h}})}{\gamma_{\mathrm{p}}}\,. \tag{4.6}$$

Since $F(s_{\mathrm{h}})$ is a monotonic increasing function of s_{h}, it is possible to write s_{h} and p_{h} in terms of e_{h}, which can be directly tuned in experiments via the total enzyme concentration (see Eqs. (B.4), (B.5) of Appendix B.1.1 and Fig. B.1).

4.3. Results

4.3.1. Instability driven by substrate-induced cross-diffusion

The homogeneous solution as given by Eqs. (4.5), (4.6) is stable for any positive values of the parameters for a well-mixed system, i.e. a system with no diffusion and no

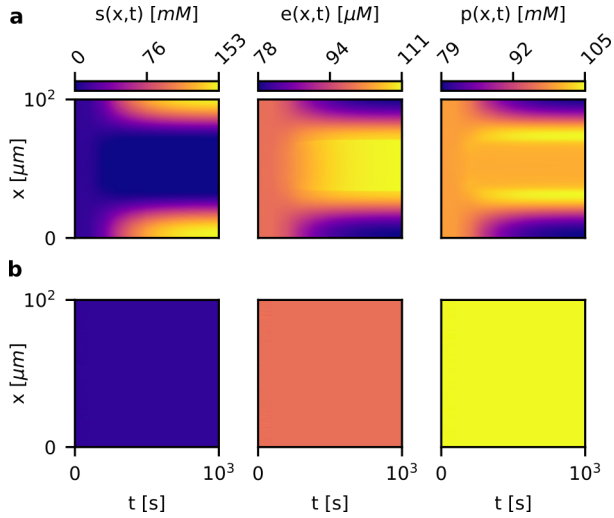


Figure 4.3.: Pattern formation for varying nonspecific interaction between substrate and enzyme. The concentration profiles for substrate $s(x,t)$, enzyme $e(x,t)$ and product $p(x,t)$ are plotted as a function of time and spatial coordinate for two different parameter sets, see Table B.1 for full parameter list. The homogeneous substrate concentration is fixed at $s_\mathrm{h} = 10^4\mu\mathrm{M}$. By varying the square of the Derjaguin length λ_s^2, we affect the nonspecific interaction between substrate and enzyme. **a** $\lambda_\mathrm{s}^2 = -1\text{Å}^2$: the interaction is repulsive and spatial patterns arise from initial homogeneous concentrations with white Gaussian noise. We can see how the enzyme accumulates in regions of low substrate concentration, as well as the product; the profiles reach the steady-state after about $\sim 10^3$ seconds, the same order of magnitude as the timescale of diffusion, which is the slowest timescale for the parameters chosen, $\tau_\mathrm{diff.} = 10^3\mathrm{s}$. **b** $\lambda_\mathrm{s}^2 = 1\text{Å}^2$: the interaction is attractive, the initial perturbation decays and the steady state profiles are homogeneous.

cross-diffusion (see Appendix B.1.2). Fig. 4.3 shows the results of two simulations of the full system of Eq. (4.4), with periodic boundary conditions and parameters as given in Table B.1. In Fig. 4.3a we see that the homogeneous solution is unstable and patterns form for a value of $\lambda_\mathrm{f}^{\mathrm{s}\,2} = \lambda_\mathrm{c}^{\mathrm{s}\,2} = -1\text{Å}^2$. A steady-state is reached after about 10^3s, corresponding to the timescale of enzyme diffusion that for the chosen parameters is the slowest timescale $\tau_\mathrm{diff.} = L^2 D_\mathrm{f}^{-1} = 10^3$s. In Fig. 4.3b we see that for $\lambda_\mathrm{f}^{\mathrm{s}\,2} = \lambda_\mathrm{c}^{\mathrm{s}\,2} = 1\text{Å}^2$ the homogeneous solution is stable. Hence, depending on the parameters, the system will spontaneously form patterns.

To characterize the instability of the homogeneous steady state solution, $\mathbf{v}_\mathrm{h} = (e_\mathrm{h}, s_\mathrm{h}, p_\mathrm{h})$, we linearize Eqs. (4.4), $\mathbf{v} \to \mathbf{v}_\mathrm{h} + \boldsymbol{\delta}\mathbf{v}$, and make the exponential ansatz $\boldsymbol{\delta}\mathbf{v} = \mathbf{v}_0\, e^{\sigma t} e^{iqx}$, with q the spatial frequency of the linear perturbation and σ the perturbation growth rate. If $\sigma > 0$, the perturbation grows with a timescale as given

by σ^{-1} and patterns form (see Appendix B.1.5 for the determination of σ^{-1} for the patterns of Fig. 4.3a). Note that since the system of Eq. (4.4) is isotropic, the results below will also be valid for systems of higher physical dimensions. A positive σ can be found provided that (see Appendix B.1.3)

$$\left[\frac{D'_{\mathrm{e}}(s_{\mathrm{h}})}{D_{\mathrm{e}}(s_{\mathrm{h}})} - \frac{F'(s_{\mathrm{h}})}{F(s_{\mathrm{h}})}\right] + \frac{D^{\mathrm{s}}_{\mathrm{xd}}(s_{\mathrm{h}}, e_{\mathrm{h}})}{e_{\mathrm{h}} D_{\mathrm{e}}(s_{\mathrm{h}})} > \frac{\gamma_{\mathrm{s}} + D_{\mathrm{s}} q^2}{k_{\mathrm{cat}} e_{\mathrm{h}} F(s_{\mathrm{h}})}\,, \tag{4.7}$$

with $D'_{\mathrm{e}}(s_{\mathrm{h}}) = \partial D_{\mathrm{e}}(s)/\partial_s|_{s_{\mathrm{h}}}$ and $F'(s_{\mathrm{h}}) = \partial F(s)/\partial_s|_{s_{\mathrm{h}}}$. Using the functional form of D_{e} in Eq. (4.7), the term given in square brackets equals $-\frac{D_{\mathrm{f}}F'}{D_{\mathrm{e}}F}$ which is negative. Hence for the homogeneous solution to be unstable we must have $D^{\mathrm{s}}_{\mathrm{xd}}(s_{\mathrm{h}}, e_{\mathrm{h}}) > 0$, i.e. enzyme molecules drift downstream the substrate gradients, performing antichemotaxis. We can have positive cross-diffusion if the nonspecific interactions between enzyme and substrate molecules are repulsive, i.e. for negative cross-diffusion strengths $C^{\mathrm{s}}_{\mathrm{c/f}} < 0$. However, how repulsive do these interactions need to be? By using the definitions of $D_{\mathrm{e}}(s)$, $F(s)$ and $D^{\mathrm{s}}_{\mathrm{xd}}(s, e)$, it is possible to show (see Eqs. (B.20)-(B.25) in the Appendix B.1.3) that relation (4.7) is fulfilled if

$$0 < q < \sqrt{-\frac{\beta[D_{\mathrm{f}}K_{\mathrm{M}} + s_{\mathrm{h}}(C^{\mathrm{s}}_{\mathrm{f}}K_{\mathrm{M}} + C^{\mathrm{s}}_{\mathrm{c}}s_{\mathrm{h}})]}{D_{\mathrm{s}}s_{\mathrm{h}}(D_{\mathrm{f}}K_{\mathrm{M}} + D_{\mathrm{c}}s_{\mathrm{h}})}}\,, \tag{4.8}$$

$$D_{\mathrm{f}} + s_{\mathrm{h}}\left(C^{\mathrm{s}}_{\mathrm{f}} + \frac{s_{\mathrm{h}}}{K_{\mathrm{M}}}C^{\mathrm{s}}_{\mathrm{c}}\right) < 0\,, \tag{4.9}$$

with $\beta = \gamma_{\mathrm{s}} s_{\mathrm{R}}$. This result holds in the strong depletion regime ($s_{\mathrm{h}} \ll s_{\mathrm{R}}$), where the effect of the reaction on the substrate dominates over the outflow to the reservoir. In this regime, the expressions are simpler and it is easier to pinpoint the driving mechanism behind the instability observed in Fig. 4.3. We refer the reader to Appendix B.1.3 for the full analysis. The argument of the square root in inequality (4.8) is positive if relation (4.9) is fulfilled. The instability is a type II instability [201] (see Appendix B.1.4), meaning that $\sigma = 0$ at $q = 0$, see Fig. B.2. This is natural as the total amount of enzyme in our system is conserved and homogeneously increasing or decreasing perturbations, i.e. perturbations at $q = 0$, would correspond to changes in the total enzyme amount. By looking at the inequalities (4.9), (4.8), we can see that negative cross-diffusion strengths $C^{\mathrm{s}}_{\mathrm{c/f}} < 0$ are needed to have instabilities. Given that $C^{\mathrm{s}}_{\mathrm{c/f}} := N_{\mathrm{A}} k_{\mathrm{B}} T \lambda^{\mathrm{s}\,2}_{\mathrm{c/f}} \eta^{-1}$, this means that repulsive interactions ($\lambda^{\mathrm{s}\,2}_{\mathrm{c/f}} < 0$) are necessary. Diffusion tends to homogenize the concentration profiles and contributes with a positive term to the left hand side of inequality (4.9). In the case where the enzyme-substrate interaction is attractive, i.e. $C^{\mathrm{s}}_{\mathrm{f}} > 0$, also the second term on the left hand side of inequality (4.9) is positive. Hence the interaction between enzyme and substrate would need to change

sign upon substrate binding to have an instability ($C_c^s < 0$). This could happen if, upon substrate binding, the electrostatic surface charge distribution of the enzyme changes significantly, for instance due to a conformational transition. There can be cases in which the change in nonspecific interactions is less abrupt and both $C_f^s < 0$ and $C_c^s < 0$. Even in the case for which $C_c^s = C_f^s = C_s$, i.e. there is no change in interaction upon substrate binding, it is possible to have an unstable homogeneous solution for

$$\lambda_s^2 < -\frac{1}{6\pi R_f N_A} \frac{K_M}{s_h(s_h + K_M)}, \tag{4.10}$$

where we rewrote relation (4.9) with the use of the Stokes-Einstein relation $k_B T = 6\pi\eta D_f R_f$, and the definition of C_s.

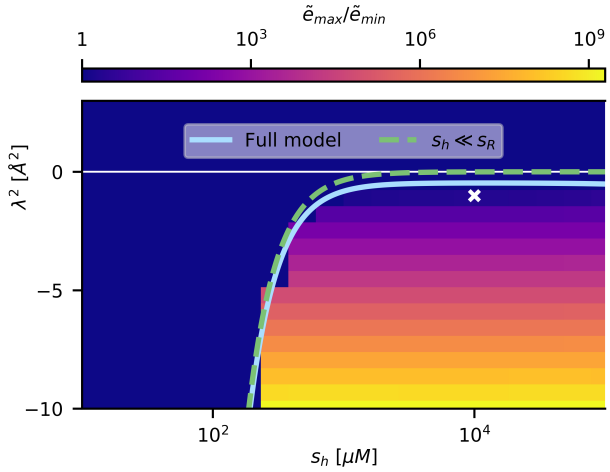


Figure 4.4.: Phase diagram of pattern formation. We plot the instability curve of the full model given by the system of Eq. (4.4) (light blue line) and the approximated curve valid in the strong depletion regime, where the reservoir concentration s_R is much larger than the homogeneous substrate concentration s_h, as given by relation (4.10) (green dashed line). The instability curves are plotted in a diagram where on the abscissa we have s_h and on the ordinate we have the square of the Derjaguin length λ_s^2 that determines the cross-diffusion strength. As a proxy to measure the presence of patterns, we use the ratio of the maximum and the minimum of the steady state profile of the enzyme concentration $\tilde{e}_{max}/\tilde{e}_{min}$. Below the curve, the system of (4.4) is unstable and patterns arise, the ratio $\tilde{e}_{max}/\tilde{e}_{min}$ can be as high as 10^9. Above the curve, the homogeneous solution is stable and the ratio $\tilde{e}_{max}/\tilde{e}_{min} = 1$. The white cross corresponds to the simulation shown in Fig. 4.3a.

It is interesting to note how relations (4.9), (4.10) depend on the substrate concentration. One could ask, given certain nonspecific interactions between enzyme and substrate: at which substrate concentration s_h^* should we begin to observe instabilities?

From relations (4.9), (4.10), we find that

$$s_{\mathrm{h}}^{*}=\frac{1}{2}\left(-K_{\mathrm{M}}+\sqrt{K_{\mathrm{M}}^{2}+\frac{4D_{\mathrm{f}}K_{\mathrm{M}}}{|C_{\mathrm{s}}|}}\right)=\frac{1}{2}\left(-K_{\mathrm{M}}+\sqrt{K_{\mathrm{M}}^{2}+\frac{2K_{\mathrm{M}}}{3\pi R_{\mathrm{f}}|\lambda_{\mathrm{s}}^{2}|N_{\mathrm{A}}}}\right), \tag{4.11}$$

with $\lambda_{\mathrm{s}}^{2}<0$. The value of s_{h}^{*} given by Eq. (4.11) is similar to the value of s_{h} above which the phoresis generated by the nonspecific interactions $v_{\mathrm{ph}}(s,\partial_x s)$ dominates over the drift induced by the enhanced diffusion $v_{\mathrm{bi}}:=-\partial_x D_{\mathrm{e}}(s)$ [162]. They differ only by a prefactor $(D_{\mathrm{c}}-D_{\mathrm{f}})/D_{\mathrm{f}}$ in the second term inside the square root of Eq. (4.11). Having $v_{\mathrm{ph}}>v_{\mathrm{bi}}$ corresponds to $D_{\mathrm{xd}}^{\mathrm{s}}(s,e)>eD_{\mathrm{e}}'(s)$. However, s_{h}^{*} is not only determined by this inequality, but also from inequality (4.7), where reaction also plays a role. By considering biologically relevant ranges, such as $R_{\mathrm{f}}\sim$ 1-10nm [202, 203], $K_{\mathrm{M}}\sim 10^{-2}$-$10^{6}$µM [30], $|\lambda_{\mathrm{s}}|$ being smaller than the Debye length $|\lambda_{\mathrm{s}}|\approx 10^{-2}$-10Å, we find that $s_{\mathrm{h}}^{*}\lesssim K_{\mathrm{M}}$ if $|C_{\mathrm{s}}|K_{\mathrm{M}}(4D_{\mathrm{f}})^{-1}\gg 1$, or equivalently $3/2\pi R_{\mathrm{f}}|\lambda_{\mathrm{s}}^{2}|N_{\mathrm{A}}K_{\mathrm{M}}\gg 1$. This can happen only for enzymes for which $K_{\mathrm{M}}=$ 0.1-1M, which is only a small fraction of enzymes [30]. For the vast majority of enzymes $|C_{\mathrm{s}}|K_{\mathrm{M}}(4D_{\mathrm{f}})^{-1}\ll 1$ and $s_{\mathrm{h}}^{*}\gg K_{\mathrm{M}}$, meaning that we are in the saturated regime of the reaction, for which $F(s)\approx 1$. In such a regime we find that $s_{\mathrm{h}}^{*}\approx\sqrt{D_{\mathrm{f}}K_{\mathrm{M}}|C_{\mathrm{s}}|^{-1}}=\sqrt{K_{\mathrm{M}}(6\pi R_{\mathrm{f}}N_{\mathrm{A}}|\lambda_{\mathrm{s}}^{2*}|)^{-1}}$.

In Fig. 4.4 we plot the phase diagram of the system of Eq. (4.4) where on the abscissa we have s_{h} and on the ordinate we have λ_{s}^{2}. The green dashed line corresponds to the instability curve (4.10) in the strong depletion regime ($s_{\mathrm{h}}\ll s_{\mathrm{R}}$) and the light blue line represents the instability curve derived for any s_{h} (see Eq. (B.23) in Appendix B.1.3). As a proxy for the determination of patterns we plot the ratio of the maximum over the minimum of the enzyme profile at steady-state. Above the instability lines the homogeneous solution is stable. Below the line the system is unstable and patterns similar to the one shown in Fig. 4.3a arise.

4.3.2. Positive feedback causing pattern formation

What is the physical mechanism underlying the instability given by the inequalities (4.9),(4.8)? We illustrate the feedback mechanism generating the pattern in Fig. 4.5. We have already seen that to have instabilities the cross-diffusion $D_{\mathrm{xd}}^{\mathrm{s}}(s,e)$ must be positive, see inequality (4.7). The enzymatic current induced by the cross-diffusion is given by $J_{\mathrm{xd}}^{\mathrm{e}}=-D_{\mathrm{xd}}^{\mathrm{s}}(s,e)\partial_x s$, which for a negative slope of substrate concentration generates a positive current for the enzyme, i.e. the enzyme moves away from a high substrate concentration. The current generated by the enhanced diffusion $J_{\mathrm{D}}^{\mathrm{e}}=-\partial_x[D_{\mathrm{e}}(s)e]$ also consists of a motion of the enzyme away from high substrate concentrations. In Fig. 4.3a we can see how in a regime where patterns form, more product is generated in locations where the substrate concentration is low and the enzyme concentration is

high. Repulsive nonspecific interactions and enhanced diffusion cause the enzyme to accumulate in such regions (Fig. 4.5 from (a) to (b)). This accumulation then generates a higher reaction flux in these regions. Having a stronger reaction flux where substrate is already low, compared to regions where substrate is abundant causes substrate gradients to become steeper (Fig.4.5 from (b) to (c)). A steeper substrate gradient, in turn, causes both J^{e}_{xd} and J^{e}_{D} to increase, hence generating a further accumulation of the enzyme in substrate depleted regions (Fig. 4.5 from (c) to (d)). This positive feedback between reaction and enzyme accumulation leads to the formation of patterns. The feedback cycle halts when the substrate concentration is too low and the reaction is balanced by the influx of substrate from the reservoir. Then the substrate gradient stops getting steeper and the system approaches a steady-state.

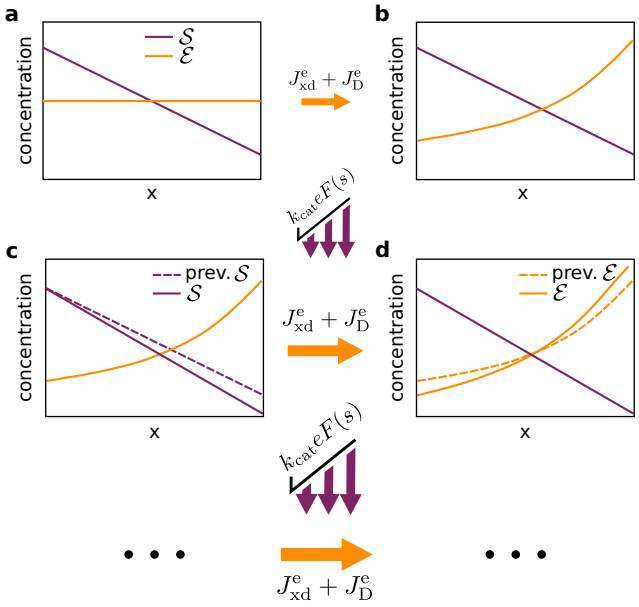


Figure 4.5.: Positive feedback mechanism behind the pattern forming process. **a** We start with a substrate gradient and a homogeneous enzyme profile. **b** The enhanced diffusive current J^{e}_{D} and the cross-diffusive one J^{e}_{xd} cause the enzyme to accumulate where substrate is low, moving from left to right as indicated by the orange arrow. **c** The substrate is depleted at a rate $k_{cat}eF(s)$ and the enzyme accumulated in the region of low substrate cause the substrate gradient to get steeper, as illustrated by the purple arrows. **d** This leads to a further increase of J^{e}_{D} and J^{e}_{xd} causing further accumulation of the enzyme, as indicated by the thicker orange arrow. This process repeats itself determining the patterns.

4.3.3. The role of enhanced diffusion

The currents $J_{\mathrm{D}}^{\mathrm{e}}$ and $J_{\mathrm{xd}}^{\mathrm{e}}$ both contribute to the accumulation of enzymes in regions of low substrate concentrations. However, the phase diagram of Fig. 4.4 suggests that for given concentrations only the cross-diffusive strength characterize the instability of the system. Hence, one could ask: What is the role played by the enhanced diffusion? To answer the above question, we first consider the case of a constant $D_{\mathrm{e}}(s)$ in Eq. (4.4), i.e. $D_{\mathrm{e}} = D_{\mathrm{c}}$. The inequality (4.9) characterizing the instability of the system is unaffected, while the interval (4.8) for the unstable wave vector q slightly changes.

Why cannot patterns be generated simply by enhanced diffusion? By considering the strong depletion regime ($s_{\mathrm{h}} \ll s_{\mathrm{R}}$) and $D_{\mathrm{xd}}^{\mathrm{s}}(s, e) = 0$, condition (4.7) becomes

$$\frac{D_{\mathrm{e}}'(s_{\mathrm{h}})}{D_{\mathrm{e}}(s_{\mathrm{h}})} - \frac{F'(s_{\mathrm{h}})}{F(s_{\mathrm{h}})} > \frac{q^2 D_{\mathrm{s}}}{\beta}\,. \tag{4.12}$$

Both $D_{\mathrm{e}}(s)$ and $F(s)$ have a Michaelis-Menten dependence on s for all models of enzyme motion proposed so far. They differ only in the prefactors and a nonzero offset for $D_{\mathrm{e}}(s)$. Inequality (4.12) is never fulfilled for such $D_{\mathrm{e}}(s)$ and $F(s)$ and consequently any initial perturbation of the concentrations is smoothed out by diffusion (see Appendix B.2.1).

However, inequality (4.12) will apply to any model with a spatially dependent diffusion coefficient that is coupled to another diffusing and reacting species that is being depleted. In case the diffusing species is being generated, the sign of the above inequality is flipped. Systems of this type are used to study bacterial motion. Interestingly synthetic bacterial populations show stripe patterns as they grow on semi-solid agar plates [204]. These bacteria produce a certain signaling molecule through which they sense their own concentration and regulate their mobility. If we neglect bacterial growth, we have a system of equations similar to the equations describing the enzyme and the substrate motion in Eq. (4.4). The enzyme would correspond to the bacteria and the substrate to the signaling molecule. For this system $F(s) = \text{const.} < 0$ and $D_{\mathrm{e}}(s)$ is a Hill function with $D_{\mathrm{e}}'(s) < 0$. Hence patterns can form for $D_{\mathrm{e}}'(s) D_{\mathrm{e}}(s)^{-1} < q^2 D_{\mathrm{s}} \beta^{-1}$, where now $\beta < 0$. This shows that such synthetic population of bacteria can form patterns even in the absence of growth. Inequality (4.12) specifies the minimal ingredients for pattern formation for such systems. It implies that, if $F(s) > 0$, patterns form whenever $D_{\mathrm{e}}(s)$ is more sensitive to perturbations in the substrate concentration than $F(s)$. Having a more sensitive $D_{\mathrm{e}}(s)$ than $F(s)$ causes a more sensitive response in the enzyme motion than depletion due to reaction. Consider a local increase in substrate concentration and that both $D_{\mathrm{e}}(s)$ and $F(s)$ are monotonically increasing functions of s. Having a more sensitive $D_{\mathrm{e}}(s)$ than $F(s)$ implies a higher increase in the current $J_{\mathrm{D}}^{\mathrm{e}}$ due to enhanced diffusivity away from the substrate, as compared to the depletion of substrate due to reaction. Molecules $\mathcal{E}$

then migrate to regions with low substrate and if they do so with a high enough rate they can cause substrate gradients to get steeper, as in step (c) of Fig 4.5. However, for the enzyme model Eq. (4.4), $D_e(s)$ and $F(s)$ alone cannot generate instabilities (see Appendix B.2.1). The accumulation of the enzyme in low substrate regions at a high enough rate can be guaranteed only via the cross-diffusive term $D^s_{xd}(s,e)$. This is because $D_e(s)$ and $F(s)$ are Michaelis-Menten functions with the same K_M and therefore are equally sensitive to substrate perturbations. In Appendix B.2.3 we show that if the Michaelis-Menten constant associated to D_e is much larger than the one associated to $F(s)$, relation (4.12) can be fulfilled, see also Fig. B.3. In this scenario, there can be a range of s_h concentrations for which $D_e(s_h)$ is in the linear regime and sensitive to substrate perturbations, while $F(s_h)$ is in the saturated regime and therefore insensitive to perturbations in s. In Appendix B.2.2 we also show that if $D_e(s)$ is a Hill function and $F(s)$ a Michaelis-Menten reaction, we can have spontaneous pattern formation for a Hill coefficient > 1.

4.3.4. Inclusion of enzyme-product interaction

Up to now we focused on the effects that short-range nonspecific interactions between substrate and enzyme have on the enzyme motion. We neglected the product-enzyme interaction. This approximation was helpful to derive relations (4.9)-(4.11) and check the validity of pattern formation for biologically relevant parameter ranges.

Now we ask: How is pattern formation affected by the short-range interaction between product and enzyme? The enzyme dynamics becomes affected by gradients of product (see Appendix B.3) and instead of Eq. (4.1) we get

$$\partial_t e(x,t) = \partial_x^2[D_e(s)e] + \partial_x[D^s_{xd}(s,e)\partial_x s] + \partial_x[D^p_{xd}(s,e)\partial_x p]\,, \tag{4.13}$$

where $D^p_{xd}(s,e)$ is the cross-diffusion due to nonspecific short-range and hydrodynamic interactions between product and enzyme. Similar to the substrate-induced cross-diffusion, $D^p_{xd}(s,e) = -[C^p_f + (C^p_c - C^p_f)F(s)]e$, where $C^p_{c/f} = N_A k_B T \lambda^{p\,2}_{c/f}\eta^{-1}$ is the cross-diffusive strength corresponding to the product-enzyme interactions for the complexed and free form respectively. The enzyme drifts upstream (downstream) the product gradients for an attractive (repulsive) interaction, for which $D^p_{xd}(s,e) < 0$ ($D^p_{xd}(s,e) > 0$). Note that $D^p_{xd}(s,e)$ depends on the substrate concentration, because the fraction of enzymes in the complexed form is given by $F(s)$.

It is possible to derive a condition characterizing the instability of the homogeneous solution in the presence of product-induced cross-diffusion, similar to inequality (4.7)

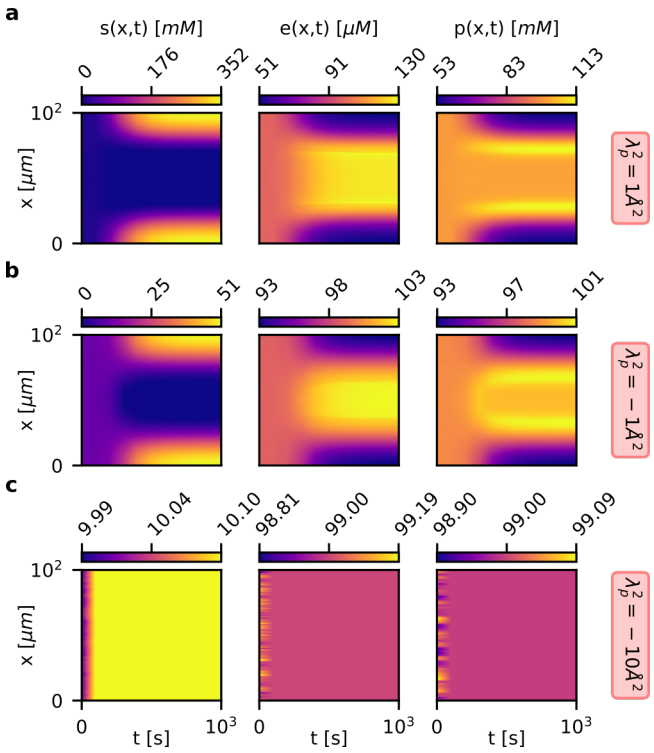


Figure 4.6.: Pattern formation for varying nonspecific interaction between product and enzyme. The concentration profiles of substrate $s(x,t)$, enzyme $e(x,t)$, and product $p(x,t)$ are plotted as a function of time and spatial coordinate. By varying the square of the Derjaguin length λ_{p}^2, we affect the nonspecific interaction between enzyme and product. **a** $\lambda_{\mathrm{p}}^2 = 1\text{Å}^2$: the interaction is attractive and patterns are more pronounced as compared to the case where no interaction is present, cf. 4.3a. **b** $\lambda_{\mathrm{p}}^2 = \lambda_{\mathrm{s}}^2 = -1\text{Å}^2$: the interaction is repulsive and equal in strength to the repulsive interaction between substrate and enzyme, given by λ_{s}^2; we can see that patterns are still present but less pronounced. **c** $\lambda_{\mathrm{p}}^2 = -10\text{Å}^2$: the interaction is so repulsive that the homogeneous solution is stable.

(see Appendix B.3):

$$\left[\frac{D_{\mathrm{e}}'(s_{\mathrm{h}})}{D_{\mathrm{e}}(s_{\mathrm{h}})} - \frac{F'(s_{\mathrm{h}})}{F(s_{\mathrm{h}})}\right] + \frac{D_{\mathrm{xd}}^{\mathrm{s}}(s_{\mathrm{h}}, e_{\mathrm{h}})}{e_{\mathrm{h}} D_{\mathrm{e}}(s_{\mathrm{h}})} - \left[\frac{\gamma_{\mathrm{s}} + D_{\mathrm{s}} q^2}{\gamma_{\mathrm{p}} + D_{\mathrm{p}} q^2}\right] \frac{D_{\mathrm{xd}}^{\mathrm{p}}(s_{\mathrm{h}}, e_{\mathrm{h}})}{e_{\mathrm{h}} D_{\mathrm{e}}(s_{\mathrm{h}})} > \frac{\gamma_{\mathrm{s}} + D_{\mathrm{s}} q^2}{k_{\mathrm{cat}} e_{\mathrm{h}} F(s_{\mathrm{h}})}. \tag{4.14}$$

As compared to inequality (4.7), in relation (4.14) there is an extra term that involves the cross-diffusion $D_{\mathrm{xd}}^{\mathrm{p}}(s,e)$. The prefactor of this term goes to zero in the limit of large γ_{p} as expected. The product-induced cross-diffusion contributes to the instability with opposite sign as compared to $D_{\mathrm{xd}}^{\mathrm{s}}(s,e)$. We have seen in Fig. 4.5 that patterns form if

the interaction between substrate and enzyme is repulsive (transition from (a) to (b)) and $D^{\mathrm{s}}_{\mathrm{xd}}(s,e) > 0$. Then the reaction has the role of steepening pre-formed substrate gradients (transition from (b) to (c)), i.e. more products are formed where substrate is already low. This generates a product gradient that has opposite sign to the substrate gradient, see Fig. 4.3a. Hence we can expect that patterns are favored by attractive nonspecific interactions between product and enzyme, for which $D^{\mathrm{p}}_{\mathrm{xd}}(s,e) < 0$. This favors the accumulation of enzymes in regions of low substrate, further steepening the substrate gradient. In Fig. 4.6a we present the results of simulations where $\lambda^{\mathrm{p}}_{\mathrm{f}} = \lambda^{\mathrm{p}}_{\mathrm{c}} = \lambda_{\mathrm{p}}$. We vary λ_{p} and fix the other parameters to be the same as the ones considered for Fig 4.3a, given in Table B.1. For $\lambda^2_{\mathrm{p}} = 1\text{Å}^2$ (Fig. 4.6a), i.e. for an attractive interaction between product and enzyme ($D^{\mathrm{p}}_{\mathrm{xd}}(s,e) < 0$), we see that the steady-state pattern is more pronounced as compared to the case in which $D^{\mathrm{p}}_{\mathrm{xd}} = 0$ (Fig. 4.3a). For $\lambda^2_{\mathrm{p}} = \lambda^2_{\mathrm{s}} = -1\text{Å}^2$ (Fig. 4.6b) the pattern is still present, however less pronounced. Note that we still get a pattern because we considered different permeation rates $\gamma_{\mathrm{s}} = 1\mathrm{s}^{-1}$ and $\gamma_{\mathrm{p}} = 10\mathrm{s}^{-1}$. Finally, for $\lambda^2_{\mathrm{p}} = -10\text{Å}^2$ (Fig. 4.6c) the repulsion between product and enzyme is so strong that the pattern disappears. In this case $D^{\mathrm{p}}_{\mathrm{xd}}(s,e) > 0$ and the cross-diffusive term induced by the product prevails over the cross-diffusive term induced by the substrate. Therefore the inequality (4.14) is not fulfilled. Note that the inequality (4.14) can also be fulfilled if both cross-diffusions are negative, i.e. both substrate and product attract the enzyme. In this case, patterns form if the product attraction is stronger than the substrate one. As a result, enzyme would still antichemotax from the substrate and the intuitive mechanism leading to patterns illustrated in Fig. 4.5 still holds. Note that one could think of the intuitive mechanism leading to patterns in terms of the feedback between product-induced cross-diffusion and product formation. Given an initial product gradient and a uniform enzyme profile, the enzyme would chemotax to the region of high product. Having more enzymes in this region would cause the formation of even more product. This would cause the steepening of the product gradient and a further accumulation of enzymes. This feedback mechanism is the "time reversal" mechanism of the one illustrated in Fig. 4.5.

4.4. Discussion

We have seen that spontaneous patterns arising from homogeneous concentration profiles can form for a single-step catalytic reaction if cross-diffusive effects are present. Patterns form given sufficiently strong repulsive nonspecific interactions between enzyme and substrate, and/or sufficiently strong nonspecific attractive interactions between enzyme and product as indicated by the inequality (4.14). These interactions

cause the enzyme to move away from regions of high substrate concentrations and accumulate in regions of low substrate, performing antichemotaxis. The accumulated enzymes then deplete the substrate, steepening substrate and product gradients. Steeper gradients further drive the accumulation of enzymes as illustrated in Fig. 4.5. This positive feedback cycle between enzyme accumulation and reaction is what generates the patterns.

The enzyme accumulation is driven by the antichemotaxis with respect to substrate gradients due to nonspecific interactions. The enzyme chemotaxis considered in some of the models [152, 159, 91] has a stabilizing effect. Enzymes accumulate in regions of high substrate concentrations. Reactions then flatten substrate gradients, breaking the feedback that leads to patterns. Hence for such systems spontaneous patterns cannot form. Even in the absence of short-range nonspecific interactions, antichemotaxis exists due to enhanced diffusivity [151, 142]. Nevertheless, we have seen that enhanced diffusivity alone cannot generate spontaneous patterns for a simple enzymatic reaction. Hence, among all models proposed so far for enzyme motion, only the model given by Eq. (4.1), first proposed by Agudo-Canalejo et al. [162], can lead to spontaneous pattern formation. Analysis of the patterns generated can shed light onto the microscopic interactions between enzyme and substrate. In fact, we have seen in Fig. 4.4 how the ratio of the maximum over the minimum of the steady-state enzyme profile is affected by the Derjaguin length $|\lambda_s^2|$ and it does not depend on the substrate concentration s_h.

The patterns observed in Fig. 4.3 are not generated via the common short-range activation and long-range inhibition mechanism [205], as neither the enzyme nor the substrate have autocatalytic activity. It is also not a motility-induced phase separation (MIPS) mechanism [206], which relies on the slowing down of active particles in regions of high particle concentrations. Here a positive feedback mechanism between particle accumulation and reaction leads to the pattern formation. Note that for the system of equations (4.4), patterns can form even if $D_e \sim D_s$, because inequality (4.9) does not depend on D_s, whereas for classical Turing patterns large differences in the diffusion coefficients of the different species are required [197]. Moreover, it is surprising to see that patterns can form for a single-step enzymatic reaction with no autocatalytic activity nor allosteric regulation. In fact, for a system where species have a constant diffusion coefficient and enzymatic reactions follow a simple Michaelis-Menten scheme, patterns form for a minimal network of three states, where forward and backward reactions are catalyzed by two different enzymes respectively [198]. In our system patterns form for a single-step catalytic reaction because of cross-diffusion.

Our findings are consistent with recent studies analyzing the effects of cross-diffusion in pattern formation [207]. Moreover it has been shown that phase separation, formation of static or self-propelled aggregates can be observed in mixtures of cross-diffusive species interacting via a fast diffusing chemical that can be produced or consumed [208].

The scenario considered here corresponds to the case of a single cross-diffusive species, i.e. the enzyme, that is able to consume the fast diffusing chemical, i.e. the substrate. Here we considered the full nonlinear forms of enhanced diffusion and cross-diffusion for the enzyme motion, whereas the other studies considered constant diffusion and constant cross-diffusion [207, 208]. The nonlinear model permitted us to address the question why enhanced diffusion alone is not able to generate patterns for a simple enzymatic reaction. We found that enhanced diffusion would need to be more sensitive to perturbations in substrate concentrations than the reaction, see the inequality (4.12). Although this is not the case for enzymes, we believe that relation (4.12) can characterize the pattern formation of species presenting different enhanced diffusion functions, such as synthetic bacteria [204]. Moreover, the cross-diffusion terms characterizing the dynamics of $\mathcal{E}$ in Eq. (4.13) can correspond to chemotactic responses of organisms to gradients of $\mathcal{S}$ and $\mathcal{P}$, which can be chemoattractants or chemorepellants [209, 210]. Hence the inequality (4.14) can be relevant in the characterization of spontaneous pattern formation of species performing chemotaxis. As similarly shown in Fig. 4.5, patterns can form if a chemotacting organism is repelled from a chemorepellent and such chemorepellent is also depleted from the organism. Moreover, patterns can also form if the organism is attracted to a chemoattractant and the chemoattractant is produced by the organism. The latter scenario corresponds to the "time reversed" mechanism of the one shown in Fig. 4.5 and is at the base of aggregation phenomena in chemotacting *Amoebae* [211].

5. Summary and outlook

The research presented in this thesis investigates two distinct aspects of the spatial and temporal organization of a single-step enzymatic reaction.

The first one concerns the spatial organization of localized (static) enzymes so as to maximize the steady-state reaction rate for a wide class of reaction-diffusion systems. These systems are characterized by the fact that the substrate enters either through the external boundary or the boundaries of internal compartments. Once in the system, the substrate diffuses and can either react with the localized enzymes or is exposed to different loss mechanisms. Systems of this type can be encountered in natural systems or in synthetic bioreactors, e.g. microfluidic chambers and synthetic droplets. Hence, understanding how enzymes can be spatially arranged to maximize the reaction flux, is relevant not only for an improved understanding of natural strategies, but also for the optimal design of synthetic bioreactors. In Ch. 3, we derive an optimal enzyme allocation principle for the class of reaction-diffusion systems that we just described. We show that the optimal allocation strategy consists in distributing enzymes to have homogeneous marginal returns (HMR) in terms of reaction flux at all positions where enzymes are placed. This HMR principle allows us to analytically characterize transitions in the optimal enzyme profile from configurations where enzymes are all bound to the boundaries where substrate is sourced, to configurations where enzymes are also present in the interior of the system. Moreover, by using the HMR criterion we derive a deterministic construction algorithm for the determination of the optimal enzyme profile. Optimally arranged enzymes generate reaction fluxes that are several-fold higher than the ones given by uniform enzyme concentrations. The construction algorithm can be used for the design of transgenic cells and synthetic bioreactors. In fact, different synthetic strategies can regulate the spatial arrangement of enzymes both *in vivo* and *in vitro* (cf. Sec. 2.5). The above results apply to systems of any dimensionality and geometry. The substrate can have advective motion, be subject to different loss mechanisms and present nonlinear reaction functions, e.g. Michaelis-Menten or Hill functions. Although the class of systems considered is quite general and also includes the possibility of having multiple sources of substrate, it applies to the maximization of a single-step enzymatic reaction. It remains interesting to study whether the HMR criterion also applies to the maximization of the reaction flux of a product coming from a multi-step enzymatic reaction. Moreover, the variational formalism used to derive the

HMR criterion could use other "utility functions" having different regulatory purposes than the one of maximizing the reaction flux. For example, these functions could aim at keeping homeostatic concentration levels of metabolites or regulating the ratio of different product fluxes at metabolic branch points.

While the first research project presented in this thesis deals with the spatial organization of enzymes, the second one focuses more on their motion. In fact, recent experiments show that the enzyme motion can depend on the substrate concentration profile. Several enzymes exhibit an apparent diffusion coefficient that increases with substrate concentration. What causes this phenomenon, known as "enzyme enhanced diffusion", is still unclear. A first question is whether conformational changes of the enzyme upon substrate or inhibitor binding can explain the phenomenon alone, or whether catalysis also plays a role. Moreover, the enzyme motion appears to be directed in substrate gradients. However, it is unclear whether enzymes drift upstream or downstream the gradients and thus whether they perform "chemotaxis" or "antichemotaxis". To explain the experimental results, different models of enzyme motion have been proposed. Selecting the correct one is important for a better understanding of enzymatic reactions and biological systems. For instance, the motion of enzymes in response to substrate gradients has been shown to lead to spontaneous co-localization of consecutive metabolic enzymes, providing a possible explanation for the spontaneous fromation of metabolons [91] (cf. Sec. 2.2.3). In Ch. 4, we show that spontaneous pattern formation experiments can be used to discern between the different models proposed. Patterns form for nonspecific repulsive interactions between the enzyme and the substrate, or nonspecific attractive interactions between the enzyme and the product. These nonspecific interactions drive the accumulation of the enzyme in regions of low substrate or high product concentrations. In these regions, the enzymatic reaction causes the further depletion of substrate and increase of product, i.e. the reaction causes the substrate and the product gradients to get steeper. This further drives the accumulation of the enzyme in such regions, causing the emergence of spontaneous spatial patterns. Spontaneous patterns can only form for the model of enzyme motion used in Ch. 4 that accounts for the effects of nonspecific interactions. Hence, if patterns are detected experimentally, this model becomes the most likely one. However, in case patterns do not form, different criteria are needed to further select for the correct model. A possibility is to consider sets of varying initial configurations for the substrate and the enzyme profiles. The different dynamics may then lead to significant differences in the concentration profiles at given times, or in the transient times required to reach the steady-state. Experiments can then be designed based on the set of initial configurations that lead to the largest measurable differences in enzyme, substrate or product profile to further select for the correct model of enzyme motion.

Overall, the work presented in this thesis provides two key results. The first is a

spatial allocation strategy for the optimal arrangement of enzymes that can be used in the design of synthetic bio-reactors. The second shows that spontaneous pattern formation can be used to select for the correct model of enzyme motion among the different ones proposed so far.

Appendices

A. Design principles for the optimal spatial arrangement of enzymes

A.1. General Lagrangian with multiple subspaces

As noted in Ch. 3, the functional derivatives of the Lagrangian are only well-defined if the function with respect to which the derivative is taken is continuous and smooth across its entire domain. However, optimal enzyme distributions may feature discontinuities. For example, there may be a finite enzyme density in one region of the system, and no enzymes (zero density) in another region. This issue can be resolved by defining separate subdomains within the system. The subdomains should be picked in the proximities of the substrate sources. For example if we have patches of sources on the surface domains, for each source patch we would consider a surface subdomain around the patch and a correspondent volume subdomain in proximity of the source. These are the regions where it is optimal to distribute the enzymes as diffusion does not allow for peak of substrate concentration away from the sources. Within each such subdomain i, the substrate and enzyme densities $\rho^{(i)}(\mathbf{r})$ and $e^{(i)}(\mathbf{r})$ are smooth and their local gradients are well defined. However, at the interfaces between such subdomains, $e(\mathbf{r})$ need not be continuous (although $\rho(\mathbf{r})$ still must be). The positions of the interfaces, and therefore the extents of the different domains, should be optimized over and as we have seen in Ch. 3 they depend on the E_{tot} considered.

In this formulation, the total reaction flux can be written as a sum of contributions from each surface and boundary domain, and the corresponding reaction-diffusion or boundary condition acts as a constraint that must be satisfied within each domain. In addition, we introduce explicit matching conditions on $\rho(\mathbf{r})$ at the interfaces I. The

resulting Lagrangian has the following form:

$$\begin{aligned}\mathcal{L} = &\sum_i \int_{S_i} k_{\text{cat}} e_S^{(i)}(\mathbf{s}) F[\rho^{(i)}(\mathbf{s})] \mathrm{d}\mathbf{s} + \sum_j \int_{V_j} k_{\text{cat}} e^{(j)}(\mathbf{r}) F[\rho^{(j)}(\mathbf{r})] \mathrm{d}\mathbf{r} \\ &- \lambda_e \left[\sum_i \int_{S_i} e_S^{(i)}(\mathbf{s}) \mathrm{d}\mathbf{s} + \sum_j \int_{V_j} e^{(j)}(\mathbf{r}) \mathrm{d}\mathbf{r} - E_{\text{tot}} \right] \\ &+ \sum_j \int_{V_j} \lambda_V^{(j)}(\mathbf{r}) \Big\{ D\nabla^2 \rho^{(j)}(\mathbf{r}) - \nabla \cdot \left[\rho^{(j)}(\mathbf{r}) \mathbf{v}(\mathbf{r}) \right] - \sigma \rho^{(j)}(\mathbf{r}) - k_{\text{cat}} e^{(j)}(\mathbf{r}) F[\rho^{(j)}(\mathbf{r})] \Big\} \mathrm{d}\mathbf{r} \\ &+ \sum_i \int_{S_i} \lambda_S^{(i)}(\mathbf{s}) \Big\{ k(\mathbf{s}) \left[j(\mathbf{s}) - D \frac{\partial \rho^{(l)}}{\partial \mathbf{n}(\mathbf{s})} + \rho^{(l)}(\mathbf{s}) \mathbf{v}(\mathbf{s}) \cdot \hat{\mathbf{n}}(\mathbf{s}) - k_{\text{cat}} e_S^{(i)}(\mathbf{s}) F[\rho^{(l)}(\mathbf{s})] \right] \\ &\qquad\qquad - h(\mathbf{s}) \rho^{(l)}(\mathbf{s}) \Big\} \mathrm{d}\mathbf{s} \\ &+ \sum_k \int_{I_k} \Big\{ \lambda_I^{(k)}(\mathbf{r}) \left[\rho^{(k_1)}(\mathbf{r}) - \rho^{(k_2)}(\mathbf{r}) \right] + \boldsymbol{\lambda}'^{(k)}_I(\mathbf{r}) \cdot \left[\nabla \rho^{(k_1)}(\mathbf{r}) - \nabla \rho^{(k_2)}(\mathbf{r}) \right] \Big\} \mathrm{d}\mathbf{r}. \end{aligned} \tag{A.1}$$

where l represents the index of the volume domain adjacent to the boundary at $\mathbf{s}$, and k_1, k_2 represent the two domains on either side of the interface I_k. Here $\lambda_I(\mathbf{r})$ is a Lagrange multiplier that enforces smoothness of the density profile at each point $\mathbf{r}$ on an interface between two domains, and $\boldsymbol{\lambda}'_I(\mathbf{r})$ is a vector of Lagrangian multipliers where each component enforces smoothness of $\rho(\mathbf{r})$ along one of the system dimensions.

It is convenient to consider first some manipulations of the general Lagrangian, Eq. (A.1). Using the vector identity $\nabla \cdot (f\mathbf{a}) = \nabla f \cdot \mathbf{a} + f \nabla \cdot \mathbf{a}$ we can rearrange the diffusive and advective transport terms to read

$$\lambda_V(\mathbf{r}) \nabla^2 \rho(\mathbf{r}) = \nabla \cdot [\lambda_V(\mathbf{r}) \nabla \rho(\mathbf{r}) - \rho(\mathbf{r}) \nabla \lambda_V(\mathbf{r})] + \rho(\mathbf{r}) \nabla^2 \lambda_V(\mathbf{r}), \tag{A.2}$$

$$\lambda_V(\mathbf{r}) \nabla \cdot [\rho(\mathbf{r}) \mathbf{v}(\mathbf{r})] = \nabla \cdot [\rho(\mathbf{r}) \lambda_V(\mathbf{r}) \mathbf{v}(\mathbf{r})] - \rho(\mathbf{r}) \mathbf{v}(\mathbf{r}) \cdot \nabla \lambda_V(\mathbf{r}) \tag{A.3}$$

Applying Gauss theorem ($\int_V \nabla \cdot \mathbf{a}\, \mathrm{d}\mathbf{r} = \oint \mathbf{a} \cdot \hat{\mathbf{n}}(\mathbf{s})\mathrm{d}\mathbf{s}$), Eq. (A.1) becomes

$$\begin{aligned}
\mathcal{L} = &\sum_i \int_{S_i} k_{\mathrm{cat}} e_S^{(i)}(\mathbf{s}) F[\rho^{(i)}(\mathbf{s})]\mathrm{d}\mathbf{s} + \sum_j \int_{V_j} k_{\mathrm{cat}} e^{(j)}(\mathbf{r}) F[\rho^{(j)}(\mathbf{r})]\mathrm{d}\mathbf{r} \\
&- \lambda_e \left[\sum_i \int_{S_i} e_S^{(i)}(\mathbf{s})\mathrm{d}\mathbf{s} + \sum_j \int_{V_j} e^{(j)}(\mathbf{r})\mathrm{d}\mathbf{r} - E_{\mathrm{tot}} \right] \\
&+ \sum_j \int_{V_j} \Big\{ \rho^{(j)}(\mathbf{r}) \left[D\nabla^2 \lambda_V^{(j)}(\mathbf{r}) + \mathbf{v}(\mathbf{r}) \cdot \nabla \lambda_V^{(j)}(\mathbf{r}) - \sigma \lambda_V^{(j)}(\mathbf{r}) \right] \\
&\qquad\qquad - k_{\mathrm{cat}} \lambda_V^{(j)}(\mathbf{r}) e^{(j)}(\mathbf{r}) F[\rho^{(j)}(\mathbf{r})] \Big\} \mathrm{d}\mathbf{r} \\
&+ \sum_i \int_{S_i} \lambda_S^{(i)}(\mathbf{s}) \Big\{ k(\mathbf{s}) \left[j(\mathbf{s}) - D \frac{\partial \rho^{(l)}}{\partial \mathbf{n}(\mathbf{s})} + \rho^{(l)}(\mathbf{s}) \mathbf{v}(\mathbf{s}) \cdot \hat{\mathbf{n}}(\mathbf{s}) - k_{\mathrm{cat}} e_S^{(i)}(\mathbf{s}) F[\rho^{(l)}(\mathbf{s})] \right] \\
&\qquad\qquad - h(\mathbf{s}) \rho^{(l)}(\mathbf{s}) \Big\} \mathrm{d}\mathbf{s} \\
&+ \sum_k \int_{I_k} \left\{ \lambda_I^{(k)}(\mathbf{r}) \left[\rho^{(k_1)}(\mathbf{r}) - \rho^{(k_2)}(\mathbf{r}) \right] + \boldsymbol{\lambda}'^{(k)}_I(\mathbf{r}) \cdot \left[\nabla \rho^{(k_1)}(\mathbf{r}) - \nabla \rho^{(k_2)}(\mathbf{r}) \right] \right\} \mathrm{d}\mathbf{r} \\
&+ \sum_i \int_{S_i} \left\{ D \lambda_V^{(l)}(\mathbf{s}) \frac{\partial \rho^{(l)}}{\partial \mathbf{n}(\mathbf{s})} - D \rho^{(l)}(\mathbf{s}) \frac{\partial \lambda_V^{(l)}}{\partial \mathbf{n}(\mathbf{s})} - \lambda_V^{(l)}(\mathbf{s}) \rho^{(l)}(\mathbf{s}) \mathbf{v}(\mathbf{s}) \cdot \hat{\mathbf{n}}(\mathbf{s}) \right\} \mathrm{d}\mathbf{s} \\
&+ \sum_k \int_{I_k} \Big\{ D \lambda_V^{(k_1)}(\mathbf{r}) \frac{\partial \rho^{(k_1)}}{\partial \mathbf{n}(\mathbf{r})} - D \rho^{(k_1)}(\mathbf{r}) \frac{\partial \lambda_V^{(k_1)}}{\partial \mathbf{n}(\mathbf{r})} - D \lambda_V^{(k_2)}(\mathbf{r}) \frac{\partial \rho^{(k_2)}}{\partial \mathbf{n}(\mathbf{r})} + D \rho^{(k_2)}(\mathbf{r}) \frac{\partial \lambda_V^{(k_2)}}{\partial \mathbf{n}(\mathbf{r})} \\
&\qquad\qquad - \left[\lambda_V^{(k_1)}(\mathbf{r}) \rho^{(k_1)}(\mathbf{r}) - \lambda_V^{(k_2)}(\mathbf{r}) \rho^{(k_2)}(\mathbf{r}) \right] \mathbf{v}(\mathbf{r}) \cdot \hat{\mathbf{n}}(\mathbf{r}) \Big\} \mathrm{d}\mathbf{r}
\end{aligned} \tag{A.4}$$

Here the last three lines contain the contributions from the boundary integral.

A.2. Optimal enzyme arrangement in a one-dimensional system

In this section we go through the full analytic optimization of the simplest exactly-solvable model. We consider a one-dimensional system with a source of substrate at one boundary at $x = 0$, ($j(0) = j_0$, $k(0) = 1$, $h(0) = 0$) and an absorbing boundary at $x = L$ ($j(L) = k(L) = 0$, $h(L) = 1$). Further we assume a linear reaction functional, no drift and no instability of $\mathcal{S}$, i.e. $F[\rho] = \rho/K_{\mathrm{M}}$ and $v = \sigma = 0$.

In our formalism we divide the system into two boundary domains (S_1: $x = 0$ and S_2: $x = L$) and two "volume" domains (V_1: $0 \leq x \leq x_0$ and V_2: $x_0 \leq x \leq L$,

with $0 \leq x_0 \leq L$ a parameter to be optimized). For simplicity of presentation we will assume that enzymes can be present at $x = 0$ or in the domain V_1, while the boundary at $x = L$ and the domain V_2 are free of enzymes. Whilst one can include an arbitrary number of additional domains within the system, we found that such domains turned out to be degenerate with the two specified above. Therefore, we will not include these possibilities here. The effective enzyme profile is therefore of the form $e_S\delta(x) + e(x)\Theta(x_0 - x)$, where $\delta(x)$ is the Dirac delta function and $\Theta(x)$ is the Heaviside function.

The Lagrangian for this specific model is

$$\begin{aligned}\mathcal{L} = {} & \kappa e_S \rho^{(1)}(0) + \int_0^{x_0} \kappa e(x)\rho^{(1)}(x)\mathrm{d}x - \lambda_e \left(e_S + \int_0^{x_0} e(x)\mathrm{d}x - E_{\mathrm{tot}} \right) \\ & + \int_0^{x_0} \lambda^{(1)}(x) \left[D\partial_x^2 \rho^{(1)}(x) - \kappa e(x)\rho^{(1)}(x) \right] \mathrm{d}x - \int_{x_0}^{L} \lambda^{(2)}(x) D\partial_x^2 \rho^{(2)}(x)\mathrm{d}x \\ & + \lambda_0 \left[j_0 + D\partial_x \rho^{(1)}(0) - \kappa e_S \rho^{(1)}(0) \right] - \lambda_L \rho^{(2)}(L) \\ & + \lambda_{x_0} \left[\rho^{(1)}(x_0) - \rho^{(2)}(x_0) \right] + \lambda'_{x_0} \left[\partial_x \rho^{(1)}(x_0) - \partial_x \rho^{(2)}(x_0) \right] ,\end{aligned} \tag{A.5}$$

where we have introduced the catalytic efficiency $\kappa = k_{\mathrm{cat}}/K_{\mathrm{M}}$. Here λ_e is the Lagrange multiplier corresponding to the constraint of having a fixed amount of enzymes E_{tot}; $\lambda^{(1)}(x)$ and $\lambda^{(2)}(x)$ those corresponding the constrained dynamics of the substrate in the two domains; λ_0 and λ_L to the boundary conditions; and λ_{x_0} and λ'_{x_0} to the continuity and smoothness that must be imposed on $\rho(x)$ at the interface x_0.

To calculate the optimal enzyme arrangement e_S^*, $e^*(x)$ we must compute the various functional derivatives of $\mathcal{L}$. Let us first consider the derivatives with respect to $e(x)$ and $\rho^{(1)}(x)$. From the form in Eq. (A.4) we find

$$\frac{\delta \mathcal{L}}{\delta e(x)} = 0 \quad \Leftrightarrow \quad \rho^{(1)}(x) = \frac{\lambda_e}{\kappa[1 - \lambda^{(1)}(x)]} \tag{A.6}$$

$$\frac{\delta \mathcal{L}}{\delta \rho^{(1)}(x)} = 0 \quad \Leftrightarrow \quad e(x) = -\frac{D\partial_x^2 \lambda^{(1)}(x)}{\kappa[1 - \lambda^{(1)}(x)]} . \tag{A.7}$$

We can compute the derivatives $\partial_x \rho^{(1)}(x)$ and $\partial_x^2 \rho^{(1)}(x)$ starting from the expressions above and we can substitute the results into the reaction-diffusion equation (given also by $\frac{\delta \mathcal{L}}{\delta \lambda^{(1)}(x)} = 0$) to obtain an ordinary differential equation for $\lambda^{(1)}(x)$:

$$\frac{\partial_x^2 \lambda^{(1)}(x)}{\partial_x \lambda^{(1)}(x)} = -\frac{\partial_x \lambda^{(1)}(x)}{1 - \lambda^{(1)}(x)} . \tag{A.8}$$

The solution to Eq. (A.8) can be found by direct integration to be

$$\lambda^{(1)}(x) = 1 - A_1 e^{B_1 x}, \tag{A.9}$$

where A_1 and B_1 are constants of integration still to be determined. Using Eqs. (A.6)-(A.7) we obtain

$$\rho^{(1)}(x) = \frac{\lambda_e}{\kappa A_1} e^{-B_1 x} \tag{A.10}$$

$$e(x) = \frac{D B_1^2}{\kappa}, \tag{A.11}$$

Equation (A.11) immediately shows that the enzyme density in the region $0 \leq x < x_0$ is constant, in agreement with previous numerical studies [167, 168].

Evaluating the functional derivative of $\mathcal{L}$ with respect to $\lambda^{(2)}$, we recover the diffusion equation $\partial_x^2 \rho^{(2)}(x) = 0$ in the domain V_2, from which

$$\rho^{(2)}(x) = A_2 + B_2 x, \tag{A.12}$$

where A_2 and B_2 are again constants of integration.

The constants $A_{1,2}$ and $B_{1,2}$ can be evaluated using the boundary conditions at $x = 0$ and $x = L$, and the continuity and smoothness conditions at $x = x_0$. Additionally, the dependence on e_S can be eliminated using $\frac{\delta \mathcal{L}}{\delta \lambda_e} = 0$, which is simply the constraint on the total enzyme number. This leaves us with an expression for the Lagrangian that, since all constraints have been satisfied equals the constrained reaction flux $J^{\mathcal{P}}$, in terms of the single optimization variable x_0,

$$\mathcal{L} = J^{\mathcal{P}} = j_0 \left[1 - \frac{1 - \frac{x_0}{L}}{1 - 2\frac{x_0}{L} + \alpha E_{\text{tot}} (1 - \frac{x_0}{L})^2} e^{-\frac{x_0}{L - x_0}} \right] \tag{A.13}$$

where we define $\alpha = \kappa L / D$. Finally, we maximize over x_0 to find

$$x_0^* = \begin{cases} 0 & \alpha E_{\text{tot}} \leq 1 \\ L \left[1 - (\alpha E_{\text{tot}})^{-1/2} \right] & \alpha E_{\text{tot}} > 1 \end{cases}. \tag{A.14}$$

By substitution we can then obtain expressions for the enzyme and substrate densities as shown in Table A.1.

Regime	e_S^*	$e^*(x)$	$\rho^{(1)}(x)$	$\rho^{(2)}(x)$	$J^{\mathcal{P}*}$
$\alpha E_{\text{tot}} \leq 1$	E_{tot}	0	N/A	$\frac{j_0(L-x)}{D(1+\alpha E_{\text{tot}})}$	$\frac{j_0}{1+(\alpha E_{\text{tot}})^{-1}}$
$\alpha E_{\text{tot}} > 1$	$\sqrt{\frac{E_{\text{tot}}}{\alpha}}$	$\frac{E_{\text{tot}}}{L}$	$\frac{j_0}{2}\frac{L}{D\sqrt{\alpha E_{\text{tot}}}}e^{-\sqrt{\alpha E_{\text{tot}}}\frac{x}{L}}$	$\frac{j_0}{2}\frac{L-x}{D\sqrt{\alpha E_{\text{tot}}}}e^{-\sqrt{\alpha E_{\text{tot}}}\frac{x_0}{L}}$	$j_0\left(1-\frac{1}{2}e^{1-\sqrt{\alpha E_{\text{tot}}}}\right)$

Table A.1.: Optimal enzyme distribution and corresponding substrate profile and flux for a one-dimensional system. Below the transition point $\alpha E_{\text{tot}} \leq 1$, the optimal enzyme profile is a fully bound configuration, where all the enzymes are placed at the source of substrate. For $\alpha E_{\text{tot}} > 1$, the optimal profile is given by a fraction of enzymes located at the source of substrate and by a constant profile in the domain $[0, x_0^*]$, see Eq.(A.14). The substrate profile is given by $\rho^{(1)}(x)$, $\rho^{(2)}(x)$ in the domains $[0, x_0^*]$ and $[x_0^*, L]$, respectively. The expressions for the optimal reaction flux $J^{\mathcal{P}*}$ are given in the last column.

A.3. Marginal returns landscape

Given a particular enzyme arrangement, the marginal returns as a function of position can be calculated directly by considering the addition of an amount δe of enzymes at position x', and solving the new modified reaction-diffusion system. Here we present an example of such a calculation. We consider the same one-dimensional system analyzed above, and suppose that we are in the regime $\alpha E_{\text{tot}} > 1$ and the enzyme profile is the optimal one of Table A.1. For other enzyme profiles, such as those shown in Fig. 3.3a, the calculation proceeds in the same way.

We now have to consider two cases: $0 \leq x' \leq x_0$ and $x_0 < x' \leq L$.

If $0 \leq x' < x_0$, we solve the reaction-diffusion equation in the three domains: $0 \leq x < x'$, $x' < x \leq x_0$, and $x_0 \leq x \leq L$. The resulting density profiles are

$$\rho_{[0,x']}(x) = a_1 e^{\sqrt{\frac{E_{\text{tot}}\kappa}{DL}}x} + b_1 e^{-\sqrt{\frac{E_{\text{tot}}\kappa}{DL}}x} \tag{A.15}$$

$$\rho_{[x',x_0]}(x) = a_2 e^{\sqrt{\frac{E_{\text{tot}}\kappa}{DL}}x} + b_2 e^{-\sqrt{\frac{E_{\text{tot}}\kappa}{DL}}x} \tag{A.16}$$

$$\rho_{[x_0,L]}(x) = a_3 x + b_3. \tag{A.17}$$

To fix the constants of integration a_i, b_i, we make use of the boundary conditions at $x = 0$ and $x = L$ as well as matching conditions on ρ at x' and x_0. Specifically, we require continuity of ρ at x' and x_0, and smoothness at x_0,

$$\rho_{[0,x']}(x') = \rho_{[x',x_0]}(x') \tag{A.18}$$

$$\rho_{[x',x_0]}(x_0) = \rho_{[x_0,L]}(x_0) \tag{A.19}$$

$$\partial_x \rho_{[x',x_0]}(x_0) = \partial_x \rho_{[x_0,L]}(x_0). \tag{A.20}$$

Additionally, conservation of mass at $x = x'$ requires that

$$D\partial_x \rho_{[x',x_0]}(x') - D\partial_x \rho_{[0,x']}(x') = \kappa \delta e \rho_{[x',x_0]}(x'). \tag{A.21}$$

The total flux $J^{\mathcal{P}}$ is then given by

$$J^{\mathcal{P}}(x', \delta e) = \kappa \left[e_S^* \rho(0) + \delta e \rho(x') + \int_0^{x_0} e^*(x)\rho(x)\mathrm{d}x \right], \tag{A.22}$$

into which we substitute the final density profile $\rho(x)$ as well as e_S^* and $e^*(x)$ from Table A.1. Finally, the marginal returns at position x' can be calculated as

$$\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(x')} = \lim_{\delta e \to 0} \frac{\partial J^{\mathcal{P}}}{\partial \delta e} = \frac{j_0}{4}\sqrt{\frac{\alpha}{E_{\mathrm{tot}}}} e^{1-\sqrt{\alpha E_{\mathrm{tot}}}}. \tag{A.23}$$

Therefore, for $0 \leq x' \leq x_0^*$ the marginal returns do not depend on x', and all positions at which we have enzymes give the same returns. It is straightforward to verify that this agrees with the value of λ_e derived from the optimal reaction flux that appears in Table A.1,

$$\lambda_e = \frac{\mathrm{d}J^{\mathcal{P}*}}{\mathrm{d}E_{\mathrm{tot}}} = \frac{j_0}{4}\sqrt{\frac{\alpha}{E_{\mathrm{tot}}}} e^{1-\sqrt{\alpha E_{\mathrm{tot}}}}. \tag{A.24}$$

In the case $x_0 \leq x' < L$, the same procedure can be applied, except with the three domains $0 \leq x \leq x_0$, $x_0 \leq x \leq x'$, and $x' \leq x \leq L$. We find that the marginal returns is given by

$$\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e(x')} = \frac{j_0 \alpha}{4} e^{1-\sqrt{\alpha E_{\mathrm{tot}}}} \left(1 - \frac{x'}{L}\right)\left(1 + \sqrt{\alpha E_{\mathrm{tot}}}\frac{x' - x_0}{L}\right). \tag{A.25}$$

As expected, this approaches Eq. (A.23) as $x' \to x_0 = L\left[1 - (\alpha E_{\mathrm{tot}})^{-1/2}\right]$.

A.4. Derivation of the transition condition

In this section we derive the inequality condition, Eq. (3.20) presented in Sec. 3.3.3, that determines whether or not enzymes should be placed in the volume adjacent to a boundary position $\mathbf{s}$. We begin by examining the functional derivative of $\mathcal{L}$ with respect to substrate profiles at the system boundary, $\rho(\mathbf{s})$. In the vicinity of a reflecting boundary (i.e. $h(\mathbf{s}) = 0$, $k(\mathbf{s}) = 1$) we obtain from Eq. (A.4)

$$\frac{\delta \mathcal{L}}{\delta \rho(\mathbf{s})} = k_{\mathrm{cat}} e_S(\mathbf{s}) \frac{\delta F[\rho]}{\delta \rho}\{1 - \lambda_S(\mathbf{s})\} - D\frac{\partial \lambda_V}{\partial \mathbf{n}(\mathbf{s})} - [\lambda_V(\mathbf{s}) - \lambda_S(\mathbf{s})]\,\mathbf{v}(\mathbf{s}) \cdot \hat{\mathbf{n}}(\mathbf{s}) = 0, \tag{A.26}$$

where equality with zero assumes that the enzyme distribution is the optimal one. We can also consider the functional derivative with respect to the gradient of $\rho(\mathbf{r})$ at the boundaries of the system,

$$\frac{\delta \mathcal{L}}{\delta \frac{\partial \rho}{\partial \mathbf{n}(\mathbf{s})}} = D\left[\lambda_V(\mathbf{s}) - \lambda_S(\mathbf{s})\right] = 0. \tag{A.27}$$

Equation (A.27) requires $\lambda_V(\mathbf{s}) = \lambda_S(\mathbf{s})$, which simplifies Eq. (A.26) to

$$k_{\text{cat}} e_S(\mathbf{s}) F'[\rho(\mathbf{s})]\{1 - \lambda_V(\mathbf{s})\} - D\frac{\partial \lambda_V}{\partial \mathbf{n}(\mathbf{s})} = 0\,, \tag{A.28}$$

where $F'[\rho(\mathbf{s})] = \frac{\delta F[\rho]}{\delta \rho}$.

We next examine variations of $\mathcal{L}$ with respect to the enzyme densities,

$$\frac{\delta \mathcal{L}}{\delta e_S(\mathbf{s})} = k_{\text{cat}} F[\rho(\mathbf{s})]\{1 - \lambda_S(\mathbf{s})\} - \lambda_e \leq 0 \tag{A.29}$$

$$\frac{\delta \mathcal{L}}{\delta e(\mathbf{r})} = k_{\text{cat}} F[\rho(\mathbf{r})]\{1 - \lambda_V(\mathbf{r})\} - \lambda_e \leq 0\,. \tag{A.30}$$

Both $\lambda_V(\mathbf{r})$ and $\rho(\mathbf{r})$ must be continuous, since they must both satisfy (generally non-linear) diffusion equations. Therefore, it follows from Eq. (A.30) that $\frac{\delta \mathcal{L}}{\delta e(\mathbf{r})}$ is itself continuous and that we can define the spatial gradient $\nabla \frac{\delta \mathcal{L}}{\delta e(\mathbf{r})}$. At all points where it is optimal to allocate a non-zero density of enzymes (both in the interior of the system or on the boundaries) we have that $\frac{\delta \mathcal{L}}{\delta e_S(\mathbf{s})} = \frac{\delta \mathcal{L}}{\delta e(\mathbf{r})} = 0$, implying that $\nabla \frac{\delta \mathcal{L}}{\delta e(\mathbf{r})} = 0$. In regions where enzymes are not present because it is suboptimal to add them, we get that $\frac{\delta \mathcal{L}}{\delta e(\mathbf{r})} < 0$, and the gradient $\nabla \frac{\delta \mathcal{L}}{\delta e(\mathbf{r})}$ will in general be finite. Furthermore, since $\rho(\mathbf{r}) \to \rho(\mathbf{s})$ and $\lambda_V(\mathbf{r}) \to \lambda_V(\mathbf{s}) = \lambda_S(\mathbf{s})$ smoothly as $\mathbf{r} \to \mathbf{s}$, $\nabla \frac{\delta \mathcal{L}}{\delta e(\mathbf{r})}$ must remain well defined as we approach the system boundaries.

If it is optimal to have enzymes at the boundaries of the system but not locally in the interior, $\frac{\delta \mathcal{L}}{\delta e_S(\mathbf{s})} = 0$ and $\frac{\delta \mathcal{L}}{\delta e(\mathbf{r})} < 0$ for a small displacement $\mathbf{r} = \mathbf{s} + \boldsymbol{\epsilon}$. In this case, evaluating the gradient of Eq. (A.30) and taking $\mathbf{r} \to \mathbf{s}$ we find

$$\frac{\partial}{\partial \mathbf{n}(\mathbf{s})} \frac{\delta \mathcal{L}}{\delta e(\mathbf{r})} = k_{\text{cat}} \left[F'[\rho(\mathbf{s})]\left(1 - \lambda_V(\mathbf{s})\right) \frac{\partial \rho}{\partial \mathbf{n}(\mathbf{s})} - F[\rho(\mathbf{s})] \frac{\partial \lambda_V}{\partial \mathbf{n}(\mathbf{s})} \right] > 0\,. \tag{A.31}$$

Substituting for $\frac{\partial \lambda_V}{\partial \mathbf{n}(\mathbf{s})}$ from Eq. (A.28), this inequality becomes

$$k_{\text{cat}} F'[\rho(\mathbf{s})]\left(1 - \lambda_V(\mathbf{s})\right) \left\{ \frac{\partial \rho}{\partial \mathbf{n}(\mathbf{s})} - \frac{k_{\text{cat}}}{D} e_S(\mathbf{s}) F[\rho(\mathbf{s})] \right\} > 0. \tag{A.32}$$

Since we assume that $F[\rho]$ is monotonically increasing, $F'[\rho] > 0$. Furthermore, setting Eq. (A.29) equal to zero implies that $1 - \lambda_V(\mathbf{s}) > 0$, since $\lambda_e > 0$. Therefore, Eq. (A.32) reduces to

$$D\frac{\partial\rho}{\partial\mathbf{n}(\mathbf{s})} > k_{\text{cat}} e_S(\mathbf{s}) F[\rho(\mathbf{s})]\,, \tag{A.33}$$

which corresponds to Eq. (3.20) of Sec. 3.3.3.

A.5. Reduction of numerical complexity in a one-dimensional system with drift

In this section we demonstrate how, even in cases for which the full analytical solution is not tractable, it is nevertheless possible to use the Lagrangian formalism to express the optimization problem as a single equation, which can then be solved numerically.

We consider a one-dimensional system, similar to that considered in Sec. A.2 but with the addition of a constant drift v. If $v > 0$ substrate molecules tend to flow away from the source at the origin, whereas for $v < 0$ they tend to flow towards it. The corresponding Lagrangian is

$$\begin{aligned}
\mathcal{L} = {} & \kappa e_S \rho^{(1)}(0) + \int_0^{x_0} \kappa e(x)\rho^{(1)}(x)\mathrm{d}x - \lambda_e \left(e_S + \int_0^{x_0} e(x)\mathrm{d}x - E_{\text{tot}} \right) \\
& + \int_0^{x_0} \lambda^{(1)}(x) \left[D\partial_x^2 \rho^{(1)}(x) - v\partial_x \rho^{(1)}(x) - \kappa e(x)\rho^{(1)}(x) \right] \mathrm{d}x \\
& - \int_{x_0}^{L} \lambda^{(2)}(x) \left[D\partial_x^2 \rho^{(2)}(x) - v\partial_x \rho^{(2)}(x) \right] \mathrm{d}x \\
& + \lambda_0 \left[j_0 + D\partial_x \rho^{(1)}(0) - v\rho^{(1)}(0) - \kappa e_S \rho^{(1)}(0) \right] - \lambda_L \rho^{(2)}(L) \\
& + \lambda_{x_0} \left[\rho^{(1)}(x_0) - \rho^{(2)}(x_0) \right] + \lambda'_{x_0} \left[\partial_x \rho^{(1)}(x_0) - \partial_x \rho^{(2)}(x_0) \right] .
\end{aligned} \tag{A.34}$$

The derivation proceeds in the same way as in Section A.2. We again use the functional derivatives to find expressions for $e(x)$ and $\rho^{(1)}(x)$ in terms of the Lagrange multiplier $\lambda^{(1)}(x)$,

$$\frac{\delta\mathcal{L}}{\delta e(x)} = 0 \quad \Leftrightarrow \quad \rho^{(1)}(x) = \frac{\lambda_e}{\kappa[1 - \lambda^{(1)}(x)]} \tag{A.35}$$

$$\frac{\delta\mathcal{L}}{\delta \rho^{(1)}(x)} = 0 \quad \Leftrightarrow \quad e(x) = -\frac{D\partial_x^2 \lambda^{(1)}(x) + v\partial_x \lambda^{(1)}(x)}{\kappa[1 - \lambda^{(1)}(x)]}\,. \tag{A.36}$$

Surprisingly, when we combine these expressions with the reaction-diffusion equation with drift, we find that the resulting differential equation for $\lambda^{(1)}(x)$ is again exactly

Eq. (A.8), which we can solve to find

$$\lambda^{(1)}(x) = 1 - \tilde{A}_1 e^{\tilde{B}_1 x}, \tag{A.37}$$

$$\rho^{(1)}(x) = \frac{\lambda_e}{\kappa \tilde{A}_1} e^{-\tilde{B}_1 x} \tag{A.38}$$

$$e(x) = \frac{D\tilde{B}_1^2 + v\tilde{B}_1}{\kappa}\,. \tag{A.39}$$

Notably the optimal enzyme density $e(x)$ is again constant, but now depends on the drift velocity v.

In the region with no enzymes, $\rho^{(2)}(x)$ obeys a diffusion equation with drift, $D\partial_x^2\rho^{(2)}(x) - v\partial_x\rho^{(2)}(x) = 0$, which has the solution

$$\rho^{(2)}(x) = \tilde{A}_2 + \tilde{B}_2 e^{\mathrm{Pe}\, x/L}\,, \tag{A.40}$$

where $\mathrm{Pe} = vL/D$ is the Péclet number. As before, we evaluate the integration constants $\tilde{A}_{1,2}$, $\tilde{B}_{1,2}$ using the boundary conditions at $x = 0$ and $x = L$, and the continuity and smoothness conditions at x_0. We again eliminate the dependence on e_S by using the constraint on the total enzyme number. We are left with an expression for $J^{\mathcal{P}}$ in terms of the single optimization variable x_0,

$$\mathcal{L} = J^{\mathcal{P}} = j_0 \left\{ 1 - \frac{\exp\left\{ -\mathrm{Pe}\frac{x_0}{L}\left[e^{\mathrm{Pe}(1-x_0/L)} - 1 \right]^{-1} \right\}}{1 + \frac{\alpha E_{\mathrm{tot}}}{\mathrm{Pe}}\left[1 - e^{-\mathrm{Pe}(1-x_0/L)} \right] - \mathrm{Pe}\frac{x_0}{L}\left[e^{\mathrm{Pe}(1-x_0/L)} - 1 \right]^{-1}} \right\}, \tag{A.41}$$

where α is defined as in Sec. A.2. It can be shown that Eq. (A.41) reduces to Eq. (A.13) in the limit $\mathrm{Pe} \to 0$.

It is not possible to find the x_0^* that maximizes $J^{\mathcal{P}}$ analytically. However, it is straightforward solve $\frac{dJ^{\mathcal{P}}}{dx_0} = 0$ numerically for specified values of the other parameters appearing in Eq. (A.41). The results for the rescaled reaction flux and the rescaled net diffusive flux at the origin are shown in Fig. 3.4 in Sec. 3.3.3.

A.6. One-dimensional system with two sources of substrate

Here we present an example of a scenario of multiple sources of substrate with different strengths. We considered a one-dimensional system in the domain $0 \leq x \leq L$, with reflecting boundaries at both ends. At $x = 0$ we apply a source of substrate with $j(0) = 1$, and at $x = L$ we apply a source with $j(L) = 0.3$. Instead of at the boundary, loss of substrate occurs via decay with rate σ. We again assume linear reactions,

$F[\rho] = \rho/K_M$. For this system we have identified a number of different regimes of the optimal enzyme arrangement as a function of E_{tot}, shown in Fig. A.1a.

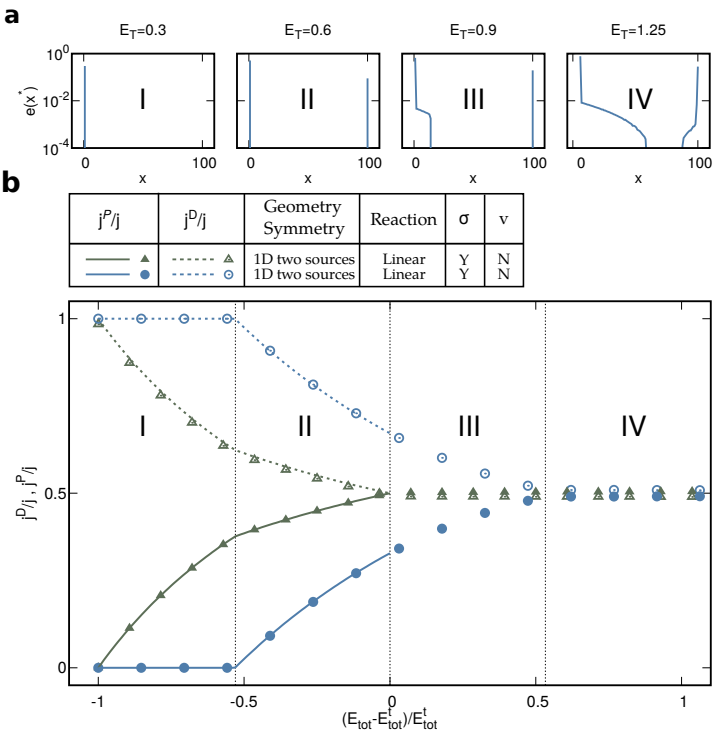


Figure A.1.: Transitions in the optimal enzyme arrangement for a one-dimensional system with two sources of substrate. The system is defined by $\sigma = 1$, $\alpha = Lk_{cat}/(K_M D) = 1$, $\nu = 0$ and boundary conditions $j(0) = 1$, $k(0) = -1$, $h(0) = 0$ and $j(L) = 0.3$, $k(L) = -1$, $h(L) = 0$. **a** Examples of the optimal enzyme profiles for different values of E_{tot}. **b** Diffusion and reaction fluxes at $x = 0$ and $x = L$, normalized by the local influxes of substrate, are plotted as a function of $(E_{tot} - E^{t2}_{tot})/E^{t2}_{tot}$. For both poles of the system, we confirm that $j^D > j^{\mathcal{P}}$ when the optimal enzyme distribution consists of only a single boundary cluster at that pole, while $j^D = j^{\mathcal{P}} = j/2$ when the optimal enzyme arrangement extends a finite distance away from the pole. Lines show analytical solutions. Points show results of numerical optimization. Fig. 3.4 of the main text shows the same data with the fluxes at $x = L$ plotted against $(E_{tot} - E^{t3}_{tot})/E^{t3}_{tot}$.

At the lowest E_{tot} values, the optimal arrangement consists of a single cluster at $x = 0$, the site of strongest substrate influx. As E_{tot} crosses a first transition level, E^{t1}_{tot}, the optimal configuration features clusters of enzymes at both ends of the system: a larger one at $x = 0$ (the stronger source) and a smaller one at $x = L$ (the weaker source). In this regime the optimal distribution of enzymes between the two poles can be calculated exactly by solving the reaction-diffusion equation with arbitrary cluster densities e_0 and e_L, and then optimizing over e_0 and e_L subject to the constraint $e_0 + e_L = E_{tot}$. The

transition position E_{tot}^{t1} is then the value at which the resulting optimal $e_L^* = 0$. However, since the expressions are large and complex we do not include them here.

As E_{tot} is increased further we reach a second transition at E_{tot}^{t2}, with enzymes being placed within the system in the vicinity of $x = 0$. This transition occurs when the diffusive and reaction fluxes at $x = 0$ become equal, $j^D(0) = j^{\mathcal{P}}(0) = j(0)/2$ (see Fig. A.1b). At E_{tot}^{t2}, the corresponding fluxes at $x = L$ are not equal and still satisfy the inequality condition Eq. (A.33) (Eq. (3.20) of Sec. 3.3.3). Therefore, around $x = L$ the optimal configuration is still purely clustered, with no enzymes at $x < L$. The transition threshold E_{tot}^{t2} can be determined exactly by calculating the fluxes $j^D(0)$ and $j^{\mathcal{P}}(0)$ from the full solution of the system with the two optimal polar clusters.

At a larger still transition value E_{tot}^{t3}, the diffusive and reactive fluxes at $x = L$ also become equal, $j^D(L) = j^{\mathcal{P}}(L) = j(L)/2$ (see Fig. A.1b). For $E_{\text{tot}} > E_{\text{tot}}^{t3}$, enzymes are present away from the boundaries at both ends of the system. In the regime $E_{\text{tot}} > E_{\text{tot}}^{t2}$ it is not possible to solve for the form of the optimal profile $e^*(x)$, or therefore to calculate $j^D(L)$ and $j^{\mathcal{P}}(L)$. The value of E_{tot}^{t3} must therefore be determined numerically.

Finally, if E_{tot} is increased further, the enzyme density profiles extending from the two ends of the system will eventually coalesce, and enzymes will we present at all positions $0 \leq x \leq L$.

It is also possible for these transitions to occur in a different order for different parameter combinations. For example, if $j(L) \ll j(0)$, the order of the first two transitions may be reversed. In this case, the condition $j^D(0) = j^{\mathcal{P}}(0) = j(0)/2$ is satisfied before the transition value E_{tot}^{t1} is reached. The first transition will then be from a single cluster at $x = 0$ to a regime in which there is a finite extension of enzymes from $x = 0$, but no cluster at $x = L$. The appearance of the second cluster at $x = L$ may then occur subsequently, at larger E_{tot}, resulting in an optimal enzyme profile than resembles type III in Fig. A.1a.

A.7. Optimal enzyme allocation algorithm for discretized systems

In this section we derive Eqs. (3.21), (3.23) of Sec. 3.3.5, which define the algorithm used for the construction of the optimal enzyme profile as enzymes are iteratively added into the system. We consider a system with discretized space. The densities of substrate and enzyme change from functions of position to a discrete set of variables, $\{e_i\} = \{e(\mathbf{r}_i)\}$ for $i = 1..N$, where N is the number of lattice sites. Instead of being a functional of $e(\mathbf{r})$, $\rho(\mathbf{r})$, the corresponding discrete Lagrangian then becomes a function of $\{e_i\}$, $\{\rho_i\}$, and functional derivatives $\frac{\delta \mathcal{L}}{\delta f(\mathbf{r})}$ become normal derivatives $\frac{\partial \mathcal{L}}{\partial f_i}$.

The derivative of $\mathcal{L}$ with respect to the enzyme density has an analogous form as in

the continuous system,

$$\frac{\partial \mathcal{L}}{\partial e_i} = k_{\text{cat}} F[\rho_i]\{1 - \lambda_i\} - \lambda_e, \tag{A.42}$$

where λ_i is the Lagrange multiplier associated with the discrete reaction-diffusion equation at site i. We again denote the first term above by $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i}$, with the same interpretation as in the continuous case. This quantity is generally easily accessible in any numerical analysis: one simply increases the amount of enzymes at position i by a small amount, solves the resulting modified reaction-diffusion system, and calculates the resulting change in the total reaction flux. In the same way as for the continuous system, for the optimal enzyme arrangement $\{e_i^*\}$ we must have that $\frac{\partial \mathcal{L}}{\partial e_i} = 0$ at all sites i where $e_i^* > 0$, and $\frac{\partial \mathcal{L}}{\partial e_i} \leq 0$ where $e_i^* = 0$.

We now consider the optimal amount of enzymes at site i to be a function of E_{tot}, $e_i^*(E_{\text{tot}})$. That is, there is a line $\mathbf{e}^*(E_{\text{tot}})$ in the space of $\mathbf{e} = (e_1, e_2, \ldots, e_N)$, that we can parameterize by the value of E_{tot}, for which the reaction flux $J_{\mathcal{P}}(\mathbf{e}^*(E_{\text{tot}}))$ is the highest that can be achieved for that value of E_{tot}. We begin by considering the derivative of $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i}$ with respect to E_{tot} along the line defined by $\mathbf{e}^*(E_{\text{tot}})$, $\frac{\mathrm{d}}{\mathrm{d}E_{\text{tot}}}\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i}$. We now expand the total derivative $\frac{\mathrm{d}}{\mathrm{d}E_{\text{tot}}}$ in terms of derivatives with respect to the underlying variables $\{e_i\}$,

$$\frac{\mathrm{d}}{\mathrm{d}E_{\text{tot}}}\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i} = \sum_{j=1}^{N} \frac{\mathrm{d}e_j}{\mathrm{d}E_{\text{tot}}}\frac{\mathrm{d}}{\mathrm{d}e_j}\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i}. \tag{A.43}$$

Note that here $\frac{\mathrm{d}}{\mathrm{d}e}$ takes into account the changes in the flux both due directly to altering the enzyme density, and the resulting change in the substrate density profile $\{\rho_i\}$. Since $\frac{\mathrm{d}e_j}{\mathrm{d}E_{\text{tot}}}$ is understood to be in the direction of the line of optimal profiles, it coincides with $\frac{\mathrm{d}e_j^*}{\mathrm{d}E_{\text{tot}}}$. Writing the set of N equations Eq. (A.43) in vector form we have

$$\frac{\mathrm{d}}{\mathrm{d}E_{\text{tot}}}\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}\mathbf{e}} = \mathbf{H}\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\text{tot}}}, \tag{A.44}$$

where $\mathbf{H}$ is the Hessian of $J^{\mathcal{P}}$, $H_{ij} = \frac{d^2 J^{\mathcal{P}}}{de_j de_i}$.

From Eq. (A.42) we know that if there are $1 \leq n \leq N$ positions at which the optimal

enzyme density is non-zero, then

$$\begin{pmatrix} \frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_1} \\ \vdots \\ \frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_n} \\ \frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_{n+1}} \\ \vdots \\ \frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_N} \end{pmatrix} = \lambda_e \begin{pmatrix} 1 \\ \vdots \\ 1 \\ c_{n+1} \\ \vdots \\ c_N \end{pmatrix}, \tag{A.45}$$

where all c's lie in the range $0 < c_i < 1$, and we have numbered lattice sites in descending order of marginal returns. Therefore, Eq. (A.44) becomes

$$\mathbf{H}\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\mathrm{tot}}} = \frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}} \begin{pmatrix} 1 \\ \vdots \\ 1 \\ c_{n+1} + \frac{\mathrm{d}c_{n+1}}{\mathrm{d}E_{\mathrm{tot}}} \frac{\lambda_e}{\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}} \\ \vdots \\ c_N + \frac{\mathrm{d}c_N}{\mathrm{d}E_{\mathrm{tot}}} \frac{\lambda_e}{\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}} \end{pmatrix}. \tag{A.46}$$

In general it is not possible to directly compute the right hand side of Eq. (A.46), because without knowing the direction of the optimal line it is not possible to calculate the derivatives $\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}$ and $\frac{\mathrm{d}c_i}{\mathrm{d}E_{\mathrm{tot}}}$. However, exploiting the fact that no enzymes should be added to the system at positions where $\frac{\mathrm{d}J^{\mathcal{P}}}{\mathrm{d}e_i} < \lambda_e$, the vector $\frac{\mathrm{d}\mathbf{e}^*}{\mathrm{d}E_{\mathrm{tot}}}$ must have zeros at all positions $i > n$. Therefore, the N-dimensional system in Eq. (A.46) can be split into two parts. Primarily, we have the n-dimensional system

$$\mathbf{H}_{(n)}\frac{\mathrm{d}\mathbf{e}^*_{(n)}}{\mathrm{d}E_{\mathrm{tot}}} = \frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}\mathbf{1}_{(n)}, \tag{A.47}$$

where the subscript (n) denotes that only the first n positions of the lattice are considered. Importantly, in order to solve Eq. (A.47) it is not necessary to know $\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}$ in advance. Instead, one can solve for any constant prefactor, and then subsequently use the normalization condition $\sum_{i=1}^{n} \frac{\mathrm{d}e_i^*}{\mathrm{d}E_{\mathrm{tot}}} = 1$ to express $\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}$ in terms of this constant. For self-consistency, the solution to Eq. (A.47) must also solve the residual

$(N-n)$-dimensional system

$$\begin{pmatrix} \frac{\mathrm{d}^2 J^{\mathcal{P}}}{\mathrm{d}e_1 de_{n+1}} & \frac{\mathrm{d}^2 J^{\mathcal{P}}}{\mathrm{d}e_2 de_{n+1}} & \cdots & \frac{\mathrm{d}^2 J^{\mathcal{P}}}{\mathrm{d}e_n de_{n+1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mathrm{d}^2 J^{\mathcal{P}}}{\mathrm{d}e_1 de_N} & \frac{\mathrm{d}^2 J^{\mathcal{P}}}{\mathrm{d}e_2 de_N} & \cdots & \frac{\mathrm{d}^2 J^{\mathcal{P}}}{\mathrm{d}e_n de_N} \end{pmatrix} \frac{\mathrm{d}\mathbf{e}_n^*}{\mathrm{d}E_{\mathrm{tot}}} = \frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}} \begin{pmatrix} c_{n+1} + \frac{\mathrm{d}c_{n+1}}{\mathrm{d}E_{\mathrm{tot}}} \frac{\lambda_e}{\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}} \\ \vdots \\ c_N + \frac{\mathrm{d}c_N}{\mathrm{d}E_{\mathrm{tot}}} \frac{\lambda_e}{\frac{\mathrm{d}\lambda_e}{\mathrm{d}E_{\mathrm{tot}}}} \end{pmatrix}, \tag{A.48}$$

which can be interpreted as defining the evolution of the marginal returns at positions without enzymes, $\frac{\mathrm{d}c_i}{\mathrm{d}E_{\mathrm{tot}}}$.

B. Cross-diffusion induced patterns for a single-step enzymatic reaction

B.1. Linear instability analysis for enzyme interacting with substrate only

In this section, we perform the linear instability analysis of the PDE system (4.4) of Ch. 4 to derive the condition for which patterns form for a single-step enzymatic reaction. For this system, we assume that only substrate gradients are affecting enzyme motion. We will include cross-diffusion caused by product gradients in Appendix B.3. As a reminder, the starting model for our analysis (Eq. (4.4) of Ch. 4) is:

$$\begin{cases} \partial_t e(x,t) = \partial_x^2[D_{\mathrm{e}}(s)e(x,t)] + \partial_x[D^{\mathrm{s}}_{\mathrm{xd}}(s,e)\partial_x s(x,t)] \\ \partial_t s(x,t) = D_{\mathrm{s}}\partial_x^2 s(x,t) - k_{\mathrm{cat}}e(x,t)F(s) - \gamma_{\mathrm{s}}(s - s_{\mathrm{R}}) \\ \partial_t p(x,t) = D_{\mathrm{p}}\partial_x^2 p(x,t) + k_{\mathrm{cat}}e(x,t)F(s) - \gamma_{\mathrm{p}}p\,, \end{cases} \tag{B.1}$$

where $D^{\mathrm{s}}_{\mathrm{xd}}(s,e) = -[C^{\mathrm{s}}_{\mathrm{f}} + (C^{\mathrm{s}}_{\mathrm{c}} - C^{\mathrm{s}}_{\mathrm{f}})F(s)]e(x,t)$, with $C^{\mathrm{s}}_{\mathrm{c/f}} = N_{\mathrm{A}}k_{\mathrm{B}}T{\lambda^{\mathrm{s}}_{\mathrm{c/f}}}^2/\eta$, η is the viscosity of the fluid, k_{B} the Boltzmann constant, N_{A} the Avogadro number, T the temperature and $\lambda^{\mathrm{s}}_{\mathrm{c/f}}$ is the Derjaguin length [199, 162].

B.1.1. Homogeneous steady state solution

The homogeneous steady state solution $\mathbf{v}_{\mathrm{h}} = (e_{\mathrm{h}}, s_{\mathrm{h}}, p_{\mathrm{h}})$ of (B.1) is given by the following expressions:

$$e_{\mathrm{h}} = \gamma_{\mathrm{s}}\frac{s_{\mathrm{R}} - s_{\mathrm{h}}}{k_{\mathrm{cat}}F(s_{\mathrm{h}})}, \tag{B.2}$$

$$p_{\mathrm{h}} = \frac{k_{\mathrm{cat}}e_{\mathrm{h}}F(s_{\mathrm{h}})}{\gamma_{\mathrm{p}}}, \tag{B.3}$$

where e_{h}, s_{h}, and p_{h} are the homogeneous concentrations of enzyme, substrate, and product respectively. In this work for simplicity we consider e_{h} and p_{h} as functions of s_{h} but it is also possible to write s_{h} and p_{h} in terms of e_{h}, which is a quantity directly

tunable in the experiments:

$$s_{\mathrm{h}} = \frac{1}{2}\left[s_{\mathrm{R}} - K_{\mathrm{M}} - \frac{k_{\mathrm{cat}}}{\gamma_{\mathrm{s}}}e_{\mathrm{h}} + \sqrt{\left(s_{\mathrm{R}} - K_{\mathrm{M}} - \frac{k_{\mathrm{cat}}}{\gamma_{\mathrm{s}}}e_{\mathrm{h}}\right)^2 + 4s_{\mathrm{R}}K_{\mathrm{M}}}\right], \tag{B.4}$$

$$p_{\mathrm{h}} = \frac{\gamma_{\mathrm{s}}}{2\gamma_{\mathrm{p}}}\left[s_{\mathrm{R}} + K_{\mathrm{M}} + \frac{k_{\mathrm{cat}}}{\gamma_{\mathrm{s}}}e_{\mathrm{h}} - \sqrt{\left(s_{\mathrm{R}} - K_{\mathrm{M}} - \frac{k_{\mathrm{cat}}}{\gamma_{\mathrm{s}}}e_{\mathrm{h}}\right)^2 + 4s_{\mathrm{R}}K_{\mathrm{M}}}\right], \tag{B.5}$$

which for $K_{\mathrm{M}} \ll s_{\mathrm{h}} < s_{\mathrm{R}}$ become

$$s_{\mathrm{h}} = s_{\mathrm{R}} - \frac{k_{\mathrm{cat}}}{\gamma_{\mathrm{s}}}e_{\mathrm{h}}, \tag{B.6}$$

and

$$p_{\mathrm{h}} = \frac{k_{\mathrm{cat}}}{\gamma_{\mathrm{p}}}e_{\mathrm{h}}. \tag{B.7}$$

In Fig. B.1, we plot e_{h} versus s_{h} for both the exact and the approximated relation. One can use Fig. B.1 to read the enzyme concentration e_{h} from the corresponding value of the substrate concentration s_{h}.

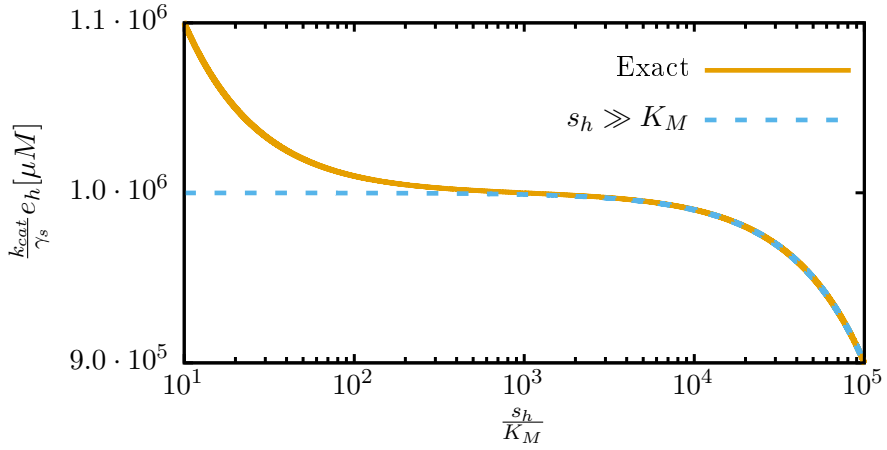


Figure B.1.: The relation between the homogeneous enzyme concentration e_{h} and the homogeneous substrate concentration s_{h}. Exact curve is for Eqs. (B.2) or (B.4) and the curve for $s_{\mathrm{h}} \gg K_{\mathrm{M}}$ is for Eq. (B.7). Parameters are $s_{\mathrm{R}} = 10^6\mu\mathrm{M}$ and $K_{\mathrm{M}} = 1\mu\mathrm{M}$.

B.1.2. Instability of the well-mixed system

We first study the instability of the solution for the well-mixed system. In the well-mixed system, the spatial derivatives and any spatial dependence of the concentrations are neglected. The system of Eq. (B.1) becomes:

$$\begin{cases} \partial_t e(t) = 0 \\ \partial_t s(t) = -k_{\text{cat}} e(t) F(s) - \gamma_{\text{s}}(s - s_{\text{R}}) \\ \partial_t p(t) = k_{\text{cat}} e(t) F(s) - \gamma_{\text{p}} p\,. \end{cases} \tag{B.8}$$

The fixed-point of the system of Eq. (B.8) is still given by $\mathbf{v}_{\text{h}}$. To perform a linear instability analysis, we perturb the system around its fixed point. Any perturbation around $\mathbf{v}_{\text{h}}$ can be written as $e(t) = e_{\text{h}} + \delta e(t)$, $s(t) = s_{\text{h}} + \delta s(t)$, $p(t) = p_{\text{h}} + \delta p(t)$ or alternatively $\mathbf{v}(t) = \mathbf{v}_{\text{h}} + \boldsymbol{\delta}\mathbf{v}(t)$ where $\boldsymbol{\delta}\mathbf{v}(t) = (\delta e(t), \delta s(t), \delta p(t))$. The amount of enzymes in the system is fixed and does not change over time because $\partial_t e(t) = 0$, hence any perturbation $\delta e(t)$ just shifts the total enzyme amount. The linearized system of equations takes the form:

$$\begin{cases} \partial_t \delta e(t) = 0 \\ \partial_t \delta s(t) = -k_{\text{cat}} e_{\text{h}} F'(s_{\text{h}}) \delta s - k_{\text{cat}} F(s_{\text{h}}) \delta e(t) - \gamma_{\text{s}} \delta s(t) \\ \partial_t \delta p(t) = k_{\text{cat}} e_{\text{h}} F'(s_{\text{h}}) \delta s + k_{\text{cat}} F(s_{\text{h}}) \delta e(t) - \gamma_{\text{p}} \delta p(t)\,, \end{cases} \tag{B.9}$$

where $F'(s_{\text{h}})$ is the derivative of $F(s)$ with respect to s at $s = s_{\text{h}}$

$$F'(s_{\text{h}}) = \left.\frac{\partial F(s)}{\partial s}\right|_{s=s_{\text{h}}}. \tag{B.10}$$

We rewrite the linear system of equations (B.9) in matrix form

$$\frac{\partial \boldsymbol{\delta}\mathbf{v}(t)}{\partial t} = \mathbf{J}_0 \cdot \boldsymbol{\delta}\mathbf{v}(t), \tag{B.11}$$

where

$$\mathbf{J}_0 = \begin{pmatrix} 0 & 0 & 0 \\ -k_{\text{cat}} F(s_{\text{h}}) & -k_{\text{cat}} e_{\text{h}} F'(s_{\text{h}}) - \gamma_{\text{s}} & 0 \\ k_{\text{cat}} F(s_{\text{h}}) & k_{\text{cat}} e_{\text{h}} F'(s_{\text{h}}) & -\gamma_{\text{p}} \end{pmatrix}. \tag{B.12}$$

The eigenvalues of $\mathbf{J}_0$ determine the fate of the perturbations. The eigenvalues are $\sigma_1 = 0$, $\sigma_2 = -k_{\text{cat}} e_{\text{h}} F'(s_{\text{h}}) - \gamma_{\text{s}}$ and $\sigma_3 = -\gamma_{\text{p}}$ and all of them are non-positive. This indicates that the perturbations do not grow with time. $\sigma_1 = 0$ corresponds to the perturbation in the total enzyme amount. Once the total amount of enzymes is

perturbed, its level does not change over time because $\partial_t e(t) = 0$. The other eigenvalues are negative, i.e. any perturbations $\delta s(t)$, $\delta p(t)$ decay over time as long as $F'(s_\mathrm{h}) > 0$. Hence the homogeneous solution of a well-mixed system is stable against perturbations for any value of the parameters.

B.1.3. Instability of the reaction-diffusion system

Similar to the well-mixed system, we again study the dynamics of small perturbations around the homogeneous steady state solution but the perturbations are now temporally and spatially dependent: $e(x,t) = e_\mathrm{h} + \delta e(x,t)$, $s(x,t) = s_\mathrm{h} + \delta s(x,t)$, $p(x,t) = p_\mathrm{h} + \delta p(x,t)$ or alternatively $\mathbf{v}(x,t) = \mathbf{v}_\mathrm{h} + \boldsymbol{\delta}\mathbf{v}(x,t)$ where $\boldsymbol{\delta}\mathbf{v}(x,t) = (\delta e(x,t), \delta s(x,t), \delta p(x,t))$. We insert the perturbed concentrations in Eq. (B.1) and linearize the system by assuming that perturbations are small to find:

$$\begin{cases} \partial_t \delta e(x,t) = D_\mathrm{e}(s_\mathrm{h})\partial_x^2 \delta e(x,t) + \left[e_\mathrm{h} D'_\mathrm{e}(s_\mathrm{h}) + D^\mathrm{s}_\mathrm{xd}(s_\mathrm{h}, e_\mathrm{h})\right] \partial_x^2 \delta s(x,t) \\ \partial_t \delta s(x,t) = D_\mathrm{s}\partial_x^2 \delta s(x,t) - k_\mathrm{cat} e_\mathrm{h} F'(s_\mathrm{h})\delta s - k_\mathrm{cat} F(s_\mathrm{h})\delta e(x,t) - \gamma_\mathrm{s}\delta s(x,t) \\ \partial_t \delta p(x,t) = D_\mathrm{p}\partial_x^2 \delta p(x,t) + k_\mathrm{cat} e_\mathrm{h} F'(s_\mathrm{h})\delta s + k_\mathrm{cat} F(s_\mathrm{h})\delta e(x,t) - \gamma_\mathrm{p}\delta p(x,t)\,, \end{cases} \tag{B.13}$$

where $D'_\mathrm{e}(s_\mathrm{h})$, $F'(s_\mathrm{h})$ are the derivatives of $D_\mathrm{e}(s)$, $F(s)$ with respect to s at $s = s_\mathrm{h}$

$$D'_\mathrm{e}(s_\mathrm{h}) = \left.\frac{\partial D_\mathrm{e}(s)}{\partial s}\right|_{s=s_\mathrm{h}}, \tag{B.14}$$

$$F'(s_\mathrm{h}) = \left.\frac{\partial F(s)}{\partial s}\right|_{s=s_\mathrm{h}}. \tag{B.15}$$

We write the linear system of Eq. (B.13) in Fourier space to partially diagonalize the equations. The result is that the dynamics of each Fourier mode is independent from other modes and is governed by the following equation:

$$\frac{\partial \mathbf{v}(q,t)}{\partial t} = \mathbf{J}(q)\cdot\boldsymbol{\delta}\mathbf{v}(q,t), \tag{B.16}$$

where

$$\boldsymbol{\delta}\mathbf{v}(q,t) = \begin{pmatrix} \delta e(q,t) \\ \delta s(q,t) \\ \delta p(q,t) \end{pmatrix} \tag{B.17}$$

and

$$\mathbf{J}(q) = \begin{pmatrix} -D_{\mathrm{e}}(s_{\mathrm{h}})q^2 & -[D^{\mathrm{s}}_{\mathrm{xd}}(s_{\mathrm{h}}, e_{\mathrm{h}}) + e_{\mathrm{h}}D'_{\mathrm{e}}(s_{\mathrm{h}})]q^2 & 0 \\ -k_{\mathrm{cat}}F(s_{\mathrm{h}}) & -k_{\mathrm{cat}}e_{\mathrm{h}}F'(s_{\mathrm{h}}) - D_{\mathrm{s}}q^2 - \gamma_{\mathrm{s}} & 0 \\ k_{\mathrm{cat}}F(s_{\mathrm{h}}) & k_{\mathrm{cat}}e_{\mathrm{h}}F'(s_{\mathrm{h}}) & -D_{\mathrm{p}}q^2 - \gamma_{\mathrm{p}} \end{pmatrix}. \tag{B.18}$$

It is not surprising that $\mathbf{J}(0)$ is the same as the Jacobian of the well-mixed system $\mathbf{J}_0$, (B.12). At $q = 0$ we are neglecting all the effects due to diffusion and cross-diffusion, cf. Eq. (B.12). Moreover perturbations with $q = 0$ correspond to homogeneous shifts in the concentrations, which are the same as considered for the well-mixed system.

The stability of the homogeneous solution can be determined by the eigenvalues of $\mathbf{J}(q)$. If the real part of all the eigenvalues of $\mathbf{J}(q)$ is negative, the homogeneous steady state solution is linearly stable; otherwise it is unstable. One of the eigenvalues of $\mathbf{J}(q)$, $\sigma_3 = -D_{\mathrm{p}}q^2 - \gamma_{\mathrm{p}}$, is always negative and corresponds to the relaxation rate of any perturbation that only perturbs the product concentration. From Eq. (B.1) we can see that the product dynamics has no feedback on the substrate and enzyme equations. This implies that the instability is characterized by feedback in the (e, s) subspace. The other two eigenvalues σ_1 and σ_2 are the solution of the following quadratic equation:

$$\begin{aligned} &\sigma^2 + \sigma\left[k_{\mathrm{cat}}e_{\mathrm{h}}F'(s_{\mathrm{h}}) + D_{\mathrm{s}}q^2 + D_{\mathrm{e}}(s_{\mathrm{h}})q^2 + \gamma_{\mathrm{s}}\right] \\ &+ k_{\mathrm{cat}}e_{\mathrm{h}}\left[D_{\mathrm{e}}(s_{\mathrm{h}})F'(s_{\mathrm{h}}) - D'_{\mathrm{e}}(s_{\mathrm{h}})F(s_{\mathrm{h}})\right]q^2 \\ &- k_{\mathrm{cat}}F(s_{\mathrm{h}})D^{\mathrm{s}}_{\mathrm{xd}}(s_{\mathrm{h}}, e_{\mathrm{h}})q^2 + \gamma_{\mathrm{s}}D_e(s_{\mathrm{h}})q^2 + D_{\mathrm{s}}D_{\mathrm{e}}(s_{\mathrm{h}})q^4 = 0\,. \end{aligned} \tag{B.19}$$

By rewriting the above expression as $(\sigma - \sigma_1)(\sigma - \sigma_2) = 0$, we see that the summation of the eigenvalues is negative, therefore the smaller eigenvalue σ_2 is always negative. The larger eigenvalue σ_1 can be positive if and only if the product of the two eigenvalues is negative. Hence the following condition must be fulfilled:

$$\left[\frac{D'_{\mathrm{e}}(s_{\mathrm{h}})}{D_{\mathrm{e}}(s_{\mathrm{h}})} - \frac{F'(s_{\mathrm{h}})}{F(s_{\mathrm{h}})}\right] + \frac{D^{\mathrm{s}}_{\mathrm{xd}}(s_{\mathrm{h}}, e_{\mathrm{h}})}{e_{\mathrm{h}}D_{\mathrm{e}}(s_{\mathrm{h}})} > \frac{\gamma_{\mathrm{s}} + D_{\mathrm{s}}q^2}{k_{\mathrm{cat}}e_{\mathrm{h}}F(s_{\mathrm{h}})}. \tag{B.20}$$

Note that the above inequality corresponds to relation (4.7) of Ch. 4. Using the relations for D_{e} and $D^{\mathrm{s}}_{\mathrm{xd}}$ and Eq. (B.2), i.e. $k_{\mathrm{cat}}e_{\mathrm{h}} = \gamma_{\mathrm{s}}(s_{\mathrm{R}} - s_{\mathrm{h}})/F(s_{\mathrm{h}})$, we obtain:

$$\frac{-D_{\mathrm{f}}F'(s_{\mathrm{h}}) - F(s_{\mathrm{h}})[C^{\mathrm{s}}_{\mathrm{f}} + (C^{\mathrm{s}}_{\mathrm{c}} - C^{\mathrm{s}}_{\mathrm{f}})F(s_{\mathrm{h}})]}{F(s_{\mathrm{h}})[D_{\mathrm{f}} + (D_{\mathrm{c}} - D_{\mathrm{f}})F(s_{\mathrm{h}})]} > \frac{1}{s_{\mathrm{R}} - s_{\mathrm{h}}} + \frac{D_{\mathrm{s}}q^2}{\gamma_{\mathrm{s}}(s_{\mathrm{R}} - s_{\mathrm{h}})}. \tag{B.21}$$

Equation (B.21) is the condition for instability of the homogeneous solution with a general reaction efficiency $F(s)$. We can see that a Michaelis-Menten form is not necessary for observing the instability. Generally a larger $F(s_{\mathrm{h}})$ and smaller $F'(s_{\mathrm{h}})$ favors

the instability, implying that weaker nonspecific interactions can generate instabilities. When $F(s_h)$ has a Michaelis-Menten form, a relatively strong repulsive interaction between substrate and enzyme can produce patterns even if $s_h \lesssim K_M$.

If we implement the Michaelis-Menten form of $F(s_h)$ in Eq. (B.21) we can find a threshold wave number q^*, for which any perturbation with smaller wave number ($q < q^*$) is unstable:

$$q^{*2} = -\frac{\gamma_s \left[K_M s_R D_f + s_h^2 D_c + s_h(s_R - s_h)(K_M C_f^s + s_h C_c^s)\right]}{D_s s_h [K_M D_f + s_h D_c]}. \tag{B.22}$$

In order to have a positive q^{*2}, it is required that

$$D_f + \left(\frac{s_h}{K_M}\right)\left(\frac{s_h}{s_R}\right) D_c + \frac{s_h}{K_M}\left(1 - \frac{s_h}{s_R}\right)(K_M C_f^s + s_h C_c^s) < 0. \tag{B.23}$$

Equations (B.22) and (B.23) are the instability conditions for the system of equations (B.1) (Eq. (4.4) of the main text). In the regime $s_h \ll s_R$, with $\beta = \gamma_s s_R$, we get

$$0 < q < q^* = \sqrt{-\frac{\beta \left[K_M D_f + s_h(K_M C_f^s + s_h C_c^s)\right]}{D_s s_h [K_M D_f + s_h D_c]}}\,, \tag{B.24}$$

$$D_f + s_h \left(C_f^s + \frac{s_h}{K_M} C_c^s\right) < 0\,, \tag{B.25}$$

which are identical to relations (4.8) and (4.9) of the main text.

By substituting $D_f = k_B T / 6\pi\eta R_f$, where R_f is the hydrodynamic radius of the free enzyme, and $C_{c/f}^s = N_A k_B T \lambda_{c/f}^{s\,2} / \eta$ into relation (B.25), we obtain the analogous expression:

$$\frac{1}{6\pi N_A R_f}\frac{K_M}{s_h^2} + \frac{K_M}{s_h}\lambda_f^{s\,2} + \lambda_c^{s\,2} < 0. \tag{B.26}$$

In case $\lambda_c^{s\,2} = \lambda_f^{s\,2}$, by rearranging, we obtain the inequality (4.10) of Ch. 4

$$\lambda^2 < -\frac{1}{6\pi R_f N_A}\frac{K_M}{s_h(s_h + K_M)}\,. \tag{B.27}$$

To be able to see the patterns, the size of the system needs to be large enough to accommodate Fourier modes with $q < q^*$. Alternatively, the minimum system size to see patterns $L_{min} = 2\pi/q^*$ can be obtained from Eq. (B.22):

$$L_{min} = 2\pi\sqrt{-\frac{D_s s_h [K_M D_f + s_h D_c]}{\gamma_s \left[K_M s_R D_f + s_h^2 D_c + s_h(s_R - s_h)(K_M C_f^s + s_h C_c^s)\right]}}\,, \tag{B.28}$$

where we considered the q^* as given by Eq. (B.22). For the parameters given in Table B.1 (the same used for the simulation shown in Fig. 4.3a of Ch. 4), $L_{\min} = 59.2\mu m$.

B.1.4. Instability type II

In this section we determine the type of instability for the system of Eq. (4.4). As can be seen by looking at inequality (B.26), we need repulsive nonspecific interactions to have an unstable system. Let us consider for example $C_c^s < 0$, then the inequality (B.26) is fulfilled when $s_h > s_h^*$ where

$$s_h^* = \frac{-K_M {\lambda_f^s}^2 - \sqrt{K_M^2 {\lambda_f^s}^4 - \frac{2K_M {\lambda_c^s}^2}{3\pi R_f N_A}}}{2{\lambda_c^s}^2}. \tag{B.29}$$

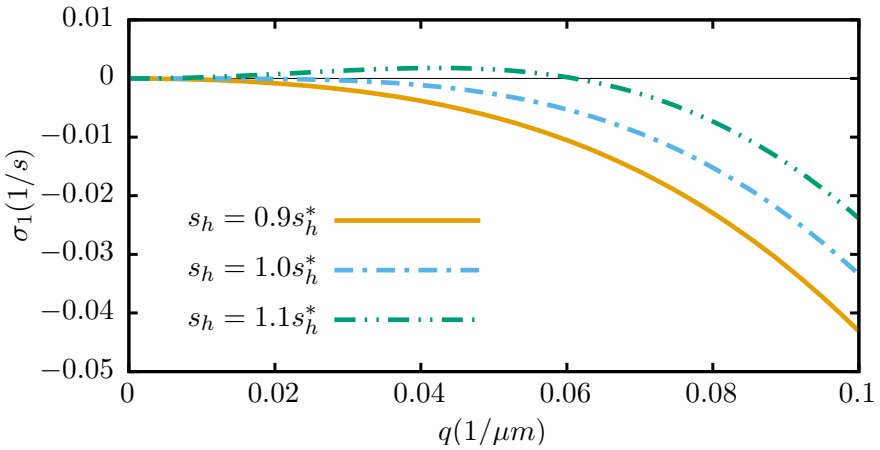


Figure B.2.: Largest eigenvalue of the Jacobian σ_1 versus the wave vector q. The horizontal black line corresponds to $\sigma_1 = 0$. By increasing the initial homogeneous concentration of substrate, s_h, from below to above critical concentration, s_h^*, the eigenvalues become positive. The shape of the eigenvalue curves indicates that the instability is of type II. The parameters for the graph are $T = 300K$, $D_f = 10\mu m^2\,s^{-1}$, $D_c = 13\mu m^2\,s^{-1}$, $D_s = 100\mu m^2\,s^{-1}$, $K_M = 1\mu M$, $\beta = 10^6 \mu M/s$, ${\lambda_f^s}^2 = 1\text{Å}^2$, ${\lambda_c^s}^2 = -1\text{Å}^2$.

Fig. B.2 shows the largest eigenvalue of the linearized equations. When $\sigma_1 > 0$, the system of equations is linearly unstable with respect to perturbations of the homogeneous solution. Fig. B.2 indicates that the system becomes unstable by increasing s_h above its critical value, s_h^*, which is related indirectly to the initial amount of enzymes in the system. The form of curves in Fig. B.2 correspond to a type II instability [201]. A type II instability is typical of systems with conserved quantities. In our case the total amount of enzymes is conserved, this implies that $\sigma_1 = 0$ at $q = 0$, otherwise we

would have homogeneous change in e_h over time, corresponding to changes in the total amount of enzymes.

B.1.5. Timescale of pattern formation

The timescale at which patterns form is determined by the inverse of the growth rate of perturbations $\tau_\sigma = \sigma^{-1}$. From Eq. (B.19), we get:

$$\begin{aligned}\sigma = &-\frac{1}{2}\left(k_{cat}e_h F'(s_h) + [D_s + D_e(s_h)]q^2 + \gamma_s\right) \\ &+ \frac{1}{2}\sqrt{\left(k_{cat}e_h F'(s_h) + [D_s - D_e(s_h)]q^2 + \gamma_s\right)^2 + 4k_{cat}F(s_h)(D^s_{xd} + e_h D'_e(s_h))q^2}\,.\end{aligned} \tag{B.30}$$

For simplicity we focus on the saturated regime where $s_h \gg K_M$, $F(s_h) \approx 1$, $F'(s_h) \approx 0$, and $D'_e(s_h) \approx 0$:

$$\sigma = -\frac{1}{2}\left([D_s + D_c]q^2 + \gamma_s\right) + \frac{1}{2}\sqrt{\left([D_s - D_c]q^2 + \gamma_s\right)^2 + 4k_{cat}D^s_{xd}q^2}\,. \tag{B.31}$$

The wave number for a box with size $L = 100\mu m$ is $q \approx 0.1\mu m^{-1}$. According to the parameters in Table B.1, the timescale of the increase of the perturbations is found to be $\tau_\sigma = \sigma^{-1} \approx 120s$. However this is the timescale associated to a linear response. As perturbations grow, the nonlinear effects become more and more important. This is the reason why the system in Fig. 4.3a of Ch. 4 reaches the steady state after 10^3s.

B.2. No short-range interactions

In this section we consider $D^s_{xd}(s, e) = 0$. Substituting this assumption into Eq. (B.20) and using Eq. (B.2), i.e. $k_{cat}e_h = \gamma_s(s_R - s_h)/F(s_h)$, we get:

$$\left[\frac{D'_e(s_h)}{D_e(s_h)} - \frac{F'(s_h)}{F(s_h)}\right] > \frac{1}{s_R - s_h} + \frac{D_s q^2}{\gamma_s(s_R - s_h)}. \tag{B.32}$$

In the regime $s_h \ll s_R$, with $\beta = \gamma_s s_R$, we get:

$$\left[\frac{D'_e(s_h)}{D_e(s_h)} - \frac{F'(s_h)}{F(s_h)}\right] > \frac{D_s q^2}{\beta}, \tag{B.33}$$

which is the same as the inequality (4.12) of Ch. 4. We use the relation (B.33) in the following subsections to show the possibility of pattern formation for three different functional forms of $D_e(s_h)$.

B.2.1. Michaelis Menten form

By using the definition of $D_e(s)$ used by Agudo-Canalejo et al. [162], $D_e = D_f + (D_c - D_f)F(s)$, we find that the inequality (B.32) takes the form:

$$-\frac{D_f F'(s_h)}{D_e(s_h)F(s_h)}(s_R - s_h) > 1 + \frac{D_s q^2}{\gamma_s}. \tag{B.34}$$

Because $s_h < s_R$ and $F'(s_h) > 0$, the left hand side of the inequality above is negative and cannot be larger than the always positive right hand side. Therefore enhanced diffusion alone cannot drive instabilities. The cross-diffusion generated by repulsive interactions is key in this respect. This also holds in the simpler case of $s_h \ll s_R$, with $\beta = \gamma_s s_R$.

Similarly it is possible to show that we cannot have instabilities for other enhanced diffusion definitions [142, 152].

B.2.2. Hill function

A Hill function for $D_e(s)$ has been considered to model enhanced diffusion in synthetic bacterial populations [204]. Here we show that condition (B.33) can be satisfied when $D_e(s)$ is a Hill function:

$$D_e(s) = D_f + (D_c - D_f)\frac{s^n}{s^n + K_M^n}. \tag{B.35}$$

By substituting this form of $D_e(s)$ in (B.33), one can show that the instability is possible with $n > 1$ under the following condition:

$$\frac{n(D_c - D_f)\left[\left(\frac{s_h}{K_M}\right)^n + \left(\frac{s_h}{K_M}\right)^{n+1}\right]}{D_c\left(\frac{s_h}{K_M}\right)^{2n} + (D_f + D_c)\left(\frac{s_h}{K_M}\right)^n + D_f} - 1 > \frac{D_s q^2}{\beta} s(1 + \frac{s}{K_M}). \tag{B.36}$$

For example at $s_h = K_M$ we find that for the instability

$$0 < q < q^* = \sqrt{\frac{\beta}{2D_s K_M}}\sqrt{\frac{n(D_c - D_f)}{D_c + D_f} - 1}, \tag{B.37}$$

$$D_c > \frac{n+1}{n-1} D_f, \tag{B.38}$$

which means that the closer n is to one, the larger D_c is required to be. In the limit of the Michaelis-Menten function ($n \to 1$), the condition leads to $D_c \to \infty$.

B.2.3. Different affinity constant

Another possible form of enhanced diffusion is assuming the Michaelis-Menten function with a different constant, K_d, i.e. $D_e(s) = D_f + (D_c - D_f)G(s)$, where $G(s) = s/(s+K_d)$. To obtain the condition for instability in the strong depletion regime, we substitute this form of $D_e(s)$ in (B.33) and find

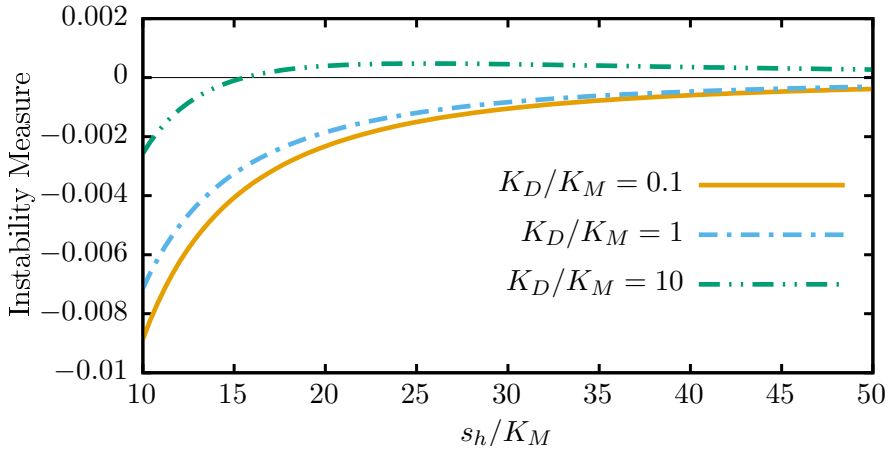


Figure B.3.: Instability measure as a function of s_h/K_M. Instability measure or the left hand side of inequality (B.40) versus s_h/K_M. The parameters for the graph are $D_f = 10\mu m^2\,s^{-1}$, $D_c = 13\mu m^2\,s^{-1}$.

$$0 < q < q^* = \sqrt{\frac{\beta}{D_s}} \left[\frac{K_d}{\frac{D_f}{(D_c - D_f)}(s_h + K_d)^2 + (s_h + K_d)s_h} - \frac{K_M}{s_h(s_h + K_M)} \right]^{\frac{1}{2}}, \tag{B.39}$$

$$\frac{K_d/K_M}{\frac{D_f}{(D_c - D_f)}\left(\frac{s_h}{K_M} + \frac{K_d}{K_M}\right)^2 + \frac{s_h}{K_M}\left(\frac{s_h}{K_M} + \frac{K_d}{K_M}\right)} - \frac{1}{\frac{s_h}{K_M}\left(\frac{s_h}{K_M} + 1\right)} > 0\,. \tag{B.40}$$

We define the left hand side of inequality (B.40) as "instability measure". Fig. B.3 shows how the Instability Measure changes with respect to s_h/K_M for different ratios of K_d/K_M. We see that for large enough K_d/K_M it is possible to fulfill the above condition. Larger K_d/K_M means slower saturation of $D_e(s_h)$ compared to $F(s_h)$ or alternatively $D_e(s_h)$ is more sensitive than $F(s_h)$ to perturbations in the substrate concentration. Interestingly it is also possible to observe instabilities in a regime where both $D_e(s)$

and $F(s)$ are saturated, i.e. for $s_\mathrm{h} \gg K_\mathrm{d} > K_\mathrm{M}$. In this limit the instability measure can be approximated with $\left(\frac{D_\mathrm{c}-D_\mathrm{f}}{D_\mathrm{c}}\right)\frac{K_\mathrm{d}}{K_\mathrm{M}} - 1 > 0$. From this approximated relation we see that for the instability we must have

$$\frac{K_\mathrm{d}}{K_\mathrm{M}} > \frac{D_\mathrm{c}}{D_\mathrm{c} - D_\mathrm{f}}\,. \tag{B.41}$$

B.3. Cross-diffusion of enzyme by product gradients

Here we add a cross diffusion term for enzymes in response to product gradients in Eq. (B.1).

$$\begin{cases} \partial_t e(x,t) = \partial_x^2[D_\mathrm{e}(s)e(x,t)] + \partial_x[D^\mathrm{s}_\mathrm{xd}(s,e)\partial_x s(x,t)] + \partial_x[D^\mathrm{p}_\mathrm{xd}(s,e)\partial_x p(x,t)] \\ \partial_t s(x,t) = D_\mathrm{s}\partial_x^2 s(x,t) - k_\mathrm{cat}e(x,t)F(s) - \gamma_\mathrm{s}(s - s_\mathrm{R}) \\ \partial_t p(x,t) = D_\mathrm{p}\partial_x^2 p(x,t) + k_\mathrm{cat}e(x,t)F(s) - \gamma_\mathrm{p}p \end{cases}, \tag{B.42}$$

where $D^\mathrm{p}_\mathrm{xd}(s,e) = -[C^\mathrm{p}_\mathrm{f} + (C^\mathrm{p}_\mathrm{c} - C^\mathrm{p}_\mathrm{f})F(s)]e(x,t)$, with $C^\mathrm{p}_\mathrm{c/f} = N_\mathrm{A}k_\mathrm{B}T{\lambda^\mathrm{p}_\mathrm{c/f}}^2/\eta$, where $\lambda^\mathrm{p}_\mathrm{c/f}$ is the Derjaguin length of the interaction between enzyme and product.

The corresponding Jacobian of the linearized equations is the following:

$$\mathbf{J}(q) = \begin{pmatrix} -D_\mathrm{e}(s_\mathrm{h})q^2 & -[D^\mathrm{s}_\mathrm{xd}(s_\mathrm{h},e_\mathrm{h}) + e_\mathrm{h}D'_\mathrm{e}(s_\mathrm{h})]q^2 & -D^\mathrm{p}_\mathrm{xd}(s_\mathrm{h},e_\mathrm{h})q^2 \\ -k_\mathrm{cat}F(s_\mathrm{h}) & -k_\mathrm{cat}e_\mathrm{h}F'(s_\mathrm{h}) - D_\mathrm{s}q^2 - \gamma_\mathrm{s} & 0 \\ k_\mathrm{cat}F(s_\mathrm{h}) & k_\mathrm{cat}e_\mathrm{h}F'(s_\mathrm{h}) & -D_\mathrm{p}q^2 - \gamma_\mathrm{p} \end{pmatrix}, \tag{B.43}$$

and the eigenvalue equation of the Jacobian is:

$$\begin{aligned} &\sigma^3 + \sigma^2\left[(D_\mathrm{e} + D_\mathrm{s} + D_\mathrm{p})q^2 + \gamma_\mathrm{s} + \gamma_\mathrm{p} + k_\mathrm{cat}e_\mathrm{h}F'(s_\mathrm{h})\right] \\ &\sigma\Big[- k_\mathrm{cat}F(s_\mathrm{h})q^2(D^\mathrm{s}_\mathrm{xd} - D^\mathrm{p}_\mathrm{xd} + e_\mathrm{h}D'_\mathrm{e}(s_\mathrm{h})) + k_\mathrm{cat}e_\mathrm{h}F'(s_\mathrm{h})(D_\mathrm{e}q^2 + D_\mathrm{p}q^2 + \gamma_\mathrm{p}) \\ &\qquad + D_\mathrm{e}q^2(D_\mathrm{s}q^2 + D_\mathrm{p}q^2 + \gamma_\mathrm{s} + \gamma_\mathrm{p}) + (D_\mathrm{p}q^2 + \gamma_\mathrm{p})(D_\mathrm{s}q^2 + \gamma_\mathrm{s})\Big] \\ &- (D_\mathrm{p}q^2 + \gamma_\mathrm{p})\left[k_\mathrm{cat}Fq^2(D^\mathrm{s}_\mathrm{xd} + e_\mathrm{h}D'_\mathrm{e}) - D_\mathrm{e}q^2(k_\mathrm{cat}e_\mathrm{h}F' + D_\mathrm{s}q^2 + \gamma_\mathrm{s})\right] \\ &+ k_\mathrm{cat}FD^\mathrm{p}_\mathrm{xd}q^2(D_\mathrm{s}q^2 + \gamma_\mathrm{s}) = 0\,. \end{aligned} \tag{B.44}$$

According to the sign of the coefficients in the above polynomial only two scenarios for the roots are possible. (i) All roots are negative, (ii) two roots are negative and one is positive. Therefore the product of the three roots can be negative (i) or positive (ii). If we rewrite Eq. (B.44) as $(\sigma - \sigma_1)(\sigma - \sigma_2)(\sigma - \sigma_3)$ then we see that the zeroth order term (the term without σ) in Eq. (B.44) is equal to $-\sigma_1\sigma_2\sigma_3$. Therefore the sign of this

term determines whether the homogeneous solution is positive or negative:

$$\left[\frac{D'_{\mathrm{e}}(s_{\mathrm{h}})}{D_{\mathrm{e}}(s_{\mathrm{h}})}-\frac{F'(s_{\mathrm{h}})}{F(s_{\mathrm{h}})}\right]+\frac{D^{\mathrm{s}}_{\mathrm{xd}}(s_{\mathrm{h}},e_{\mathrm{h}})}{e_{\mathrm{h}}D_{\mathrm{e}}(s_{\mathrm{h}})}-\left[\frac{\gamma_{\mathrm{s}}+D_{\mathrm{s}}q^2}{\gamma_{\mathrm{p}}+D_{\mathrm{p}}q^2}\right]\frac{D^{\mathrm{p}}_{\mathrm{xd}}(s_{\mathrm{h}},e_{\mathrm{h}})}{e_{\mathrm{h}}D_{\mathrm{e}}(s_{\mathrm{h}})}>\frac{\gamma_{\mathrm{s}}+D_{\mathrm{s}}q^2}{k_{\mathrm{cat}}e_{\mathrm{h}}F(s_{\mathrm{h}})}. \quad \text{(B.45)}$$

The inequality (B.45) corresponds to relation 4.14 of Ch. 4 and it is similar to the inequality (B.20) but with a correction due to the product-induced cross-diffusion. Consequently if $D^{\mathrm{p}}_{\mathrm{xd}}=0$, relation (B.45) is the same as relation (B.20). The effect of product interaction on the inequality reduces when γ_{p} increases.

B.4. Simulation parameters

The numerical simulations of the system of Eq. (4.4) have been carried out using COMSOL Multiphysics v5.3. We used the parameters listed in Table B.1 for the simulations shown in Fig. 4.3. We used ${\lambda^{\mathrm{s}}_{\mathrm{f}}}^2={\lambda^{\mathrm{s}}_{\mathrm{c}}}^2=-1\,\text{Å}^2$ for the unstable homogeneous solution (Fig. 4.3a) and ${\lambda^{\mathrm{s}}_{\mathrm{f}}}^2={\lambda^{\mathrm{s}}_{\mathrm{c}}}^2=1\,\text{Å}^2$ for the stable homogeneous solution (Fig. 4.3b). The initial homogeneous concentrations e_{h}, s_{h}, p_{h} were perturbed by adding white Gaussian noise with variance $e_{\mathrm{h}}\cdot 10^{-2}$, $s_{\mathrm{h}}\cdot 10^{-2}$, $p_{\mathrm{h}}\cdot 10^{-2}$ respectively. We considered periodic boundary conditions and used the "time dependent" solver of COMSOL with a relative tolerance of 10^{-6} and an element size of 0.01 µm, i.e. 10^4 lattice points.

For the results shown in Fig. 4.4 we again used the "time dependent" solver of COMSOL. We considered a grid of 20×20 values for ${\lambda^{\mathrm{s}}_{\mathrm{c}}}^2={\lambda^{\mathrm{s}}_{\mathrm{f}}}^2=\lambda^2$ in the interval $[-10,3]\text{Å}^2$ (evenly spaced), and s_{h} in $[10^1,10^5]$µM (evenly spaced on a log-scale). For the COMSOL specific parameters, we used a relative tolerance of 10^{-6} and a maximum element size of 0.02 µm, except when two peaks were observed as a final result of the simulation. In these cases, we repeated the simulations with a finer grid (element size of 0.01 µm) and again observed a single peak for the concentrations as the final result of the simulations. All other parameters were the same as in Table B.1 and the boundary conditions were periodic.

For the simulations shown in Fig. 4.6 we modified the dynamics of the enzyme, following Eq. (4.13). We then considered the same simulation parameters as shown in Table B.1, the same COMSOL specific parameters used for the simulations depicted in Fig. 4.3 and we considered the λ_{p} values shown in Fig. 4.6.

Par.	Description	Value
D_f	Diffusion constant of free enzyme	$10\mu\mathrm{m}^2\,\mathrm{s}^{-1}$
D_c	Diffusion constant of complexed enzyme	$13\mu\mathrm{m}^2\,\mathrm{s}^{-1}$
D_s	Diffusion constant of substrate	$100\mu\mathrm{m}^2\,\mathrm{s}^{-1}$
D_p	Diffusion constant of product	$100\mu\mathrm{m}^2\,\mathrm{s}^{-1}$
k_cat	Catalytic rate	$10^4\mathrm{s}^{-1}$
K_M	K_M of the enzyme	$1\mu\mathrm{M}$
s_h	Homogeneous steady-state substrate concentration	$10^4\mu\mathrm{M}$
s_R	Substrate concentration in the reservoir	$10^6\mu\mathrm{M}$
${\lambda_\mathrm{f}^\mathrm{s}}^2$	Derjaguin length of interaction between substrate and free enzyme	homogeneous: $1\mathring{\mathrm{A}}^2$ pattern: $-1\mathring{\mathrm{A}}^2$
${\lambda_\mathrm{c}^\mathrm{s}}^2$	Derjaguin length of interaction between substrate and complexed enzyme	homogeneous: $1\mathring{\mathrm{A}}^2$ pattern: $-1\mathring{\mathrm{A}}^2$
C_f^s	Corresponding cross-diffusion strength between substrate and free enzyme	homogeneous: $28\mu\mathrm{m}^2\,\mathrm{s}^{-1}\mathrm{M}^{-1}$ pattern: $-28\mu\mathrm{m}^2\,\mathrm{s}^{-1}\mathrm{M}^{-1}$
C_c^s	Corresponding cross-diffusion strength between substrate and complexed enzyme	homogeneous: $28\mu\mathrm{m}^2\,\mathrm{s}^{-1}\mathrm{M}^{-1}$ pattern: $-28\mu\mathrm{m}^2\,\mathrm{s}^{-1}\mathrm{M}^{-1}$
γ_s	Substrate permeation rate	$1\mathrm{s}^{-1}$
γ_p	Product permeation rate	$10\mathrm{s}^{-1}$
η	Viscosity of the solution	$8.9\times10^4\mathrm{Pa\,s}$
T	Temperature of the system	300K
L	Length of the reaction chamber	$100\mu\mathrm{m}$

Table B.1.: Parameters used for the COMSOL simulations shown in Fig. 4.3 of the main text.

List of Figures

List of Tables

Bibliography

[1] W. Kühne. "Über das Verhalten verschiedener organisirter und sog. ungeformter Fermente." In: *Verhandlungen des Naturhistorisch-medicinischen Vereins zu Heidelberg* 1 (Feb. 4, 1874), (3):333–369.

[2] E. Buchner. "Alkoholische Gärung ohne Hefezellen." In: *Berichte der deutschen chemischen Gesellschaft* 30.1 (Jan. 1, 1897), pp. 1110–1113. ISSN: 0365-9496. DOI: 10.1002/cber.189703001215.

[3] J. B. Sumner. "The Isolation and Crystallization of the Enzyme Urease Preliminary Paper." In: *Journal of Biological Chemistry* 69.2 (Aug. 1, 1926), pp. 435–441. ISSN: 0021-9258, 1083-351X.

[4] J. H. Northrop. "Crystalline Pepsin." In: *Science* 69.1796 (May 31, 1929), pp. 580–580. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.69.1796.580.

[5] J. H. Northrop. "The Preparation of Pure Enzymes and Virus Proteins." In: *Nobel Lecture* (Dec. 12, 1946).

[6] B. C. Stark, R. Kole, E. J. Bowman, and S. Altman. "Ribonuclease P: an enzyme with an essential RNA component." In: *Proceedings of the National Academy of Sciences* 75.8 (Aug. 1, 1978), pp. 3717–3721. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.75.8.3717.

[7] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech. "Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena." In: *Cell* 31.1 (Nov. 1, 1982), pp. 147–157. ISSN: 0092-8674, 1097-4172. DOI: 10.1016/0092-8674(82)90414-7.

[8] W. Gilbert. "Origin of life: The RNA world." In: *nature* 319.6055 (1986), pp. 618–618.

[9] G. M. Cooper. "The Central Role of Enzymes as Biological Catalysts." In: *The Cell: A Molecular Approach*. 2nd. Sinauer Associates, 2000. ISBN: 978-0-87893-106-4.

[10] J. M. Berg, J. L. Tymoczko, and L. Stryer. "Free Energy Is a Useful Thermodynamic Function for Understanding Enzymes." In: *Biochemistry. 5th edition*. 5th. W. H. Freeman, 2002.

[11] J. M. Berg, J. L. Tymoczko, and L. Stryer. "Enzymes Accelerate Reactions by Facilitating the Formation of the Transition State." In: *Biochemistry. 5th edition*. 5th. W. H. Freeman, 2002.

[12] R. Wolfenden and M. J. Snider. "The Depth of Chemical Time and the Power of Enzymes as Catalysts." In: *Accounts of Chemical Research* 34.12 (Dec. 1, 2001), pp. 938–945. ISSN: 0001-4842. DOI: 10.1021/ar000058i.

[13] S. Lindskog. "Structure and mechanism of carbonic anhydrase." In: *Pharmacology & Therapeutics* 74.1 (Jan. 1, 1997), pp. 1–20. ISSN: 0163-7258. DOI: 10.1016/S0163-7258(96)00198-2.

[14] L. H. Chen, G. L. Kenyon, F. Curtin, S. Harayama, M. E. Bembenek, G. Hajipour, and C. P. Whitman. "4-Oxalocrotonate tautomerase, an enzyme composed of 62 amino acid residues per monomer." In: *The Journal of Biological Chemistry* 267.25 (Sept. 5, 1992), pp. 17716–17721. ISSN: 0021-9258.

[15] C. P. Whitman. "The 4-oxalocrotonate tautomerase family of enzymes: how nature makes new enzymes using a β–α–β structural motif." In: *Archives of Biochemistry and Biophysics* 402.1 (June 1, 2002), pp. 1–13. ISSN: 0003-9861. DOI: 10.1016/S0003-9861(02)00052-8.

[16] J. A. Kelly, A. R. Sielecki, B. D. Sykes, M. N. James, and D. C. Phillips. "X-ray crystallography of the binding of the bacterial cell wall trisaccharide NAM-NAG-NAM to lysozyme." In: *Nature* 282.5741 (Dec. 20, 1979), pp. 875–878. ISSN: 0028-0836. DOI: 10.1038/282875a0.

[17] J. M. Berg, J. L. Tymoczko, and L. Stryer. "Enzymes Are Powerful and Highly Specific Catalysts." In: *Biochemistry. 5th edition*. 5th. W. H. Freeman, 2002.

[18] E. Fischer. "Einfluss der Configuration auf die Wirkung der Enzyme." In: *Berichte der deutschen chemischen Gesellschaft* 27.3 (Oct. 1, 1894), pp. 2985–2993. ISSN: 0365-9496. DOI: 10.1002/cber.18940270364.

[19] D. E. Koshland. "Application of a Theory of Enzyme Specificity to Protein Synthesis." In: *Proceedings of the National Academy of Sciences* 44.2 (Feb. 1, 1958), pp. 98–104. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.44.2.98.

[20] Y. Savir and T. Tlusty. "Conformational Proofreading: The Impact of Conformational Changes on the Specificity of Molecular Recognition." In: *PLOS ONE* 2.5 (May 23, 2007), e468. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0000468.

[21] H. Nishimasu, S. Fushinobu, H. Shoun, and T. Wakagi. "Crystal Structures of an ATP-dependent Hexokinase with Broad Substrate Specificity from the Hyperthermophilic Archaeon Sulfolobus tokodaii." In: *Journal of Biological Chemistry* 282.13 (Mar. 30, 2007), pp. 9923–9931. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.M610678200.

[22] F. Zeller and M. Zacharias. "Substrate Binding Specifically Modulates Domain Arrangements in Adenylate Kinase." English. In: *Biophysical Journal* 109.9 (Nov. 2015), pp. 1978–1985. ISSN: 0006-3495. DOI: 10.1016/j.bpj.2015.08.049.

[23] B. Pelz, G. Žoldák, F. Zeller, M. Zacharias, and M. Rief. "Subnanometre enzyme mechanics probed by single-molecule force spectroscopy." en. In: *Nature Communications* 7.1 (Feb. 2016), p. 10848. ISSN: 2041-1723. DOI: 10.1038/ncomms10848.

[24] A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, and M. H. M. Olsson. "Electrostatic Basis for Enzyme Catalysis." In: *Chemical Reviews* 106.8 (Aug. 1, 2006), pp. 3210–3235. ISSN: 0009-2665. DOI: 10.1021/cr0503106.

[25] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. John Wiley & Sons, Feb. 22, 2013. 609 pp. ISBN: 978-3-527-66549-5.

[26] L. Michaelis and M. L. Menten. "Die kinetik der invertinwirkung." In: *Biochemische Zeitschrift* 49 (Feb. 4, 1913), pp. 333–369.

[27] C. M. Guldberg and P. Waage. *Etudes sur les affinités chimiques*. Brøgger & Christie, 1867.

[28] G. E. Briggs and J. B. S. Haldane. "A Note on the Kinetics of Enzyme Action." In: *Biochemical Journal* 19.2 (1925), pp. 338–339. ISSN: 0264-6021.

[29] J. M. Berg, J. L. Tymoczko, and L. Stryer. "The Michaelis-Menten Model Accounts for the Kinetic Properties of Many Enzymes." In: *Biochemistry. 5th edition*. 5th. W. H. Freeman, 2002.

[30] A. Bar-Even, E. Noor, Y. Savir, W. Liebermeister, D. Davidi, D. S. Tawfik, and R. Milo. "The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters." In: *Biochemistry* 50.21 (May 31, 2011), pp. 4402–4410. ISSN: 0006-2960. DOI: 10.1021/bi2002289.

[31] H. P. Lu, L. Xun, and X. S. Xie. "Single-Molecule Enzymatic Dynamics." en. In: *Science* 282.5395 (Dec. 1998), pp. 1877–1882. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.282.5395.1877.

[32] B. P. English, W. Min, A. M. van Oijen, K. T. Lee, G. Luo, H. Sun, B. J. Cherayil, S. C. Kou, and X. S. Xie. "Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited." en. In: *Nature Chemical Biology* 2.2 (Feb. 2006), pp. 87–94. ISSN: 1552-4469. DOI: 10.1038/nchembio759.

[33] R. Benesch and R. E. Benesch. "The effect of organic phosphates from the human erythrocyte on the allosteric properties of hemoglobin." In: *Biochemical and Biophysical Research Communications* 26.2 (Jan. 23, 1967), pp. 162–167. ISSN: 0006-291X. DOI: 10.1016/0006-291X(67)90228-8.

[34] J. Monod, J.-P. Changeux, and F. Jacob. "Allosteric proteins and cellular control systems." In: *Journal of Molecular Biology* 6.4 (Apr. 1, 1963), pp. 306–329. ISSN: 0022-2836. DOI: 10.1016/S0022-2836(63)80091-1.

[35] A. V. Hill. "The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves." In: *The Journal of Physiology* 40 (1910), pp. 4–7. DOI: 10.1113/jphysiol.1910.sp001386.

[36] S. J. Pollack, J. R. Atack, M. R. Knowles, G. McAllister, C. I. Ragan, R. Baker, S. R. Fletcher, L. L. Iversen, and H. B. Broughton. "Mechanism of inositol monophosphatase, the putative target of lithium therapy." In: *Proceedings of the National Academy of Sciences* 91.13 (June 21, 1994), p. 5766. DOI: 10.1073/pnas.91.13.5766.

[37] E. M. Wise and J. T. Park. "Penicillin: its basic site of action as an inhibitor of a peptide cross-linking reaction in cell wall mucopeptide synthesis." In: *Proceedings of the National Academy of Sciences* 54.1 (July 1, 1965), pp. 75–81. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.54.1.75.

[38] A. Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen." en. In: *Annalen der Physik* 322.8 (1905), pp. 549–560. ISSN: 1521-3889. DOI: 10.1002/andp.19053220806.

[39] M. v. Smoluchowski. "Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen." In: *Zeitschrift für Physikalische Chemie* 92U.1 (Nov. 1, 1918), pp. 129–168. ISSN: 2196-7156, 0942-9352. DOI: 10.1515/zpch-1918-9209.

[40] F. C. Collins and G. E. Kimball. "Diffusion-controlled reaction rates." In: *Journal of Colloid Science* 4.4 (Aug. 1, 1949), pp. 425–437. ISSN: 0095-8522. DOI: 10.1016/0095-8522(49)90023-9.

[41] J. M. Schurr. "The Role of Diffusion in Bimolecular Solution Kinetics." In: *Biophysical Journal* 10.8 (Aug. 1, 1970), pp. 700–716. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(70)86330-5.

[42] D. Shoup and A. Szabo. "Role of diffusion in ligand binding to macromolecules and cell-bound receptors." In: *Biophysical Journal* 40.1 (Oct. 1, 1982), pp. 33–39. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(82)84455-X.

[43] D. Shoup, G. Lipari, and A. Szabo. "Diffusion-controlled bimolecular reaction rates. The effect of rotational diffusion and orientation constraints." In: *Biophysical Journal* 36.3 (Dec. 1981), pp. 697–714. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(81)84759-5.

[44] P. Debye. "Reaction Rates in Ionic Solutions." In: *Transactions of the Electrochemical Society* 82.1 (Oct. 1, 1942), p. 265. ISSN: 1945-6859. DOI: 10.1149/1.3071413.

[45] A. D. Riggs, S. Bourgeois, and M. Cohn. "The lac represser-operator interaction: III. Kinetic studies." In: *Journal of Molecular Biology* 53.3 (Nov. 14, 1970), pp. 401–417. ISSN: 0022-2836. DOI: 10.1016/0022-2836(70)90074-4.

[46] R. A. Alberty and G. G. Hammes. "Application of the Theory of Diffusion-controlled Reactions to Enzyme Kinetics." In: *The Journal of Physical Chemistry* 62.2 (Feb. 1, 1958), pp. 154–159. ISSN: 0022-3654. DOI: 10.1021/j150560a005.

[47] S. H. Koenig and R. D. Brown. "H2CO3 as Substrate for Carbonic Anhydrase in the Dehydration of HCO3-." In: *Proceedings of the National Academy of Sciences* 69.9 (Sept. 1, 1972), pp. 2422–2425. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.69.9.2422.

[48] E. W. Miles, S. Rhee, and D. R. Davies. "The Molecular Basis of Substrate Channeling." In: *Journal of Biological Chemistry* 274.18 (Apr. 30, 1999), pp. 12193–12196. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.274.18.12193.

[49] H. O. Spivey and J. Ovádi. "Substrate Channeling." In: *Methods* 19.2 (Oct. 1, 1999), pp. 306–321. ISSN: 1046-2023. DOI: 10.1006/meth.1999.0858.

[50] I. Wheeldon, S. D. Minteer, S. Banta, S. C. Barton, P. Atanassov, and M. Sigman. "Substrate channelling as an approach to cascade reactions." In: *Nature Chemistry* 8.4 (Apr. 2016), pp. 299–309. ISSN: 1755-4349. DOI: 10.1038/nchem.2459.

[51] L. J. Sweetlove and A. R. Fernie. "The role of dynamic enzyme assemblies and substrate channelling in metabolic regulation." en. In: *Nature Communications* 9.1 (May 2018), p. 2136. ISSN: 2041-1723. DOI: 10.1038/s41467-018-04543-8.

[52] P. A. Srere. "The metabolon." In: *Trends in Biochemical Sciences* 10.3 (Mar. 1, 1985), pp. 109–110. ISSN: 0968-0004. DOI: 10.1016/0968-0004(85)90266-X.

[53] J. Ovádi. "Physiological significance of metabolic channelling." In: *Journal of Theoretical Biology* 152.1 (Sept. 7, 1991), pp. 1–22. ISSN: 0022-5193. DOI: 10.1016/S0022-5193(05)80500-4.

[54] T. R. Manney. "Physiological Advantage of the Mechanism of the Tryptophan Synthetase Reaction." In: *Journal of Bacteriology* 102.2 (May 1970), pp. 483–488. ISSN: 0021-9193.

[55] B. D. Davis. "On the importance of being ionized." In: *Archives of Biochemistry and Biophysics* 78.2 (Dec. 1, 1958), pp. 497–509. ISSN: 0003-9861. DOI: 10.1016/0003-9861(58)90374-6.

[56] J. M. Krahn, J. H. Kim, M. R. Burns, R. J. Parry, H. Zalkin, and J. L. Smith. "Coupled Formation of an Amidotransferase Interdomain Ammonia Channel and a Phosphoribosyltransferase Active Site," in: *Biochemistry* 36.37 (Sept. 1, 1997), pp. 11061–11068. ISSN: 0006-2960. DOI: 10.1021/bi9714114.

[57] E. M. Sampson and T. A. Bobik. "Microcompartments for B12-dependent 1,2-propanediol degradation provide protection from DNA and cellular damage by a reactive metabolic intermediate." In: *Journal of Bacteriology* 190.8 (Apr. 2008), pp. 2966–2971. ISSN: 1098-5530. DOI: 10.1128/JB.01925-07.

[58] E. L. Maynard and P. A. Lindahl. "Catalytic Coupling of the Active Sites in Acetyl-CoA Synthase, a Bifunctional CO-Channeling Enzyme." In: *Biochemistry* 40.44 (Nov. 1, 2001), pp. 13262–13267. ISSN: 0006-2960. DOI: 10.1021/bi015604+.

[59] M. Castellana, M. Z. Wilson, Y. Xu, P. Joshi, I. M. Cristea, J. D. Rabinowitz, Z. Gitai, and N. S. Wingreen. "Enzyme clustering accelerates processing of intermediates through metabolic channeling." In: *Nature Biotechnology* 32.10 (Oct. 2014), pp. 1011–1018. ISSN: 1546-1696. DOI: 10.1038/nbt.3018.

[60] F. Wu and S. Minteer. "Krebs Cycle Metabolon: Structural Evidence of Substrate Channeling Revealed by Cross-Linking and Mass Spectrometry." In: *Angewandte Chemie International Edition* 54.6 (2015), pp. 1851–1854. ISSN: 1521-3773. DOI: 10.1002/anie.201409336.

[61] J. Müller and C. M. Niemeyer. "DNA-directed assembly of artificial multienzyme complexes." In: *Biochemical and Biophysical Research Communications* 377.1 (Dec. 5, 2008), pp. 62–67. ISSN: 0006-291X. DOI: 10.1016/j.bbrc.2008.09.078.

[62] O. I. Wilner, Y. Weizmann, R. Gill, O. Lioubashevski, R. Freeman, and I. Willner. "Enzyme cascades activated on topologically programmed DNA scaffolds." In: *Nature Nanotechnology* 4.4 (Apr. 2009), pp. 249–254. ISSN: 1748-3395. DOI: 10.1038/nnano.2009.50.

[63] J. Fu, M. Liu, Y. Liu, N. W. Woodbury, and H. Yan. "Interenzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures." In: *Journal of the American Chemical Society* 134.12 (Mar. 28, 2012), pp. 5516–5519. ISSN: 0002-7863. DOI: 10.1021/ja300897h.

[64] Y. Zhang, S. Tsitkov, and H. Hess. "Proximity does not contribute to activity enhancement in the glucose oxidase–horseradish peroxidase cascade." In: *Nature Communications* 7.1 (Dec. 22, 2016), p. 13982. ISSN: 2041-1723. DOI: 10.1038/ncomms13982.

[65] P. Bauler, G. Huber, T. Leyh, and J. A. McCammon. "Channeling by Proximity: The Catalytic Advantages of Active Site Colocalization Using Brownian Dynamics." In: *The Journal of Physical Chemistry Letters* 1.9 (May 6, 2010), pp. 1332–1335. DOI: 10.1021/jz1002007.

[66] F. Hinzpeter, F. Tostevin, A. Buchner, and U. Gerland. "Trade-offs and design principles in the spatial organization of catalytic particles." en. In: *bioRxiv* (June 2020), p. 2020.06.14.146076. DOI: 10.1101/2020.06.14.146076.

[67] P. A. Srere. "Complexes of Sequential Metabolic Enzymes." In: *Annual Review of Biochemistry* 56.1 (1987), pp. 89–124. DOI: 10.1146/annurev.bi.56.070187.000513.

[68] C. C. Hyde, S. A. Ahmed, E. A. Padlan, E. W. Miles, and D. R. Davies. "Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium." In: *Journal of Biological Chemistry* 263.33 (Nov. 25, 1988), pp. 17857–17871. ISSN: 0021-9258, 1083-351X.

[69] E. J. Brignole, S. Smith, and F. J. Asturias. "Conformational flexibility of metazoan fatty acid synthase enables catalysis." en. In: *Nature Structural & Molecular Biology* 16.2 (Feb. 2009), pp. 190–197. ISSN: 1545-9985. DOI: 10.1038/nsmb.1532.

[70] K. S. Anderson, E. W. Miles, and K. A. Johnson. "Serine modulates substrate channeling in tryptophan synthase. A novel intersubunit triggering mechanism." In: *Journal of Biological Chemistry* 266.13 (May 5, 1991), pp. 8020–8033. ISSN: 0021-9258, 1083-351X.

[71] P. M. Anderson and A. Meister. "Bicarbonate-Dependent Cleavage of Adenosine Triphosphate and Other Reactions Catalyzed by Escherichia coli Carbamyl Phosphate Synthetase." In: *Biochemistry* 5.10 (Oct. 1, 1966), pp. 3157–3163. ISSN: 0006-2960. DOI: 10.1021/bi00874a012.

[72] J. B. Thoden, H. M. Holden, G. Wesenberg, F. M. Raushel, and I. Rayment. "Structure of carbamoyl phosphate synthetase: a journey of 96 Å from substrate to product." In: *Biochemistry* 36.21 (1997), pp. 6305–6316.

[73] A. H. Elcock and J. A. McCammon. "Evidence for Electrostatic Channeling in a Fusion Protein of Malate Dehydrogenase and Citrate Synthase." In: *Biochemistry* 35.39 (Jan. 1996), pp. 12652–12658. ISSN: 0006-2960. DOI: 10.1021/bi9614747.

[74] R. W. Guynn, H. J. Gelberg, and R. L. Veech. "Equilibrium Constants of the Malate Dehydrogenase, Citrate Synthase, Citrate Lyase, and Acetyl Coenzyme A Hydrolysis Reactions under Physiological Conditions." en. In: *Journal of Biological Chemistry* 248.20 (Oct. 1973), pp. 6957–6965. ISSN: 0021-9258, 1083-351X.

[75] D. R. Knighton, C.-C. Kan, E. Howland, C. A. Janson, Z. Hostomska, K. M. Welsh, and D. A. Matthews. "Structure of and kinetic channelling in bifunctional dihydrofolate reductase–thymidylate synthase." en. In: *Nature Structural Biology* 1.3 (Mar. 1994), pp. 186–194. ISSN: 1545-9985. DOI: 10.1038/nsb0394-186.

[76] R. N. Perham. "Swinging Arms and Swinging Domains in Multifunctional Enzymes: Catalytic Machines for Multistep Reactions." In: *Annual Review of Biochemistry* 69.1 (2000), pp. 961–1004. DOI: 10.1146/annurev.biochem.69.1.961.

[77] Z. H. Zhou, D. B. McCarthy, C. M. O'Connor, L. J. Reed, and J. K. Stoops. "The remarkable structural and functional organization of the eukaryotic pyruvate dehydrogenase complexes." en. In: *Proceedings of the National Academy of Sciences* 98.26 (Dec. 2001), pp. 14802–14807. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.011597698.

[78] M. Smolle, A. E. Prior, A. E. Brown, A. Cooper, O. Byron, and J. G. Lindsay. "A New Level of Architectural Complexity in the Human Pyruvate Dehydrogenase Complex." en. In: *Journal of Biological Chemistry* 281.28 (July 2006), pp. 19772–19780. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.M601140200.

[79] F. Hinzpeter, F. Tostevin, and U. Gerland. "Regulation of reaction fluxes via enzyme sequestration and co-clustering." In: *Journal of The Royal Society Interface* 16.156 (July 2019), p. 20190444. DOI: 10.1098/rsif.2019.0444.

[80] S. An, R. Kumar, E. D. Sheets, and S. J. Benkovic. "Reversible Compartmentalization of de Novo Purine Biosynthetic Complexes in Living Cells." en. In: *Science* 320.5872 (Apr. 2008), pp. 103–106. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1152241.

[81] Y. Deng, J. Gam, J. B. French, H. Zhao, S. An, and S. J. Benkovic. "Mapping Protein-Protein Proximity in the Purinosome." en. In: *Journal of Biological Chemistry* 287.43 (Oct. 2012), pp. 36201–36207. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.M112.407056.

[82] J. B. French, H. Zhao, S. An, S. Niessen, Y. Deng, B. F. Cravatt, and S. J. Benkovic. "Hsp70/Hsp90 chaperone machinery is involved in the assembly of the purinosome." en. In: *Proceedings of the National Academy of Sciences* 110.7 (Feb. 2013), pp. 2528–2533. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1300173110.

[83] F. Verrier, S. An, A. M. Ferrie, H. Sun, M. Kyoung, H. Deng, Y. Fang, and S. J. Benkovic. "GPCRs regulate the assembly of a multienzyme complex for purine biosynthesis." en. In: *Nature Chemical Biology* 7.12 (Dec. 2011), pp. 909–915. ISSN: 1552-4469. DOI: 10.1038/nchembio.690.

[84] H. Zhao, C. R. Chiaro, L. Zhang, P. B. Smith, C. Y. Chan, A. M. Pedley, R. J. Pugh, J. B. French, A. D. Patterson, and S. J. Benkovic. "Quantitative Analysis of Purine Nucleotides Indicates That Purinosomes Increase de Novo Purine Biosynthesis." en. In: *Journal of Biological Chemistry* 290.11 (Mar. 2015), pp. 6705–6713. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.M114.628701.

[85] N. Miura, M. Shinohara, Y. Tatsukami, Y. Sato, H. Morisaka, K. Kuroda, and M. Ueda. "Spatial Reorganization of Saccharomyces cerevisiae Enolase To Alter Carbon Metabolism under Hypoxia." en. In: *Eukaryotic Cell* 12.8 (Aug. 2013), pp. 1106–1119. ISSN: 1535-9778, 1535-9786. DOI: 10.1128/EC.00093-13.

[86] M. Jin, G. G. Fuller, T. Han, Y. Yao, A. F. Alessi, M. A. Freeberg, N. P. Roach, J. J. Moresco, A. Karnovsky, M. Baba, J. R. Yates, A. D. Gitler, K. Inoki, D. J. Klionsky, and J. K. Kim. "Glycolytic Enzymes Coalesce in G Bodies under Hypoxic Stress." en. In: *Cell Reports* 20.4 (July 2017), pp. 895–908. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2017.06.082.

[87] R. Narayanaswamy, M. Levy, M. Tsechansky, G. M. Stovall, J. D. O'Connell, J. Mirrielees, A. D. Ellington, and E. M. Marcotte. "Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation." en. In: *Proceedings of the National Academy of Sciences* 106.25 (June 2009), pp. 10147–10152. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0812771106.

[88] S. Jang, J. C. Nelson, E. G. Bend, L. Rodríguez-Laureano, F. G. Tueros, L. Cartagenova, K. Underwood, E. M. Jorgensen, and D. A. Colón-Ramos. "Glycolytic Enzymes Localize to Synapses under Energy Stress to Support Synaptic Function." en. In: *Neuron* 90.2 (Apr. 2016), pp. 278–291. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2016.03.011.

[89] M. E. Campanella, H. Chu, and P. S. Low. "Assembly and regulation of a glycolytic enzyme complex on the human erythrocyte membrane." en. In: *Proceedings of the National Academy of Sciences* 102.7 (Feb. 2005), pp. 2402–2407. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.0409741102.

[90] E. Puchulu-Campanella, H. Chu, D. J. Anstee, J. A. Galan, W. A. Tao, and P. S. Low. "Identification of the Components of a Glycolytic Enzyme Metabolon on the Human Red Blood Cell Membrane." In: *The Journal of Biological Chemistry* 288.2 (Jan. 2013), pp. 848–858. ISSN: 0021-9258. DOI: 10.1074/jbc.M112.428573.

[91] X. Zhao, H. Palacci, V. Yadav, M. M. Spiering, M. K. Gilson, P. J. Butler, H. Hess, S. J. Benkovic, and A. Sen. "Substrate-driven chemotactic assembly in an enzyme cascade." en. In: *Nature Chemistry* 10.3 (Mar. 2018), pp. 311–317. ISSN: 1755-4349. DOI: 10.1038/nchem.2905.

[92] Y. Zhang, K. F. M. Beard, C. Swart, S. Bergmann, I. Krahnert, Z. Nikoloski, A. Graf, R. G. Ratcliffe, L. J. Sweetlove, A. R. Fernie, and T. Obata. "Protein-protein interactions and metabolite channelling in the plant tricarboxylic acid cycle." In: *Nature Communications* 8 (May 2017). ISSN: 2041-1723. DOI: 10.1038/ncomms15212.

[93] T. Laursen, J. Borch, C. Knudsen, K. Bavishi, F. Torta, H. J. Martens, D. Silvestro, N. S. Hatzakis, M. R. Wenk, T. R. Dafforn, C. E. Olsen, M. S. Motawia, B. Hamberger, B. L. Møller, and J.-E. Bassard. "Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum." en. In: *Science* 354.6314 (Nov. 2016), pp. 890–893. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aag2347.

[94] K. Jørgensen, A. V. Rasmussen, M. Morant, A. H. Nielsen, N. Bjarnholt, M. Zagrobelny, S. Bak, and B. L. Møller. "Metabolon formation and metabolic channeling in the biosynthesis of plant natural products." en. In: *Current Opinion in Plant Biology*. Physiology andmetabolism 8.3 (June 2005), pp. 280–291. ISSN: 1369-5266. DOI: 10.1016/j.pbi.2005.03.014.

[95] G. M. Cooper. "Mitochondria." en. In: *The Cell: A Molecular Approach. 2nd edition* (2000).

[96] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. "The Compartmentalization of Cells." en. In: *Molecular Biology of the Cell. 4th edition* (2002).

[97] G. M. Cooper. "Peroxisomes." en. In: *The Cell: A Molecular Approach. 2nd edition* (2000).

[98] F. A. G. Reubsaet, J. H. Veerkamp, M. L. P. Brückwilder, J. M. F. Trijbels, and L. A. H. Monnens. "Peroxisomal oxidases and catalase in liver and kidney homogenates of normal and di(ethylhexyl)phthalate-fed rats." en. In: *International Journal of Biochemistry* 23.9 (Jan. 1991), pp. 961–967. ISSN: 0020-711X. DOI: 10.1016/0020-711X(91)90086-3.

[99] M. R. Badger and G. D. Price. "CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution." en. In: *Journal of Experimental Botany* 54.383 (Feb. 2003), pp. 609–622. ISSN: 0022-0957. DOI: 10.1093/jxb/erg076.

[100] Z. Dou, S. Heinhorst, E. B. Williams, C. D. Murin, J. M. Shively, and G. C. Cannon. "CO2 Fixation Kinetics of Halothiobacillus neapolitanus Mutant Carboxysomes Lacking Carbonic Anhydrase Suggest the Shell Acts as a Diffusional Barrier for CO2." en. In: *Journal of Biological Chemistry* 283.16 (Apr. 2008), pp. 10377–10384. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.M709285200.

[101] F. Cai, B. B. Menon, G. C. Cannon, K. J. Curry, J. M. Shively, and S. Heinhorst. "The Pentameric Vertex Proteins Are Necessary for the Icosahedral Carboxysome Shell to Function as a CO2 Leakage Barrier." en. In: *PLOS ONE* 4.10 (Oct. 2009), e7521. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0007521.

[102] F. Hinzpeter, U. Gerland, and F. Tostevin. "Optimal Compartmentalization Strategies for Metabolic Microcompartments." English. In: *Biophysical Journal* 112.4 (Feb. 2017), pp. 767–779. ISSN: 0006-3495. DOI: 10.1016/j.bpj.2016.11.3194.

[103] S. R. Brinsmade, T. Paldon, and J. C. Escalante-Semerena. "Minimal Functions and Physiological Conditions Required for Growth of Salmonella enterica on Ethanolamine in the Absence of the Metabolosome." In: *Journal of Bacteriology* 187.23 (Dec. 2005), pp. 8039–8046. ISSN: 0021-9193. DOI: 10.1128/JB.187.23.8039-8046.2005.

[104] T. F. Moraes and R. A. F. Reithmeier. "Membrane transport metabolons." en. In: *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1818.11 (Nov. 2012), pp. 2687–2706. ISSN: 0005-2736. DOI: 10.1016/j.bbamem.2012.06.007.

[105] C. M. Fontes and H. J. Gilbert. "Cellulosomes: Highly Efficient Nanomachines Designed to Deconstruct Plant Cell Wall Complex Carbohydrates." In: *Annual Review of Biochemistry* 79.1 (June 2010), pp. 655–681. ISSN: 0066-4154. DOI: 10.1146/annurev-biochem-091208-085603.

[106] J. B. French, S. A. Jones, H. Deng, A. M. Pedley, D. Kim, C. Y. Chan, H. Hu, R. J. Pugh, H. Zhao, Y. Zhang, T. J. Huang, Y. Fang, X. Zhuang, and S. J. Benkovic. "Spatial colocalization and functional link of purinosomes with mitochondria." en. In: *Science* 351.6274 (Feb. 2016), pp. 733–737. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aac6054.

[107] H. Chu, E. Puchulu-Campanella, J. A. Galan, W. A. Tao, P. S. Low, and J. F. Hoffman. "Identification of cytoskeletal elements enclosing the ATP pools that fuel human red blood cell membrane cation pumps." en. In: *Proceedings of the National Academy of Sciences* 109.31 (July 2012), pp. 12794–12799. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1209014109.

[108] R. H. Doi and A. Kosugi. "Cellulosomes: plant-cell-wall-degrading enzyme complexes." en. In: *Nature Reviews Microbiology* 2.7 (July 2004), pp. 541–551. ISSN: 1740-1534. DOI: 10.1038/nrmicro925.

[109] E. A. Bayer, J.-P. Belaich, Y. Shoham, and R. Lamed. "The Cellulosomes: Multienzyme Machines for Degradation of Plant Cell Wall Polysaccharides." In: *Annual Review of Microbiology* 58.1 (Sept. 2004), pp. 521–554. ISSN: 0066-4227. DOI: 10.1146/annurev.micro.57.030502.091022.

[110] S. John, J. N. Weiss, and B. Ribalet. "Subcellular Localization of Hexokinases I and II Directs the Metabolic Fate of Glucose." en. In: *PLOS ONE* 6.3 (Mar. 2011), e17674. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0017674.

[111] J. W. A. Graham, T. C. R. Williams, M. Morgan, A. R. Fernie, R. G. Ratcliffe, and L. J. Sweetlove. "Glycolytic Enzymes Associate Dynamically with Mitochondria in Response to Respiratory Demand and Support Substrate Channeling." en. In: *The Plant Cell* 19.11 (Nov. 2007), pp. 3723–3738. ISSN: 1040-4651, 1532-298X. DOI: 10.1105/tpc.107.053371.

[112] S. An, Y. Deng, J. W. Tomsho, M. Kyoung, and S. J. Benkovic. "Microtubule-assisted mechanism for functional metabolic macromolecular complex formation." en. In: *Proceedings of the National Academy of Sciences* 107.29 (July 2010), pp. 12872–12876. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1008451107.

[113] C. Y. Chan, A. M. Pedley, D. Kim, C. Xia, X. Zhuang, and S. J. Benkovic. "Microtubule-directed transport of purine metabolons drives their cytosolic transit to mitochondria." en. In: *Proceedings of the National Academy of Sciences* 115.51 (Dec. 2018), pp. 13009–13014. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1814042115.

[114] V. Pareek, H. Tian, N. Winograd, and S. J. Benkovic. "Metabolomics and mass spectrometry imaging reveal channeled de novo purine synthesis in cells." en. In: *Science* 368.6488 (Apr. 2020), pp. 283–290. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aaz6465.

[115] H. G and W. Gj. "Metabolic pathways as enzyme complexes: evidence for the synthesis of phenylpropanoids and flavonoids on membrane associated enzyme complexes." English. In: *Archives of Biochemistry and Biophysics* 237.1 (Feb. 1985), pp. 88–100. ISSN: 0003-9861, 1096-0384. DOI: 10.1016/0003-9861(85)90257-7.

[116] B. Winkel-Shirley. "Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways." en. In: *Physiologia Plantarum* 107.1 (1999), pp. 142–149. ISSN: 1399-3054. DOI: 10.1034/j.1399-3054.1999.100119.x.

[117] B. S. Winkel. "Metabolic channeling in plants." In: *Annual Review of Plant Biology* 55.1 (Apr. 2004), pp. 85–107. ISSN: 1543-5008. DOI: 10.1146/annurev.arplant.55.031903.141714.

[118] L. J. Sweetlove and A. R. Fernie. "The Spatial Organization of Metabolism Within the Plant Cell." In: *Annual Review of Plant Biology* 64.1 (Apr. 2013), pp. 723–746. ISSN: 1543-5008. DOI: 10.1146/annurev-arplant-050312-120233.

[119] C. M. Agapakis, P. M. Boyle, and P. A. Silver. "Natural strategies for the spatial optimization of metabolism in synthetic biology." en. In: *Nature Chemical Biology* 8.6 (June 2012), pp. 527–535. ISSN: 1552-4469. DOI: 10.1038/nchembio.975.

[120] H. Rabeharindranto, S. Castaño-Cerezo, T. Lautier, L. F. Garcia-Alles, C. Treitz, A. Tholey, and G. Truan. "Enzyme-fusion strategies for redirecting and improving carotenoid synthesis in S. cerevisiae." en. In: *Metabolic Engineering Communications* 8 (June 2019), e00086. ISSN: 2214-0301. DOI: 10.1016/j.mec.2019.e00086.

[121] T. Thomik, I. Wittig, J.-y. Choe, E. Boles, and M. Oreb. "An artificial transport metabolon facilitates improved substrate utilization in yeast." en. In: *Nature Chemical Biology* 13.11 (Nov. 2017), pp. 1158–1163. ISSN: 1552-4469. DOI: 10.1038/nchembio.2457.

[122] S. Moraïs, Y. Barak, J. Caspi, Y. Hadar, R. Lamed, Y. Shoham, D. B. Wilson, and E. A. Bayer. "Contribution of a Xylan-Binding Module to the Degradation of a Complex Cellulosic Substrate by Designer Cellulosomes." en. In: *Applied and Environmental Microbiology* 76.12 (June 2010), pp. 3787–3796. ISSN: 0099-2240, 1098-5336. DOI: 10.1128/AEM.00266-10.

[123] S.-L. Tsai, J. Oh, S. Singh, R. Chen, and W. Chen. "Functional Assembly of Minicellulosomes on the Saccharomyces cerevisiae Cell Surface for Cellulose Hydrolysis and Ethanol Production." en. In: *Applied and Environmental Microbiology* 75.19 (Oct. 2009), pp. 6087–6093. ISSN: 0099-2240, 1098-5336. DOI: 10.1128/AEM.01538-09.

[124] S.-L. Tsai, N. A. DaSilva, and W. Chen. "Functional Display of Complex Cellulosomes on the Yeast Surface via Adaptive Assembly." In: *ACS Synthetic Biology* 2.1 (Jan. 2013), pp. 14–21. DOI: 10.1021/sb300047u.

[125] F. Liu, S. Banta, and W. Chen. "Functional assembly of a multi-enzyme methanol oxidation cascade on a surface-displayed trifunctional scaffold for enhanced NADH production." en. In: *Chemical Communications* 49.36 (2013), pp. 3766–3768. DOI: 10.1039/C3CC40454D.

[126] J.-L. Lin, J. Zhu, and I. Wheeldon. "Synthetic Protein Scaffolds for Biosynthetic Pathway Colocalization on Lipid Droplet Membranes." In: *ACS Synthetic Biology* 6.8 (Aug. 2017), pp. 1534–1544. DOI: 10.1021/acssynbio.7b00041.

[127] C. J. Delebecque, A. B. Lindner, P. A. Silver, and F. A. Aldaye. "Organization of Intracellular Reactions with Rationally Designed RNA Assemblies." en. In: *Science* 333.6041 (July 2011), pp. 470–474. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1206938.

[128] J. E. Dueber, G. C. Wu, G. R. Malmirchegini, T. S. Moon, C. J. Petzold, A. V. Ullal, K. L. J. Prather, and J. D. Keasling. "Synthetic protein scaffolds provide modular control over metabolic flux." en. In: *Nature Biotechnology* 27.8 (Aug. 2009), pp. 753–759. ISSN: 1546-1696. DOI: 10.1038/nbt.1557.

[129] Y. Zhang, S.-Z. Li, J. Li, X. Pan, R. E. Cahoon, J. G. Jaworski, X. Wang, J. M. Jez, F. Chen, and O. Yu. "Using Unnatural Protein Fusions to Engineer Resveratrol Biosynthesis in Yeast and Mammalian Cells." In: *Journal of the American Chemical Society* 128.40 (Oct. 2006), pp. 13030–13031. ISSN: 0002-7863. DOI: 10.1021/ja0622094.

[130] R. J. Conrado, G. C. Wu, J. T. Boock, H. Xu, S. Y. Chen, T. Lebar, J. Turnšek, N. Tomšič, M. Avbelj, R. Gaber, T. Koprivnjak, J. Mori, V. Glavnik, I. Vovk, M. Benčina, V. Hodnik, G. Anderluh, J. E. Dueber, R. Jerala, and M. P. DeLisa. "DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency." en. In: *Nucleic Acids Research* 40.4 (Feb. 2012), pp. 1879–1889. ISSN: 0305-1048. DOI: 10.1093/nar/gkr888.

[131] P. W. K. Rothemund. "Folding DNA to create nanoscale shapes and patterns." en. In: *Nature* 440.7082 (Mar. 2006), pp. 297–302. ISSN: 1476-4687. DOI: 10.1038/nature04586.

[132] Y. Fu, D. Zeng, J. Chao, Y. Jin, Z. Zhang, H. Liu, D. Li, H. Ma, Q. Huang, K. V. Gothelf, and C. Fan. "Single-Step Rapid Assembly of DNA Origami Nanostructures for Addressable Nanoscale Bioreactors." In: *Journal of the American Chemical Society* 135.2 (Jan. 2013), pp. 696–702. ISSN: 0002-7863. DOI: 10.1021/ja3076692.

[133] V. Linko, M. Eerikäinen, and M. A. Kostiainen. "A modular DNA origami-based enzyme cascade nanoreactor." en. In: *Chemical Communications* 51.25 (2015), pp. 5351–5354. DOI: 10.1039/C4CC08472A.

[134] J. Fu, Y. R. Yang, A. Johnson-Buck, M. Liu, Y. Liu, N. G. Walter, N. W. Woodbury, and H. Yan. "Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm." en. In: *Nature Nanotechnology* 9.7 (July 2014), pp. 531–536. ISSN: 1748-3395. DOI: 10.1038/nnano.2014.100.

[135] M. Feng and M. K. Gilson. "Enhanced Diffusion and Chemotaxis of Enzymes." In: *Annual Review of Biophysics* 49.1 (2020), pp. 87–105. DOI: 10.1146/annurev-biophys-121219-081535.

[136] M. Börsch, P. Turina, C. Eggeling, J. R. Fries, C. A. M. Seidel, A. Labahn, and P. Gräber. "Conformational changes of the H+-ATPase from Escherichia coli upon nucleotide binding detected by single molecule fluorescence." en. In: *FEBS Letters* 437.3 (1998), pp. 251–254. ISSN: 1873-3468. DOI: 10.1016/S0014-5793(98)01247-2.

[137] H. Yu, K. Jo, K. L. Kounovsky, J. J. d. Pablo, and D. C. Schwartz. "Molecular Propulsion: Chemical Sensing and Chemotaxis of DNA Driven by RNA Polymerase." In: *Journal of the American Chemical Society* 131.16 (Apr. 2009), pp. 5722–5723. ISSN: 0002-7863. DOI: 10.1021/ja900372m.

[138] S. H. Muddana, S. Sengupta, E. T. Mallouk, A. Sen, and J. P. Butler. "Substrate Catalysis Enhances Single-Enzyme Diffusion." In: *Journal of the American Chemical Society* 132.7 (Feb. 2010), pp. 2110–2111. ISSN: 0002-7863. DOI: 10.1021/ja908773a.

[139] S. Sengupta, K. K. Dey, H. S. Muddana, T. Tabouillot, M. E. Ibele, P. J. Butler, and A. Sen. "Enzyme Molecules as Nanomotors." In: *Journal of the American Chemical Society* 135.4 (Jan. 2013), pp. 1406–1414. ISSN: 0002-7863. DOI: 10.1021/ja3091615.

[140] C. Riedel, R. Gabizon, M. C. A. Wilson, K. Hamadani, K. Tsekouras, S. Marqusee, S. Presse, and C. Bustamante. "The heat released during catalytic turnover enhances the diffusion of an enzyme." In: *Nature* 517 (Jan. 2015), pp. 227–230. ISSN: 0028-0836. DOI: 10.1038/nature14043.

[141] P. Illien, X. Zhao, K. K. Dey, J. P. Butler, A. Sen, and R. Golestanian. "Exothermicity Is Not a Necessary Condition for Enhanced Diffusion of Enzymes." In: *Nano Letters* 17.7 (July 2017), pp. 4415–4420. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.7b01502.

[142] A.-Y. Jee, S. Dutta, Y.-K. Cho, T. Tlusty, and S. Granick. "Enzyme leaps fuel antichemotaxis." In: *Proceedings of the National Academy of Sciences* 115.1 (Jan. 2018). DOI: 10.1073/pnas.1717844115.

[143] J.-P. Günther, M. Börsch, and P. Fischer. "Diffusion Measurements of Swimming Enzymes with Fluorescence Correlation Spectroscopy." In: *Accounts of Chemical Research* (Aug. 2018). ISSN: 0001-4842. DOI: 10.1021/acs.accounts.8b00276.

[144] Y. Zhang, M. J. Armstrong, N. M. B. Kazeruni, and H. Hess. "Aldolase does not show enhanced diffusion in dynamic light scattering experiments." en. In: *Nano Letters* (Nov. 2018). DOI: 10.1021/acs.nanolett.8b04240.

[145] Z. Chen, A. Shaw, H. Wilson, M. Woringer, X. Darzacq, S. Marqusee, Q. Wang, and C. Bustamante. "Single-molecule diffusometry reveals no catalysis-induced diffusion enhancement of alkaline phosphatase as proposed by FCS experiments." en. In: *Proceedings of the National Academy of Sciences* 117.35 (Sept. 2020), pp. 21328–21335. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2006900117.

[146] M. Xu, J. L. Ross, L. Valdez, and A. Sen. "Direct Single Molecule Imaging of Enhanced Enzyme Diffusion." In: *Physical Review Letters* 123.12 (Sept. 2019), p. 128101. DOI: 10.1103/PhysRevLett.123.128101.

[147] A.-Y. Jee, K. Chen, T. Tlusty, J. Zhao, and S. Granick. "Enhanced diffusion and oligomeric enzyme dissociation." In: *Journal of the American Chemical Society* (Nov. 2019). ISSN: 0002-7863. DOI: 10.1021/jacs.9b06949.

[148] H. Wang, M. Park, R. Dong, J. Kim, Y.-K. Cho, T. Tlusty, and S. Granick. "Boosted molecular mobility during common chemical reactions." en. In: *Science* 369.6503 (July 2020), pp. 537–541. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.aba8425.

[149] R. Golestanian. "Enhanced Diffusion of Enzymes that Catalyze Exothermic Reactions." In: *Physical Review Letters* 115 (Sept. 2015). DOI: 10.1103/PhysRevLett.115.108102.

[150] M. Feng and M. K. Gilson. "A Thermodynamic Limit on the Role of Self-Propulsion in Enhanced Enzyme Diffusion." en. In: *Biophysical Journal* 116.10 (May 2019), pp. 1898–1906. ISSN: 0006-3495. DOI: 10.1016/j.bpj.2019.04.005.

[151] A.-Y. Jee, Y.-K. Cho, S. Granick, and T. Tlusty. "Catalytic enzymes are active matter." en. In: *Proceedings of the National Academy of Sciences* (Nov. 2018), p. 201814180. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1814180115.

[152] F. Mohajerani, X. Zhao, A. Somasundar, D. Velegol, and A. Sen. "A Theory of Enzyme Chemotaxis: From Experiments to Modeling." In: *Biochemistry* 57.43 (Oct. 2018), pp. 6256–6263. ISSN: 0006-2960. DOI: 10.1021/acs.biochem.8b00801.

[153] P. Illien, T. Adeleke-Larodo, and R. Golestanian. "Diffusion of an enzyme: The role of fluctuation-induced hydrodynamic coupling." en. In: *EPL (Europhysics Letters)* 119.4 (Aug. 2017), p. 40002. ISSN: 0295-5075. DOI: 10.1209/0295-5075/119/40002.

[154] S. Kondrat and M. Popescu. "Brownian dynamics assessment of enhanced diffusion exhibited by 'fluctuating-dumbbell enzymes'." en. In: *Physical Chemistry Chemical Physics* (2019). DOI: 10.1039/C9CP02842K.

[155] Y. Zhang and H. Hess. "Enhanced Diffusion of Catalytically Active Enzymes." In: *ACS Central Science* (May 2019). ISSN: 2374-7943. DOI: 10.1021/acscentsci.9b00228.

[156] A. S. Mikhailov and R. Kapral. "Hydrodynamic collective effects of active protein machines in solution and lipid bilayers." en. In: *Proceedings of the National Academy of Sciences* 112.28 (July 2015), E3639–E3644. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1506825112.

[157] M. Dennison, R. Kapral, and H. Stark. "Diffusion in systems crowded by active force-dipole molecules." en. In: *Soft Matter* 13.20 (2017), pp. 3741–3749. DOI: 10.1039/C7SM00400A.

[158] N. B. Shah, M. L. Hutcheon, B. K. Haarer, and T. M. Duncan. "F1-ATPase of Escherichia coli THE ϵ-INHIBITED STATE FORMS AFTER ATP HYDROLYSIS, IS DISTINCT FROM THE ADP-INHIBITED STATE, AND RESPONDS DYNAMICALLY TO CATALYTIC SITE LIGANDS." en. In: *Journal of Biological Chemistry* 288.13 (Mar. 2013), pp. 9383–9395. ISSN: 0021-9258, 1083-351X. DOI: 10.1074/jbc.M113.451583.

[159] J. M. Schurr, B. S. Fujimoto, L. Huynh, and D. T. Chiu. "A Theory of Macromolecular Chemotaxis." In: *The Journal of Physical Chemistry B* 117.25 (June 2013), pp. 7626–7652. ISSN: 1520-6106. DOI: 10.1021/jp302587d.

[160] A. V. Kiselev. "Non-specific and specific interactions of molecules of different electronic structures with solid surfaces." en. In: *Discussions of the Faraday Society* 40.0 (Jan. 1965), pp. 205–218. ISSN: 0366-9033. DOI: 10.1039/DF9654000205.

[161] M. J. Schnitzer. "Theory of continuum random walks and application to chemotaxis." In: *Physical Review E* 48.4 (Oct. 1993), pp. 2553–2568. DOI: 10.1103/PhysRevE.48.2553.

[162] J. Agudo-Canalejo, P. Illien, and R. Golestanian. "Phoresis and Enhanced Diffusion Compete in Enzyme Chemotaxis." In: *Nano Letters* 18.4 (Apr. 2018), pp. 2711–2717. ISSN: 1530-6984. DOI: 10.1021/acs.nanolett.8b00717.

[163] S. Ebbens, M.-H. Tu, J. R. Howse, and R. Golestanian. "Size dependence of the propulsion velocity for catalytic Janus-sphere swimmers." In: *Physical Review E* 85.2 (Feb. 2012), p. 020401. DOI: 10.1103/PhysRevE.85.020401.

[164] J. L. Anderson. "Colloid Transport by Interfacial Forces." In: *Annual Review of Fluid Mechanics* 21.1 (1989), pp. 61–99. DOI: 10.1146/annurev.fl.21.010189.000425.

[165] D. McQuarrie. *Statistical mechanics*. Harper's chemistry series. Harper & Row, 1975.

[166] J. L. Kelly. "A New Interpretation of Information Rate." en. In: *Bell System Technical Journal* 35.4 (1956), pp. 917–926. ISSN: 1538-7305. DOI: 10.1002/j.1538-7305.1956.tb03809.x.

[167] A. Buchner, F. Tostevin, and U. Gerland. "Clustering and Optimal Arrangement of Enzymes in Reaction-Diffusion Systems." In: *Physical Review Letters* 110.20 (May 2013), p. 208104. DOI: 10.1103/PhysRevLett.110.208104.

[168] A. Buchner, F. Tostevin, F. Hinzpeter, and U. Gerland. "Optimization of collective enzyme activity via spatial localization." In: *The Journal of Chemical Physics* 139.13 (Oct. 2013), p. 135101. ISSN: 0021-9606. DOI: 10.1063/1.4823504.

[169] N. Katoh and T. Ibaraki. "Resource Allocation Problems." en. In: *Handbook of Combinatorial Optimization: Volume1–3*. Ed. by D.-Z. Du and P. M. Pardalos. Boston, MA: Springer US, 1998, pp. 905–1006. ISBN: 978-1-4613-0303-9. DOI: 10.1007/978-1-4613-0303-9_14.

[170] K. Kumar, R. A. Mella-Herrera, and J. W. Golden. "Cyanobacterial Heterocysts." en. In: *Cold Spring Harbor Perspectives in Biology* 2.4 (Apr. 2010), a000315. ISSN: , 1943-0264. DOI: 10.1101/cshperspect.a000315.

[171] M. Scott, C. W. Gunderson, E. M. Mateescu, Z. Zhang, and T. Hwa. "Interdependence of Cell Growth and Gene Expression: Origins and Consequences." en. In: *Science* 330.6007 (Nov. 2010), pp. 1099–1102. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1192588.

[172] C. You, H. Okano, S. Hui, Z. Zhang, M. Kim, C. W. Gunderson, Y.-P. Wang, P. Lenz, D. Yan, and T. Hwa. "Coordination of bacterial proteome with metabolism by cyclic AMP signalling." en. In: *Nature* 500.7462 (Aug. 2013), pp. 301–306. ISSN: 1476-4687. DOI: 10.1038/nature12446.

[173] A. Maitra and K. A. Dill. "Bacterial growth laws reflect the evolutionary importance of energy efficiency." en. In: *Proceedings of the National Academy of Sciences* 112.2 (Jan. 2015), pp. 406–411. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1421138111.

[174] M. Y. Pavlov and M. Ehrenberg. "Optimal control of gene expression for fast proteome adaptation to environmental change." en. In: *Proceedings of the National Academy of Sciences* 110.51 (Dec. 2013), pp. 20527–20532. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1309356110.

[175] D. W. Erickson, S. J. Schink, V. Patsalo, J. R. Williamson, U. Gerland, and T. Hwa. "A global resource allocation strategy governs growth transition kinetics of Escherichia coli." en. In: *Nature* 551.7678 (Nov. 2017), pp. 119–123. ISSN: 1476-4687. DOI: 10.1038/nature24299.

[176] D. L. Schmitt and S. An. "Spatial Organization of Metabolic Enzyme Complexes in Cells." In: *Biochemistry* 56.25 (June 2017), pp. 3184–3196. ISSN: 0006-2960. DOI: 10.1021/acs.biochem.7b00249.

[177] J. Zhou and B. Xu. "Enzyme-Instructed Self-Assembly: A Multistep Process for Potential Cancer Therapy." In: *Bioconjugate Chemistry* 26.6 (June 2015), pp. 987–999. ISSN: 1043-1802. DOI: 10.1021/acs.bioconjchem.5b00196.

[178] J. Li, Y. Kuang, J. Shi, J. Zhou, J. E. Medina, R. Zhou, D. Yuan, C. Yang, H. Wang, Z. Yang, J. Liu, D. M. Dinulescu, and B. Xu. "Enzyme-Instructed Intracellular Molecular Self-Assembly to Boost Activity of Cisplatin against Drug-Resistant Ovarian Cancer Cells." en. In: *Angewandte Chemie International Edition* 54.45 (2015), pp. 13307–13311. ISSN: 1521-3773. DOI: 10.1002/anie.201507157.

[179] B. Hove-Jensen, K. R. Andersen, M. Kilstrup, J. Martinussen, R. L. Switzer, and M. Willemoës. "Phosphoribosyl Diphosphate (PRPP): Biosynthesis, Enzymology, Utilization, and Metabolic Significance." en. In: *Microbiology and Molecular Biology Reviews* 81.1 (Mar. 2017). ISSN: 1092-2172, 1098-5557. DOI: 10.1128/MMBR.00040-16.

[180] R. E. Goldstein and J.-W. van de Meent. "A physical perspective on cytoplasmic streaming." In: *Interface Focus* 5.4 (Aug. 2015), p. 20150030. DOI: 10.1098/rsfs.2015.0030.

[181] J. F. Presley, N. B. Cole, T. A. Schroer, K. Hirschberg, K. J. M. Zaal, and J. Lippincott-Schwartz. "ER-to-Golgi transport visualized in living cells." en. In: *Nature* 389.6646 (Sept. 1997), pp. 81–85. ISSN: 1476-4687. DOI: 10.1038/38001.

[182] O. Grimm, M. Coppey, and E. Wieschaus. "Modelling the Bicoid gradient." en. In: *Development* 137.14 (July 2010), pp. 2253–2264. ISSN: 0950-1991, 1477-9129. DOI: 10.1242/dev.032409.

[183] D. Oh, C.-H. Yu, and D. J. Needleman. "Spatial organization of the Ran pathway by microtubules in mitosis." en. In: *Proceedings of the National Academy of Sciences* 113.31 (Aug. 2016), pp. 8729–8734. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1607498113.

[184] Y. E. Chen, C. Tropini, K. Jonas, C. G. Tsokos, K. C. Huang, and M. T. Laub. "Spatial gradient of protein phosphorylation underlies replicative asymmetry in a bacterium." en. In: *Proceedings of the National Academy of Sciences* 108.3 (Jan. 2011), pp. 1052–1057. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1015397108.

[185] J. R. Robbins, D. Monack, S. J. McCallum, A. Vegas, E. Pham, M. B. Goldberg, and J. A. Theriot. "The making of a gradient: IcsA (VirG) polarity in Shigella flexneri." en. In: *Molecular Microbiology* 41.4 (2001), pp. 861–872. ISSN: 1365-2958. DOI: 10.1046/j.1365-2958.2001.02552.x.

[186] T. E. Saunders, K. Z. Pan, A. Angel, Y. Guan, J. V. Shah, M. Howard, and F. Chang. "Noise Reduction in the Intracellular Pom1p Gradient by a Dynamic Clustering Mechanism." en. In: *Developmental Cell* 22.3 (Mar. 2012), pp. 558–572. ISSN: 1534-5807. DOI: 10.1016/j.devcel.2012.01.001.

[187] O. Idan and H. Hess. "Origins of Activity Enhancement in Enzyme Cascades on Scaffolds." In: *ACS Nano* 7.10 (Oct. 2013), pp. 8658–8665. ISSN: 1936-0851. DOI: 10.1021/nn402823k.

[188] E. O. Thorp. "Portfolio choice and the Kelly criterion." In: *Stochastic Optimization Models in Finance*. Elsevier, 1975, pp. 599–619.

[189] B. Ho, A. Baryshnikova, and G. W. Brown. "Unification of Protein Abundance Datasets Yields a Quantitative Saccharomyces cerevisiae Proteome." en. In: *Cell Systems* 6.2 (Feb. 2018), 192–205.e3. ISSN: 2405-4712. DOI: 10.1016/j.cels.2017.12.004.

[190] H. Markowitz. "Portfolio Selection." en. In: *The Journal of Finance* 7.1 (1952), pp. 77–91. ISSN: 1540-6261. DOI: 10.1111/j.1540-6261.1952.tb01525.x.

[191] C. Runge. "Ueber die numerische Auflösung von Differentialgleichungen." de. In: *Mathematische Annalen* 46.2 (June 1895), pp. 167–178. ISSN: 1432-1807. DOI: 10.1007/BF01446807.

[192] W. Kutta. "Beitrag zur naherungsweisen Integration totaler Differentialgleichungen." In: *Z. Math. Phys.* 46 (1901), pp. 435–453.

[193] S. K. Kufer, E. M. Puchner, H. Gumpp, T. Liedl, and H. E. Gaub. "Single-Molecule Cut-and-Paste Surface Assembly." en. In: *Science* 319.5863 (Feb. 2008), pp. 594–596. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1151424.

[194] K. R. Erlich, S. M. Sedlak, M. A. Jobst, L. F. Milles, and H. E. Gaub. "DNA-free directed assembly in single-molecule cut-and-paste." en. In: *Nanoscale* 11.2 (2019), pp. 407–411. DOI: 10.1039/C8NR08636B.

[195] G. Giunta, H. Seyed-Allaei, and U. Gerland. "Cross-diffusion induced patterns for a single-step enzymatic reaction." en. In: *Communications Physics* 3.1 (Sept. 2020), pp. 1–9. ISSN: 2399-3650. DOI: 10.1038/s42005-020-00427-w.

[196] C. Weistuch and S. Pressé. "Spatiotemporal Organization of Catalysts Driven by Enhanced Diffusion." In: *The Journal of Physical Chemistry B* (Sept. 2017). ISSN: 1520-6106. DOI: 10.1021/acs.jpcb.7b06868.

[197] A. M. Turing. "The chemical basis of morphogenesis." en. In: *Bulletin of Mathematical Biology* 52.1 (Jan. 1990), pp. 153–197. ISSN: 1522-9602. DOI: 10.1007/BF02459572.

[198] S. S. Sugai, K. L. Ode, and H. R. Ueda. "A Design Principle for an Autonomous Post-translational Pattern Formation." English. In: *Cell Reports* 19.4 (Apr. 2017), pp. 863–874. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2017.03.081.

[199] B. Derjaguin, G. Sidorenkov, E. Zubashchenkov, and E. Kiseleva. "Kinetic phenomena in boundary films of liquids." In: *Kolloidn. zh* 9 (1947), pp. 335–347.

[200] R. Milo and R. Phillips. *Cell biology by the numbers*. Garland Science, 2015.

[201] M. Cross and H. Greenside. *Pattern Formation and Dynamics in Nonequilibrium Systems by Michael Cross*. en. July 2009. DOI: 10.1017/CBO9780511627200.

[202] M. Arrio-Dupont, G. Foucault, M. Vacher, P. F. Devaux, and S. Cribier. "Translational Diffusion of Globular Proteins in the Cytoplasm of Cultured Muscle Cells." en. In: *Biophysical Journal* 78.2 (Feb. 2000), pp. 901–907. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(00)76647-1.

[203] H. P. Erickson. "Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy." In: *Biological Procedures Online* 11 (May 2009), pp. 32–51. ISSN: 1480-9222. DOI: 10.1007/s12575-009-9008-x.

[204] C. Liu, X. Fu, L. Liu, X. Ren, C. K. L. Chau, S. Li, L. Xiang, H. Zeng, G. Chen, L.-H. Tang, P. Lenz, X. Cui, W. Huang, T. Hwa, and J.-D. Huang. "Sequential Establishment of Stripe Patterns in an Expanding Cell Population." en. In: *Science* 334.6053 (Oct. 2011), pp. 238–241. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1209042.

[205] A. J. Koch and H. Meinhardt. "Biological pattern formation: from basic mechanisms to complex structures." In: *Reviews of Modern Physics* 66.4 (Oct. 1994), pp. 1481–1507. DOI: 10.1103/RevModPhys.66.1481.

[206] M. E. Cates and J. Tailleur. "Motility-Induced Phase Separation." In: *Annual Review of Condensed Matter Physics* 6.1 (Mar. 2015), pp. 219–244. ISSN: 1947-5454. DOI: 10.1146/annurev-conmatphys-031214-014710.

[207] V. K. Vanag and I. R. Epstein. "Cross-diffusion and pattern formation in reaction–diffusion systems." en. In: *Physical Chemistry Chemical Physics* 11.6 (Jan. 2009), pp. 897–912. ISSN: 1463-9084. DOI: 10.1039/B813825G.

[208] J. Agudo-Canalejo and R. Golestanian. "Active Phase Separation in Mixtures of Chemically Interacting Particles." In: *Physical Review Letters* 123.1 (July 2019), p. 018101. DOI: 10.1103/PhysRevLett.123.018101.

[209] D. E. Woodward, R. Tyson, M. R. Myerscough, J. D. Murray, E. O. Budrene, and H. C. Berg. "Spatio-temporal patterns generated by Salmonella typhimurium." In: *Biophysical Journal* 68.5 (May 1995), pp. 2181–2189. ISSN: 0006-3495. DOI: 10.1016/S0006-3495(95)80400-5.

[210] L. Tweedy, D. A. Knecht, G. M. Mackay, and R. H. Insall. "Self-Generated Chemoattractant Gradients: Attractant Depletion Extends the Range and Robustness of Chemotaxis." en. In: *PLOS Biology* 14.3 (Mar. 2016), e1002404. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.1002404.

[211] E. F. Keller and L. A. Segel. "Initiation of slime mold aggregation viewed as an instability." en. In: *Journal of Theoretical Biology* 26.3 (Mar. 1970), pp. 399–415. ISSN: 0022-5193. DOI: 10.1016/0022-5193(70)90092-5.

Acknowledgments

The research presented in this dissertation would have not been possible without the guidance, the support and the encouragement of many people. They all contributed to make the PhD the most formative period of my life.

First, I want to thank my supervisor Uli Gerland for encouraging me to knock at his door anytime to discuss project ideas. I thank him also for guiding me while letting me the freedom to follow my own curiosities in the design of the projects. And, above all, I am grateful to him for having been a fatherly and supportive person all the times that I was facing difficulties.

I am deeply grateful to Filipe Tostevin for the patience demonstrated in listening to the countless questions I had over the years. Professionally, he taught me a lot on how to program, how to pose, judge and analyze research questions. There is only one thing I did not learn from him and this is how to be quicker in taking turns at board games. I am sure he wishes I had learned that. Above all, I thank Filipe for having become a mentor-like figure and a close friend.

I also want to thank Hamid Seyed-Allaei for the fun had while working together on the project on enzyme motion. I especially thank him for the patience shown while correcting all the wrong estimates of parameters that I was computing. I found in Hamid a great collaborator and I am sure we will keep working and discuss science together.

I am also grateful to Sorin Tănase-Nicola for contributing to the enzyme arrangement project by providing deep mathematical insights.

I am also very thankful to the Quantitative Bioscience Munich (QBM) school for taking me as one of their graduate students and for supporting me financially during the first three years of PhD. I also thank QBM for the great scientific environment created and for the very inspiring retreats and lecture series organized.

I also want to thank former and current group members for creating a welcoming and warm working atmosphere. Especially I want to thank Florian for the many stimulating discussions had about enzymes, cuisines and travels. Stephan for discussing exciting future possible project ideas and for teaching me cross-country skating without having my neck broken. Linda for being the best Stammtisch buddy and for preparing me the best setteveli ever. Elena for always helping me in the typical Italian expat bureaucracy problems and for having become a good friend and a confidant. Zara for sharing my

passion for sushi and for the nice afterwork activities. Severin for always transmitting his passion about science and for having become a great friend. Michael and Manon for the Awesome Theory Journal Clubs, the many scientific discussions and the fun had in the SUP sessions. Tobi and Joachim for being great office neighbors and being always willing to help with any scientific and personal question. Bernhard for the feedback provided in writing the manuscript about the effects of the cross-diffusion of enzymes and for sharing, together with Filipe, the passion for playing Twilight Imperium (which no, it has nothing to do with the TV series). Mareike for organizing together with Tobi amazing group retreats at the Albert-Link-Hütte and Cesar for exchanging very interesting ideas and references. Special thanks also to Giovanni, Cecilia and Maryam for being great office mates and good companions. I am also grateful to former group members: Johannes, Thiago, Vladimir and Patrick for their guidance and their feedback provided at the beginning of the PhD. Thanks also to all the other group members for the great time.

At the department I also had the chance to meet a great scientist and an even better friend and boulder coach, Kasia Tych. I would like to thank her for the exciting scientific conversations, the many lunches and boulder sessions had together.

I also want to thank Federica, Stefano, Karl, Leila, Linda, Stephan and Kasia for the many inputs given me for the thesis and helping me in the proofreading. I also thank Ben, Felix and Linda for helping me in the translation of the German parts of the thesis and the documents associated to it.

I am also grateful to the awesome friends I met during my time in Munich. I especially want to thank Leila for having been always cheerful and supportive throughout the past seven years, including the writing of this thesis. My long time uni colleagues and awesome friends Ben and Mauro for the fun times together. My travel buddies Luca, beautiful Mina and Mariana for the many trips necessary to recharge the energies in between the intense periods of research. Aurora for the joint decision of moving from Catania to Munich after our Bachelor studies. Carmen, Karl, Rui, Giovanni, Fede and Ste for the many family-like Sunday lunches, dinners and aperitifs. All the Italian friends part of the MCC for the daily instructive and inspiring conversations. Special thanks to the awesome tandem group for never giving up on my German: Susanna, Felix, Theresa, Fede and Ste; and to the people that always made Munich one of my homes from the very first day and continue to do so: Uta, Walther, Magic, Consti, Veronica and Xavi.

I am also grateful to my former Bachelor colleagues for their occasional visits and great support: Mario, Pietro, Enzo, Musci, Luca, Emiliano, Ale. Together with friends that are also bit more far away: Fausto, Claudio, Gianluca, Sebi, Giulia, Gloria, Emi and my private English teacher Heather.

I am deeply grateful to my family for their continuous encouragement and support.

I especially want to thank my parents for always supporting and reassuring me and my brothers while showing us important examples of collective social responsibility. Their closeness and their strength transmitted me the confidence and serenity necessary to carry out the research presented in this dissertation.

Last I would like to thank Maria for always being next to me over the last months. I thank her for being there when celebrating for the first PhD publication. And I thank her for the infinite patience, cheerfulness, love and peace transmitted me while I was writing this dissertation.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die bei der promotionsführenden Einrichtung

Fakultät für Physik

der TUM zur Promotionsprüfung vorgelegte Arbeit mit dem Titel:

Spatio-temporal Organization of Enzymatic Reactions

in

Fakultät für Physik, Lehrstuhl für Theoretische Physik - Theorie komplexer Biosysteme

unter der Anleitung und Betreuung durch: **Prof. Dr. Ulrich Gerland**
ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Ab. 6 und 7 Satz 2 angebotenen Hilfsmittel benutzt habe.

- ☒ Ich habe keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht, oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.
- ☒ Ich habe die Dissertation in dieser oder ähnlicher Form in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.
- ☐ Die vollständige Dissertation wurde in _____ veröffentlicht. Die promotionsführende Einrichtung _____ hat der Veröffentlichung zugestimmt.
- ☒ Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.
- ☐ Ich habe bereits am _____ bei der Fakultät für _____ der Hochschule _____ unter Vorlage einer Dissertation mit dem Thema _____ die Zulassung zur Promotion beantragt mit dem Ergebnis: _____

Die öffentlich zugängliche Promotionsordnung der TUM ist mir bekannt, insbesondere habe ich die Bedeutung von § 28 (Nichtigkeit der Promotion) und § 29 (Entzug des

Doktorgrades) zur Kenntnis genommen. Ich bin mir der Konsequenzen einer falschen Eidesstattlichen Erklärung bewusst.
Mit der Aufnahme meiner personenbezogenen Daten in die Alumni-Datei bei der TUM bin ich

- ☒ einverstanden
- ☐ nicht einverstanden

Ort, Datum, Unterschrift