Measuring kinetic rate constants of multiple-component reactions with optical biosensors

David A. Edwards a, Ryan M. Evans b,*, Wenbin Li a

a Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA
b Applied and Computational Mathematics Division, Information and Technology Lab, National Institute for Standards and Technology, Gaithersburg, MD 20899, USA

A R T I C L E I N F O

Article history:
Received 20 March 2017
Received in revised form 14 June 2017
Accepted 16 June 2017
Available online 22 June 2017

Keywords:
Optical biosensors
Kinetic rate constants
Mathematical modeling
Parameter identification

A B S T R A C T

One may measure the kinetic rate constants associated with biochemical reactions using an optical biosensor: an instrument in which ligand molecules are convected through a flow cell over a surface to which receptors are immobilized. If there are multiple reactants, one is faced with the problem of fitting multiple kinetic rate constants to one signal, since data from all of the reacting species is lumped together. Even in the presence of ambiguous data, one may use a series of experiments to accurately determine the rate constants. Moreover, the true set of rate constants may be identified by either postprocessing the signals or adjusting the ligand inflow concentrations.

P1 : E + L2 \rightarrow_{j_D}^j{k_a} EL_2.

Published by Elsevier Inc.

Introduction

Many biochemical reactions in nature involve a stream of ligand molecules flowing through a fluid-filled volume over a surface to which receptors are confined. Such surface-volume reactions occur during platelet adhesion [1], drug absorption [2], and antigen-antibody interactions [3]. Fundamental to understanding these reactions is obtaining accurate quantitative measurements of the underlying kinetic rate constants. As seen in Fig. 1, scientists use optical biosensors for measuring rate constants associated with surface-volume reactions.

In a typical biosensor experiment, ligand molecules are convected through a flow cell over a surface to which receptors are immobilized. Mass changes on the surface due to ligand binding are averaged over a portion of the channel floor to produce sensogram readings of the form

\[ S(t) = sB(t), \]

where \( B(t) \) denotes the concentration of bound ligand, and \( s \) is proportional to the molecular weight of the ligand.

Recent advances in biosensor technology have enabled researchers to measure kinetic rate constants associated with multiple simultaneous independent reactions. Examples of such technology includes array-based [4,5] and multiple-channel [6–8] biosensors, and while these devices typically yield multiple independent measurements for each array-spot or channel, scientists are currently attempting to infer multiple kinetic rate constants associated with coupled reactions from a single signal.

In particular, chemists are now using biosensor experiments to study how cells cope with DNA damage. Harmful DNA lesions can impair a cell’s ability to replicate DNA. One way a cell may respond to such a lesion is through DNA translesion synthesis. For a complete description of this process we refer the interested reader to [9–11]; however, for our purposes it is sufficient to know that DNA translesion synthesis involves multiple interacting components: a Proliferating Cell Nuclear Antigen (PCNA) molecule, polymerase \( \eta \), and polymerase \( \delta \). Further, in order for a successful DNA translesion synthesis event to occur, polymerase \( \eta \) must bind with a PCNA molecule. A central question surrounding DNA translesion synthesis is whether polymerase \( \eta \) directly binds with the PCNA, or whether the polymerase \( \eta \)-PCNA complex forms as a result of a catalysis-type ligand switching process [12].

The former scenario is depicted in Fig. 2, which shows polymerase \( \eta \) directly binding with a PCNA molecule, i.e. the reaction

\[ P_1 : E + L_2 \rightarrow_{j_D}^j{k_a} EL_2. \]
Here, the PCNA and polymerase η molecules are denoted as E and L₃ respectively. Additionally, kₐ denotes the rate at which L₃ binds with a PCNA molecule E, and kₐ denotes the rate at which L₃ dissociates from a PCNA molecule E. We will refer to this as pathway one, or simply P₁ as in (1.2a).

The ligand switching process is shown in Fig. 3 and stated precisely as

\[
P₂ : E + L₁ \overset{kₐ}{\underset{jₐ}{\rightleftharpoons}} EL₁, EL₁ \overset{jₐ}{\underset{kₐ}{\rightleftharpoons}} EL₂ + L₁.
\]  

Fig. 1. Cross-sectional schematic of an optical biosensor experiment. The instrument has length I and height h. Ligand molecules are convected into instrument in a Poiseuille flow profile and transported to the surface to bind with receptors. The receptors are limited to the reacting zone \(x = x_{\text{min}}\) to \(x = x_{\text{max}}\).

Here, the PCNA and polymerase η molecules are denoted as E and L₂ respectively. Additionally, kₐ denotes the rate at which L₂ binds with a PCNA molecule E, and kₐ denotes the rate at which L₂ dissociates from a PCNA molecule E. We have labeled polymerase δ, polymerase η, and the PCNA molecule as L₁, L₂, and E respectively.

By measuring the rate constants associated with this multiple-component process, one could determine whether EL₂ forms as a result of ligand binding or the ligand switching process described above. To measure the kinetic rate constants, scientists first immobilize PCNA molecules on the surface of the biosensor. After receptor immobilization, scientists inject L₁ and L₂ into the instrument at concentrations C₁ and C₂. The two ligands are then transported to the surface to bind with available receptor sites E, creating the three reacting species EL₁, EL₁L₂, and EL₂ at concentrations B₁(t), B₁(t), and B₂(t).

The presence of multiple reacting species changes the interpretation of the sensogram reading. Most biosensors, including the Biacore, measure only mass changes at the surface due to ligand binding; hence (1.1) becomes

\[
\mathcal{S}(t) = s₁B₁(t) + (s₁ + s₂)B₁₂(t) + s₂B₂(t),
\]  

where \(sᵢ\) is proportional to the molecular weight of \(Lᵢ\), for \(i = 1, 2\). The lumped signal (1.3) complicates parameter estimation. Since one is faced with the challenge of fitting multiple rate constants to one signal, it may be difficult for standard algorithms to find a unique solution to the corresponding least-squares problem.

Indeed, the problem of inferring multiple kinetic rate constants from one signal is an interesting problem which has received little attention to date. This problem is addressed [13,14], although both of these papers are on multiple receptor kinetics. Multiple-ligand
kinetics are discussed briefly in [15, pp.101–102]. Svitel et al. propose a method for inferring a distribution of kinetic rate constants in a model that could be used to describe multiple-ligand kinetics or multiple receptor sites for the same ligand in Ref. [16]. However, their results differ from the present work in at least three respects. In Ref. [16], the authors consider a continuous distribution of functionally distinct ligand molecules. By contrast, since we are concerned with elucidating the role of polymerase $\eta$ during DNA translesion synthesis, we are considering two distinct ligands. Additionally, their approach is based upon Tikhonov regularization, which introduces a bias into parameter estimation. Since this may lead to erroneous results, it is important that Tikhonov regularization is combined with a priori knowledge about the system at hand. Since conventional techniques such as fluorescent microscopy may modify protein activity, a priori estimates of the rate constants are unavailable, and methods not based on Tikhonov regularization are necessary. Finally, in Ref. [16] the authors consider systems in which reactions associated with different ligand molecules proceed independently, which is clearly not the case for the application that we have in mind. In fact, if the reactions (1.2) decoupled then one could use standard techniques to measure the kinetic rate constants and the role of polymerase $\eta$ in DNA translesion synthesis would be understood.

We explore the difficulty of fitting the rate constants in (1.3) by developing a mathematical model for the multiple-ligand biosensor experiment described herein. This mathematical model is used to show that there are two different sets of kinetic rate constants which correspond to very similar sensogram signals. Moreover, by proposing and numerically simulating a set of four experiments, it is shown that one can recover the true rate constants associated with each of the signals. Furthermore, it is shown that the sensogram readings can be disambiguated by post-processing or varying the ligand inflow concentrations.

**Mathematical model, numerical experiments, and results**

The surface reactions (1.2) are modeled with a set of Ordinary Differential Equations (ODEs):

$$\frac{d B_1}{dt} = \frac{1}{2} k_3 (R_t - B_x) C_1 + \frac{1}{2} k_4 B_1 - \frac{1}{2} k_4 B_1 C_2.$$  \hspace{1cm} (2.1a)

$$\frac{d B_2}{dt} = \frac{1}{2} k_3 B_1 C_2 + \frac{1}{2} k_4 B_2 C_1 - \frac{1}{2} k_4 B_2 C_2 - \frac{1}{2} k_4 B_2.$$  \hspace{1cm} (2.1b)

$$\frac{d B_3}{dt} = \frac{1}{2} k_3 (R_t - B_x) C_2 + \frac{1}{2} k_4 B_3 C_1 - \frac{1}{2} k_4 B_3 C_2 - \frac{1}{2} k_4 B_3 C_1.$$  \hspace{1cm} (2.1c)

Here $B_3 = B_1 + B_1 + B_2$. The system (2.1) is subject to the initial conditions

$$B_1(0) = B_{1.0}, B_{12}(0) = B_{12.0}, B_2(0) = B_{2.0}.$$  \hspace{1cm} (2.2)

The system (2.1) and (2.2) is expressed more compactly in matrix form as

$$\frac{d B}{dt} = -A B + f, \quad B(0) = B_0.$$  \hspace{1cm} (2.3)

$$B(t) = (B_1(t), B_{12}(t), B_2(t))^T.$$  \hspace{1cm} (2.4)

$B_0 = (B_{1.0}, B_{12.0}, B_{2.0})^T$ is a vector containing the three initial conditions and

$$A = \begin{pmatrix}
(1 \cdot k_3 C_1 + \frac{1}{2} k_4 C_2)
-\frac{1}{2} k_4 C_2
\frac{1}{2} k_4 C_2

(1 \cdot k_3 C_1 + \frac{1}{2} k_4 C_2)
\frac{1}{2} k_4 C_2
-\frac{1}{2} k_4 C_2

(1 \cdot k_3 C_1 + \frac{1}{2} k_4 C_2)
\frac{1}{2} k_4 C_2
-\frac{1}{2} k_4 C_2
\end{pmatrix}.$$  \hspace{1cm} (2.5)

$$f = (1 \cdot k_4 R_t C_1, 0, 2 \cdot k_4 R_t C_2)^T.$$  \hspace{1cm} (2.6)

The ODE model (2.3) assumes that reaction kinetics are accurately described by the well-stirred approximation. Since mass transport effects on ligand binding are typically minimal in experimental regimes, they are not considered herein. For a discussion of transport effects on multiple-component reactions in optical biosensors one may see Refs. [13, 17].

The gold standard for verifying the model is comparison with experimental data. At press time, the PCNA experiments had not been completed. Therefore, we use synthetic data from numerical simulations of the model (2.3) using an ODE solver in MATLAB. This data was then used to generate the signal (1.3), which is written concisely as

$$S(t) = s^T B(t).$$  \hspace{1cm} (2.7)

$$s = (S_1, S_{12}, S_2)^T.$$  \hspace{1cm} (2.8)

Even though the expression for $B$ from (2.3) is in some sense “exact”, the synthetic sensogram data $S$ which results still lumps the solutions into a single signal from which the rate constants must be estimated, either with or without postprocessing for disambiguation.

In the numerical experiments, ligands $L_1$ and $L_2$ were injected sequentially. First, $L_1$ was injected until (2.5) reached a steady state. The injection of $L_1$ was stopped simultaneously with starting the injection of $L_2$, which continued until the signal sensogram signal reached a steady state again. There was no bound ligand at the start of the first injection (so $B_0 = 0), while the initial condition for the second injection was the steady-state solution to (2.3) during the first injections. This numerical experiment was performed for two sets of parameters, which are labeled Signal 1 and Signal 2 in the exact column of Table 1. The corresponding signals are depicted in Figs. 4 and 5, and are labeled Signal 1 and Signal 2 in accordance with the parameter values in Table 1.

First consider signal one, and notice the increase in the sensogram reading upon the start of the second injection. Examining the bound state concentrations depicted in Fig. 4, it is evident that the increase is due to a large influx of $EL_2$. Since the rate constant $k_4$ is an order of magnitude larger than the others, $L_2$ molecules almost immediately bind with available receptor sites at the start of the second injection.

Now consider signal two; again, there is a sharp and sudden increase in the sensogram reading at the start of the second injection. Since the rate constant $k_3$ is an order of magnitude larger than the others, $L_2$ molecules almost immediately bind with the $EL_1$ present upon the start of the second injection. This behavior is evident in the bound state concentrations depicted in Fig. 5: once $L_2$ is injected there is a sudden decrease in $B_1$, and a sudden increase in $B_1$. Although the sensogram signals correspond to different sets of rate constants, they appear quite similar to the eye. Indeed, Fig. 6 indicates that the maximum relative difference between the two readings is approximately five percent. This is significant in light of the fact that true physical experimental data may be contaminated by noise.
Table 1  
The values in the column labeled “Exact” were used to generate the signals 2.1 and 2.2. The notation $C_1 = 10^{-11}$ mol/cm³ indicates that a value of $C_1 = 10^{-11}$ mol/cm³ was used during the first injection, and a value of $C_2 = 0$ mol/cm³ was used during the second; analogous notation is used for $C_2$. The values in the column labeled “Fitted” are those recovered from our parameter estimation algorithm described on pages six and seven.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Signal 1</th>
<th>Signal 2</th>
<th>Signal 1</th>
<th>Signal 2</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_a$ (cm³/(mol·s))</td>
<td>$10^8$</td>
<td>$10^8$</td>
<td>$1.00 \times 10^8$</td>
<td>$1.00 \times 10^8$</td>
<td>[4, 18, 19]</td>
</tr>
<tr>
<td>$k_d$ (s⁻¹)</td>
<td>$10^{-3}$</td>
<td>$3 \times 10^9$</td>
<td>$10^{-3}$</td>
<td>$3 \times 10^9$</td>
<td>[4, 18, 19]</td>
</tr>
<tr>
<td>$C_1$ (mol/cm³)</td>
<td>$10^{-11}, 0$</td>
<td>$0, 10^{-11}$</td>
<td>$10^{-11}$</td>
<td>$10^{-11}$</td>
<td>[4]</td>
</tr>
<tr>
<td>$s_1$ (kDa)</td>
<td>220.2</td>
<td>71.5</td>
<td>220.2</td>
<td>71.5</td>
<td>[20–22]</td>
</tr>
</tbody>
</table>

Fig. 4. Left: Sensogram signal 1. Ligand $L_1$ was injected until the signal (2.5) reached steady state at $t = 5000$. At this point we stopped injecting $L_1$, and injected $L_2$ until the signal reached steady state again at $t = 15000$. The time $t$ is measured in seconds. Right: corresponding bound ligand concentrations.

Fig. 5. Left: Sensogram signal 2, generated in the same manner as the reading in Fig. 4. Right: Corresponding bound ligand concentrations.
Relative difference between signal 1 and signal 2. Since the signals are identical during the first injection, we have depicted the relative difference during the second injection only; i.e., from $t = 5000$ to $t = 15000$.

Since the two parameter regimes identified in Table 1 lead to nearly identical sensogram signals, it may be difficult for standard algorithms to identify the true set of rate constants. Therefore, a set of four numerical experiments has been performed to identify the true set of rate constants. Since true experimental data is still forthcoming, the efficacy of this procedure is demonstrated by numerically simulating these four experiments with our mathematical model.

The first experiment consisted of injecting $L_1$ only, and was simulated by solving the ODE

$$\frac{dB_1}{dt} = j_{ka}(R_1 - B_1)C_1 - j_{kd}B_1.$$  

We used lsqcurvefit in MATLAB to fit the parameters $j_{ka}$ and $j_{kd}$ in (2.6) to data from the first numerical experiment. Then, an experiment which consisted of injecting of $L_2$ only was simulated in an analogous manner to the first experiment. The second experiment generated estimates of $j_{ka}$ and $j_{kd}$. Next, a third experiment was done by simulating the injection of $L_1$ until the signal reached a steady-state, and then halting injection of $L_1$ and immediately starting injection of $L_2$ until the signal reached a steady-state again (analogous to the simulations in Figs. 4 and 5). The third experiment generated estimates of $j_{ka}, \frac{1}{2}j_{ka}$, and $\frac{1}{2}j_{kd}$. A fourth experiment analogous to the third can be done to determine $\frac{3}{2}j_{kd}$ by switching the roles of $L_1$ and $L_2$ in experiment three. Due to the similarity of this experiment to experiment three, this computation has been omitted for simplicity.

Despite the closeness of sensogram signals 1 and 2, the column labeled “Fitted” in Table 1 reveals the results from our parameter estimation algorithm agree quite well with the original values recorded in the “Exact” column—i.e., those used in our numerical experiments depicted in Figs. 4 and 5. Our parameter estimation algorithm is robust with respect to noise; however, the results included herein were computed without the addition of noise.

Although the proposed parameter estimation procedure correctly identifies the set of rate constants corresponding to each of the signals, in practice it may be desirable to disambiguate a potentially unclear sensogram signal before parameter estimation. For the two signals considered herein, this can be done through postprocessing. To motivate how this is done, recall that the jump in signal one results from $L_2$ binding with available receptor sites. By excluding this from the signal (2.5), the resulting sensogram reading will not exhibit a sudden increase at the start of the second injection. Hence, the signal was postprocessed by subtracting the time-dependent quantity $s_2 \tilde{B}_2(t)$ from (2.5) for $t \geq 5000$ (i.e., during the second injection). The function $\tilde{B}_2$ satisfies

$$\frac{dB_2}{dt} = 2j_{ka}\left(\frac{1}{j_{ka}C_1 + j_{kd}} - \frac{1}{2}C_2 - 2j_{kd}B_2\right),$$  

$\tilde{B}_2(0) = 0.$

Note that in (2.7) the factor

$$R_1\frac{j_{ka}C_1}{j_{ka}C_1 + j_{kd}}$$

has been subtracted from the total number of receptors $R_1$. The factor (2.8) is simply the steady state of (2.6), and hence corresponds to the number of receptors filled by $B_1$ at the beginning of the second state. Hence the parenthetical quantity in (2.7) represents the receptors available at the start of the second injection.

This subtraction excludes $EL_2$ forming as a result of $L_2$ binding with empty receptors at the start of the second injection. The postprocessed signals are depicted in Fig. 7. By excluding the influx of $EL_2$ at the start of the second injection, signal one no longer exhibits an obvious increase at the start of the second injection. Conversely, even after postprocessing a sudden increase in sensogram two is still evident—this is due to the fact that the jump in signal two results from a large influx of $EL_1L_2$, not an influx $EL_2$.

The two signals in Figs. 4 and 5 can also be disambiguated by adjusting the ligand inflow concentrations. To demonstrate this, the same numerical experiments as in Figs. 4 and 5 were conducted, however the ligand concentrations were simultaneously reduced to $C_1 = 4.5 \times 10^{-12}$ mol/cm$^3$ and $C_2 = 4.5 \times 10^{-12}$ mol/cm$^3$. These values are 45 percent of those given in Table 1, and all of the other values remained the same. The resulting sensogram readings are depicted in Figs. 8 and 9.

First consider signal 1: notice that the increase seen at the start of the second injection is more pronounced than in signal 2. Since $C_1$ is less than half of its original value, there are fewer unbound $L_1$ molecules at the surface of the biosensor available for binding; in turn $B_1$ reaches a lower steady-state value. This implies that the quantity $R_1 - B_1$ is larger at $t = 5000$, and there are more empty receptor sites for $L_2$ molecules to bind with at the start of the second injection. The latter results in the pronounced jump, because the rate constant $j_{ka}$ is an order of magnitude larger than the other association constants and pathway one (1.2a) is prominent.

In sensogram two, there was no obvious increase at the start of the second injection. A comparison with Fig. 4 shows that this is because less $EL_1L_2$ forms. Since less $EL_1$ is present for $L_2$ molecules to bind with, the jump at the start of the second injection is not as pronounced. This is due to the fact the rate constant $\frac{1}{2}j_{ka}$ is an order of magnitude larger than the other association constants, so pathway two (1.2b) is prominent.

Conclusions

Scientists are currently attempting to determine the pathway of polymerase η during DNA translesion synthesis using optical biosensor experiments. The presence of multiple reacting species on the sensor surface complicates the associated parameter estimation problem. Using a mathematical model, two parameters regimes that lead to nearly identical sensogram signals have been identified. By proposing and numerically simulating a set of four experiments, it has been shown that it is possible to recover the true set of rate constants.

Furthermore, it has been shown that the signals may be
These results are motivated by the study of the role of polymerase \( \gamma \) during DNA translesion synthesis; however, they may be generic enough to be applied to other multiple-component biosensor experiments.

Disambiguated in two ways. An estimate for binding from one pathway can be subtracted from the data, revealing the binding from the other pathway. Alternatively, one may use differing values of the inflow concentrations to isolate binding in each pathway.

**Fig. 7.** The signals depicted in Figs. 4 and 5 after postprocessing. In both figures the postprocessed signal is depicted by the dotted line, and the original signal (shown in Figs. 4 and 5) is depicted by the dashed line.

**Fig. 8.** Left: Signal 1 after adjusting ligand inflow concentrations. Right: Corresponding bound state concentrations.

**Fig. 9.** Left: Signal 2 after adjusting ligand inflow concentrations. Right: Corresponding bound state concentrations.
Future work will include incorporating mass-transport effects into our parameter estimation algorithm. Moreover, we will seek to develop MATLAB software to fit the rate constants in (1.2) to true experimental data, and use this software to verify the post-processing strategy discussed herein.

Funding

The authors were supported by the National Science Foundation, under award number NSF-DMS 1312529. In addition, the second author was partially supported by the National Research Council, in the form of a postdoctoral fellowship.

Disclaimer

This paper is an official contribution of the National Institute of Standards and Technology and is not subject to copyright in the United States. Commercial products are identified in order to adequately specify certain procedures. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the identified products are necessarily the best available for the purpose.

References