PAPER

How many enzyme molecules are needed for discrimination oriented applications?

Jerzy Gorecki,*[a] Joanna N. Gorecka,[b] Bogdan Nowakowski,[ac] Hiroshi Ueno[d] and Kenichi Yoshikawa[d]

Chemical reactions establish a molecular mechanism for information processing in living organisms. Here we consider a simple enzymatic reaction model that can be used to discriminate parameters characterizing periodic reagent inflow. Numerical simulations based on the kinetic equations show that there exist a range of inflow frequencies and amplitudes in which the time evolution of the system is very sensitive to small changes in the values of these parameters. However, the kinetic equations are derived for the thermodynamic limit, whereas in a real biological medium, like a cell, the number of enzyme molecules is an integer and finite. We use stochastic simulations to estimate discriminator reliability as a function of the number of enzyme molecules involved. For systems with 10 000 molecules the functionality predicted by kinetic equations is confirmed. If the number of molecules is decreased to 100, discrimination becomes unreliable.

1 Introduction

Living organisms use chemical reactions to acquire and process information about the environment they live in. High-level organisms have developed a spatial structure (the nerve system) that performs information processing tasks. The functionality of a nerve system is based on reaction–diffusion type phenomena proceeding in an excitable medium.[1] A structure similar to the nerve system is missing in simple organisms. Therefore, it can be expected that the necessary information processing ability is incorporated into the metabolic pathway. This observation motivates investigation on computational strategies based on enzymatic processes characterized by a complex, nonlinear chemical kinetics and exhibiting a rich variety of stationary spatio-temporal patterns.[2–4]

Enzymatic systems can accept and transform information in the form of chemical substrates with particular properties that meet the binding specificity criteria of a selected enzyme. The input information can be coded in concentrations of specific molecules that undergo catalyzing reactions leading to products with different properties than that of input reagents. The output information is coded in concentrations of products. Of course, products can participate in subsequent reaction steps.

Transformations of products into other molecules can be regarded as the following information processing operations. Studies on networks of connected enzyme-catalyzed reactions with additional chemical and enzymatic processes that incorporate filtering steps into the reaction cascades[5] demonstrated that scaled logic variables for the inputs, outputs and intermediate products can be useful for describing the functionality of an enzyme cascade.

The majority of experimental studies on (bio-)chemical information processing have focused on the binary logic.[6] A convex line representing the relationship between the substrate concentration and the rate of product generation, observed for example for the Michaelis–Menten kinetics,[7] is typical for enzymatic reactions. For such a relationship there is an identifiable linear regime, typically near the zero concentration, as well as the saturation regime for higher concentrations. In such a case the relationship between reagent concentrations and binary logic values is straightforward: the high concentrations represent the logic TRUE state, the low concentrations are interpreted as the logic FALSE. However, enzymatic reactions with a convex relationship between the substrate concentration and the product production rate are not the most suitable for enzymatic information processing applications. A sigmoid, filter like relationship between the substrate concentration and the product generation rate seems to be more suited, because it gives a better balance between low and high values of concentrations.

The sigmoidal relationship often appears for appropriately chained enzymatic reactions.[8,9] Therefore, it can be designed on demand by coupling an enzymatic process characterized by a convex kinetics with other reactions[10] (see also Section 2).

[a] *Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44/52, 01-224 Warsaw, Poland. E-mail: jgorecki@ichf.edu.pl*

[b] *Institute of Physics, Polish Academy of Sciences, al. Lotnikow 32/46, 02-668 Warsaw, Poland*

[c] *SGGW, Warsaw University of Life Sciences, Nowoursynowska 159, 02-776 Warsaw, Poland*

[d] *Faculty of Life and Medical Sciences, Doshisha University, Kyoto 610-0394, Japan*

Various physical or chemical stimuli can be applied to adjust the relationship between the substrate concentration and the product generation rate. It was demonstrated that in the enzymatic reaction involving glucose oxidase system response depends on the initial concentration of hexokinase and ATP.[10] Even more interesting is that the transition between convex and sigmoidal kinetics can be achieved by system illumination.[11]

The sigmoidal relationship between the substrate concentration and the rate of product production is expected for allosteric enzymatic reactions.[12–14] Allosteric regulation is one of the many ways in which enzyme activity can be controlled. A typical enzyme contains binding sites where substrate molecules can be attached and the catalytic site where the activation energy is reduced and the reaction, specific for a given enzyme, proceeds. The allosteric enzymes contain another region, separated from the substrate binding site and the catalytic site, to which small, regulatory molecules (called allosteric regulatory molecules) can attach to and thereby control the catalytic activity. The allosteric regulatory molecules modify the conformation and vibrational modes of the enzyme. Such changes are transduced to the active site and affect the reaction rate. Allosteric interactions can either inhibit or activate enzymes. The regulating reactions can couple with the main enzymatic process. In Section 2 we give an example of allosteric reaction for which the sigmoidal rate is observed if the rates of reactions are appropriately selected. A biocatalytic system with a double-sigmoid characteristic (sigmoidal relationship with respect to two types of molecules) is especially interesting because it can directly operate on two input signals.[15–17] Enzymatic reactions capable of recognizing two specific molecules or ions in the solution (for example Mg^{2+} and Ca^{2+} [15,18]) were identified and their application in the construction of logic gates (the AND binary gate,[19] the OR binary gate[20]) was reported. Enzymatic systems can also perform more complex computational tasks like three-input logic gates or molecular full adders.[21,22]

The theoretical analysis of information processing using enzymatic reactions, presented in the papers cited above, was based on the deterministic kinetic equations valid in the thermodynamic limit. However, the volume of a real living cell is finite and the number of enzyme molecules is limited. For a small number of enzyme molecules involved, the internal fluctuations can have a significant influence on the chemical kinetics. In this paper we study the influence of the system size on a prototype of an enzymatic binary discriminator in order to estimate the minimum number of molecules needed to perform such a task reliably.

An enzymatic system that can distinguish which of two classes a given case belongs to is important because a significant part of the computational activity of living organisms is related to decision making. Information has to be processed in order to find the optimum condition for their existence in a time dependent environment. The time dependent environment includes many periodic stimuli, like the temperature or illumination in a day–night cycle or the frequency related to a circadian rhythm. Therefore, the discrimination of the amplitude or frequency of an external periodic stimulus can be important in the search for the optimum living conditions. Many decision making problems require a binary (YES or NO) answer. Having in mind that information processing ability of living organisms is based on biochemical reactions we introduce an enzymatic reaction model that can be applied for decision making problems in which the input information is coded in periodically changing inflow of reagents.

The papers on enzymatic computing cited above were based on information coded in time-independent, stationary concentrations of selected reagents. In this paper we discuss a new approach to enzymatic computing with information coded in periodic oscillations of reagent concentrations. The input variables are parameters describing the inflow of a selected reagent X into a medium in which an enzymatic reaction proceeds (the amplitude A and the frequency f). The output information is coded in the type of oscillations of another reagent Y. The information and the matter flows through the discriminator are schematically illustrated in Fig. 1.

We have shown[23,24] that a chemical system characterized by a sigmoidal kinetics is especially suitable for discrimination of frequency and amplitude characterizing the inflow of reagents. For such a system there is a range of parameters in which the system is bistable and its stable states belong to two separate classes S_1 and S_2. Periodically perturbed bistable system oscillates. The location of system oscillations in the $X \times Y$ space depends on the amplitude and frequency of the inflow. In particular we can distinguish oscillations covering the states in class S_1 and oscillations between states in class S_2. Numerical simulations have shown that for a sigmoidal reaction kinetics a marginal change of inflow parameters (amplitude or frequency) can force a sharp transition between oscillations of Y belonging to different classes.[23,24] The transition occurs in a narrow range of perturbation parameters. The difference between various oscillation types can be easily observed because they occur in various regions of Y concentration. Therefore, an enzymatic reaction exhibiting hysteresis seems to be a good candidate for a discriminator of inflow properties.

This paper is organized as follows: in Section 2 we introduce a simple model of enzymatic reaction and we demonstrate that it can operate as a binary discriminator of the inflow parameters. In Section 3 we introduce a stochastic model for the considered reaction and describe the simulation algorithm that mixes a discrete description of enzyme complexes with a continuous description of other reagents. The results of numerical simulations are presented in the next section.
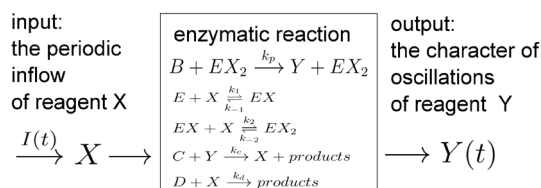

Fig. 1 A schematic illustration of the discriminator of inflow parameters based on an enzymatic reaction. Depending on the character of oscillation of Y reagent concentration one can distinguish two separate classes of reagent X inflow.

This journal is © the Owner Societies 2016

*Phys. Chem. Chem. Phys.*, 2016, **18**, 20518–20527 | **20519**

# 2 Determination of inflow properties with enzymatic reactions

We consider a reaction model in which the enzyme molecule E with two attached molecules of reagent X forms the complex EX_2 which catalyzes the transformation of the reactant B into the final product Y:

$$B + EX_2 \xrightarrow{k_p} Y + EX_2$$
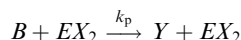
We assume that the formation of EX_2 complex is described by two chained reversible reactions:

$$E + X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EX$$

$$EX + X \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EX_2$$
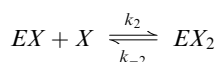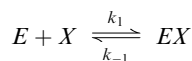
The model based on equations given above would lead to unbounded growth in concentration of Y. In order to have a stationary state with finite concentrations we take into account the transformation of product Y into reactant X:

$$C + Y \xrightarrow{k_c} X + \text{products}_C$$
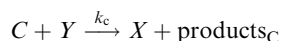
and the spontaneous decomposition of the reagent X:

$$D + X \xrightarrow{k_d} \text{products}_D$$
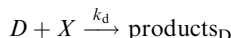
In the considered model we also include the inflow of X molecules with the time dependent rate $I(t)$:

$$\xrightarrow{I(t)} X$$

For the following analysis we assume that concentrations of B, C and D molecules are large if compared to the other reagents and their changes in time are negligible. The evolution of the system is described by time dependent concentrations of E, EX, EX_2, X and Y denoted as $e(t)$, $f(t)$, $g(t)$, $x(t)$ and $y(t)$, respectively. The total number of enzyme molecules in its free and complexed forms remains constant in time. Let us denote it as e_0. This constrain can be used to calculate $e(t)$ if $f(t)$ and $g(t)$ are known. The time evolution of the medium with reactions listed above is described by the set of equations:

$$e(t) = e_0 - f(t) - g(t) \tag{1}$$

$$\begin{aligned}\frac{dx(t)}{dt} &= -k_1 e(t)x(t) + k_{-1}f(t) - k_2 f(t)x(t) \\ &\quad + k_{-2}g(t) - k_d dx(t) + k_c cy(t) + I(t)\end{aligned} \tag{2}$$

$$\frac{df(t)}{dt} = k_1 e(t)x(t) - k_{-1}f(t) - k_2 f(t)x(t) + k_{-2}g(t) \tag{3}$$

$$\frac{dg(t)}{dt} = k_2 f(t)x(t) - k_{-2}g(t) \tag{4}$$

$$\frac{dy(t)}{dt} = k_p bg(t) - k_c cy(t) \tag{5}$$

For the following analysis we can introduce a new time scale that is linked to the real one by the equation: $t' = k_d \times t$.

Such transformation of time is equivalent to the assumption that $k_d = 1$ and that all other rate constants are scaled by k_d. Moreover, without the loss of generality we can assume that: $b = c = d = 1$ because the values of concentrations b, c and d can be incorporated into the rate constants k_p, k_c and k_d, respectively.

If we assume that the reactions involving an enzyme and its complexes are faster than those leading to products$_C$ and products$_D$ then the concentrations of complexes EX and EX_2 can be regarded as quasistationary ones that immediately adjust to the local values of $x(t)$ and $y(t)$. The corresponding values of $e(t)$, $f(t)$ and $g(t)$ are:

$$e = \frac{e_0 k_{-1}k_{-2}}{k_{-1}k_{-2} + k_1 k_{-2}x(t) + k_1 k_2 x(t)^2} \tag{6}$$

$$f = \frac{e_0 k_1 k_{-2}x(t)}{k_{-1}k_{-2} + k_1 k_{-2}x(t) + k_1 k_2 x(t)^2} \tag{7}$$

$$g = \frac{e_0 k_1 k_2 x(t)^2}{k_{-1}k_{-2} + k_1 k_{-2}x(t) + k_1 k_2 x(t)^2} \tag{8}$$

In such a case the time evolution of the system is described by two kinetic equations for $x(t)$ and $y(t)$:

$$\begin{aligned}\frac{dx(t)}{dt} &= Q(x(t), y(t), t) \\ &= -k_d dx(t) + k_c cy(t) + I(t)\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{dy(t)}{dt} &= R(x(t), y(t)) \\ &= \frac{k_1 k_2 k_p b e_0 x^2(t)}{k_{-1}k_{-2} + k_1 k_{-2}x(t) + k_1 k_2 x(t)^2} - k_c cy(t)\end{aligned} \tag{10}$$

The reduced model leads to a kinetic term in a Hill-like form[25] with the Hill coefficient equal to 2. Such a value has been observed for allosteric reactions characterized by a positive cooperativity.[13,26] The reaction rates and constant concentrations can be adjusted such that the kinetic term in eqn (10) has a sigmoidal form. As we know[23,24] a chemical system characterized by a sigmoidal kinetics can be used for discrimination of periods and amplitudes characterizing the inflow of reagents.

For the system described by eqn (9) and (10) we obtain a sigmoid-shaped nullcline $R(x, y) = 0$ if the rate constants are $k_1 = 2$, $k_{-1} = 2$, $k_2 = 4$, $k_{-2} = 2$, and $k_p = 1820$, the rate constants for the slow reactions lead to products $k_c = 0.1$ and $k_d = 1$ and the total concentration of enzyme is $e_0 = 0.001$. The locations of the $R(x, y) = 0$ nullcline and the $Q(x, y, t) = 0$ nullcline are illustrated in Fig. 2. The nullcline $Q(x, y, t) = 0$ is shown for three different values of reagent inflow: $I(t) \equiv 0$ (no inflow of X), $I(t) \equiv 0.0335807 = A_2$ and $I(t) \equiv 0.0862972 = A_1$. The last two values of inflow correspond to cases when the nullcline $Q(x, y, t)$ is tangential to the nullcline $R(x, y)$. At the beginning we consider an inflow that remains constant after it is switched on at $t = 0$, i.e.: $I(t) = I_0 \cdot \Theta(t)$, where $\Theta(t)$ is the Heaviside step function. Let the initial ($t = 0$) state of the system is $x(t = 0) = y(t = 0) = 0$, which is the steady state if there is no inflow of X. If the inflow $I_0 < A_2$
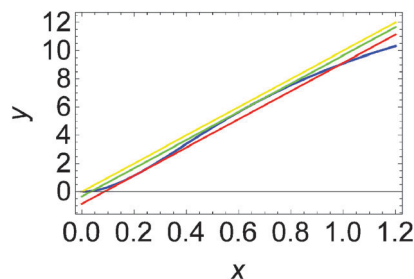
20520 | *Phys. Chem. Chem. Phys.*, 2016, **18**, 20518–20527

This journal is © the Owner Societies 2016

Fig. 2 Positions of nullclines for the dynamical system defined by eqn (9) and (10) in the phase space (x, y). The nullcline $R(x, y) = 0$ is plotted with the blue line. The nullcline $Q(x, y, t) = 0$ is shown for a few cases: $I(t) \equiv 0$ (the yellow line), $I(t) \equiv 0.0335807 = A_2$ (the green line) and $I(t) \equiv 0.0862972 = A_1$ (the red line). For the selected parameters of the model, the concentrations at tangential points are $(x_1, y_1) = (0.202441, 1.16144)$ and $(x_2, y_2) = (0.657363, 6.23782)$.

then the system has a single, stable stationary state (x_s, y_s) for which $x_s < 0.202441$ and $y_s < 1.16144$. For $I_0 > A_1$ the system has a single, stable stationary state (x_s, y_s) for which $x_s > 0.657363$ and $y_s > 6.23782$. If $A_2 < I_0 < A_1$ then the system of eqn (9) and (10) is bistable and shows hysteresis. Having in mind results of our recent papers[23,24] it is expected that the considered reaction model can function as a discriminator of frequency and amplitude of a time dependent inflow.

Upon selecting the model parameters we intentionally made enzyme concentration much smaller than concentrations of other reagents to agree with the stochastic model described in Section 3. In the stochastic model we assume discrete numbers of the enzyme and its complexes, whereas average concentrations describe the time evolution of other reagents.

In the considered set of rate constants the rates k_c and k_d are not significantly lower than the other rates thus we do not expect that the reduced model (eqn (9) and (10)) is a good approximation for the full one (eqn (1)–(5)). Nevertheless, we found the reduced model useful for the estimation of reaction parameters for which a sharp transition in the types of oscillations of $y(t)$ as a function of amplitude and frequency of periodic inflow of X molecules is observed.

In the following we consider a periodic inflow of X described by the formula:

$$I(t) = A \cdot (\sin(2\pi\nu t + \phi_0) + 1) \cdot \Theta(t)$$

It is parametrized by the amplitude A, the frequency ν and the initial (for $t = 0$) phase ϕ_0. The inflow $I(t)$ is always non-negative. The Heaviside step function $\Theta(t)$ describes the case where there is no inflow for $t < 0$ and it is switched on at $t = 0$. If $\phi_0 = 3 \cdot \pi/2$ then $I(t)$ is a continuous function. In this case $I(t \leq 0) \equiv 0$, next it increases and finally oscillates. For any other phase the inflow term is not continuous at $t = 0$; for example if $\phi_0 = \pi/2$ then $I(t = 0) = 2 \cdot A$. For $t > 0$ the time average of $I(t)$ equals to A and it is independent of the frequency and the initial phase. If the inflow amplitude $A = 0$ then $(x = 0, y = 0)$ is a only steady state of eqn (1)–(5) and it is stable. This stable state is the initial state for all numerical simulations of system evolution.

Fig. 3 illustrates the discrimination potential of the considered enzymatic model. The time evolution of $y(t)$ in the interval
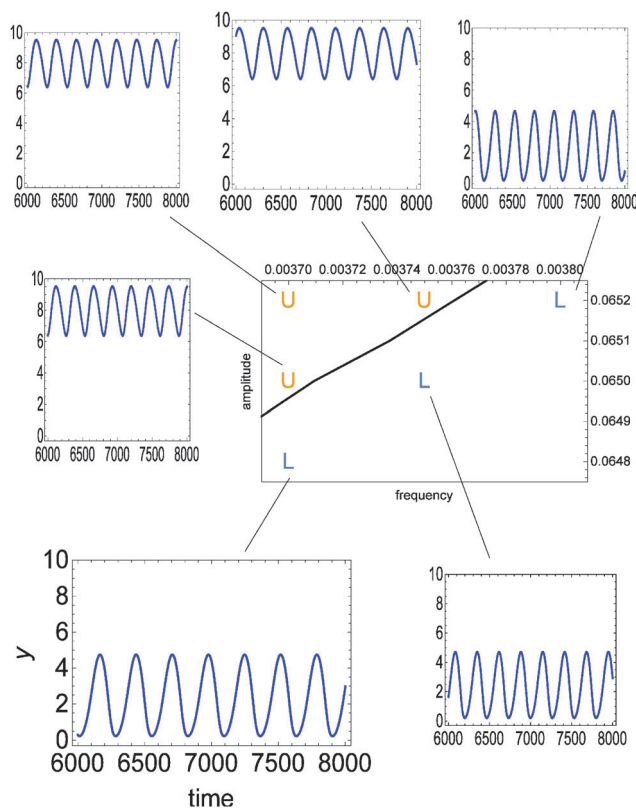

Fig. 3 The time evolution of $y(t)$ in the interval $t \in [6000, 8000]$ calculated from the full reaction model (eqn (1)–(5)) for a few values of amplitude A and frequency ν. In all cases $\phi_0 = 3 \cdot \pi/2$. The considered pairs of $(\nu, A) \in \{(0.00370, 0.0648), (0.00370, 0.0650), (0.00370, 0.0652), (0.00375, 0.0650), (0.00375, 0.0652), (0.00380, 0.0652)\}$ are marked in the phase space. The solid line separates the region where oscillations at low (L) and high (U) values of y are observed.

$t \in [6000, 8000]$ is shown for a few values of A and ν. The time interval in which the evolution is studied has been selected such that the system arrives to a stationary type of oscillations after the time $t_{min} = 6000$. Time $t_{max} = 8000$ is long enough to confirm that oscillations are stationary. Depending on the values of ν and A we observe two significantly different oscillation types: either $\forall_{t \in [t_{min}, t_{max}]} y(t) < 5$ or $\forall_{t \in [t_{min}, t_{max}]} y(t) > 6$. On the basis of this property we can introduce the following classification of oscillations. Let us define:

$$y_{min} = \min_{t \in [t_{min}, t_{max}]} y(t) \qquad (11)$$

and

$$y_{max} = \max_{t \in [t_{min}, t_{max}]} y(t) \qquad (12)$$

Depending on y_{min} and y_{max} we distinguish three types of oscillations. The oscillations of type-I extend over a large range of Y concentrations. They are expected for $2A > I_2$ and for a small frequency of inflow. In such a case the system dynamics is fast enough to adjust the concentrations of X and Y to the local value of inflow that changes between 0 and $2A$. Therefore $y_{min} \approx 0$ and $y_{max} \geq y_2$. The type-U oscillations are limited to the upper stable branch of the $R(x, y) = 0$ nullcline, so that $y_{min} > 6$.
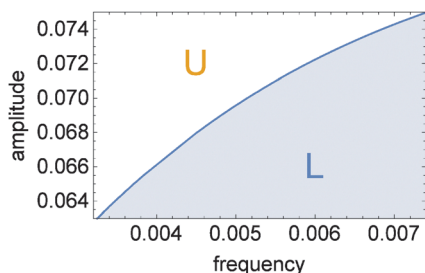
This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, **18**, 20518–20527 | **20521**

Fig. 4 The phase diagram showing the oscillation type as a function of inflow parameters (ν, A). The solid line corresponds to the full reaction model.

The type-L oscillations are on the lower stable branch of the $R(x, y) = 0$ nullcline and we have $y_{max} \leq 6$.

The line separating type-L from type-U oscillations in the phase space ν, A is shown in Fig. 3 and 4. It is calculated from a numerical solution of eqn (1)–(5). For classification oriented tasks it is important that the boundary region between domains corresponding to different oscillation types is very narrow. In Fig. 3 the location of points (ν, A) for which the transition between type-L and type-U oscillations occurs inside the interval $[t_{min}, t_{max}]$ is smaller than the line width. Therefore, the considered enzymatic system can precisely detect small changes in amplitude or frequency if they occur near the line separating oscillations with different character. The relative changes in these parameters of the order of 10^{-3} lead to distinctly different oscillation types (U or L). A significant difference in values of y can be used to couple the result of discrimination with the subsequent steps of chemical information processing.

The proposed discrimination method is based on the location of a given oscillation type with respect to the line separating different types of oscillations. Let us denote points on this line as (ν_c, A_c). As seen in Fig. 4 ν_c treated as a function of A_c is an increasing function $\nu_c = Z(A_c)$. For example, if we like to determine if the frequency of inflow ν is higher than ν_0 then we should select the inflow amplitude equal to $A = Z^{-1}(\nu_0)$ and observe the character of oscillations. If we observe type-U oscillations then $\nu < \nu_0$. Observation of type-L oscillations indicates that $\nu > \nu_0$. The derivative of $A = Z^{-1}(\nu)$ decreases with ν thus the sensitivity of the method changes with the frequency ν_0. In order to detect a high frequency inflow, the inflow amplitude should be selected with much higher precision than for $0.0032 \leq \nu_c \leq 0.004$. One can also use the transition between type-L and type-U oscillations to determine the inflow amplitude. Unlike for frequency, the range of discriminated amplitudes does not extend outside the interval $[A_1/2, A_1]$. If the amplitude $A > A_1$ then, at a high frequency ν the inflow of control molecules is close to the average inflow A. For such inflow the steady state of the system is located at the upper stable branch of the $R(x, y) = 0$ nullcline and no transition between oscillation character is expected.

3 Stochastic model of time evolution for a small number of enzyme molecules

In the stochastic analysis of the time evolution of the considered enzymatic system we assume that the internal fluctuations are related to a small number of enzyme molecules. The numbers of

molecules of all other reactants (B, C, D, X and Y) are large and these reagents are described by their average concentrations. This assumption is consistent with selected parameters of the reaction model. The sum of enzyme concentrations in the free and complexed forms is $e_0 = 0.001$ concentration unit, whereas concentrations of X and Y at oscillation minima are over 50 times larger.

The state of the system is described by three stochastic variables representing the numbers of molecules of E, EX, and EX_2 as functions of time (denoted as $N_E(t)$, $N_{EX}(t)$ and $N_{EX_2}(t)$, respectively) and by concentrations $x(t)$ and $y(t)$. Let us assume that the total number of enzyme molecules in free and all complexed forms is N_0 and that the volume of the system is $\Omega (= N_0/e_0 = 1000 N_0)$. At time $t = 0$ the state of the system is characterized by $N_{EX}(t = 0) = N_{EX}^0$, $N_{EX_2}(t = 0) = N_{EX_2}^0$, $N_E(t = 0) = N_E^0 = N_0 - N_{EX}^0 - N_{EX_2}^0$, $x(t = 0)$ and $y(t = 0)$. The time evolution of such a system is described by a master equation for the conditional probability distribution $P\big(N_{EX}, N_{EX_2}|N_E^0, N_{EX}^0, N_{EX_2}^0, x(t = 0), y(t = 0)\big)(t)$ coupled with the stochastic differential equations for $x(t)$ and $y(t)$:

$$\frac{dy}{dt} = k_p b n_0 \frac{N_{EX_2}(t)}{N_0} - k_c c y(t) \tag{13}$$

$$\begin{aligned}\frac{dx}{dt} = & -k_1 x(t) n_0 \frac{N_E(t)}{N_0} + k_{-1} n_0 \frac{N_{EX}(t)}{N_0} \\ & - k_2 x(t) n_0 \frac{N_{EX}(t)}{N_0} + k_{-2} n_0 \frac{N_{EX_2}(t)}{N_0} \\ & - k_d d x(t) + k_c c y(t) + I(t)\end{aligned} \tag{14}$$

$$N_E(t) = N_0 - N_{EX}(t) - N_{EX_2}(t) \tag{15}$$

$$\begin{aligned}&\frac{d}{dt} P\big(N_{EX}, N_{EX_2}|N_E^0, N_{EX}^0, N_{EX_2}^0, x(t = 0), y(t = 0)\big)(t) \\ &= -\big(k_1 x(t) N_E + k_{-1} N_{EX} + k_2 x(t) N_{EX} + k_{-2} N_{EX_2}\big) \\ &\quad \times P\big(N_{EX}, N_{EX_2}|N_E^0, N_{EX}^0, N_{EX_2}^0, x(t = 0), y(t = 0)\big)(t) \\ &\quad + k_1 x(t)(N_E + 1) \\ &\quad \times P\big(N_{EX} - 1, N_{EX_2}|N_E^0, N_{EX}^0, N_{EX_2}^0, x(t = 0), y(t = 0)\big)(t) \\ &\quad + k_{-1}(N_{EX} + 1) \\ &\quad \times P\big(N_{EX} + 1, N_{EX_2}|N_E^0, N_{EX}^0, N_{EX_2}^0, x(t = 0), y(t = 0)\big)(t) \\ &\quad + k_2 x(t)(N_{EX} + 1) \\ &\quad \times P\big(N_{EX} + 1, N_{EX_2} - 1|N_E^0, N_{EX}^0, N_{EX_2}^0, x(t = 0), y(t = 0)\big)(t) \\ &\quad + k_{-2}(N_{EX_2} + 1) \\ &\quad \times P\big(N_{EX} - 1, N_{EX_2} + 1|N_E^0, N_{EX}^0, N_{EX_2}^0, x(t = 0), y(t = 0)\big)(t)\end{aligned}$$
$$\tag{16}$$

where $n_0 = N_0/\Omega$.

The kinetic Monte Carlo method[27] can be applied for numerical simulations of system time evolutions. According to eqn (16) the probability that a system remains in its initial state: $\left(N_E^0, N_{EX}^0, N_{EX_2}^0\right)$ at time t is described by the equation:

$$\frac{d}{dt}P\left(N_{EX}^0, N_{EX_2}^0 | N_E^0, N_{EX}^0, N_{EX_2}^0, x(t=0), y(t=0)\right)(t)$$

$$= -\omega(t) \cdot P\left(N_{EX}^0, N_{EX_2}^0 | N_E^0, N_{EX}^0, N_{EX_2}^0, x(t=0), y(t=0)\right)(t) \tag{17}$$

where $\omega(t) = k_1 x(t) N_E^0 + k_{-1} N_{EX}^0 + k_2 x(t) N_{EX}^0 + k_{-2} N_{EX_2}^0$.

The solution of this equation is:

$$P\left(N_{EX}^0, N_{EX_2}^0 | N_E^0, N_{EX}^0, N_{EX_2}^0, x(t=0), y(t=0)\right)(t)$$

$$= \exp\left(-\int_0^t \omega(s)ds\right) \tag{18}$$

Thus, the cumulative distribution function that a next reaction changing the number of enzymes occurs at a time shorter than t is:

$$D\left(N_E^0, N_{EX}^0, N_{EX_2}^0, x(t=0), y(t=0)\right)(t) = 1 - \exp\left(-\int_0^t \omega(s)ds\right) \tag{19}$$

On the basis of eqn (19) we can apply the following method of numerical stochastic simulations of the system, based on the kinetic Monte Carlo algorithm:[27]

(1) At time t the system state is $N_E^0, N_{EX}^0, N_{EX_2}^0, x(t), y(t)$,

(2) Find the time δt at which the first reaction involving the enzyme occurs by selecting a random number r from the uniform distribution in the interval $[0, 1]$ and solving the equation:

$$r = 1 - \exp\left(-\int_t^{t+\delta t} \omega(s)ds\right)$$

The values of $x(s)$ and $y(s)$ come from solution of equations:

$$\frac{dy(s)}{ds} = k_p b n_0 \frac{N_{EX_2}(t)}{N_0} - k_c c y(s) \tag{20}$$

and

$$\frac{dx(s)}{ds} = -k_d d x(s) + k_c c y(s) + I(s) \tag{21}$$

These equations describe the time evolution of $x(s)$ and $y(s)$ in the time interval $[t, t + \delta t]$ when no reaction that changes the number of enzyme molecules occurs.

(3) Identify the process involving the enzyme that occurred at δt by selecting the specific process with probabilities equal to:

$$N_E^0 k_1 \int_t^{t+\delta t} x(s)ds \bigg/ \int_t^{t+\delta t} \omega(s)ds,$$

$$\left(N_{EX}^0 k_{-1}\int_t^{t+\delta t} ds = N_{EX}^0 k_{-1}\delta t\right) \bigg/ \int_t^{t+\delta t} \omega(s)ds,$$

$$\left(N_{EX}^0 k_2 \int_t^{t+\delta t} x(s)ds\right) \bigg/ \int_t^{t+\delta t} \omega(s)ds$$

and

$$\left(N_{EX_2}^0 k_{-2}\int_t^{t+\delta t} ds = N_{EX_2}^0 k_{-2}\delta t\right) \bigg/ \int_t^{t+\delta t} \omega(s)ds$$

respectively.

(4) Modify the numbers N_E, N_{EX} and N_{EX_2} according to the process selected in step (3) and change the value of x by $\pm e_0/N_0$ to take into account production or consumption of a single X molecule in the selected reaction.

(5) Update time ($t = t + \delta t$) and repeat (1) until the selected time of system simulation (t_{max}) is reached.

4 Results

The algorithm described in the previous section generates a stochastic evolution that takes into account the discrete character of reactions involving the enzyme. Having the value of $y(t)$ we can define the probability p_U that the evolution belongs to type-U oscillations as:

$$p_U = \mu(y(s) > 6 | s \in [t_{min}, t_{max}])/(t_{max} - t_{min}) \tag{22}$$

where $\mu(y(s) > 6 | s \in [t_{min}, t_{max}])$ is the Lebesgue measure[29] of the set of points s in the interval $[t_{min}, t_{max}]$ at which $y(s) > 6$. Similarly the probability p_L that a trajectory belongs to type-L oscillations can be calculated as:
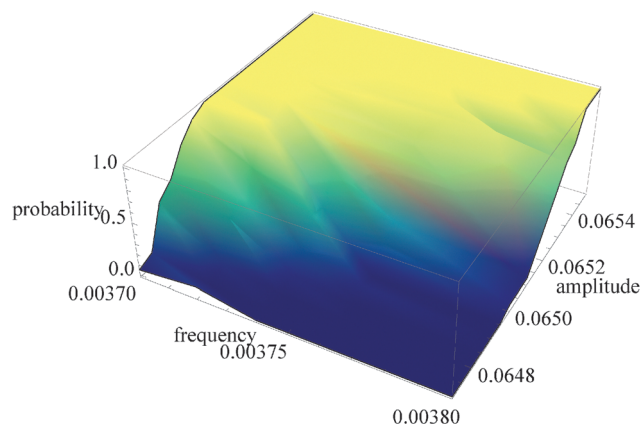
$$p_L = \mu(y(s) \leq 6 | s \in [t_{min}, t_{max}])/(t_{max} - t_{min}) \tag{23}$$

Of course $p_L + p_U = 1$. In the deterministic case such a definition is in a full agreement with the previously introduced criteria. It returns $p_L = 1$ for type-L oscillations and $p_U = 1$ for type-U oscillations.

Fig. 5a, 6a and 7a show $p_U(\nu, A)$ in the form of 3D plots for $N_0 = 10\,000$, 1000 and 100, respectively. For each pair (ν, A) the result is averaged over 50 independent evolutions that are initiated from: $N_{EX}(t = 0) = N_{EX_2}(t = 0) = 0$, $N_E(t = 0) = N_0$, $x(t = 0) = 0$ and $y(t = 0) = 0$. The range of ν and A is the same as that used in Fig. 3. The same result, showing the value of $p_U(\nu, A)$ coded in color, is given in Fig. 5b, 6b and 7b. The black solid line in these figures separates regions of type-L and type-U oscillations obtained from phenomenological kinetic equations describing the full reaction model (eqn (1)–(5)) and it is the same as in Fig. 3 and 4.

In order to make a qualitative estimation of discriminator reliability let us simplify the device and assume that the binary input information is introduced in the form of two different inflows. One of them is characterized by the frequency ν_{01} and the amplitude A_{01} (state 0) and the other is described by (ν_{02}, A_{02}) (state 1). The discriminator is supposed to recognize the input and produce type-U oscillations in the first case and type-L oscillations in state 1. At the beginning let us consider $\nu_{01} = \nu_{02} = 0.0037$, $A_{01} = 0.0652$ and $A_{02} = 0.0648$. According to the results shown in Fig. 2 the reaction with a large number of enzyme molecules performs the required operation, so it discriminates the input inflow according to its amplitude. The discriminator accuracy is better than 0.0004 amplitude unit. Similarly for $A_{01} = A_{02} = 0.065$, $\nu_{01} = 0.0037$ and

This journal is © the Owner Societies 2016

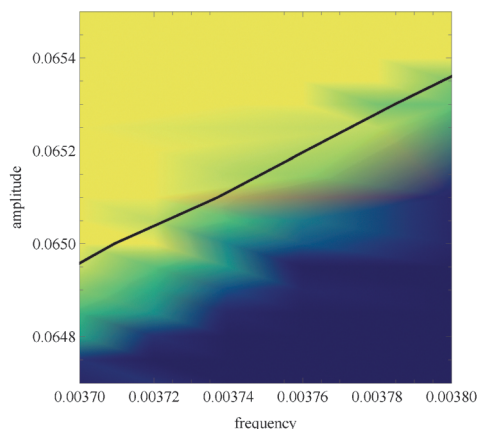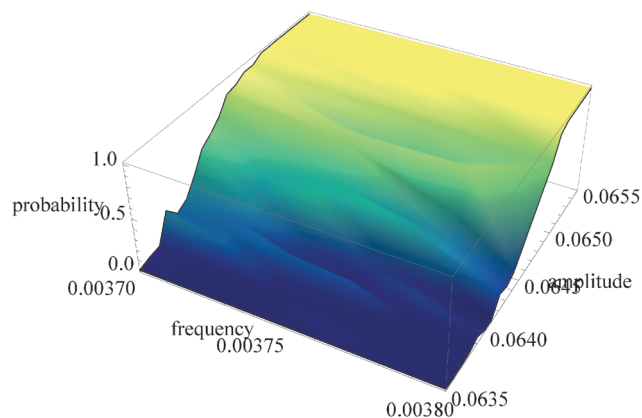*Phys. Chem. Chem. Phys.*, 2016, **18**, 20518–20527 | **20523**

Fig. 5 The function $p_U(\nu, A)$ for $N_0 = 10\,000$. (a) 3D plot of probability; (b) the same data as a surface plot; the solid line separates regions of type-L and type-U oscillations and it is the same as on Fig. 3 and 4.
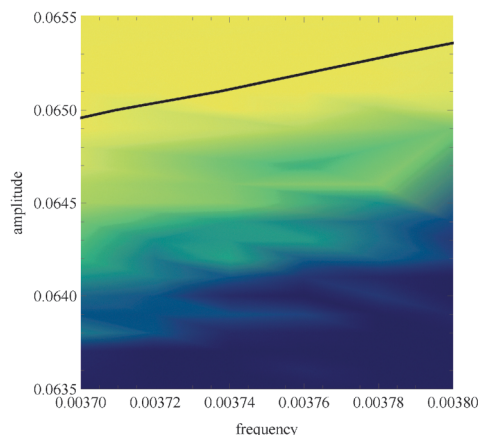

Fig. 6 The function $p_U(\nu, A)$ for $N_0 = 1000$. (a) 3D plot of probability; (b) the same data as a surface plot; the solid line separates regions of type-L and type-U oscillations and it is the same as in Fig. 3 and 4.

$\nu_{02} = 0.0038$ the reaction functions as a discriminator of the inflow frequency. Such simplified discrimination activity can be interpreted as a signal transmission through a communication channel. The idea is illustrated in Fig. 8. The basic quantity characterizing a communication channel is the channel capacity.[28] It represents the maximum amount of information that can be transmitted through the channel within a single operation. If the number of enzyme molecules is very large then the reaction works as a noiseless binary communication channel and for both amplitude and frequency discrimination its capacity is $C_{amp} = C_{freq} = 1$ bit. In the case of a small number of molecules the reaction represents a noisy communication channel with probabilities of correct discrimination p_1 and p_2. For the considered numbers of molecules the value of $p_1 = p_U(\nu, A)$ is given in Fig. 5–7. The value of $p_2 = 1 - p_U(\nu, A)$. The capacity of asymmetric binary channel is:

$$C = \beta h(p_1) + (1 - \beta)h(p_2) - h(\beta p_1 + (1 - \beta)(1 - p_2)) \quad (24)$$

where

$$h(z) = -z \log_2(z) - (1 - z) \log_2(1 - z)$$

$$\beta = \frac{p_2 + \alpha \times (p_2 - 1)}{(1 + \alpha)(p_1 + p_2 - 1)}$$

and

$$\alpha = 2^{\frac{h(p_2) - h(p_1)}{1 - p_1 - p_2}}$$

If $p_1 = 1 - p_2$ then the channel capacity $C = 0$. It reflects an obvious fact that a channel that gives exactly the same answer to both input values does not transmit any information.

For $N_0 = 10\,000$ (Fig. 5) the system response to an oscillating inflow, predicted by the deterministic kinetic equations agrees well with the stochastic simulations. The phenomenological line separating regions corresponding to different oscillation types is located in the region where the gradient of $p_U(\nu, A)$ is large. It gives a reasonable estimation for separation of the region where $p_U(\nu, A)$ is small (blue color) from the region where $p_U(\nu, A) \sim 1$ (yellow color). A more precise separation of these regions can be obtained if the separating line is shifted by 0.0001 unit towards smaller amplitudes. Therefore, it can be expected that the system with $N_0 = 10\,000$ enzyme molecules should work as a reliable discriminator of the periodic inflow, just as described in Section 2. The reliability of amplitude and frequency discriminators described above, measured as the capacity of asymmetric channels, is equal to: $C_{amp} = 0.355$ bit and $C_{freq} = 0.882$ bit. The capacity of the
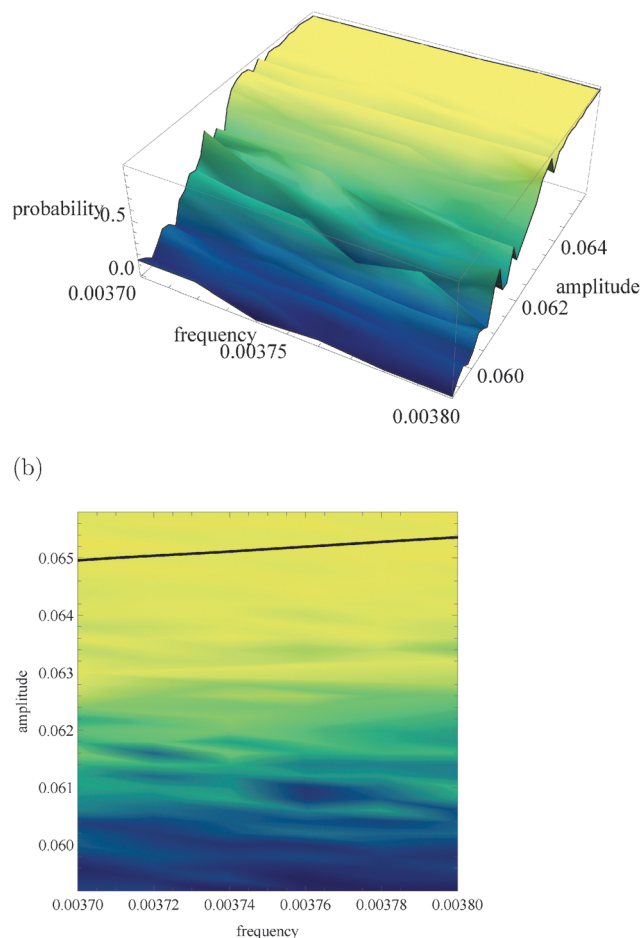
Fig. 7 The function $p_U(\nu, A)$ for $N_0 = 100$. (a) 3D plot of probability, (b) the same data as a surface plot, the solid line separates regions of type-L and type-U oscillations and it is the same as on Fig. 3 and 4.

(a)

$$(\nu_{01}, A_{01}) \xrightarrow{1} \text{type-U oscillations}$$

$$(\nu_{02}, A_{02}) \xrightarrow{1} \text{type-L oscillations}$$

(b)

$$(\nu_{01}, A_{01}) \xrightarrow{p_1} \text{type-U oscillations}$$

$$1 - p_2$$

$$1 - p_1$$

$$(\nu_{02}, A_{02}) \xrightarrow{p_2} \text{type-L oscillations}$$
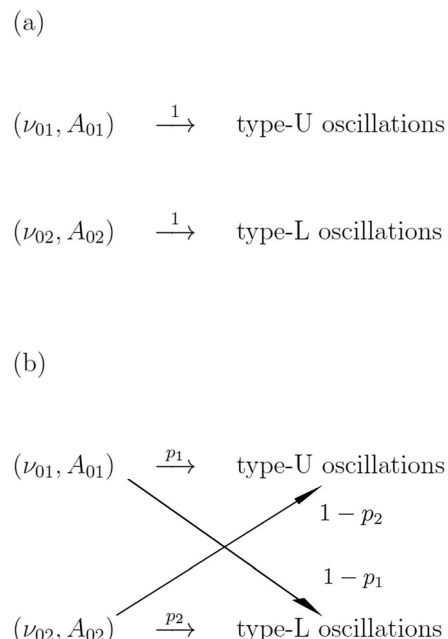
Fig. 8 Schematic representation of the discriminator as a binary communication channel. (a) For a large number of molecules the system acts as a noiseless binary symmetric channel; (b) for a small number of molecules the system can be regarded as a noisy binary asymmetric channel.

amplitude discriminator can be increased if the input amplitudes are decreased by 0.0001 unit. For $\nu_{01} = \nu_{02} = 0.0037$, $A_{01} = 0.0651$ and $A_{02} = 0.0647$ we obtain $C_{amp2} = 0.849$ bit. Therefore, for $N_0 = 10\,000$ the presence of internal fluctuations does not significantly change the functionality of a discriminator based on an enzymatic reaction. For the considered reaction parameters it can operate as a reliable discriminator of amplitudes that differ by 0.0004 unit or frequencies that differ by 0.0001 unit.

For $N_0 = 1000$ (Fig. 6) the region of type-L oscillations is even more shifted towards smaller values of A if compared to predictions of phenomenological kinetic equations. Qualitatively a similar trend was observed for a thermochemical discriminator of temperature described in ref. 24. Now the line separating regions corresponding to different types of oscillations are located in the region, where the stochastic model predicts a high probability of type-U oscillations. Therefore, if $N_0 = 1000$ the analysis presented in Section 2 fails to describe correctly the discriminating potential of the considered reaction. The capacity of channels representing amplitude and frequency discriminators considered above is $C_{amp} = 0.038$ bit and $C_{freq} = 0.094$ bit. For $N_0 = 1000$ the transition between type-L

and type-U oscillations is still well pronounced but it is shifted towards low amplitudes by 0.001 unit if compared with phenomenological kinetic equations. In order to find the optimum conditions for a discriminator of amplitudes that differ by 0.0004 unit let us consider: $\nu_{01} = \nu_{02} = 0.0037$, $A_{01} = 0.0645$ and $A_{02} = 0.0641$ The capacity of the corresponding information channel is $C_{amp3} = 0.157$ bit, and thus it is over 4 times larger than C_{amp}. In order to discriminate frequency one can use $A_{01} = A_{02} = 0.0645$, $\nu_{01} = 0.0037$ and $\nu_{02} = 0.0038$ and $C_{freq2} = 0.238$ bit. For such inputs the discriminator needs at least $1/C_{freq2} \sim 1/C_{amp3} \sim 5$ observations for a correct estimation. In conclusion, the system with $N_0 = 1000$ can operate as a discriminator, but the phenomenological approach is useless for estimation of the range of frequencies or amplitudes where it can operate. The accurate analysis should take internal fluctuations into account.

For $N_0 = 100$ (cf. Fig. 7) the region where a gradient of $p_U(\nu, A)$ is observed is shifted towards a smaller amplitude by ~ 0.004 unit with respect to kinetic equations. The parameters of input signals considered for $N_0 = 10\,000$ or $N_0 = 1000$ are located in the region where for $N_0 = 100$ $p_U(\nu, A)$ is almost constant and close to 1. The capacity of corresponding communication channels is close to zero. If $N_0 = 100$ then the reaction can be still useful as an amplitude discriminator if the amplitudes of input signals are significantly reduced and the resolution of the discriminator is much smaller. For example if $\nu_{01} = \nu_{02} = 0.0037$, $A_{01} = 0.064$ and $A_{02} = 0.060$ then the corresponding channel capacity is $C_{amp4} = 0.433$ bit. The function $p_U(\nu, A)$ does not depend on the frequency in the considered range of (ν, A), so for $N_0 = 100$ the reaction does not work as a frequency discriminator.

This journal is © the Owner Societies 2016

Phys. Chem. Chem. Phys., 2016, **18**, 20518–20527 | **20525**

5 Conclusions

In this paper we presented arguments that an enzymatic reaction can function as a medium for analyzing information coded in the frequency and amplitude of the reagent inflow. There are many parameters (rate constants and concentrations of reagents in excess) in the considered reaction model. Therefore, there is a broad range of parameter values, at which the reaction can be used as a discriminator. We restricted our attention to the cases in which reactions involving the enzyme and its complexes can be regarded as faster than the other processes. We reduced the full reaction model to two kinetic equations and selected parameter values such that in the reduced model one of its nullclines has a sigmoidal form. Using arguments given in ref. 23 we estimated the range of input frequencies or amplitudes in which a sharp transition between different forms of oscillations in product concentration, necessary for discrimination, is observed. The precise calculations of discriminator characteristics were based on the complete reaction model, which provides a more realistic description of the dynamic response than the reduced model. On the basis of phenomenological kinetic equations we have shown that the considered reaction separates the space of input states into two output classes. The output classes can be easily distinguished, because the corresponding concentrations of reaction products are significantly different. The considered enzymatic reaction significantly enhances differences between inputs, especially for the input states located close to the line separating different answers of the system. According to the kinetic equations the reaction can transform small relative differences in the input amplitude of the order of 10^{-3} into relative differences in output concentrations of the order of 1.

Having in mind that information processing in simple living organisms is incorporated into their metabolic cycle we analyzed the case where the number of enzyme molecules is small and the internal fluctuations related to reactions involving enzyme molecules and its complexes should be taken into account. Such analysis was performed using a master equation that describes the probability of finding a specific state of a system as a function of time. We considered a narrow range of frequencies and amplitudes and studied discriminators operating with $N_0 = 10\,000$, 1000 and 100 enzyme molecules. The conclusions coming from phenomenological kinetic equations were qualitatively confirmed by the stochastic analysis based on kinetic Monte-Carlo simulations. For $N_0 = 10\,000$ the results of stochastic simulations match quite well the phenomenology. The discrepancy between kinetic equations and stochastic simulations becomes important when N_0 gets smaller. For $N_0 = 1000$ the reaction still operates as a discriminator, but the range of frequencies and amplitudes for which it is useful is different than predicted by the phenomenology. For yet smaller N_0 (=100) the considered range of frequencies was too small and the medium was not able to discriminate between inputs corresponding to different frequencies. The reaction with $N_0 = 100$ can still function as the amplitude discriminator, but its resolution is an order of magnitude smaller than that for $N_0 = 1000$.

It is interesting to notice that the estimation of the number of particles needed for information processing with an enzymatic reaction gives similar results as for the recently analyzed thermochemical discriminator.[24] It was found that such a discriminator of the ambient temperature needs more than 10^5 molecules for a reliable operation. Here the number is around 1000 molecules of the enzyme only. Including the other reagents the total number of active particles involved in information processing is of the order of millions, which fairly agree with the previous estimation.

References

1 H. Haken, *Brain Dynamics: An Introduction to Models and Simulations*, Springer Series in Synergetics, Springer, 2008.
2 *Biomolecular Information Processing. From Logic Systems to Smart Sensors and Actuators*, ed. E. Katz, Wiley-VCH Verlag, 2012.
3 *Molecular and Supramolecular. From Molecular Switches to Logic Systems Information Processing*, ed. E. Katz, Wiley-VCH Verlag, 2012.
4 F. Muzika and I. Schreiber, *J. Chem. Phys.*, 2013, **139**, 164108.
5 V. Privman, O. Zavalov, L. Halamkova, F. Moseley, J. Halamek and E. Katz, *J. Phys. Chem. B*, 2013, **117**, 14928–14939.
6 E. Katz and V. Privman, *Chem. Soc. Rev.*, 2010, **39**, 1835–1857.
7 L. Michaelis and M. L. Menten, Die Kinetik der Invertinwirkung, *Biochem. Z.*, 1913, **49**, 333–369, the recent English translation: K. A. Johnson and R. S. Goody, *Biochemistry*, 2011, **50**, 8264–8269.
8 V. Privman, J. Halamek, M. A. Arugula, D. Melnikov, V. Bocharova and E. Katz, *J. Phys. Chem. B*, 2010, **114**, 14103–14109.
9 M. Pita, V. Privman, M. A. Arugula, D. Melnikov, V. Bocharova and E. Katz, *Phys. Chem. Chem. Phys.*, 2011, **13**, 4507–4513.
10 S. Domansky and V. Privman, *J. Phys. Chem. B*, 2012, **116**, 13690–13695.
11 V. Privman, B. E. Fratto, O. Zavalov, J. Halamek and E. Katz, *J. Phys. Chem. B*, 2013, **117**, 7559–7568.
12 R. A. Laskowski, F. Gerick and J. M. Thornton, *FEBS Lett.*, 2009, **583**, 1692–1698.
13 T. Tsuruyama, T. Nakamura, G. Jin, M. Ozeki, Y. Yamada and H. Hiai, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 8253–8258.
14 T. Tsuruyama, *PLoS One*, 2014, **9**, e102911.
15 K.-P. Zauner and M. Conrad, *Biotechnol. Prog.*, 2001, **17**, 553–559.
16 J. D. Rabinowitz, J. J. Hsiao, K. R. Gryncel, E. R. Kantrowitz, X.-J. Feng, G. Li and H. Rabitz, *Biochemistry*, 2008, **47**, 5881–5888.

17 S. Bakshi, O. Zavalov, J. Halamek, V. Privman and E. Katz, *J. Phys. Chem. B*, 2013, **117**, 9857–9865.

18 H. C. Barrett, *Mind and Language*, 2005, **20**, 259–287.

19 J. Halamek, O. Zavalov, L. Halamkova, S. Korkmaz, V. Privman and E. Katz, *J. Phys. Chem. B*, 2012, **116**, 4457–4464.

20 O. Zavalov, V. Bocharova, V. Privman and E. Katz, *J. Phys. Chem. B*, 2012, **116**, 9683–9689.

21 H. Lederman, J. Macdonald, D. Stefanovic and M. N. Stojanovic, *Biochemistry*, 2006, **45**, 1194–1199.

22 R. Baron, O. Lioubashevski, E. Katz, T. Niazov and I. Willner, *Angew. Chem., Int. Ed.*, 2006, **45**, 1572–1576.

23 H. Ueno, T. Tsuruyama, B. Nowakowski, J. Gorecki and K. Yoshikawa, *Chaos*, 2015, **25**, 103115.

24 J. Gorecki, B. Nowakowski, J. N. Gorecka and A. Lemarchand, *Phys. Chem. Chem. Phys.*, 2016, **18**, 4952–4960.

25 A. V. Hill, *J. Physiol.*, 1910, **40**, 389–403.

26 S. Goutelle, M. Maurin, F. Rougier, X. Barbaut, L. Bourguignon, M. Ducher and P. Maire, *Fundam. Clin. Pharmacol.*, 2008, **22**, 633–648.

27 D. T. Gillespie, *Annu. Rev. Phys. Chem.*, 2007, **58**, 35.

28 T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley and Sons, Hoboken, NJ, 2nd edn, 2006.

29 K. Maurin, *Analysis, part II*, Springer, Netherlands, 1980.